Investigation of the phenolic and antioxidant content in Australian grains using traditional and non-invasive analytical techniques

by

Joel Benjamin Johnson

Thesis

Submitted in fulfillment of the requirements for the degree of

Master of Applied Science

Central Queensland University

School of Health, Medical and Applied Sciences

Supervisory panel: Dr Mani Naiker, Prof Kerry Walsh

July 2022

RHD Thesis Declaration

CANDIDATE'S STATEMENT

By submitting this thesis for formal examination at CQUniversity Australia, I declare that it meets all requirements as outlined in the Research Higher Degree Theses Policy and Procedure.

STATEMENT OF AUTHORSHIP AND ORIGINALITY

By submitting this thesis for formal examination at CQUniversity Australia, I declare that all of the research and discussion presented in this thesis is original work performed by the author. No content of this thesis has been submitted or considered either in whole or in part, at any tertiary institute or university for a degree or any other category of award. I also declare that any material presented in this thesis performed by another person or institute has been referenced and listed in the reference section.

COPYRIGHT STATEMENT

By submitting this thesis for formal examination at CQUniversity Australia, I acknowledge that thesis may be freely copied and distributed for private use and study; however, no part of this thesis or the information contained therein may be included in or referred to in any publication without prior written permission of the author and/or any reference fully acknowledged.

PREVIOUS SUBMISSION STATEMENT

This paper HAS NOT been submitted for an award by another research degree candidate (Co-Author), either at CQUniversity or elsewhere.

ACKNOWLEDGEMENT OF SUPPORT PROVIDED BY THE AUSTRALIAN GOVERNMENT

This RHD candidature was supported under the Commonwealth Government's Research Training Program/Research Training Scheme. I gratefully acknowledge the financial support provided by the Australian Government.

ACKNOWLEDGEMENT OF FINANCIAL SUPPORT

I gratefully acknowledge operational cost funding received from AgriVentis Technology Ltd. which has supported this research.

ACKNOWLEDGEMENT OF OTHER SUPPORT

This research was undertaken with in-kind support of grain samples provided by AgriVentis Technology Ltd., the Australian Export Grains Innovation Centre (AEGIC) and Agriculture Victoria Research.

Abstract

Recent years have seen the emergence of the concept of “functional foods”– where the value of food products is based on their health-benefiting properties in addition to their basic nutritional value. Globally, the functional food market is worth US $170 billion and is projected to grow at 7.5% p.a. over the next 10 years. In order to capitalise on this lucrative emerging market, producers and wholesalers need to demonstrate that their products contain high levels of these desirable compounds. This is typically assessed through time-consuming, expensive analytical techniques such as high-performance liquid chromatography (HPLC) or liquid chromatography-mass spectrometry (LC-MS). While these methods provide a high level of specificity and sensitivity, more rapid analytical techniques may be better suited to the routine, near-real-time analysis of large numbers of samples. Furthermore, there is currently a lack of basic context data on the typical levels of bioactive compounds that are found in many crops grown under Australian conditions, particularly for grain crops. This lack of context data makes it challenging to know whether a particular product would be considered high or low quality from a functional food perspective.

Consequently, the first major aim of this project was to profile the typical levels of bioactive compounds present in economically significant grain crops grown in Australia – specifically faba bean, wheat, mungbean and chickpea. The major focus was on phenolic compounds, as these possess high levels of antioxidant activity and are found in relatively high levels in grain crops. Furthermore, this class of compounds is associated with a wide range of health benefits, particularly for the prevention of cardiovascular disease.

Using spectrophotometric methods and HPLC analysis, moderate differences were found in the phenolic contents and antioxidant capacity of different varieties from each crop. This was particularly noted for the ten varieties of faba bean analysed, where there was a 121% difference in total phenolic content (TPC) between the varieties with the lowest and highest contents. This pulse also contained the highest total phenolic contents (258-571 mg GAE/100 g) and ferric reducing antioxidant potential (237-531 mg TE/100 g) of all crops investigated. The five mungbean varieties showed lower levels and more minor differences in phenolic content (79-105 mg GAE/100 g; 32% variation between varieties) and cupric reducing antioxidant capacity (498-584 mg TE/100 g; 17% variation), while the while the ferric reducing antioxidant potential did not differ significantly between varieties (14-20 mg TE/100 g). However, the content of numerous phenolic compounds (*p*-hydroxybenzoic acid, vanillic acid, caffeic acid, sinapic acid, *trans*-ferulic acid, cinnamic acid and vitexin) were significantly different between the mungbean varieties investigated. Similar observations were made for the chickpea samples, where there were moderate differences in total phenolic content (73-

94 mg GAE/100 g; 29% variation) and ferric reducing antioxidant potential (25-40 mg TE/100 g; 62% variation) between varieties. Again, the content of most phenolic acids analysed by HPLC were significantly different between varieties. Although varietal differences were not examined for wheat, the TPC of the 65 samples was higher than mungbean and chickpea (130-180 mg GAE/100 g), while the ferric reducing antioxidant potential ranged from 14-64 mg TE/100 g.

In addition to the varietal differences, the impact of growing location and season on phenolic content and antioxidant capacity were investigated in faba bean. Although these variables had no effect on the total phenolic content, the growing location did alter the levels of several individual phenolic compounds (protocatechuic, vanillic and chlorogenic acids, as well as the flavonoids vitexin and rutin).

The second major aim of this project was to investigate the prospect of infrared spectroscopy as a rapid technique for the prediction of phenolic content and antioxidant capacity in Australian grain crops. Promising results were found for the estimation of total phenolic content and antioxidant capacity in faba bean and wheat flour, particularly using near-infrared spectroscopy. The NIR model for TPC showed an R^2_{test} of 0.66 and RMSEP of 76 mg GAE/100 g when applied to faba bean, and an R^2_{test} of 0.86 and RMSEP of 4 mg GAE/100 g in wheat. However, infrared spectroscopy was unable to predict the concentrations of these analytes in mungbean or chickpea flour. This may be due to additional matrix constituents obscuring the analyte signals in the infrared region, or a consequence of the lower phenolic/antioxidant contents in these crops.

Nevertheless, the overall results suggest that infrared spectroscopy could be used for the estimation of total phenolic content or antioxidant capacity (i.e., prediction of high or low contents) in certain grain crops. This technique could potentially be applied for the routine screening of bioactive constituents, helping Australian producers to capitalise on the growing domestic and international functional food markets. Monitoring bioactive compound levels in Australian grain – either through traditional or non-invasive analytical techniques – could provide an additional level of quality assurance for producers of functional food crops and help maintain Australia's global recognition as a producer of high-quality food.

List of Main Abbreviations Used

Abbreviation	Definition
AACC	American Association of Cereal Chemists
AEGIC	Australian Export Grains and Innovation Centre
ANOVA	Analysis of variance
AOAC	Association of Official Analytical Chemists
ATR	Attenuated total reflectance
AVTMB	AgriVentis mungbean
CUPRAC	Cupric reducing antioxidant potential
%CV	Percent coefficient of variation
cyd-3-glu	Cyanidin-3-glucoside
DW/DWB	Dry weight / Dry weight basis
FRAP	Ferric reducing antioxidant potential
FTIR	Fourier transform infrared
FW	Fresh weight
G × E	Genotype × Environment
GAE	Gallic acid equivalents
GC-MS	Gas chromatography mass spectroscopy
HKW	Hundred kernel weight
HPLC	High-performance liquid chromatography
HPLC-DAD	High-performance liquid chromatography with diode array detection
IR	Infrared
LC-MS	Liquid chromatography mass spectrometry
LOO	Leave-one-out
m/z	Mass to charge ratio
MIRS	Mid-infrared spectroscopy
NIRS	Near-infrared spectroscopy
PBA	Pulse Breeding Australia
PC	Principal component

PCA	Principal component analysis
PLS-DA	Partial least squares discriminant analysis
PLSR	Partial least squares regression
rcf	Relative centrifugal force
RMSECV	Root mean square error of cross-validation
RMSEP	Root mean square error of prediction
RPD	Ratio of performance to deviation
SNV	Standard normal variate
TAC	Total antioxidant capacity
TE	Trolox equivalents
TMAC	Total monomeric anthocyanin content
TPC	Total phenolic content
TPTZ	2,4,6-tris(2-pyridyl)-s-triazine
Trolox	6-hydroxy-2,5,7,8-tetramethyl-chromane-2-carboxylic acid
UV	Ultraviolet

Acknowledgements

Firstly, I would like to thank my parents for their continued support while I completed this degree. I also thank my siblings for their understanding and encouragement.

Thanks to my primary supervisor, Dr Mani Naiker, for his consistent support over the last three years. It has been a long journey, but I have learned a lot along the way.

My associate supervisor, Prof Kerry Walsh, has been a great mentor and source of knowledge throughout this degree. I am also grateful to him for supporting my participation in the 4th Fundamentals and Applications of Near Infrared Spectroscopy (NIRS) course, which I completed between May-Oct 2020. Thanks to the University of Cordoba and the International Virtual Platform for Learning and Teaching Near Infrared Spectroscopy for hosting this course.

Thanks to the CQU technical staff for their assistance with running the instruments. In particular, Tania Collins has been a great help – from the first day of my Summer Research Scholarship through to now. Thanks also to Andrew Bryant for running the CN analyses.

I am also grateful for the support from my fellow students in the CQU “Bioactives” research team, particularly Janice Mani and Elena Hoyos.

A big shout out to Jens Altvater from Shimadzu for making the trip up to Rocky to provide the LC-MS training. I’m fairly sure that more than half of it went over my head, but I eventually managed to sort out the identification of isovitexin in the mungbean samples (Chapter 5).

Thanks to Daniel Skylas from the Australian Export Grains Innovation Centre for providing the faba bean samples and wheat samples which formed the basis of Chapters 3 and 4. I would also like to thank Lewis Hunter (from AgriVentis Technology Ltd.) and Dr Surya Bhattarai for providing access to the mungbean samples used in Chapter 5. Thanks to Prof Joe Panozzo and Dr Linda McDonald from Agriculture Victoria Research for supplying the chickpea samples used in Chapter 6.

Thanks to Dr Adam Rose and my fellow RHD students for their support and ideas through the Advanced Candidature Program, run by CQU.

Part of the introduction (Chapter 1) of this thesis is adapted from material submitted as coursework for the unit RSCH20001, which forms part of the candidature process at Central Queensland University.

Finally, I would like to thank the Lord “who has enabled me” to complete this work (1 Tim 1:12).

Publications related to this study

The five following publications by the candidate are relevant to the thesis.

PAPER 1

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION

Johnson, J.B., Collins, T., Skylas, D., Quail, K., Blanchard, C. and Naiker, M., 2020. Profiling the varietal antioxidative contents and macrochemical composition in Australian faba beans (*Vicia faba* L.). *Legume Science*, *2*(2), p.e28. DOI: 10.1002/leg3.28

Status: Published 13 Jan 2020
(18 citations)

NATURE OF CANDIDATE'S CONTRIBUTION, INCLUDING PERCENTAGE OF TOTAL

In conducting the study, I was responsible for reviewing the literature, conducting the experiments, analysing the data, interpreting results, and drafting and editing the paper. This publication was written by me and my contribution was 90%.

NATURE OF CO-AUTHORS' CONTRIBUTIONS, INCLUDING PERCENTAGE OF TOTAL

My co-authors, Tania Collins, Daniel Skylas, Ken Quail, Christopher Blanchard and Mani Naiker, contributed to the paper by planning the experiments and assisting with editing the manuscript. Their overall contribution was 10%.

This paper forms around one-third of the results presented in Chapter 3.

PAPER 2

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION

Johnson, J.B., Walsh, K. and Naiker, M., 2020. Application of infrared spectroscopy for the prediction of nutritional content and quality assessment of faba bean (*Vicia faba* L.). *Legume Science*, *2*(3), p.e40. DOI: 10.1002/leg3.40

Status: Published 24 Apr 2020
(1 citation)

NATURE OF CANDIDATE'S CONTRIBUTION, INCLUDING PERCENTAGE OF TOTAL

In conducting the study, I was responsible for planning the paper, reviewing and interpreting the literature, and drafting and editing the paper. This publication was written by me and my contribution was 90%.

NATURE OF CO-AUTHORS' CONTRIBUTIONS, INCLUDING PERCENTAGE OF TOTAL

My co-authors, Kerry Walsh and Mani Naiker, contributed to the paper by reviewing and editing the manuscript. Their overall contribution was 10%.

Parts of this paper are incorporated into the introduction and discussion of Chapter 3.

PAPER 3

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION

Johnson, J.B., Mani, J.S., Skylas, D., Walsh, K.B., Bhattarai, S.P. and Naiker, M., 2021. Phenolic profiles and nutritional quality of four new mungbean lines grown in northern Australia. *Legume Science*, *3*(2), p.e70. DOI: 10.1002/leg3.70

Status: Published 30 Dec 2020
(7 citations)

NATURE OF CANDIDATE'S CONTRIBUTION, INCLUDING PERCENTAGE OF TOTAL

In conducting the study, I was responsible for planning the paper, conducting the experiments, analysing the data, interpreting results, and drafting and editing the paper. This publication was written by me and my contribution was 85%.

NATURE OF CO-AUTHORS' CONTRIBUTIONS, INCLUDING PERCENTAGE OF TOTAL

My co-authors, Janice Mani, Daniel Skylas, Kerry Walsh, Surya Bhattarai and Mani Naiker, contributed to the paper by assisting with some experiments, and reviewing and editing the manuscript. Their overall contribution was 15%.

This paper forms around one-third of the results presented in Chapter 5.

PAPER 4

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION

Johnson, J.B., Walsh, K.B., Bhattarai, S.P. and Naiker, M., 2021. Partitioning of nutritional and bioactive compounds between the kernel, hull and husk of five new chickpea genotypes grown in Australia. *Future Foods*, *4*, p.100065. DOI: 10.1016/j.fufo.2021.100065

Status: Published 26 Jul 2021
(2 citations)

NATURE OF CANDIDATE'S CONTRIBUTION, INCLUDING PERCENTAGE OF TOTAL

In conducting the study, I was responsible for planning the paper, conducting the experiments, analysing the data, interpreting results, and drafting and editing the paper. This publication was written by me and my contribution was 85%.

NATURE OF CO-AUTHORS' CONTRIBUTIONS, INCLUDING PERCENTAGE OF TOTAL

My co-authors, Kerry Walsh, Surya Bhattarai and Mani Naiker, contributed to the paper by assisting with some experiments, and reviewing and editing the manuscript. Their overall contribution was 15%.

Information from this paper is incorporated into the introduction and discussion of Chapter 6. The methods developed for this paper were also used for the analysis of the chickpea samples in Chapter 6.

PAPER 5

DECLARATION OF CO-AUTHORSHIP AND CO-CONTRIBUTION

Johnson, J.B., Skylas, D.J., Mani, J.S., Xiang, J., Walsh, K.B. and Naiker, M., 2021. Phenolic Profiles of Ten Australian Faba Bean Varieties. *Molecules*, *26*(15), p.4642. DOI: 10.3390/molecules26154642

Status: Published 31 July 2021
(Q1 journal; IF 4.412; 1 citation)

NATURE OF CANDIDATE'S CONTRIBUTION, INCLUDING PERCENTAGE OF TOTAL

In conducting the study, I was responsible for planning the paper, conducting the experiments, analysing the data, interpreting results, and drafting and editing the paper. This publication was written by me and my contribution was 90%.

NATURE OF CO-AUTHORS' CONTRIBUTIONS, INCLUDING PERCENTAGE OF TOTAL

My co-authors, Daniel Skylas, Janice Mani, Jinle Xiang, Kerry Walsh and Mani Naiker, contributed to the paper by reviewing and editing the manuscript. Their overall contribution was 10%.

This paper forms around one-third of the results presented in Chapter 3.

Additional Publications by the Candidate Relevant to the Thesis but not Forming Part of it

The following related publications were also written by the candidate, but do not form part of the thesis.

Johnson, J., Collins, T., Power, A., Chandra, S., Skylas, D., Portman, D., Panozzo, J., Blanchard, C. and Naiker, M., 2020. Antioxidative properties and macrochemical composition of five commercial mungbean varieties in Australia. *Legume Science*, *2*(1), p.e27. DOI: 10.1002/leg3.27
(19 citations)
My contribution: 90%

Johnson, J., Mani, J., Ashwath, N. and Naiker, M., 2020. Potential for Fourier transform infrared (FTIR) spectroscopy toward predicting antioxidant and phenolic contents in powdered plant matrices. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *233*, p.118228. DOI: 10.1016/j.saa.2020.118228
(Q2 journal; IF 4.098; 19 citations)
My contribution: 85%

Mani, J.S., **Johnson, J.B.***, Steel, J.C., Broszczak, D.A., Neilsen, P.M., Walsh, K.B. and Naiker, M., 2020. Natural product-derived phytochemicals as potential agents against coronaviruses: A review. *Virus Research*, *284*, p.197989. DOI: 10.1016/j.virusres.2020.197989
*equal first authorship
(Q2 journal; IF 3.303; 254 citations)
My contribution: 40%

Johnson, J., Collins, T., Walsh, K. and Naiker, M., 2020. Solvent extractions and spectrophotometric protocols for measuring the total anthocyanin, phenols and antioxidant content in plums. *Chemical Papers*, *74*(12), pp.4481-4492. DOI: 10.1007/s11696-020-01261-8
(Q2 journal; IF 2.097; 15 citations)
My contribution: 80%

Naiker, M., Anderson, S., **Johnson, J.B.**, Mani, J.S., Wakeling, L. and Bowry, V., 2020. Loss of *trans*-resveratrol during storage and ageing of red wines. *Australian Journal of Grape and Wine Research*, *26*(4), pp.385-387. DOI: 10.1111/AJGW.12449
(Q1 journal; IF 2.688; 1 citation)
My contribution: 35%

Kiran, S., **Johnson, J.B.**, Mani, J.S., Portman, A., Mizzi, T. and Naiker, M., 2021. Commercial lentils (*Lens culinaris*) provide antioxidative and broad-spectrum anti-cancerous effects. *Legume Research - An International Journal*, *44*, pp.202-206. DOI: 10.18805/LR-557
(Q2 journal; IF 0.63; 2 citations)
My contribution: 35%

Johnson, J.B., Collins, T., Mani, J.S. and Naiker, M., 2021. Nutritional quality and bioactive constituents of six Australian plum varieties. *International Journal of Fruit Science*, *21*(1), pp.115-132. DOI: 10.1080/15538362.2020.1860863
(Q3 journal; IF 1.359; 5 citations)
My contribution: 80%

Johnson, J.B., Broszczak, D.A., Mani, J.S., Anesi, J. and Naiker, M., 2021. A cut above the rest: Oxidative stress in chronic wounds and the potential role of polyphenols as therapeutics. *Journal of Pharmacy and Pharmacology*. DOI: 10.1093/jpp/rgab038
(Q1 journal; IF 3.765; 3 citations)
My contribution: 70%

Johnson, J.B., Mani, J.S. and Naiker, M., 2021. Infrared Spectroscopy for the Quality Assessment of Habanero Chilli: A Proof-of-Concept Study. *Engineering Proceedings*, *8*(1), p.19. DOI: 10.3390/engproc2021008019
My contribution: 90%

Johnson, J.B., Mani, J.S., Broszczak, D., Prasad, S.S., Ekanayake, C.P., Strappe, P., Valeris, P. and Naiker, M., 2021. Hitting the sweet spot: A systematic review of the bioactivity and health benefits of phenolic glycosides from medicinally used plants. *Phytotherapy Research*. *35*(7), pp. 3484-3508. DOI: 10.1002/PTR.7042
(Q2 journal; IF 5.882; 3 citations)
My contribution: 65%

Mangwanda, T., **Johnson, J.B.**, Mani, J.S., Jackson, S., Chandra, S., McKeown, T., White, S. and Naiker, M., 2021. Processes, challenges and optimisation of rum production from molasses—a contemporary review. *Fermentation*, *7*(1), p.21. DOI: 10.3390/fermentation7010021
(Q1 journal; IF 4.197; 5 citations)
My contribution: 40%

Johnson, J.B., Budd, C., Mani, J.S., Brown, P., Walsh, K.B. and Naiker, M., 2022. Carotenoids, ascorbic acid and total phenolic content in the root tissue from five Australian-grown sweet potato cultivars. *New Zealand Journal of Crop and Horticultural Science*, *50*(1),

pp.32-47. DOI: 10.1080/01140671.2021.1895230
(Q2 journal; IF 1.20; 1 citation)
My contribution: 75%

Mani, J.S., **Johnson, J.B.**, Hosking, H., Ashwath, N., Walsh, K.B., Neilsen, P.M., Broszczak, D.A. and Naiker, M., 2021. Antioxidative and therapeutic potential of selected Australian plants: A review. *Journal of Ethnopharmacology*, *268*, p.113580.
(Q2 journal; IF 4.36; 8 citations)
My contribution: 15%

Wilson, A., **Johnson, J.B.**, Batley, R., Lal, P., Wakeling, L. and Naiker, M., 2021. Authentication using volatile composition: A proof-of-concept study on the volatile profiles of fourteen Queensland ciders. *Beverages*, *7*(2), p.28. DOI: 10.3390/beverages7020028
(2 citations)
My contribution: 30%

Johnson, J.B., Ekanayake, C.P., Caravani, F., Mani, J.S., Lal, P., Calgaro, S.J., Prasad, S.S., Warner, R.D. and Naiker, M., 2021. A Review of Vitamin D and Its Precursors in Plants and Their Translation to Active Metabolites in Meat. *Food Reviews International*, pp.1-29. DOI: 10.1080/87559129.2021.1936006
(Q1 journal; IF 6.478)
My contribution: 70%

Johnson, J.B., Mani, J.S., White, S., Brown, P. and Naiker, M., 2021. Pungent and volatile constituents of dried Australian ginger. *Current Research in Food Science*, *4*, pp.612-618.
My contribution: 90%

Johnson, J., Mani, J. and Naiker, M., 2021. Development and Validation of a 96-Well Microplate Assay for the Measurement of Total Phenolic Content in Ginger Extracts. *Food Analytical Methods*. DOI: 10.1007/s12161-021-02127-9
(Q1 journal; IF 3.16)
My contribution: 90%

Kiran, S., Chew, G.S., **Johnson, J.B.**, Mani, J.S., Wakeling, L., Portman, A. and Naiker, M., 2021. Australian Ethnomedicinal Plant Extracts Promote Apoptosis-Mediated Cell Death in Human Hepatocellular Carcinoma in vitro. *Pharmacognosy Communications*, *11*(4), pp.210-213.
My contribution: 35%

Johnson, J.B., Steicke, M., Mani, J.S., Rao, S., Anderson, S., Wakeling, L. and Naiker, M., 2021. Changes in Anthocyanin and Antioxidant Contents during Maturation of Australian

Highbush Blueberry (*Vaccinium corymbosum* L.) Cultivars. *Engineering Proceedings*, *11*(1), p.6. DOI:
(1 citation)
My contribution: 50%

Mani, J.S., **Johnson, J.B.** and Naiker, M., 2021. The Phytochemistry and Anticarcinogenic Activity of Noni Juice. *Engineering Proceedings*, *11*(1), p.16. DOI: 10.3390/ASEC2021-11154
My contribution: 15%

Johnson, J.B., Mani, J.S., White, S., Brown, P. and Naiker, M., 2021. Quantitative profiling of gingerol and its derivatives in Australian ginger. *Journal of Food Composition and Analysis*, *104*, p.104190.
(Q1 journal; IF 4.556)
My contribution: 90%

Batley, R., **Johnson, J.**, Mani, J., Broszczak, D. and Naiker, M., Finding alternative uses for Australian rosella (Hibiscus sabdariffa) by-products: Nutritional potential and in vitro digestibility studies. *Animal Production Science*, early online.
(Q2 journal; IF 1.53)
My contribution: 35%

Johnson, J.B., Ohri, B., Walsh, K.B. and Naiker, M., 2022. A simple isocratic HPLC–UV method for the simultaneous determination of citrulline and arginine in Australian cucurbits and other fruits. *Food Analytical Methods*, *15*(1), pp.104-114. DOI: 10.1007/s12161-021-02110-4
(Q1 journal; IF 3.16; 1 citation)
My contribution: 90%

Johnson, J.B., Batley, R., Manson, D., White, S. and Naiker, M., 2022. Volatile compounds, phenolic acid profiles and phytochemical content of five Australian finger lime (*Citrus australasica*) cultivars. *LWT - Food Science and Technology*, *154*, p.112640.
(Q1 journal; IF 4.952)
My contribution: 85%

Intellectual Property Statement

The Intellectual Property Rights generated from the AgriVentis mungbean samples (Chapter 5) are vested in AgriVentis. In exchange, Agriventis has granted the author a perpetual, irrevocable, non-exclusive, worldwide and royalty free licence (including the right to sublicense) to use such New IP for any purpose. Additionally, all copyright of the material in this thesis remains the property of the author.

All of the other work described in this thesis remains the Intellectual Property of the author, as described in Section 3.6 of the CQUniversity “Intellectual Property and Moral Rights Policy”.

Contents

Table of tables

Table of figures

Chapter 1. Introduction

1.1 Background and problem statement

Australia is an agriculturally-driven nation (Henzell, 2007). The gross value of the Australian agricultural sector was AU $61 billion in 2019-20, comprising approximately 4% of GDP and around 14% of export income (Australian Bureau of Statistics, 2021). Within the cropping sector, broadacre cropping – including wheat, barley and pulses – is responsible for over AU $13 billion value (around 46% of the total value of the cropping sector). Approximately two-thirds of Australian agricultural produce by value is exported (Gunasekera et al., 2008), with Australia being the largest global exporter of certain pulse crops such as faba bean and chickpea (AEGIC, 2017). It is also a major exporter of other pulses and cereal crops, including wheat. In general, Australia's export produce is internationally recognised to be of a high standard, thanks to the stringent quality assurance measures implemented on-farm, during post-harvest processing, and prior to export, through the assistance of facilities such as the Australian Export Grains and Innovation Centre (AEGIC). With significant worldwide population growth and an estimated increase of 50-100% demand for food between 2015 and 2050 (McKenzie & Williams, 2015), there is a significant opportunity for the Australian agricultural industry to increase production levels in view of increasing their export potential in the global market share (Adamson, 2013). In particular, there is considerable interest in expanding the cropping sector in northern Australia (i.e., above the Tropic of Capricorn) (Ash et al., 2017; Chauhan & Williams, 2018; CRCNA, 2020; Petheram et al., 2018).

Additionally, Australian producers have the opportunity to add value to their commodity based products by capitalising on the expanding "functional food" market – where foods are purchased for their health-benefiting effects, rather than as a source of basic nutrition and energy (Granato et al., 2017; Urala & Lähteenmäki, 2007). For example, consumption of juice from the Queen Garnet plums has been shown to reduce oxidative stress (Netzel et al., 2012) and reduce the risk of blood clot formation in clinical trials (Santhakumar et al., 2015). Similarly, *in vitro* studies have suggested that polyphenolics isolated from chickpeas could protection against colorectal cancer (Bochenek et al., 2019). If high levels of such health-benefiting compounds can be demonstrated in a particular crop, consumers may pay a price premium for these products, particularly if the buyer is familiar with the concepts of functional foods (Di Pasquale et al., 2011). Furthermore, crops with high levels of bioactive compounds have potential for the development of value-added foods and ingredients (López-Barrios et al., 2014; Vioque et al., 2012; Zhang et al., 2012), which can also be marketed on the basis of their levels of health-benefiting compounds. Examples of foods experiencing a considerable rise in popularity due to their reported health benefits include the "ancient grains" (e.g., chia,

quinoa, millet and spelt), pulse crops (including mungbean, chickpeas, faba beans and lentils), and numerous other crops.

However, maintaining product quality through stringent quality assurance protocols is paramount in order to capitalise on this demand for functional foods and maintain a strong position in the global export market. The sale of crops as functional foods necessitates routine monitoring and quantification of the levels of bioactive compounds in produce from various growing locations and seasons. This ensures that consumers are receiving the premium product that they are paying for from a functional food-based quality perspective. However, baseline information on the typical levels of these compounds and their variation between different genotypes and environmental conditions is lacking from an Australian perspective.

In addition, the standard analytical techniques used to measure such compounds are hampered by three factors: their high running costs, the time consumption, and the amount of environmentally detrimental waste they produce (Wang et al., 2000). Due to the high levels of precision and accuracy that can be obtained, high-end instrumentation methods such as high-performance liquid chromatography (HPLC), liquid chromatography-mass spectrometry (LC-MS) and gas chromatography-mass spectrometry (GC-MS) remain the gold standard for the measurement of a vast range of bioactive analytes. Nevertheless, more rapid and non-invasive analytical techniques such as infrared spectroscopy are increasingly being used for food quality assessment (Cortés et al., 2019; Cozzolino et al., 2014; Gordon et al., 2018; Huang et al., 2008; Lu & Rasco, 2012; Pandiselvam et al., 2021). Infrared spectroscopy uses electromagnetic wavelengths with a frequency lower than that of visible light in order to characterise the functional groups present in a given matrix, allowing for the tentative characterisation and quantification of the chemical constituents present.

The bioactive compounds responsible for the health-benefiting effects of functional food crops can all be classified as phytochemicals – i.e., compounds that are produced by plants (Huang et al., 2016). The major classes of phytochemicals in food products include polyphenols, carotenoids, alkaloids, glucosinolates, polyacetylenes, polysaccharides, allium compounds, lectins, terpenes, capsaicinoids and betalains (Campos-Vega & Oomah, 2013). Of these, polyphenols have attracted the greatest interest in recent years, as they are widely present in most functional foods and display well-known health-benefiting effects (Shahidi & Ambigaipalan, 2015). Consequently, the non-invasive measurement of polyphenols has attracted attention across a number of crop types, predominantly in fruit and fruit products (Aleixandre-Tudo et al., 2018; Cobaleda-Velasco et al., 2018; Fragoso et al., 2011a; Fragoso et al., 2011b; Queji et al., 2010). However, there have been limited studies performed on grain and pulse crops using infrared spectroscopy for this purpose.

With increasing demand for food production at a lower cost, the development of such rapid, non-invasive analytical techniques for the assessment of crop quality is essential to creating a sustainable cropping sector. Consequently, this work aimed to investigate the use of infrared spectroscopy for the prediction of bioactive contents in economically significant grain crops.

1.1.1 Research design and methodology

The study followed a quantitative research design. A relatively large number of samples (>60) were analysed for each matrix type so as to allow the creation of accurate regression models. Although more samples may be required for the creation of highly robust prediction models, this sample size was considered sufficiently large for the proof-of-concept work performed in this study.

Where possible, the grain samples were chosen to encompass a wide range of phytochemical contents, including selecting samples from different varieties (for all crops) and growing locations (in the case of all crops except mungbean). In addition to this, the faba bean and chickpea samples included samples from different growing seasons.

1.2 Aims and scope of the research

The research aim of this project was to explore the typical levels of bioactive compounds of selected grain crops grown in Australia, namely faba bean, mungbean, chickpea and wheat. A particular focus was on the comparative levels of bioactive compounds found in different commercial varieties of each crop. This was conducted through the use of traditional, destructive analytical techniques, as well as rapid analytical techniques (near-infrared and mid-infrared spectroscopy) which could be applied to the ground grain samples.

The following research questions were addressed in this project:

- What are the typical levels of bioactive compounds present in the selected crops when grown under typical Australian conditions? What level of variation is present due to environmental or genotypic factors?
- Can infrared spectroscopy techniques be used to predict the levels of bioactive compounds, such as polyphenolics and anthocyanins, across the food types included in this study?
- What are the optimal data pre-processing techniques for the non-invasive prediction of bioactive compounds? Does this vary with the compound class or crop types?
- What other information can be obtained through infrared spectroscopy?

In order to address these questions, they were broken down into specific aims, as outlined in the following sections.

1.2.1 Aim one

The first aim of this research project was to characterise the typical levels of phenolic and antioxidant compounds found in Australian-grown faba bean, wheat, mungbean and chickpea. This was conducted using several traditional analytical techniques, with varying degrees of sensitivity and accuracy:

1. Spectrophotometric-based wet chemical methods (benchtop methods). This included the TPC, FRAP, CUPRAC and TMAC assays.
2. Profiling of individual phenolic compounds using high-performance liquid chromatography with diode array detection (HPLC-DAD).
3. Confirmation of the identity of one selected phenolic compound (isovitexin) using liquid chromatography-mass spectrometry (LC-MS).

1.2.2 Aim two

The second aim was to investigate the extent of variation in the aforementioned bioactive compounds found in different varieties of each crop. Included in this was the assessment of the impact of growing location and season for most crop types.

Specifically, the objectives were:

1. To determine the extent of variation in bioactive compounds found in different varieties of faba bean, mungbean and chickpea.
2. To determine the impact of growing location on the level of bioactive compounds in faba bean and chickpea.
3. To determine the impact of growing season on the level of bioactive compounds in faba bean and chickpea.

1.2.3 Aim three

The third aim was to investigate the correlations between the broad classes of bioactive analytes (TPC, FRAP, CUPRAC, TMAC), individual phenolic acids, and their relationship with other nutritive-related parameters (e.g., protein content) in each crop.

1.2.4 Aim four

The fourth and final aim was to investigate the prospect of using infrared spectroscopy (including both near-infrared and mid-infrared spectroscopy) as a rapid, non-destructive means of predicting the level of bioactive compounds in the crops investigated.

Prediction of the following bioactive compound classes was attempted:

1. Total phenolic content (TPC)
2. Antioxidant capacity (FRAP and/or CUPRAC)
3. Total monomeric anthocyanin content (TMAC)

1.3 Limitations to this research

As with any research project, the work presented here has its limitations.

Firstly, the sample size was somewhat limited due to time constraints and the challenges of obtaining grain samples from different (known) genotypes, origins and growing seasons. Larger sample sizes generally increase the prediction accuracy and robustness of infrared spectroscopy models; consequently, the sample size is one factor that may impact on the robustness of the models when applied to new samples or populations.

Additionally, not all of the samples could be sourced from controlled field trials of different varieties grown across different locations and seasons, due to the time constraints involved in the sample procurement phase of this work. Although the samples were chosen to have the widest range of phytochemical compositions feasible, the performance of some models may be improved by including a wider range of analyte concentrations.

In a similar vein, it was not feasible to sample every commercial Australian variety of each crop for the phytochemical analysis. Consequently, this work cannot be seen as an exhaustive comparison of the phytochemical composition of Australian faba bean, wheat, mungbean and chickpea genotypes. However, a sufficiently large number of samples were analysed for most crop types to provide reasonable insight into the typical phenolic contents and antioxidant capacity of grain crops grown under Australian conditions.

Another factor influencing the performance of non-invasive prediction models is the accuracy of the reference methods (Sørensen, 2002). In this case, this is the accuracy of the benchtop spectrophotometric methods used for the measurement of total phenolic content and antioxidant capacity. While they have reasonable accuracy, the coefficient of variation between replicate extracts of the same sample is typically in the order of ~5%. Most of this variation would be attributable to sampling error, while a small amount would come from error associated with the spectrophotometric methods.

Finally, regression modelling was only conducted using partial least square regression in this study. It is possible that the use of emerging non-linear regression techniques, such as neural networks or support vector regression, could provide more accurate prediction models (Ni et al., 2014). Future work could investigate these chemometric techniques for the prediction of

bioactive compounds in grain matrices and compare their performance to traditional modelling methods.

1.4 Structure of the thesis

This thesis begins with a literature review on the use of infrared spectroscopy for the prediction of bioactive compounds in other food matrices (**Chapter 2**). This provides important background information on infrared spectroscopy as an analytical technique, and on its reported use for the prediction of bioactive analytes in other crops.

Chapter 3 provides the results of a detailed investigation into the phenolic composition of faba bean. Australia is the leading exporter of this crop globally, making it of considerable economic significance. The results presented include an examination of the effects of variety, growing season and location on the levels of bioactive compounds in faba bean. Additionally, the prediction of these bioactive compounds was attempted using IR spectroscopy.

Following on from the promising results found in faba bean, the use of IR spectroscopy for the prediction of phenolic content and antioxidant capacity was investigated in a crop containing lower levels of these compounds – wheat (**Chapter 4**).

The next crop investigated was mungbean (**Chapter 5**), a pulse crop of increasing importance in Australia. Although all the samples were sourced from one location, the variation in bioactive compounds between different varieties was investigated. Infrared spectroscopy was also trialled for the prediction of these analytes.

The fourth and final grain crop investigated in this work was chickpea (**Chapter 6**), another crop for which Australia is the leading exporter worldwide. The differences in bioactive composition were examined for different varieties, growing locations and years. Again, the prediction of these analytes was attempted using IR spectroscopy.

The thesis finishes with a general discussion and recommendations (**Chapter 7**) and final conclusions (**Chapter 8**).

Chapter 2. Literature review

2.1 Overview

Infrared spectroscopy (wavelengths between 750-25,000 nm) offers a rapid means of assessing the chemical composition of a wide range of sample types, both for qualitative and quantitative work. Its use in the food industry has been increasing significantly over the past five decades and it is now an accepted analytical technique for the routine analysis of certain proximate analytes. Furthermore, it is commonly used for routine screening and quality control purposes in numerous industry settings, albeit not typically for the analysis of bioactive compounds. Using the Scopus database, a systematic search of literature of the five years between 2016-2020 identified 45 studies using near-infrared spectroscopy for the quantification of bioactive compounds in food products, and 17 studies using mid-infrared spectroscopy for this purpose. The most common bioactive compounds assessed were polyphenols, anthocyanins, carotenoids and ascorbic acid. Numerous factors affect the accuracy of the developed model, including the analyte class and concentration, matrix type, instrument geometry, wavelength selection and spectral processing/pre-processing methods. Additionally, very few of the studies were validated on independently sourced samples, thus are likely to over-estimate the accuracy of this method. Nevertheless, the results demonstrate some promise of infrared spectroscopy for the rapid estimation of a wide range of bioactive compounds in food matrices.

2.2 Infrared spectroscopy

2.2.1 Infrared light regions

Infrared (IR) spectroscopy is a well-established tool in analytical chemistry, offering a non-invasive, non-destructive and rapid means of assessing the chemical composition of a wide range of sample types. It operates on the principle that dipole-active covalent bonds can absorb electromagnetic radiation, causing excitation of the bond from a lower to a higher vibrational (or rotational) energy level. For this reason, IR spectroscopy can also be termed vibrational spectroscopy. Covalent bonds can absorb electromagnetic wavelengths with a lower frequency than that of visible light (i.e. infrared wavelengths). The relationship between the wavelength and frequency of electromagnetic waves is represented using the following equation:

$$\upsilon = c/\lambda \qquad \text{Equation 2.1}$$

where υ = frequency; c = speed of light in a vacuum; λ = wavelength

This means that infrared light has a longer wavelength (750-25,000 nm) compared to visible light (380-750 nm). Furthermore, the lower frequency of infrared light means it carries less energy compared to visible light, as can be ascertained from the Planck-Einstein relation:

$E = h\upsilon$ Equation 2.2

where E = energy; h = the Planck constant; υ = frequency

For the purposes of analytical spectroscopy, the infrared spectrum can be divided into three main regions: near-infrared (NIR), mid-infrared (MIR) and far-infrared, with the NIR region being closest to the visible light spectrum. The NIR region spans wavelengths from 750 to 2500 nm (Pasquini, 2003). Historically, NIRS has been and continues to be utilised more than mid-infrared spectroscopy in the food industry due to its lower cost and greater penetrative power (i.e., lower absorption by the sample). This allows for more representative sampling of more of the sample (Almeida et al., 2006) and reduces sample preparation requirements (Burks et al., 2000). Wavelengths in the NIR region result from the overtone and combination bands of IR-active bonds, rather than their fundamental tones; hence show weaker signals compared to those in the MIR region.

The MIR region spans the wavelengths of 2,500 to 25,000 nm and contains the fundamental absorptions of IR-active bonds. For ease of interpretation, these values are more commonly presented in units of "wavenumbers" (cm^{-1}). The wavelengths in the MIR region (2,500 to 25,000 nm) correspond to the wavenumbers 4000 and 400 cm^{-1} (Pasquini, 2003).

Finally, the far infrared region (wavenumbers of 400-10 cm^{-1}) is rarely used except in astronomy, although there are occasional applications that have been reported in the food analysis sector (Bershtein & Ryzhov, 1994; Han et al., 2018; Susi & Ard, 1973). Consequently, the remainder of this review focuses on applications of NIR and MIR spectroscopy.

2.2.2 Introduction to infrared spectroscopy

To collect an infrared spectrum from a sample, the instrument (spectrometer) emits the full spectrum of infrared wavelengths into the sample, where the chemical bonds present absorb specific wavelengths of infrared light. The degree of absorption (amplitude of the signal) is proportional to the number of corresponding bonds that interact with and absorb the infrared light, while the frequency of the peak is characteristic of the bond type. Hence from the remaining light that is reflected from or transmitted through the sample, an analyst can deduce the identity and quantity of the compounds present.

Compared to other analytical methods, the main advantages of infrared spectroscopy are its speed, relatively low price of the instrument, and the fact that it is typically non-destructive and non-invasive, lowering or eliminating sample preparation time (Bureau et al., 2019; Walsh et

al., 2020). Furthermore, IR spectroscopy is highly sensitive, requires small sample sizes and can analyse samples from a wide variety of matrix types, including solids, powders, films, gels, liquids and gases (Bureau et al., 2019). This technique does not produce any waste products, in contrast to separative methods such as liquid chromatography (Johnson & Naiker, 2019). Conversely, the disadvantages of this method include the challenge of interpreting spectra from complex mixtures containing a number of different compounds, and the need to develop and maintain robust calibration models for the quantitative analysis of analytes (Bureau et al., 2019). As infrared spectroscopy can only detect compounds that contain IR-active bonds, it cannot be used for the direct analysis of monoatomic species, ions or elements.

In addition, infrared spectroscopy – particularly NIRS – is best suited for the analysis of macroconstituents (usually those present at concentrations of ~0.5% or higher). Below this concentration range, it is difficult to separate out the signal of the analyte from the rest of the spectral peaks. In many instances where previous researchers have reported detecting analytes at much lower concentrations, it is likely that NIRS is actually detecting a different analyte present at macro-levels – the concentration of which is correlated with the targeted analyte. This is known as a surrogate or secondary correlation (Walsh et al., 2020). In some cases, this correlation may be unavoidable due to both analytes absorbing in similar regions (Velasco et al., 1998). In other situations, it may be the only way through which IR spectroscopy can be used to estimate the microconstituent concentration. The use of such secondary correlations is acceptable in many cases – as long as the correlation holds true for all samples analysed. Some publications have reported that these correlations may change between different sample populations or harvest years (Velasco et al., 1998), which may explain the poor performance of independent test sets found in some studies analysing microconstituents with IR spectroscopy.

Despite these limitations, the speed and cost-effectiveness of IR spectroscopy have led to its adoption across many sectors of the food industry. This review focuses on the application of IR spectroscopy (both MIR and NIR) for the quantitative assessment of bioactive compounds in foodstuffs. It concludes with a contemporary perspective on the future of IR spectroscopy for the analysis of bioactive compounds in the food industry and highlights key areas where further research is required.

2.2.3 A brief history of infrared spectroscopy

Although infrared light was discovered by Wilhelm Herschel in 1800 (Herschel, 1800), the development of infrared spectroscopy for analytical chemistry really began in the late 1800s and early 1900s (Thomas, 1991). Beginning in 1903, William Coblentz published three volumes of infrared spectra for hundreds of different compounds (Barnes & Bonner, 1936), a

remarkable achievement considering that it took 3-4 hours to record a single spectrum at this time (Thomas, 1991). During the next few decades, the sensitivity and accuracy of the instruments continued to be refined, although applications remained limited. The first company to manufacture infrared spectrometers (Perkin-Elmer) was set up during this period, in 1937 (Thomas, 1991). By the 1940s and 50s, dispersive IR spectrophotometers began to be more readily available and affordable, and thus were adopted by many analytical laboratories worldwide for the identification of purified organic compounds (Bureau et al., 2019).

The next step along the journey was the analysis of complex matrices containing numerous different chemical constituents, such as food products. The first use of IR spectroscopy in the food industry is not well documented; however, reports of the structural assessment of individual carbohydrate sugars in aqueous solutions date back to the 1960s (Bureau et al., 2019). The first applications of IR spectroscopy for food analysis were also in this decade (Capuano & van Ruth, 2016).

Historically, infrared spectroscopy chemists believed that it was impossible to separate out functional chemical information from the overlapped overtone peaks of the NIR spectrum; hence focused their efforts on MIR spectroscopy. However, pioneering work by Karl Norris in the 1960s for the grain industry demonstrated that useful chemical information could be teased out of NIR spectra from "real" food samples through the use of chemometric techniques (Norris, 1965; Williams & Norris, 1987). By the late 1980s, NIR spectroscopy had been used for the analysis of intact vegetables (Renfroe & Kays, 1985) and fruits (Dull et al., 1989). The ongoing progress of this analytical technique was assisted by several technological advances in the instruments available, including the development of Fourier transform instruments in the 1960s, fiber optics in the 1980s and diode array spectrophotometers late in the 20th century (Jha, 2010; McClure, 2003).

Currently, infrared spectroscopy (principally NIR spectroscopy) is used in the quality analysis of a wide variety of food products, including fruit and vegetables (Bureau et al., 2019; Walsh et al., 2020), meat (Weeranantanaphan et al., 2011) and fish (Liu et al., 2013; Power & Cozzolino, 2020), grains and cereal products (Caporaso et al., 2018; Cozzolino, 2014b; Johnson, 2020), honey (Cozzolino et al., 2011), dairy products (De Marchi et al., 2018) and beverages (Huang et al., 2008).

2.2.4 Principles of infrared spectroscopy

2.2.4.1 Back to basics: the atom

In order to understand the principles of infrared spectroscopy, it is first necessary to begin with a solid theoretical understanding of atoms and chemical bonds. Our contemporary

understanding of the atom is that it comprises a nucleus of positively charged protons and neutrally charged neutrons, surrounded by a cloud of orbiting negatively charged electrons. Although the exact location of the electron is unknown at any given time, the Schrödinger wave equation can be used to describe the area in which there is a ≥95% probability of finding the electron.

When a single covalent bond is formed between two atoms, overlap or "sharing" of the electron orbits allows a pair of electrons to be shared between the two nuclei. Similarly, a double bond results from the sharing of four electrons between two nuclei. The electrostatic attraction between the bonded electrons and the nuclei prevents the atoms from drifting apart. Conversely, electrostatic repulsion between the two positively charged nuclei prevent them from coming too close together. Consequently, the bond length (the distance between the two nuclei) is not fixed. Rather, bonds act like springs between the atoms they join, vibrating at specific frequencies depending on the atomic masses of the atoms involved and the "stiffness" of the bond itself. If the vibration of the bond changes the dipole moment of the molecule (the location of partial charge(s) within the molecule), then the bond is considered to be infrared-active. Consequently, infrared photons of the corresponding frequency are able to be absorbed by the bond and excite it to a higher vibrational level.

The fundamental vibration wavelength can be calculated using Hooke's law for a simple harmonic oscillator:

$$\bar{\nu} = \frac{1}{2\pi c}\sqrt{\frac{K}{\mu}} \qquad \text{Equation 2.3}$$

where $\bar{\nu}$ = wavenumber (in cm^{-1}), c = speed of light (in cm s^{-1}), K = force constant (in dynes cm^{-1}) and μ = reduced mass of the bond in g $atom^{-1}$ (calculated as shown in Equation 2.4)

$$\mu = \frac{m_1 \times m_2}{m_1 + \boldsymbol{m_2}} \qquad \text{Equation 2.4}$$

Based on this formula, it can be seen that stronger bonds (e.g. alkenes) will generally vibrate faster than weaker bonds (e.g. alkanes), and bonds involving lighter atoms (e.g. C-H) will vibrate faster compared to those formed from heavier atoms (e.g. C-O).

If the bond vibration was purely harmonic, then there would be only one vibrational frequency (the fundamental) and no overtones. However, bonds are more easily stretched than squeezed, thus act as anharmonic oscillators (see Figure 2-1). Hence for each IR-active bond there are a number of unequally spaced vibrational levels (overtones), each consecutively reducing in intensity. Higher frequency photons (i.e., lower wavelengths) are required to excite the bond to higher overtones. The vibrational energy of the bond for each overtone (V = 0, 1, 2…) can be calculated as shown in Equation 2.5. Hydrogen-bearing groups (e.g., C-H, O-H,

N-H) are the most anharmonic, meaning that these groups are most readily detected in the IR region.

$$E_v = \left(V + \tfrac{1}{2}\right) h\bar{\omega}_e - \left(V + \tfrac{1}{2}\right)^2 h\bar{\omega}_e\bar{x}_e + \left(V + \tfrac{1}{2}\right)^3 h\bar{\omega}_e\bar{y}_e \qquad \text{Equation 2.5}$$

where $\bar{x}_e$ and $\bar{y}_e$ are the anharmonicity constants, V = overtone (0, 1, 2…)

If a bond is sufficiently excited, it will break (this is illustrated by the dashed grey line in Figure 2-1). The bond dissociation energy (*D*) is a measure of the amount of energy required to break a covalent bond under standard conditions.

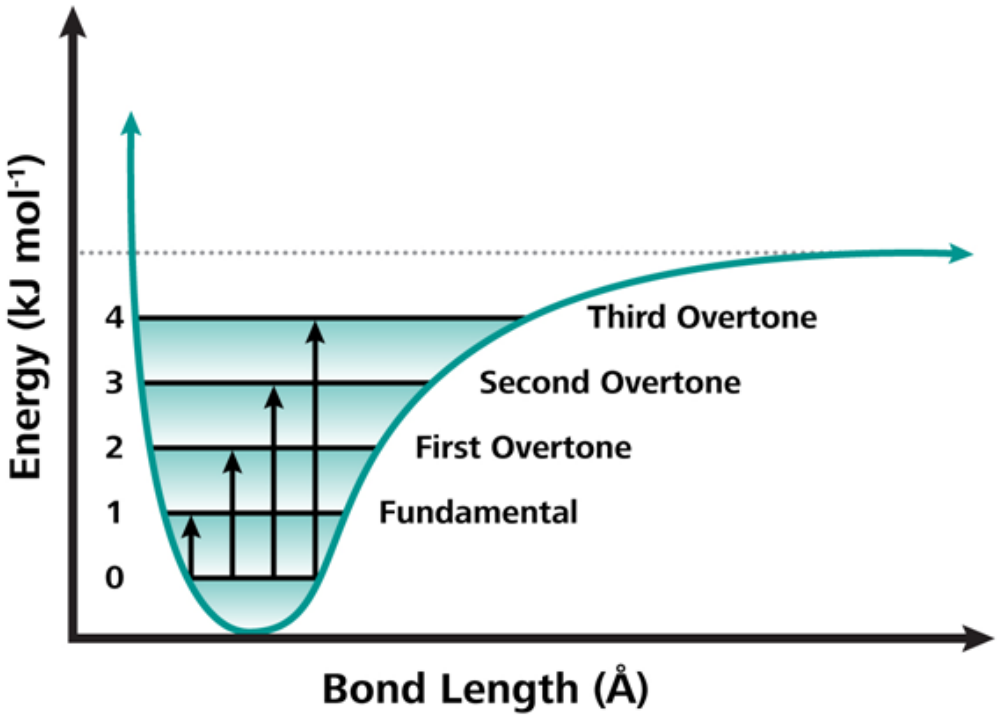

Figure 2-1: Potential energy well of a covalent bond. Source: Metrohm AG, Herisau, Switzerland. Reproduced with permission.

Furthermore, an atom bonded to two or more other atoms may experience different vibrational modes, such as symmetric stretching, antisymmetric symmetric stretching, scissoring, rocking, wagging and twisting (see Figure 2-2). Each of these vibrational modes will result in absorption at different frequencies. When all of this is taken into consideration, the total vibrational energy of a diatomic molecule can be expressed as shown in Equation 2.6:

$$E_v = \underbrace{\left(v+\frac{1}{2}\right)hv_e}_{\text{Harmonic Oscillator}} - \underbrace{\left(v+\frac{1}{2}\right)^2 X_e hv_e}_{\text{anharmonicity}} + \underbrace{B_e J(J+1)}_{\text{Rigid Rotor}} - \underbrace{D_e J^2(J+1)^2}_{\text{centrifugal stretching}} - \underbrace{\alpha_e\left(v+\frac{1}{2}\right)J(J+1)}_{\text{rovibrational coupling}}$$

Equation 2.6

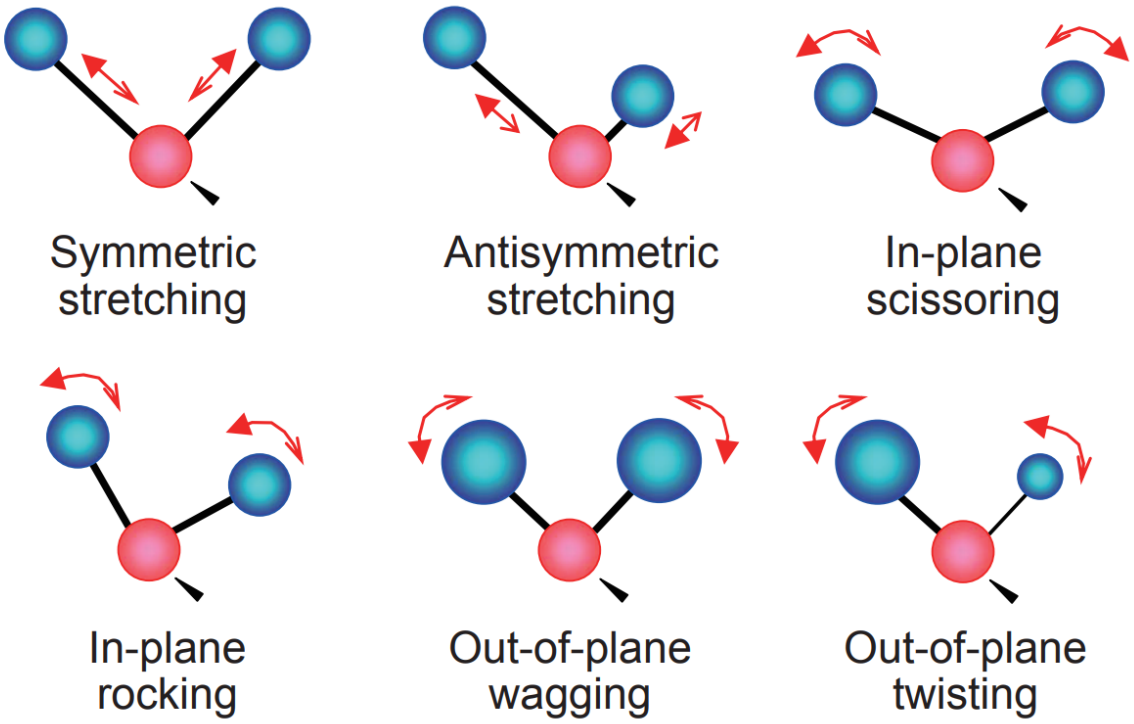

Figure 2-2: The various vibrational and rotational modes that may occur in a molecule. Source: Yokogawa Australia Pty Ltd. Reproduced with permission.

The background information presented here is essential for understanding the interaction of infrared light with covalent bonds and consequently the principles behind the use of infrared spectroscopy for the measurement of analytes in food matrices. More detailed theory behind bond vibrational levels and their interaction with IR light is beyond the scope of this review; for this, the reader is referred to previous publications on this topic (Beć & Huck, 2019; Cozzolino, 2014a; Ozaki et al., 2021).

2.2.4.2 Absorption of infrared light and band assignments

When a bond is exposed to infrared light of the correct frequency, it will absorb light photons at that frequency and the bond will be stimulated to a higher energy level for a very brief period (approximately 10^{-12} seconds) before returning to its ground state.

The corresponding frequency for each bond depends upon the bond structure, as discussed in Section 2.2.4.1. Numerous factors influence the wavelength at which infrared light is absorbed by a bond, including the mass of both atoms comprising the bond, the vibronic coupling associated with the bond, and the shape of the molecular surface surrounding the bond. Hence the same bond can absorb at slightly different wavelengths, depending on the

other bonds and atoms surrounding the bond of interest. However, the typical wavelengths at which bonds absorb infrared light are well documented.

Figure 2-3 details the absorption locations of common chemical bonds in the mid-infrared region, while Table 2-1 provides more specific information on the MIR absorption locations of important bonds for food analysis. For example, OH bonds absorb between 3700-3200 cm^{-1}, while CH bonds absorb between 3200-2800 cm^{-1}.

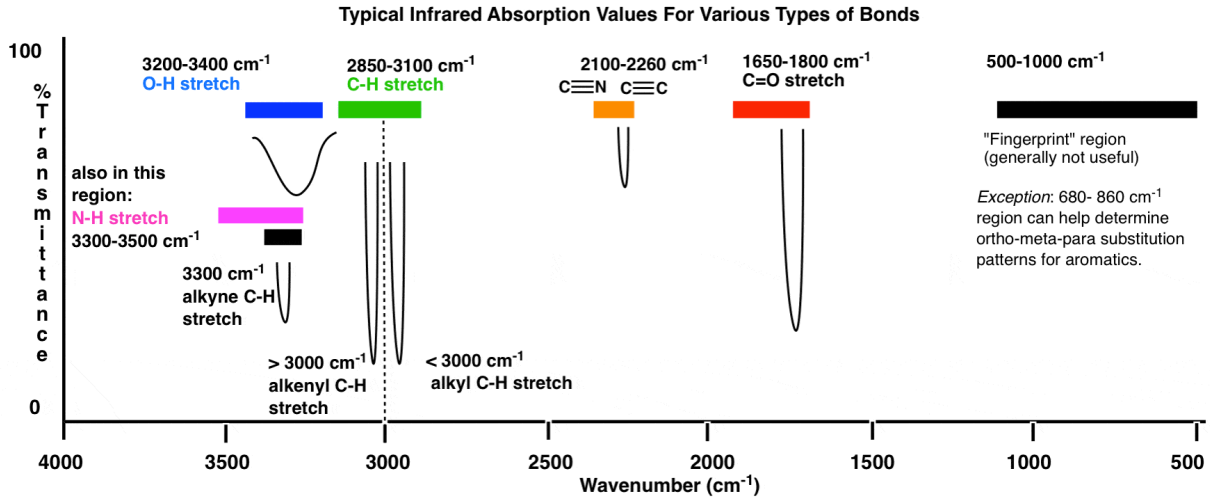

Figure 2-3: The locations of some major absorption bands in the mid-infrared region. Source: Master Organic Chemistry (https://www.masterorganicchemistry.com/2016/11/23/quick_analysis_of_ir_spectra/). Reproduced with kind permission from James Ashenhurst.

Table 2-1: Mid-infrared absorption bands for a range of bonds important for food analysis.

Bond	Compound/functional group	Wavenumbers (cm^{-1})
O-H stretch	Water, alcohol	3600-3200
C-H stretch	Alkene	3100-3000
C-H stretch	Aromatic ring	3060-3020
C-H stretch	CH_2/CH_3	2960-2860
C=O stretch	Carboxylic acid	~1750
C=O stretch	Ester	1750-1715
C=O stretch (amide I)	Amide	1700-1600
C=C stretch	Alkene	1666-1640
C=C stretch	Aromatic ring	1625-1590, 1590-1575, 1525-1470, 1465-1430
-C-H deformation vibration (asymmetric & symmetric)	Methoxy group	1470-1435
O-H deformation	Phenol	1390-1330
C-O-H deformation	Phenol	1382-1317
C-O vibration	Alkyl-aryl ether	1310-1210, 1120-1020
C-O stretch	Phenol	1260-1180
C-C stretch	Phenyl carbon	1225-1075
C-O stretch	Ester, alcohol	1230-1030
-C-H rocking vibration	Methoxy group	1200-1185
C-O stretching vibration	Phenol	1150-1040
C-H out-of-plane deformation	Aromatic ring	900-700
O-H out-of-plane deformation	Aromatic ring	~720

References: (Abbas et al., 2017; Dufour, 2009; Mecozzi & Sturchio, 2017)

As previously mentioned (Section 2.2.4.1), the presence of overtones in the NIR region means that absorption peaks from a single bond occur repeatedly throughout the NIR spectrum, at different levels of attenuation (Figure 2-4). In addition, combination bands can occur in the far-NIR region when two or more fundamental vibrations are excited simultaneously (Bokobza, 1998).

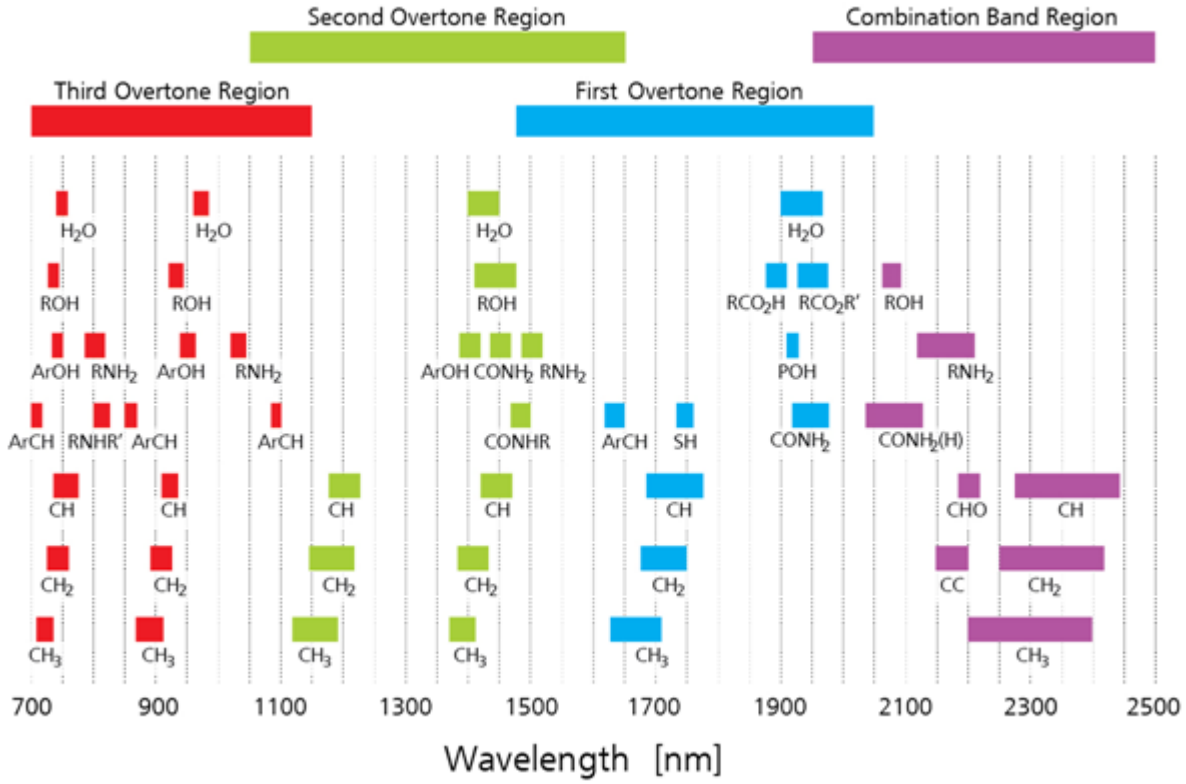

Figure 2-4: Near-infrared absorption band locations. Source: Metrohm AG, Herisau, Switzerland. Reproduced with permission.

Determination of the peak locations in a given matrix allows the qualitative identification of the type of bonds present in the sample. Furthermore, the amount of infrared light absorbed by the matrix is theoretically proportional to the number of bonds present, following the Beer-Lambert's law:

$$A = \log\left(\frac{I_0}{I}\right) = \epsilon l c \qquad \text{Equation 2.7}$$

where A = absorbance, ϵ = molar absorptivity, l = length of light path, c = concentration

However, the Beer-Lambert law is strictly valid only for transmission measurements obtained through non-scattering media. Consequently, a number of modified laws have been proposed for transmittance spectra obtained from scattering media (Mallet et al., 2021; Mamouei et al., 2021) and for reflectance spectra (Bhatt et al., 2016; Mayerhöfer et al., 2020).

Nevertheless, the underlying relationship between bond abundance and absorbance has allowed infrared spectroscopy to be used as a quantitative method for the assessment of various IR-active analytes across a wide range of sample matrices (Caporaso et al., 2018; Cortés et al., 2019; Mahesar et al., 2019; McGoverin et al., 2010; Wilson & Tapp, 1999).

2.2.5 Sample presentation

To gain an accurate assessment of the sample matrix using infrared spectroscopy techniques, it is essential that the portion of the sample that is "seen" by the instrument is representative of the whole sample. The use of an appropriate method of sample presentation is crucial in this respect. Furthermore, due to the wide range of sample types which can be analysed using IR spectroscopy (including solids, liquids, films, gels and powders), there are a variety of sample presentation methods that have been adopted.

Perhaps the simplest form of sample presentation is full transmittance mode (180° light-sample-detector). This is also the only method for which the Beer-Lambert law holds strictly true. In this presentation mode, the IR light enters one side of the sample and some wavelengths are absorbed by the sample, with the remaining light measured as it exits the other side of the sample. As long as the length of the light path is sufficiently low, the use of transmittance mode ensures that the emitted light has an opportunity to interact with nearly all of the analytes present in the light path. Consequently, it is typically quite representative of the true matrix composition. However, it is only suitable for analysing relatively thin samples due to the high level of absorbance in aqueous-based matrices. As shown by Beer-Lambert's law (Equation 2.7), increasing the light path length will proportionally increase the absorbance, making it more difficult to detect the signal of the resultant spectra. For example, a path length of only a few millimetres is often required when using transmittance NIR spectroscopy for the analysis of aqueous solutions. Due to path length limitations, the use of transmittance spectroscopy for the analysis of solid or powder substances can be more challenging compared to reflectance modes. However, analysis of products such as whole fruits is possible using higher incidence light intensities and more sensitive detectors (Clark et al., 2003; Fraser et al., 2003).

One variation of full transmittance mode is partial transmission spectroscopy, also known as interactance spectroscopy. This refers to the mode where the infrared light is partially transmitted through the sample matrix, before being detected by another sensor at the matrix surface but located adjacent to the source. These instruments utilise a physical barrier between the light source and detector to prevent the detector from receiving any IR light reflected directly from the sample surface (see Figure 2-5). The benefits of this method are a

reduced path length compared to full transmittance mode, while providing increased interaction between the IR light and the sample matrix compared to reflectance geometry.

Reflectance mode is one of the most commonly used presentation modes in infrared spectroscopy applications, particularly for NIRS. In this mode, the infrared light enters one side of the sample and interacts with the sample matrix as it penetrates into the sample. The majority of non-absorbed light is then reflected back to the surface of the sample, where it is detected by the instrument sensor. Some non-absorbance scattering of the IR light can also occur, which can bias the resultant spectra. One of the main advantages of reflectance mode is its one-dimensionality (i.e., the instrument only needs access to the sample surface in one location, as opposed to transmittance spectroscopy where both sides of the sample must be accessible), allowing it to be used in a much broader range of applications compared to transmittance spectroscopy. However, it is entirely reliant on the assumption that the composition of the surface material is representative of the entire sample matrix.

Within the food sector, reflectance NIR spectroscopy is widely reported in publications for the analysis of horticultural produce (Kumar et al., 2017; Ncama et al., 2018) and in the grains industry (Caporaso et al., 2018; Yang et al., 2011). There are no commercial instruments designed to use this geometry mode for the analysis of whole fruits, as differences in skin composition between fruit populations can reduce the robustness of the model. However, reflectance NIR spectroscopy is commonly used for the analysis of ground grain products in industry/commercial settings, as these samples are generally quite homogenous throughout.

Diffuse or body reflectance mode is also commonly used by NIR spectroscopists. It functions similarly to regular reflectance spectroscopy, but benefits from increased interaction between the IR light and the sample compared with specular (surface) reflectance modes (Figure 2-5).

A diagrammatical summary of the main sample presentation modes used for the infrared spectroscopy analysis of solid matrices is shown in Figure 2-5. As each sample presentation mode has its drawbacks and benefits (Kawano, 2008; Wilson & Tapp, 1999), the optimum method will depend on the sample matrix and intended application.

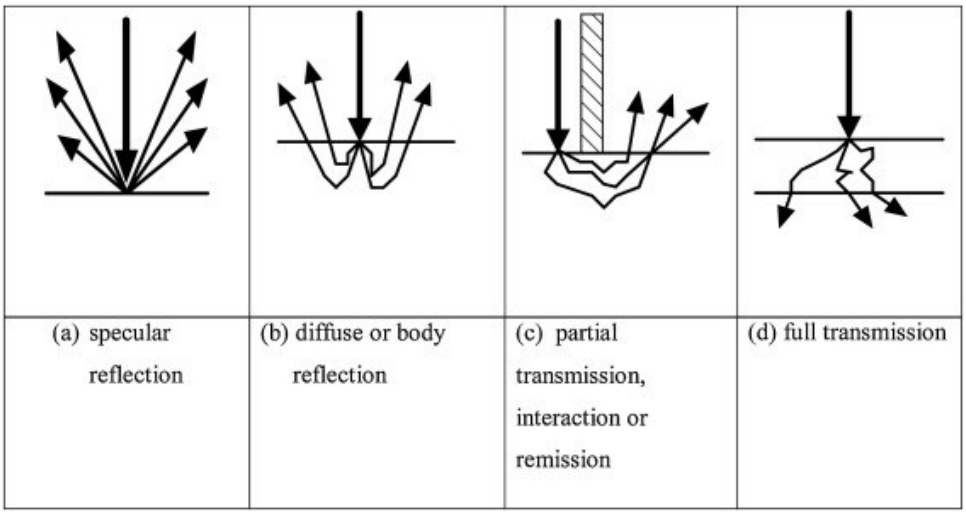

Figure 2-5: Sample presentation modes used in the infrared spectroscopy analysis of solid materials, showing the interaction of the light with the sample. Source: Walsh et al. (2020). Reproduced under Creative Commons 4.0 licence.

2.2.6 Data analysis

The final stage in the use of infrared spectroscopy for analytical purposes is the processing of the spectral data. In many cases, the signal of the desired analyte may be obscured by other matrix components present in much higher concentrations, such as water or carbohydrate-based structures. The use of modern mathematical data analysis techniques – termed chemometrics – can aid in uncovering minor analyte signals and developing optimum models for the quantification of the analytes. However, it is important to note that no amount of data analysis or chemometrics can "uncover" an analyte if the signal from the analyte is either not present or its concentration is too low to be detected by the instrument. The exception to this is if there is a secondary correlation between the analyte and a macroconstituent that can be detected by NIRS (see Section 2.2.2).

2.2.6.1 Spectral pre-processing

Typically, infrared spectra are subjected to pre-processing before they can be used for quantitative analytical purposes. The aim of this procedure is to remove spectroscopic artefacts from the measurement process, such as random noise, scatter or baseline drift (Gautam et al., 2015; Rinnan, 2014). The effects of these artefacts are particularly detrimental when attempting to analyse complex mixtures or analytes present in very low concentrations (Schoot et al., 2020).

A variety of spectral pre-processing methods are available. These include smoothing, multiplicative scatter-correction (MSC), standard normal variate (SNV), normalization by range (NBR) and the calculation of derivatives (Dotto et al., 2018). As previous authors have

reviewed the range of available spectral pre-processing methods in detail (Lee et al., 2017; Mishra et al., 2020), only a brief summary of the most commonly pre-processing methods are presented here.

Standard normal variate (SNV) is a normalisation-based pre-processing method. In this pre-processing method, the mean value of each spectrum is calculated and this constant value is subtracted across the entire spectrum, before the spectrum is divided by the standard deviation of the entire spectrum.

Calculating the derivative of spectra is another common approach to account for baseline shift or amplitude differences in the spectra. First and second derivatives are the most commonly used. Although higher order derivatives such as the third derivative have been successfully used in some applications (Orman & Schumann, 1991; Rodriguez-Otero et al., 1995; Terhoeven-Urselmans et al., 2008), there is an accompanying decrease in the signal-to-noise ratio as the derivative order increases (Rodriguez-Otero et al., 1995).

Finally, it is important to note that pre-processing methods are often combined. For example, typical pre-processing of spectra for use in analytical spectroscopy could involve calculating the SNV of the spectra, before subsequently calculating the first derivative of the SNV-processed spectra.

The choice of optimum spectral pre-processing methods is poorly defined and strongly dependent upon both the matrix type and analyte of interest. Furthermore, the need for and choice of pre-processing method may also vary with the sample size of the population (Schoot et al., 2020). In the absence of definitive guidelines, trial and error is often the best approach when developing new applications for infrared spectroscopy.

2.2.6.2 Data analysis methods

For quantitative applications of infrared spectroscopy, regression modelling is among the most commonly used data analysis methods. One of the first chemometric methods applied in quantitative infrared spectroscopy applications was multiple linear regression (MLR), which attempts to predict the analyte concentration from the spectral absorbance at several different wavelengths. However, it cannot be used for the analysis of entire spectra, due to the high multicollinearity of the adjacent datapoints comprising the spectra.

Partial least squares regression (PLSR) is a derivative of MLR suited to datasets with high levels of multicollinearity, such as infrared spectra (Mehmood & Ahmed, 2016). Through a variety of algorithms, the key contributing variables are identified and weighted such that the wavelengths most closely correlated with the analyte concentration have the greatest contribution to the PLSR model (Mehmood & Ahmed, 2016). PLSR is widely used for the

development of infrared spectroscopy models across the food science sector (Cobaleda-Velasco et al., 2018; de Oliveira et al., 2014; Hu et al., 2016).

In recent years, there has also been interest in emerging machine learning techniques such as artificial neural networks (ANNs), support vector machine (SVM) and deep learning (Gabriëls et al., 2020; Le, 2020; Rajalakshmi & Gopal, 2020; Sharabiani et al., 2019). These non-linear techniques look for patterns within the data in order to optimise model weighting and extract the desired information from the spectra. As more samples and spectra are added to the dataset, the model can update over time in order to provide more accurate prediction results.

As with spectral pre-processing, the optimum chemometric technique often depends on the sample matrix and/or analyte (Ludwig et al., 2019; Ni et al., 2014).

2.3 Bioactive compounds and their significance in functional foods

2.3.1 Functional foods

Recent years have seen an expansion of the "functional food" market – where foods are purchased for their health-benefiting effects, rather than as a source of basic nutrition and energy (Granato et al., 2017; Johnson et al., 2020f; Urala & Lähteenmäki, 2007). For example, phenolics isolated from chickpeas have been found to provide anti-cancer effects, particularly against colorectal cancer (Bochenek et al., 2019). Similarly, phenolics from the pseudocereal Quinoa (*Chenopodium quinoa*) have been shown to provide strong *in vitro* anti-inflammatory effects (Liu et al., 2020a). Consumers may pay a price premium for such health-benefiting foods (Di Pasquale et al., 2011). For example, Spanish consumers reported that they would pay ~55% extra for resveratrol-enriched wine (Barreiro-Hurlé et al., 2008), as this compound has purported benefits for cardiovascular health. This willingness to pay a premium for healthier food has been mirrored in several other studies (Hirogaki, 2013; Markosyan et al., 2009; Miškolci, 2014), albeit with typically lower price premiums reported (e.g., 10-15% higher than the regular price).

Even if there is not a market for the functional food in its unprocessed form, such produce also has potential for the development of value-added foods and ingredients (López-Barrios et al., 2014; Vioque et al., 2012; Zhang et al., 2012), marketed on the basis of their levels of health-benefiting compounds. Examples of foods experiencing a considerable rise in popularity due to their reported health benefits include the so-called ancient grains (such as chia, quinoa, millet and spelt), pulse crops (including mungbean, chickpeas, faba bean and lentils), as well as numerous other crops (Boukid et al., 2018; Cooper, 2015; Singh et al., 2017). For instance,

the plum variety Queen Garnet was developed and marketed with an emphasis on its exceptionally high levels of anthocyanins, which possess antioxidative and anti-thrombotic properties (Fanning et al., 2013; Netzel et al., 2012; Xiang et al., 2019c). Another well-known example is the açaí berry from South America, popularised due to its high anthocyanin content and antioxidant capacity (Yamaguchi et al., 2015).

2.3.2 Definition of bioactive compounds

There is no clear literature consensus on the definition of bioactive compounds, with Guaadaoui et al. (2014) proposing them to be "*compounds which have the capability and the ability to interact with one or more component(s) of living tissue by presenting a wide range of probable effects*". However, from a consumer perspective, bioactive compounds are generally understood to be compounds which will promote good health or health-benefiting effects. This concept is more similar to the consensus statement from the 23rd Hohenheim Consensus Meeting, which stated that "*bioactive compounds are essential and non-essential compounds (e.g., vitamins or polyphenols) that occur in nature, are part of the food chain, and can be shown to have an effect on human health*" (Biesalski et al., 2009). Such bioactive compounds may also be referred to as "nutraceuticals" (Kalra, 2003), which reflects their presence in the human diet.

2.3.3 Classes of bioactive compounds

The majority of bioactive compounds can be broadly classified as phytochemicals – compounds that are produced by plants – although some (such as fatty acids) are also found in animal-based foods. There are numerous classes of bioactive compounds (Figure 2-6), each with their own distinct biological activities and health benefits. These include polyphenols, flavonoids, carotenoids, phytosterols, phytoestrogens, alkaloids, glucosinolates, anthocyanins, terpenoids and others (Hamzalıoğlu & Gökmen, 2016; Mani et al., 2021). Each compound class is characterised by distinct structural features in their chemical composition. For example, polyphenols display the presence of multiple phenol groups, while all flavonoids comprise two phenyl rings and a heterocyclic ring containing an embedded oxygen heteroatom.

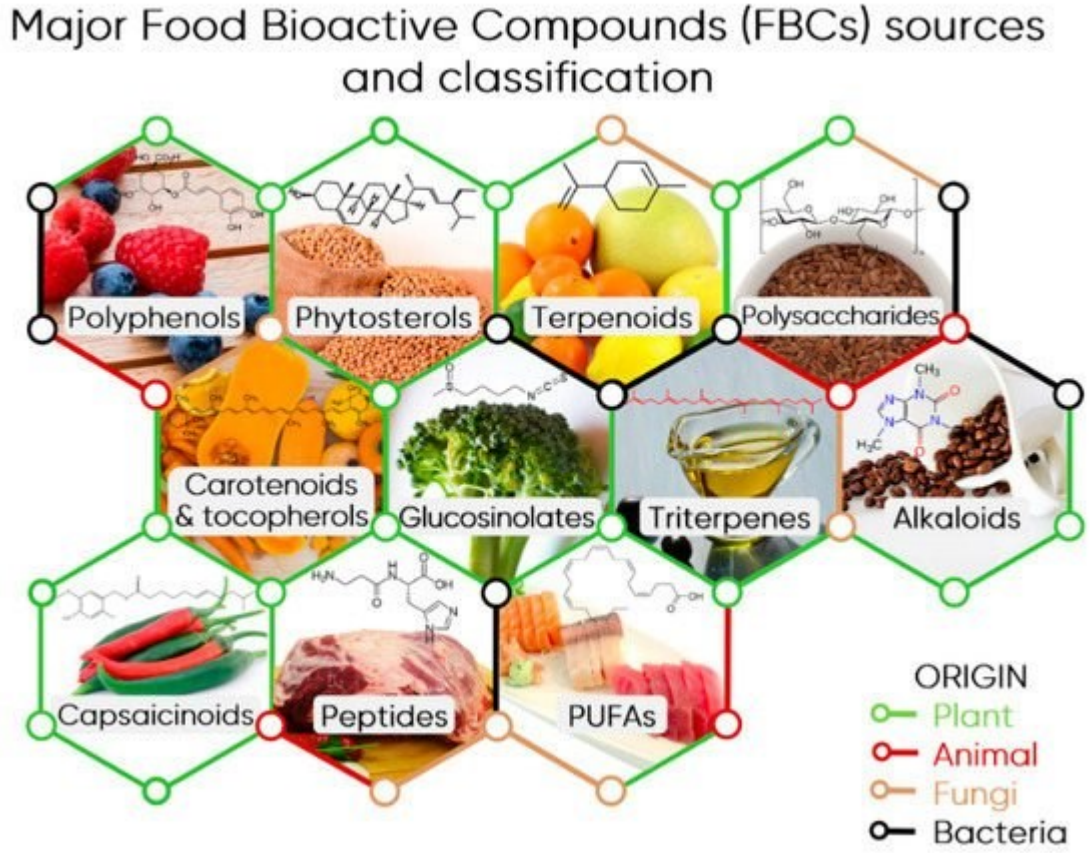

Figure 2-6: The various classes of major bioactive compounds found in food products. Source: Câmara et al. (2021). Reproduced under Creative Commons 4.0 licence.

It could be considered that compounds which show antioxidant activity form a class of bioactive compounds. However, a structurally diverse array of compounds may exhibit antioxidant activity, including polyphenols, anthocyanins, flavonoids and carotenoids. For this reason, this review excluded studies solely reporting quantification of the total antioxidant capacity (TAC) of samples, as in many cases the antioxidant activity of a matrix cannot be directly related to the concentration of a specific structural class of bioactive compounds (Pellegrini et al., 2020). Nevertheless, this does not detract from the importance of TAC as a potential indicator of crude biological activity. Although the concept of TAC has been criticised by several researchers as a result of its lack of specificity (Fraga et al., 2014; Pompella et al., 2014), numerous clinical trials have indicated a positive relationship between greater intake of antioxidants and reduced levels of oxidative stress and inflammatory markers (Detopoulou et al., 2010; Hermsdorff et al., 2011; Kobayashi et al., 2012; Wang et al., 2014b), reduced all-cause mortality (in non-elderly cohorts) (Agudo et al., 2007; Bastide et al., 2017) and reduced risk of adverse cardiovascular events, particularly ischaemic stroke (Colarusso et al., 2017; Del Rio et al., 2010; Rautiainen et al., 2012a; Rautiainen et al., 2013; Rautiainen et al., 2012b).

2.3.4 Current analytical methods

There are numerous analytical methods available for the quantification of bioactive compounds, depending on the physical and chemical properties of the specific class of compound(s) of interest.

For example, terpenoids and other volatile compounds are commonly analysed by gas chromatography coupled with mass spectrometry (GC-MS), which uses a mobile inert gas phase and a stationary column phase to separate the compounds of interest (da Silva et al., 2019; Kukula-Koch et al., 2018).

Non-volatile compounds such as polyphenols, anthocyanins, flavonoids and carotenoids are typically analysed using the related technique of liquid chromatography coupled with mass spectrometry (LC-MS) (Baysal et al., 2021; Netzel et al., 2006; Xiang et al., 2019a). As with gas chromatography, the column contains the stationary phase, while a liquid mobile phase carrying the analyte flows through the column. The relative affinity of the analyte for the mobile and stationary phases allows for its separation from other matrix constituents. Finally, the mass spectrometry module is used to identify the analyte based on its molecular mass.

In cases where the compounds of interest are known and pure standards are available for comparative purposes, high performance liquid chromatography (HPLC) with ultraviolet-visible detection may suffice (Santos et al., 2017). This method works in the same way as LC-MS but uses absorbance in the ultraviolet-visible region to detect the eluting compounds, rather than mass spectrometry.

Colorimetric methods such as the Folin-Ciocalteu assay may also be used for the analysis of total phenolics, or for the quantification of anthocyanins using the pH-differential method (Johnson et al., 2020c). These methods are less specific compared to separation-based techniques such as liquid and gas chromatography; however, they are must faster and cost-efficient.

More recently, there has been an interest in using rapid, non-invasive analytical techniques such as infrared spectroscopy for the prediction of bioactive compounds (Aleixandre-Tudo et al., 2018; Johnson et al., 2020d; Mahesar et al., 2019). This emerging area of research is the focus of this review.

2.3.5 Previous work and aims

Although several previous reviews have focused on the use of infrared spectroscopy for the estimation of specific groups of bioactive compounds, such as antioxidants (Cozzolino, 2015; Lu & Rasco, 2012) and phenolics (Ignat et al., 2011), there are no contemporary reviews in

the last decade on this technique for the quantification of bioactive compounds in food products. For instance, the review by McGoverin et al. (2010) on this topic is over ten years old, with numerous IR-related papers published during the ensuing period. Similarly, the review by Pallone et al. (2018) on the use of vibrational spectroscopy in food analysis included only 7 studies quantifying constituents which could be classified as “bioactive” compounds. Hence, this chapter aims to review the contemporary literature (last five years) reporting the estimation or quantification of bioactive compounds in food matrices.

2.4 Methods

The Scopus database (https://www.scopus.com/) was searched for articles between 2016-2020 containing the following terms in their title, abstract or keywords sections:

- Any of the following: near infrared OR mid infrared OR spectroscopy
- AND food
- AND bioactive OR phenolic OR antioxidant OR anthocyanin
- AND quantification OR determination OR measurement

In this way, articles pertaining to the quantification of bioactive constituents of functional foods using infrared spectroscopy were able to be acquired.

Articles up to and including 31st December 2020 were considered, with the search limited to articles published in the 5 years prior (i.e. 1st January 2016 to 31st December 2020). The titles and abstracts of all articles were manually screened to find relevant articles for inclusion in this review.

Inclusion criteria were:

- Original studies published in the 5 years between 2016 and 2020
- Quantified a compound or group of compounds with recognised health-benefiting effects, above that expected from basic nutritional needs
- The matrix was a food or potential food product

2.5 Scientific effort (2016-2020)

The scientific effort over the past five years is summarised in Table 2-2 (for NIRS) and Table 2-3 (for MIRS). The information presented in the tables includes the type of matrix analysed, analyte(s) investigated, sample size of the calibration and validation sets, wavelength range used in the optimised model, and statistical method used for analysis of the spectra. All fruit and vegetable samples were analysed fresh and intact, unless otherwise stated in the table. The test set column shows whether the authors used a dependent test set for the model

validation (i.e., samples from the same population as the calibration set) or independent test set (i.e., samples drawn from a different population to the calibration set, such as from a different year, season or geographic location). The cross-validation statistics (R^2_{CV} and RMSECV) are reported in the corresponding columns for all studies. In cases where the study also included an independent test set, the R^2 and RMSEP for the test set are reported in the test set column. Finally, the notes column provides information about the sample population details and notable findings of the study.

Table 2-2: Studies reporting the use of near-infrared spectroscopy for the quantification of bioactive compounds in food products (2016-2020).

Food matrix	Analyte(s)	Sample size (cal/val)	Wavelength range (nm)	Optical geometry	Statistical method	Test set	Cross validation	R^2 (CV)	RMSECV	Notes	Reference
Fruit											
Açaí and juçara	Total anthocyanin content	Variable (n=374 total)	1606-1793	Reflectance	PLS	Independent populations R^2 = 0.74-0.88; RMSEP = 5.09-6.76 g/kg	LOO	0.89-0.91	2.50-2.91 g/kg	Fruit from 2 seasons & 4 localities	(Cunha Júnior et al., 2016)
Bilberry (dried powder)	Anthocyanins	38/27	1064-1640, 1833-2354	Reflectance	PLS	Dependent test set (randomly selected samples)	LOO	0.995	0.28 % w/w	NIR analysis could identify counterfeit bilberry samples	(Gardana et al., 2018)
Blackberry	Total phenolics Total carotenoids	90/30	400-2500	Reflectance	PLS	None	n/s	0.69 0.76	1.69 0.95 mg/g		(Toledo-Martín et al., 2018)
Grapes (red)	*Trans*-resveratrol Quercetin Total phenols	15	900-1700	Reflectance	PLS	None	LOO	0.988 0.955 0.974	0.424 0.008 mg/kg 12.15 mmol/kg	3 locations; 2 seasons	(Tzanova et al., 2020)
Grapes (red and white)	Total phenolics	203/67	400-1100 900-2500	Reflectance	PLS, SVM	Dependent test population	n/s	0.872-0.914 0.697-0.726	0.15-0.22 0.28-0.31 mg/g	2 cultivars from one season and location. SVM gave better results than PLS	(Xiao et al., 2018)

Grapes (red)	Total anthocyanins Total 3-O-glucoside anthocyanins Total 3-O-(6-acetyl)glucoside anthocyanins Total 3-O-(6-p-coumaroyl)glucoside anthocyanins Malvidin 3-O-glucoside Malvidin 3-O-(6-acetyl)glucoside Malvidin 3-O-(6-p-coumaroyl)glucoside Petunidin 3-O-glucoside Petunidin 3-O-(6-acetyl)glucoside Petunidin 3-O-(6-p-coumaroyl)glucoside Delphinidin 3-O-glucoside Delphinidin 3-O-(6-p-coumaroyl) glucoside Peonidin 3-O-glucoside Peonidin 3-O-(6-acetyl)glucoside Peonidin 3-O-(6-p-coumaroyl)glucoside Cyanidin 3-O-glucoside Cyanidin 3-O-(6-acetyl)glucoside Cyanidin 3-O-(6-p-coumaroyl)glucoside	60/20	380-1028	Reflectance (hyperspectral imaging)	MPLS	Dependent test set (stratified samples)	6-fold cross validation	0.91 0.92 0.90 0.83 0.87 0.90 0.80 0.93 0.57 0.91 0.91 0.92 0.88 0.80 0.75 0.88 0.77 0.86	189.05 155.94 4.12 23.09 73.93 4.13 13.32 29.44 0.98 2.04 48.41 3.28 30.35 0.31 3.81 16.43 0.16 1.78 mg/L	8 different cultivars from 2 sites	(Diago et al., 2016)
Grapes (red)	Nonacylated anthocyanins Total anthocyanins	47/-	950-1650	Reflectance (hyperspectral imaging of single grapes)	PLS	Dependent test population	LOO	0.72 0.72	0.78 0.70 mg/ grape	Fruit from 2 dates, 2 vineyards, 1 season	(Martínez-Sandoval et al., 2016)

Grape pomace (marc)	Catechin Epicatechin Proanthocyanidin B1 Proanthocyanidin B2 Proanthocyanidin B3 Proanthocyanidin B4 Proanthocyanidin trimer 1 Proanthocyanidin trimer 2 Proanthocyanidin tetramer 1 Proanthocyanidin tetramer 2 Proanthcyanidin B2-3-O-gallate Galloyl proanthocyanidin Total flavanols Gallic acid Protocatechuic acid Caffeic acid Caftaric acid Cis-coutaric acid Trans-coutaric acid Total phenolic acids Quercetin 3-O-rutinoside Quercetin 3-O-glucuronide Quercetin 3-O-glucoside Quercetin pentoside Kaempferol 3-O-galactoside Kaempferol 3-O-glucuronide Kaempferol 3-O-glucoside	12/-	950-1650	Reflectance (hyperspectral imaging)	PLS	None	LOO	0.80 0.96 0.65 0.75 0.50 0.63 0.65 0.86 0.65 0.53 0.89 0.58 0.78 0.75 0.82 0.92 0.91 0.83 0.95 0.87 0.63 0.81 0.64 0.15 0.98 0.93 0.98	14.00 4.72 20.53 1.86 3.43 3.01 3.12 7.68 11.62 2.66 6.29 7.27 66.63 5.58 2.70 0.36 2.56 0.15 0.19 9.61 1.82 4.36 5.95 0.04 0.11 0.07 0.41	Fruit from 1 variety, season and location	(Jara-Palacios et al., 2016)

	Quercetin Kaempferol Total flavonols							0.72 0.97 0.70	0.19 0.02 14.27 mg/100g		
Grape skins (red)	Total iron-reactive phenolics Anthocyanins Tannins	40/20	977-1625	Reflectance (hyperspectral imaging)	PCR, PLS, SVR	Dependent test set	Segment validation	0.907 0.879 0.896	0.178 0.144 0.107 mg/L	Five cultivars from 4 dates in 1 growing season	(Zhang et al., 2017)
Grape seeds (red)	Total iron-reactive phenolics Tannins	40/20	977-1625	Reflectance (hyperspectral imaging)	PCR, PLS, SVR	Dependent test set	Segment validation	0.879 0.924	0.240 0.519 mg/L	Five cultivars from 4 dates in 1 growing season	(Zhang et al., 2017)
Guava (frozen pulp)	Ascorbic acid	50	1000-1892, 2007-2227	Transflectance	PLS	Dependent test set (randomly selected samples)	LOO	0.85 (test set)	6.14 mg/100g (test set)	Samples from 2 Brazilian marketplaces	(Alamar et al., 2016)
Jujube	Gallic acid Caffeic acid L-epicatechin Phloridzin Cianidanol	52/26	900-1700	Transmittance	Si-ACO-PLS	Dependent test population	n/s	0.879 0.887 0.906 0.858 0.836	3.06 6.04 16.30 0.84 16.01 µg/g	Samples from 5 regions	(Arslan et al., 2020)
Kakadu plum (powder)	Ascorbic acid	80/5	866-2532	Diffuse reflectance	PLS	Independent (commercially sourced samples) R^2 = 0.73; RMSEP = 4733 mg/ 100g	n/s	0.93	1839 mg/100g		(Cozzolino et al., 2020)
Marsh grapefruit	β-carotene Total carotenoids	240	850-2500 400-850	Reflectance	PLS	Independent (samples from a separate orchard)	Test set validation	0.99 0.92 (test set)	0.00 2.69 µg/g (test set)	Fruit from 1 season and 2 locations	(Ncama et al., 2018)
Raspberries	Total phenols Total anthocyanins TAC (FRAP)	168	950-1650	Reflectance (hyperspectral imaging)	PLS	None	n/s	0.70 0.63 0.61	127 12 39 mg/100g		(Rodríguez-Pulido et al., 2017)

Tomato	Total phenolics Lycopene Total flavonoid β-carotene	50	285-1200	Reflectance	PLS	None	LOO?	0.834 0.864 0.790 0.708	1.80 1.03 1.82 1.14 µg/g		(Alenazi et al., 2020)
Tomato	Lycopene β-carotene	180/60	500-1100	Transmittance	PLS	Semi-independent (separate harvest in same season) R^2 = 0.85, 0.77; RMSEP = 1.79, 1.00 mg/kg	LOO	0.89 0.88	1.56 0.63 mg/kg		(Tilahun et al., 2018)
Tomato (dehydrated and ground)	Lycopene Phenols TAC (DPPH) TAC (FRAP) TAC (ABTS)	61/31	800-2500	Reflectance	PLS, RBF-NN	Dependent test set	n/s	0.882 0.910 0.882 0.876 0.937	1.61 80 0.70 0.97 0.86 mg/100g	RBF-NN generally performed better than PLSR	(Ding et al., 2016)
Wax jambu	Total phenolics Total anthocyanins	50/35	1000-2400	Diffuse reflectance	PLS	Dependent test set (Kennard-Stone selection)	n/s	0.94 0.98	22.18 9.0 mg/100g		(Viegas et al., 2016)
Vegetables											
Carrot	Ascorbic acid β-carotene	24/6 (?)	420-1100	Reflectance	PLS	Dependent test set	4-fold cross validation	0.98 0.98	0.04 µg/g 0.10 µg/100g	Roots sampled over an 8-week storage period	(Rady et al., 2018)
Red cabbage (EtOH extract)	Total anthocyanins Monomeric anthocyanins Total polyphenols	1 (with 33 serial dilutions)	1000-2500	Transmittance	PLS	9 dilutions prepared from new cabbage extract	Segment validation	0.98 0.98 0.96	16.4 20.2 42.7 mg/L		(de Oliveira et al., 2018)
Potato	Total phenolics Antioxidant capacity (DPPH)	160/68	1100-2300	Reflectance	PLS	Dependent test set	Venetian blind cross-validation	0.84 0.67	1.20 1.21 mg/g	White, red, yellow & purple fleshed cultivars	(López-Maestresalas et al., 2017)

Grains/pulses											
Barley malt	Total phenolics	10	1000-2500	Reflectance (hyperspectral imaging)	SVM, SVR	Dependent test set (5% of total pixels)	n/s	0.85	1 ppm		(Tschannerl et al., 2019)
Oat, millet, buckwheat	Total phenolics	77	1596-2396 1128-2162 740-1070	Reflectance	PLS	Test set used but no information provided on its origins or size	LOO (for most models)	0.921 0.951 0.823	1.46 1.11 1.98 mg/g	Compared 3 handheld instruments (microPhazir RX, MicroNIR 2200, SCiO)	(Wiedemair & Huck, 2018)
Common bean (flour)	Total phenols *Ortho*-diphenols Flavonoids Gallic acid Catechin Myricetin-3-glucoside Quercetin-3-6''-manolyl-glucoside Kaempferol-3-glucoside Kaempferol-3-6''-manolyl-glucoside Kaempferol	42/-	1000-2500	Reflectance	PLS	Spectra randomly selected from dataset (1/3 of total spectra)	LOO	0.91 0.85 0.90 0.96 0.48 0.97 0.90 0.85 0.93 0.87	RPDs: 5.20 4.84 5.18 10.25 2.38 10.25 7.52 6.27 9.32 6.51	21 varieties; 2 seasons	(Carbas et al., 2020)
Mungbean	Catechin Chlorogenic acid Caffeic acid *p*-coumaric acid *t*-ferulic acid Vitexin Isovitexin Myricetin Quercetin Kaempferol	42/18	1600-2500	Reflectance (from whole grains)	PLS	Dependent test set	Segment validation	0.996 0.998 0.992 0.989 0.998 0.997 0.997 0.994 0.989 0.998	0.603 0.590 1.78 1.8 0.519 0.238 0.23 1.82 1.67 0.5 (% relative)		(Meenu et al., 2016a)

Quinoa (whole seed)	Total free phenolics Total betalains TAC (DPPH)	38/-	400-2500	Reflectance	PLS	None	Segment validation	n/s n/s 0.73	n/s n/s 8.6 mmol/kg	For TAC in ground seed; R^2 = 0.66; RMSECV = 9.6 mmol/kg	(Macavilca & Condezo-Hoyos, 2020)
Soybean	Total anthocyanins Cyanidin-3-glucoside Delphinidin-3-glucoside	70	1000-2500	Reflectance	PLS	Subset of spectra of samples included in calibration set	n/s	0.88 0.90 0.88	0.13 0.12 0.03 mg/g		(Amanah et al., 2020)
Oils											
Olive oil	Squalene	118/59	1100-2300	Transmittance	PLS	Dependent test set	LOO?	0.83 (test set)	2.31 g/kg (test set)	Poorer results obtained using Vis-NIRS data	(Cayuela & García, 2018)
Olive oil	Total tocopherols *α*-tocopherol β-tocopherol γ-tocopherol	197/91 189/93 197/102 195/101	350-2500	Transmittance, transflectance	PLS	Dependent test set	LOO?	0.89 0.92 0.54 0.85	SEC: 43.83 33.90 0.59 4.54 mg/kg	Vis-NIRS gave slightly better results than NIRS in most cases	(Cayuela & García, 2017)
Olive oil	Tyrosol Tyrosol secoiridoids Hydroxytyrosol Hydroxytyrosol secoiridoids Total phenolics	75/18	800-2500	Transmittance	PLS	None	LOO	0.55 0.84 0.55 0.82 0.82	5.27 41.5 4.84 43.1 76.7 mg/kg		(Mora-Ruiz et al., 2017)
Olives (as paste)	Total phenolics Oleuropein	291/53 147/53	1400-2400	Reflectance	PLS	Dependent test set	LOO?	0.71 0.73 (cal)	0.08 6.6 mg/kg	Samples obtained across 7 seasons	(Trapani et al., 2017)

Aromatic plants											
Black pepper (whole)	Piperine	132/-	950-1650	Reflectance	PLS	None	Segment validation	0.726	0.289 g/100g	For ground samples; R^2 = 0.850, RMSECV = 0.231g/ 100g	(Park et al., 2020)
Black tea	Caffeine Epigallocatechin gallate	37/19	950-1650	Reflectance	PLS, MLR	Dependent test set	LOO?	0.933 0.782	3.65 3.32 mg/g		(Wang et al., 2021b)
Black tea	Cianidanol Ferulic acid Gallic acid Rutin Phloridzin L-epicatechin	84/56 (20 replicate samples at 7 time points)	899-1724	Transmittance	CARS-PLS	Dependent test set	n/s	0.956 0.928 0.911 0.825 0.881 0.969	9.66 0.21 4.22 0.77 6.85 20.1 mg/100g	20 tea samples collected at 7 time points during fermentation process	(Zareef et al., 2019)
Cocoa bean	Total phenols Catechin Epicatechin Epigallocatechin Theobromine	74/- 76/- 75/- 72/- 75/-	400-2498	Reflectance	PLS	None	LOO?	0.71 0.62 0.04 0.02 0.77	6.09 0.65 5.24 0.09 4.55 mg/g		(Hernández-Hernández et al., 2021)
Cocoa bean	Total polyphenols	72	800-2778	Diffuse reflectance	PLS	None	LOO	0.84	0.93 mg/g	Samples from different storage and fermentation periods	(Sunoj et al., 2016)
Cocoa bean husk	Total phenols Catechin Epicatechin Epigallocatechin Theobromine	77/- 80/- 79/- 78/- 78/-	400-2498	Reflectance	PLS	None	LOO?	0.81 0.74 0.06 0.20 0.83	4.75 0.55 5.31 0.10 3.72 mg/g		(Hernández-Hernández et al., 2021)

Coffee bean	Chlorogenic acid Total phenolics	101/36	950–1650	Reflectance (hyperspectral imaging)	MPLS	Dependent test set. SEP = 15.6 and 17.6%	n/s	0.81 0.58 (cal)	0.91 4.63 mg/g		(Nogales-Bueno et al., 2020)
Ginger	Zingerone 6-gingerol 8-gingerol 10-gingerol 6-shogaol	58/22	1389-2500	Reflectance	PLS	Dependent test set	LOO	0.981 0.986 0.988 0.997 0.998 (cal)	0.076 0.072 0.078 0.077 0.084 mg/g		(Yan et al., 2021)
Beverages											
Cashew apple nectar	Ascorbic acid	49/16	1000-1903, 1971-2227	Transflectance	PLS	Dependent test set (randomly selected samples)	n/s	0.84 (cal)	4.8 mg/100g (test set)	Samples from 2 Brazilian marketplaces	(Caramês et al., 2017a)
Coffee aqueous solution	Chlorogenic acid	86	401-1871	Transmittance (1 mm path length)	PLS	None	LOO	0.556	0.76 mg/mL	Key predictor wavelength was around 1450 nm (C-H vibration; 2nd overtone)	(Shan et al., 2017)
Grape juice	Total phenolics Anthocyanins	49/16	1000-2500	Transflectance	PLS	Dependent test set (randomly selected samples)	Optimising no. of latent variables	0.96 0.84 (cal)	37 4.44 mg/100mL (test set)	Slightly worse results for phenolic content compared to MIR	(Caramês et al., 2017b)
Guava nectar	Ascorbic acid	41/13	1000-1899, 1983-2227	Transflectance	PLS	Dependent test set (randomly selected samples)	n/s (LOO?)	0.86 (cal)	7.44 mg/100g (test set)	Samples from 2 Brazilian marketplaces	(Caramês et al., 2017a)
Soft drink (grape & passion-fruit)	Ascorbic acid	~47/20	1000-2500	Reflectance	PLS	Test set of 5 samples created by diluting one sample to specific concentrations	LOO	0.70 0.76	0.67 0.56 mg/g		(Santana et al., 2020)

Wine (red)	*Trans*-resveratrol Quercetin Total phenols	20	900-1700	Transmittance	PLS	None	LOO	0.994 0.990 0.996	0.113 0.073 mg/L 0.144 mM	3 locations; 2 seasons	(Tzanova et al., 2020)
Wine (red)	Gallic acid Catechin B1 (flavonol dimer) Polymeric phenols Caftaric acid Caffeic acid Coutaric acid p-coumaric acid Quercetin-3-glucoside Quercetin Kaempherol Delphinidin-3-glucoside Cyanidin-3-glucoside Petunidin-3-glucoside Peonidin-3-glucoside Malvidin-3-glucoside Delphinidin-3-acetylglucoside Cyanidin-3-acetylglucoside Petunidin-3-acetylglucoside Peonidin-3-acetylglucoside Malvidin-3-acetylglucoside Delphinidin-3-cumarylglucoside Petunidin-3-cumarylglucoside Peonidin-3-cumarylglucoside	~387/182	800-2500	Transmittance	PLS using PRESS	Dependent test set	Segment validation	0.86 0.83 0.76 0.88 0.86 0.87 0.84 0.87 0.88 0.84 0.85 0.92 0.86 0.9 0.85 0.87 0.88 0.91 0.92 0.91 0.85 0.86 0.85 0.86	3.01 5.85 4.94 135 8.8 0.82 2.63 0.61 10.3 1.65 0.15 2.32 0.05 2.16 1.73 16.5 0.65 0.34 0.89 0.65 7.15 0.19 0.57 0.84	Wines comprised four cultivars from 13 vinifications over 2 seasons. NIRS more accurate at predicting phenolic content than ATR-MIR or transmission FT-IR.	(Aleixandre-Tudo et al., 2018)

	Malvidin-3-cumarylglucoside Polymeric pigments MCP tannins Anthocyanins							0.84 0.86 0.92 0.87	4.27 5.71 204 53.1 mg/L		
Other foods											
Honey	Phenolics Flavonoids Carotenoids Antioxidants (FRAP)	105/45	1000-2500	Reflectance	PLS	Dependent test set (randomly selected samples)	Segment validation	0.884 0.903 0.922 0.922	14.5 1.01 0.035 0.43 mg/100g	6 different floral varieties of honey	(Tahir et al., 2016)
Propolis	Flavones & flavonols Flavanones & dihydroflavonols Antioxidant capacity (ABTS)	70/29	1100-2000	Reflectance (fibre-optic) on ground sample	MPLS	Dependent test set (randomly selected samples)	Segment validation	0.63 0.68 0.87 (cal)	29.4 9.5 112 mg/g	Samples sourced from Chile and Spain	(Betances-Salcedo et al., 2017)

Abbreviations: RBF-NN = radial basis function neutral network; LOO = leave-one-out cross-validation; n/s = not specified; PLS = partial least squares; SVM = support vector machine; TAC = total antioxidant capacity

Table 2-3: Studies reporting the use of mid-infrared spectroscopy for the quantification of bioactive compounds in food products (2016-2020).

Food matrix	Analyte(s)	Sample size (cal/val)	Wavenumber range (cm^{-1})	Optical geometry/ presentation	Statistical method	Test set	Cross validation	R^2 (cv)	RMSECV	Notes	Reference
Fruit											
Kakadu plum (powder)	Ascorbic acid	80/5	4000-400	ATR	PLS	Independent (commercially sourced samples) R^2 = 0.65; RMSEP = 2367 mg/ 100g	n/s	0.91	1811 mg/100g		(Cozzolino et al., 2020)
Vegetables											
Red cabbage (EtOH extract)	Total anthocyanins Monomeric anthocyanins Total polyphenols	1 (with 33 serial dilutions)	4000-650	ATR	PLS	9 dilutions prepared from new cabbage extract	Segment validation	0.98 0.98 0.96	18.1 21.3 44.4 mg/L		(de Oliveira et al., 2018)
Grains/pulses											
Buckwheat (leaves and flowers)	Rutin Quercetin Quercitrin Sum of flavonoids	Not stated (total = 108)	4000-500	ATR (whole and ground dried samples)	PLS	Dependent test set	LOO	0.99 0.99 0.95 0.98	3.63 0.06 2.48 4.80 mg/g	Used 7 different species of buckwheat	(Kokalj Ladan et al., 2017)
Common bean (flour)	Total phenols *Ortho*-diphenols Flavonoids Gallic acid Catechin Quercetin-3-glucoside	42/-	4000-400	ATR (flour)	PLS	Spectra randomly selected from dataset (1/3 of total spectra)	LOO	0.86 0.31 0.86 0.94 0.89 0.43	RPDs: 4.36 1.54 4.30 10.12 9.47 1.91		(Carbas et al., 2020)

	Quercetin-3-6''-manolyl-glucoside Kaempferol-3-glucoside Myricetin Kaempferol-3-6''-manolyl-glucoside Kaempferol							0.73 0.38 0.35 0.39 0.84	4.25 1.23 1.81 1.79 7.03		
Soybean	Total anthocyanins Cyanidin-3-glucoside Delphinidin-3-glucoside	70/-	4000-650	ATR (whole seeds)	PLS	Spectra of samples included in calibration set	n/s	0.86 0.88 0.87	0.15 0.13 0.03 mg/g	70 different varieties	(Amanah et al., 2020)
Oils											
Olive oil	Fatty acid methyl esters Fatty acid ethyl esters Fatty acid alkyl esters Diacylglycerols: C34 1,2 C34 1,3 C36 1,2 C36 1,3 Pheophytin *a* Chlorophyll *a* Pheophytin *b* Total xanthophyll Lutein Chlorophyll *b*	59/30	4000-650	ATR	PLS	Dependent test set	n/s	0.87 0.85 0.87 0.62 0.83 0.79 0.77 0.72 0.75 0.71 0.61 0.75 0.72	41.63 27.43 60.10 1.07 1.26 4.29 4.02 2.42 0.32 0.10 0.41 0.71 0.21 mg/kg	Samples from 2 seasons. Quite poor test set validation results for colour pigments. Improved results from fusion of UV-Vis and IR spectra	(Uncu et al., 2019)
Olive oil	Tyrosol Tyrosol secoiridoids Hydroxytyrosol Hydroxytyrosol secoiridoids Total phenolics	75/18	4000-400	ATR	PLS	None	LOO	0.32 0.30 0.17 0.19 0.44	4.98 105.7 9.96 106.1 162.1 mg/kg		(Mora-Ruiz et al., 2017)

Olive oil	Total phenolics	70/30	4000-600	ATR	PLS	Dependent test set	n/s	0.998	0.072 g/L		(Hirri et al., 2016)
Beverages											
Cachaça	Total phenolics	32/16	4000-650	ATR (liquid sample)	PLS	Ranked subset of samples (60% cal; 20% val; 20% test set)	n/s	0.820	248 mg/L Test set: R^2 = 0.690; RMSE = 318 mg/L	Much poorer results than fluorescence spectroscopy	(Carvalho et al., 2020)
Grape juice	Total phenolics Anthocyanins	49/16	4000-400	ATR	PLS	Dependent test set (randomly selected samples)	Optimising no. of latent variables	0.90 0.81 (cal)	21 4.22 mg/100mL (test set)	Performed better than NIR for phenolic content	(Caramês et al., 2017b)
Shiraz wine	Total anthocyanins Total phenolics	70/30	1700-950	ATR (liquid sample)	PLS	Dependent test set	LOO	0.61 0.60	32 mg/L 5.7 au	Wines from 24 different Australian locations	(Ristic et al., 2016)
Wine (red)	Gallic acid Catechin B1 Polymeric phenols Caftaric acid Caffeic acid Coutaric acid p-coumaric acid Quercetin-3-glucoside Quercetin Kaempherol Delphinidin-3-glucoside Cyanidin-3-glucoside Petunidin-3-glucoside Peonidin-3-glucoside Malvidin-3-glucoside Delphinidin-3-acetylglucoside	~387/182	4000-600	ATR (liquid sample)	PLS using PRESS	Dependent test set	Segment validation	0.83 0.78 0.8 0.85 0.85 0.86 0.85 0.81 0.85 0.69 0.82 0.88 0.76 0.86 0.84 0.85 0.86	3.42 7.26 4.99 128 9.76 1.07 3.14 0.63 13 2.56 0.34 2.96 0.06 2.57 1.47 20.7 1.44	Wines comprised four cultivars from 13 vinifications over 2 seasons. Slightly less accurate at phenolic content compared to FT-NIR	(Aleixandre-Tudo et al., 2018)

	Cyanidin-3-acetylglucoside Petunidin-3-acetylglucoside Peonidin-3-acetylglucoside Malvidin-3-acetylglucoside Delphinidin-3-cumarylglucoside Petunidin-3-cumarylglucoside Peonidin-3-cumarylglucoside Malvidin-3-cumarylglucoside Polymeric pigments MCP tannins Anthocyanins							0.85 0.88 0.89 0.89 0.85 0.85 0.85 0.85 0.85 0.89 0.86	0.48 1.19 1.03 6.49 0.41 0.79 0.91 3.98 5.5 261 47.2 mg/L		
Wine (red)	Gallic acid Catechin B1 Polymeric phenols Caftaric acid Caffeic acid Coutaric acid p-coumaric acid Quercetin-3-glucoside Quercetin Kaempherol Delphinidin-3-glucoside Cyanidin-3-glucoside Petunidin-3-glucoside Peonidin-3-glucoside Malvidin-3-glucoside	~387/182	4000-600	Transmission	PLS using PRESS	Dependent test set	Segment validation	0.85 0.85 0.84 0.91 0.87 0.86 0.84 0.83 0.82 0.84 0.87 0.84 0.82 0.88 0.85 0.86	4.57 5.39 3.91 132 9.87 1.02 2.8 0.625 13 1.59 0.328 4.15 0.0645 4.1 2.1 24.2	Wines comprised four cultivars from 13 vinifications over two seasons. Slightly less accurate at phenolic content compared to FT-NIR	(Aleixandre-Tudo et al., 2018)

	Delphinidin-3-acetylglucoside Cyanidin-3-acetylglucoside Petunidin-3-acetylglucoside Peonidin-3-acetylglucoside Malvidin-3-acetylglucoside Delphinidin-3-cumarylglucoside Petunidin-3-cumarylglucoside Peonidin-3-cumarylglucoside Malvidin-3-cumarylglucoside Polymeric pigments MCP tannins Anthocyanins							0.84 0.86 0.88 0.9 0.84 0.88 0.85 0.87 0.85 0.82 0.92 0.89	1.28 0.513 1.24 1.12 8.85 0.463 0.831 1.01 4.7 7.49 224 56.5 mg/L		
Wine (red, rose and white)	Total phenolics Total anthocyanins	35/-	4000-650	ATR	PLS	Cross-validation only	LOO	0.91 0.86	269.2 1.79 mg/L	7 wines, each at 5 different processing points	(Canal & Ozen, 2017)
Wine (red and white)	Total polyphenols Malvidin-3-O-glucoside Peonidin-3-O-glucoside Petunidin-3-O-glucoside Delphinidin-3-O-glucoside Delphinidin-3-O-(6-acetyl)-glucoside	51/21	4000-650	ATR	PLS	Dependent test set	LOO	0.75 0.53 0.56 0.67 0.71 0.27	249.1 6.87 0.38 1.1 0.73 0.15	Samples from various locations across 2 seasons	(Sen et al., 2016)

	Petunidin-3-O-(6-acetyl)-glucoside Peonidin-3-O-(6-acetyl)-glucoside Malvidin-3-O-(6-acetyl)-glucoside Delphinidin-3-O-(6-p-coumaroyl)-glucoside Malvidin-3-O-(6-p-coumaroyl)-glucoside o-coumaric acid							0.29 0.31 0.41 0.45 0.69 0.63	0.24 0.22 2.59 0.12 0.65 0.33 mg/L		
Other foods											
Chocolate	(+)-catechin (+)-epicatechin Total phenolics TAC (DPPH) TAC (ORAC)	18/7	4000-550	ATR	PLS	Semi-independent (7 randomly selected commercial chocolate brands) R^2 = 0.86, 0.72, 0.88, 0.89, 0.90; RMSEP = 0.10, 0.57, 5.08, 13.07, 37.92 mg/g	9-fold cross validation	0.94 0.87 0.93 0.92 0.89	0.09 0.58 4.21 1.05 11.38 mg/g	18 different types of chocolate containing 35-100% cacao	(Hu et al., 2016)
Honey	Catechin Syringic acid Vanillic acid Chlorogenic acid TAC (DPPH)	64/36	3000-2800, 1800-700	ATR	PLS	Dependent test set (ranked subset of samples)	LOO	0.999 0.992 0.946 0.994 0.955	0.40 1.08 0.45 0.43 µg/g 1.63 mg/100g	Models based on Raman spectra were slightly better than FTIR	(Tahir et al., 2017)

Abbreviations: LOO = leave-one-out cross-validation; n/s = not specified; SFA = saturated fatty acids; MUFA = monounsaturated fatty acids; PUFA = polyunsaturated fatty acids; TAC = total antioxidant capacity

2.5.1 General trends

2.5.1.1 Publications by year

Between 2016 and 2020, an average of 10 studies per year were published on the use of infrared spectroscopy for the measurement of bioactive compounds in food products. The number of studies published per year over this period remained relatively constant, although only 3 published studies were found for 2019 (Table 2-4).

Table 2-4: Number of published papers reporting the use of infrared spectroscopy for the quantification of bioactive compounds in food products between 2016-2020. Note that duplicated references from Tables 2-2 and 2-3 (i.e., studies using both NIR and MIR spectroscopy on the same matrix) were only counted once.

Year	Number of published studies
2016	13
2017	12
2018	10
2019	3
2020	11

2.5.1.2 Matrix type

Interrogation of the included studies by matrix type revealed that NIR spectroscopy was most commonly used for the analysis of bioactive compounds in fruit matrices, followed by aromatic plants, grains/pulses and beverages (Table 2-5). In contrast, MIR spectroscopy was most often reported for the analysis of beverages, likely due to the ease of presentation for this sample type.

Table 2-5: Number of studies included in this review, broken down by matrix type. If the same study used both NIR and MIR spectroscopy, it was counted separately in each column.

Matrix type	Number of published studies	
	NIR	MIR
Fruit	18	1
Vegetables	3	1
Grains/pulses	6	3
Oils	4	3
Aromatic plants	7	0
Beverages	6	7
Others	2	2
Total	**46**	**18**

2.5.1.3 Optical geometry

The majority of publications (58%) using NIR spectroscopy for the prediction of bioactive compounds used reflectance or diffuse reflectance geometry. A further 16% of studies used hyperspectral imaging in reflectance mode. Only 20% of studies used transmittance and 9% used transflectance, the majority of which were performed on beverage or oil samples. However, it should be cautioned that the vast majority of studies were not validated through independent test set validation and hence have not shown their robustness in "real-world" use; consequently, the optical geometry types used in the academic studies reported here may not reflect the optical geometry of instruments used commercially.

All of the MIR spectroscopy studies except one (Aleixandre-Tudo et al., 2018) used an Attenuated Total Reflection (ATR) sampling platform, which requires samples to be placed in close contact with the ATR crystal.

2.5.1.4 Sample size and test sets

The number of calibration samples ranged from 10 to 387 (mean = 83 ± 72 samples), while the size of the validation set ranged from 5 to 182 (mean = 37 ± 32 samples). The majority of studies used a dependent test set (65%) or did not use any test set (24%), while only 9% of studies used an independent test set for validation of the developed model.

Within the four NIRS studies utilising an independent test set, one used transmittance (Tilahun et al., 2018), while the others used reflectance (Cunha Júnior et al., 2016; Ncama et al., 2018) or diffuse reflectance geometry (Cozzolino et al., 2020). Cunha Júnior et al. (2016) sourced

their test set from the following season to the calibration set, while Ncama et al. (2018) used test set samples from a geographically distinct farm (~400 km away) and Cozzolino et al. (2020) used commercially sourced samples for validation purposes. The study by Tilahun et al. (2018) could arguably be classified as using a semi-independent test set, as the authors used samples from a different harvest time point within the same season and from the same location. Interestingly, all four of these NIRS studies were performed on fruit rather than other food matrices.

In most of these studies, the test set validation statistics were moderately poorer compared to the cross-validation statistics. For example, Cunha Júnior et al. (2016) found an R^2_{CV} of 0.89-0.91 and RMSECV of 2.5-2.9 g/kg for the prediction of total anthocyanin content in açaí and juçara fruit, compared to an R^2_{test} of 0.74-0.88 and RMSEP of 5.1-6.8 g/kg. Similarly, the RMSEP for the prediction of lycopene content in tomato fruit was moderately higher at 1.79 mg/kg compared to the RMSECV of 1.56 mg/kg (Tilahun et al., 2018). However, the performance of the test set from Cozzolino et al. (2020) was much worse, with an R^2 of 0.73 and RMSEP of 4733 mg/100 g (compared to an R^2 of 0.93 and RMSECV of 1839 mg/100 g for cross-validation).

Using MIRS for the analysis of chocolate samples, Hu et al. (2016) found that the test set statistics for the prediction of (+)-catechin, (+)-epicatechin and total phenolics in chocolate using MIRS were quite comparable to the cross-validation statistics. However, the RMSEP for prediction of total antioxidant capacity (TAC) in the same samples was 3-12 times higher than the RMSECV, suggesting that MIRS was not suitable for the accurate estimation of TAC in this matrix. These few examples illustrate the level of over-optimistic results which are likely to be reported when using no test set or a dependent test set for model validation.

2.5.1.5 Chemometric techniques

Virtually all of the publications used partial least squares regression (PLSR) or some derivative of this regression technique for model development. Tschannerl et al. (2019) used Support Vector Regression (SVR) for the prediction of total phenolic content in barley malt samples using hyperspectral imaging. However, only 10 samples were investigated in that study, with no independent test set used. Zhang et al. (2017) also used SVR for the prediction of phenolic content in wine grape skins and seeds from their hyperspectral images, demonstrating that for most analytes, the use of SVR gave better results than PLSR or principal component regression (PCR). Xiao et al. (2018) used a Least Squares Support Vector Machine (LS-SVM) algorithm for the prediction of total phenolics in white and red grapes, again with better results found compared to the standard PLSR algorithm. Finally, Ding et al. (2016) compared the use of Radial Basis Function Neural Networks (RBF-NN) and PLSR in dehydrated tomato samples,

finding that RBF-NN performed better for the lycopene, total phenolic content and total antioxidant capacity measured by the DPPH and ABTS assays, while PLSR performed better for the prediction of total antioxidant capacity via the FRAP method. No studies were found using deep learning or ANN algorithms, although this is a matter of increasing interest for other areas of IR spectroscopy (Anderson et al., 2021; Beć et al., 2021).

2.5.2 Trends by analyte class

Another major aspect of interest to researchers is the types of bioactive analyte(s) that have been measured using IR spectroscopy. Consequently, Table 2-6 presents a break-down of the studies included in this review by the compound class of the reported analytes. Additionally, the major classes are discussed in the following sections.

Table 2-6: Number of studies included in this review, broken down by analyte class. Note that if the same study investigated multiple matrices or investigated more than one analyte class in the same matrix, it was counted separately.

Analyte class	Number of published studies
Total polyphenol content^	34
Specific polyphenols^	21
Total anthocyanin content	13
Specific anthocyanins	4
Total carotenoid content	2
Specific carotenoids (β-carotene, lycopene)	6
Ascorbic acid (vitamin C)	6
Alkaloids (theobromine, caffeine, piperine)	4
Fatty acid esters & other bioactive hydrocarbons	2
Chlorophylls	1
Tocopherols	1
Total	**94**

^ polyphenols includes phenolic acids and flavonoid derivatives

2.5.2.1 Polyphenols

The greatest number of studies examined for the purpose of this current review were focused on predicting the total polyphenol content, or the content of specific polyphenol compounds present in the matrix (Table 2-6), with over half of all investigations focused on these analytes.

There is an ongoing interest in biochemical characterisation and quantification of polyphenols across a wide range of food products, given that compounds from this class have been associated with a wide range of potential health-benefiting effects (Alves-Santos et al., 2020; Cassidy et al., 2020; Johnson et al., 2021a; Koch, 2019; Rasouli et al., 2017), particularly in improving cardiovascular health (Costa et al., 2017; Ed Nignpense et al., 2020; González Arbeláez et al., 2018; Rasines-Perea & Teissedre, 2017; Sanches-Silva et al., 2020). Consequently, the rapid prediction of total polyphenol content using infrared spectroscopy could have potential to greatly benefit the effectiveness and robustness of the quality assurance process for functional food products (Johnson et al., 2020f; McGoverin et al., 2010).

Ferrer-Gallego et al. (2020) provided a recent review of the use of vibrational spectroscopy in the prediction of the phenolic composition of grapes and wines, although other food matrices were not considered in that review. The authors considered that this technique showed considerable promise for this purpose, although noted that future studies on grapes and wine should incorporate a wider range of environmental and genotypic variation.

Some authors have reported difficulty in creating robust models for the prediction of total polyphenols using infrared spectroscopy. For example, Martín-Tornero et al. (2020) found that NIRS and MIRS could only be used as a screening method for the total polyphenol content in grape leaves, due to the high prediction errors associated with the models created. These authors used a dependent test set (leaves collected from different locations within the same vineyards). In blackberry fruit, the best model for total phenolics reported by Toledo-Martín et al. (2018) had a R^2_{cv} of 0.69 and RMSECV of 169 mg/100 g. Again, the cross-validation samples used in this study were randomly selected from the same population as the calibration samples; consequently, the model performance on an independent population would be lower again. Similar results in terms of model accuracy were found by Rodríguez-Pulido et al. (2017) in raspberries, Trapani et al. (2017) in olive paste and Hernández-Hernández et al. (2021) in cocoa bean, while quite poor cross-validation results were found by Nogales-Bueno et al. (2020) for the prediction of total phenolic content (TPC) in coffee bean using NIR hyperspectral imaging. As the mean TPC of the samples was 3.6% w/w, the poor performance appears more attributable to the reproducibility of sample presentation or the wavelength selection, rather than the concentration of the analyte.

In contrast, Tzanova et al. (2020) and Jara-Palacios et al. (2016) reported quite good findings for the prediction of total polyphenol content in grapes and grape pomace, respectively (R^2_{cv} = 0.87-0.97; RMSECV = 9.6-21 mg/100 g), indicating that the instrument choice, geometry and data processing techniques may have an influence in addition to the matrix type. However,

it is important to note that none of the aforementioned studies on the prediction of total phenolic content used an independent test set; therefore, the results should be taken with caution.

There do not appear to have been any studies that focused on the IR quantitation of specific phenolic compounds or total phenolic content in model systems; hence it is difficult to know what limit of detection and level of error to expect when using IR spectroscopy for this purpose. Although Abbas et al. (2017) used MIRS for the qualitative identification of 36 phenolic compounds (presented in powder form), they did not attempt the quantitation of these compounds in a model matrix.

2.5.2.2 Anthocyanins

The second-most common analyte type that has been investigated using infrared spectroscopy was anthocyanins. Most of these studies (13 out of 17) looked at the total anthocyanin content, while only 4 studies attempted the prediction of specific anthocyanins. As a class of flavonoids, anthocyanins are less abundant than total polyphenols, so would be expected to be a more challenging target for infrared spectroscopy. Anthocyanins are brightly coloured and absorb light at around 520 nm; hence it may be thought that they could be detected using the visible wavelengths of Vis-NIR instruments. However, surprisingly, all except one of the studies using NIRS for the measurement of anthocyanins did not include the visible light region in the optimised models, indicating that the infrared region actually contained most of the functional information pertaining to the anthocyanin content. Given the low concentration of anthocyanins, their prediction through NIRS is likely to rely upon secondary correlations with other matrix constituents.

Most studies using NIRS reported reasonably high accuracies for anthocyanin prediction in fresh sample matrices (R^2_{cv} = 0.72-0.98; RMSECV = 9-13 mg/100 g), while MIRS performed similarly well for the estimation of anthocyanin content in soybean, grape juice and red wine. Rodríguez-Pulido et al. (2017) found a lower model linearity using NIRS in raspberry fruit (R^2_{cv} = 0.63), although the RMSECV obtained was roughly comparable at 12 mg/100 g.

Studies attempting the prediction of individual anthocyanins in red grapes (Diago et al., 2016) and wine (Aleixandre-Tudo et al., 2018; Sen et al., 2016) reported that the concentrations of most of these compounds could be predicted with only slightly lower accuracy compared to the total anthocyanin content. Given the very low concentrations of many of these compounds, it is likely that the created models were indirectly measuring their concentration via their secondary correlations with more abundant compounds which are more readily detected using infrared spectroscopy (possibly the predominant individual anthocyanin compounds present in the sample). Somewhat confusingly, many of the studies reported the anthocyanin content in units of mg/L of the sample extracts, rather than being correctly reported in mg/g or mg/100

g of the intact fruit from which the infrared spectra were obtained. Hence the results of these models should be interpreted with some degree of prudence. Future researchers in this area should be aware of and avoid this common pitfall.

2.5.2.3 Carotenoids

In contrast to the trends observed for anthocyanins, studies investigating carotenoids using infrared spectroscopy mainly attempted the prediction of specific carotenoid compounds (β-carotene, lycopene) rather than predicting the total carotenoid content. In addition, all of the studies attempting carotenoid prediction were performed using NIRS.

Several studies in intact fresh tomato fruit reported similar results for the prediction of lycopene (R^2_{cv} = 0.85-0.86; RMSECV = 18-103 mg/100 g FW) and β-carotene (R^2_{cv} = 0.71-0.77; RMSECV = 10-114 mg/100 g FW) (Alenazi et al., 2020; Tilahun et al., 2018). Similar results were found for lycopene in dried tomato powder (Ding et al., 2016).

Toledo-Martín et al. (2018) also found acceptable results for the total carotenoid content in blackberry (R^2_{cv} = 0.76, RMSECV = 0.01 mg/100 g), with the carotenoid model outperforming that developed for total phenolic content in the same crop. Higher model accuracies (R^2_{cv} > 0.9; RMSECV <0.01 mg/100 g) were reported for β-carotene content in carrot (Rady et al., 2018) and marsh grapefruit (Ncama et al., 2018), as well as for total carotenoids in honey (Tahir et al., 2016).

2.5.2.4 Ascorbic acid

Studies using infrared spectroscopy (NIRS or MIRS) for the estimation of ascorbic acid content were performed in Kakadu plum powder (Cozzolino et al., 2020), carrot (Rady et al., 2018), frozen guava pulp (Alamar et al., 2016), cashew apple and guava nectar (Caramês et al., 2017a), and soft drinks (Santana et al., 2020). Most models showed reasonable accuracy (R^2_{cv} = 0.7-0.98; RMSECV = 4-7 mg/100 g). Due to the exceptionally high ascorbic acid content in Kakadu plum (mean content of 14,323 mg/100 g), the RMSECV values of Cozzolino et al. (2020) were much higher at 1811-1839 mg/100 g. The model linearity was quite high (R^2_{cv} = 0.91-0.93), with an RPD of 4.0-4.1, although the independent test set validation (comprising commercially purchases samples of Kakadu plum powder) gave a considerably poorer RMSEP (4733 mg/100 g). All of the other aforementioned studies did not validate their models using independent test sets, but only used dependent test sets (comprising randomly selected samples from the full dataset).

Cozzolino et al. (2020) was also the only study to compare the performance of NIRS and MIRS for predicting ascorbic acid content, finding slightly improved accuracy of NIRS compared to MIRS in dried Kakadu plum powder.

2.5.2.5 Other analytes

Other bioactive compounds assessed using infrared spectroscopy included chlorophylls, fatty acid esters, squalene and tocopherols (compounds related to vitamin E) in olive oil, piperine in black pepper, caffeine in black tea, and theobromine in cocoa bean. In general, good results were generally found for caffeine, and most tocopherols and fatty acids, while reasonable results were found for theobromine, squalene, chlorophylls and piperine. It should be noted that most of these analytes were only investigated in a single study. Nevertheless, these results support the use of infrared spectroscopy as a highly adaptable tool for the rapid estimation of a substantially wide range of bioactive compounds in food-based matrices.

2.5.3 Future directions

As found throughout this review, infrared spectroscopy shows considerable potential for the quantification and relative prediction of the levels of bioactive components in food products. Although most research to date has been proof-of-concept work and/or conducted under controlled laboratory conditions, interest and applications in this field are likely to continue to grow. A brief discussion on several particular aspects worth noting is provided here.

Hyperspectral imaging is a rapidly growing area of research in the food science sector, particularly for the determination of food quality and safety (Caporaso et al., 2021; Khan et al., 2020; Lu et al., 2020; Wang et al., 2021a), but also for authentication purposes (Feng et al., 2021; Temiz & Ulaş, 2021). This technique collects near-infrared spectra from each pixel in a photograph (creating a 'hypercube' dataset), allowing for analysis of the spatial variation of the analyte, as well as its mean concentration. Consequently, hyperspectral imaging could potentially be used for the quantification of bioactive compounds (Kiani et al., 2018). Indeed, several of the studies reviewed here applied hyperspectral imaging for the estimation of anthocyanins and phenolic acids in grapes and grape byproducts (Diago et al., 2016; Jara-Palacios et al., 2016; Martínez-Sandoval et al., 2016; Zhang et al., 2017), and for the estimation of phenolics in barley malt (Tschannerl et al., 2019) and coffee beans (Nogales-Bueno et al., 2020). However, there have been limited applications of hyperspectral imaging systems in industrial applications to date, due to its associated challenges such as obtaining reproducible sample presentation, minimising the effects of ambient light, and the complexity of data analysis (Khan et al., 2020). Furthermore, hyperspectral imaging can only be used with a reflectance geometry. Finally, the cost of these instruments remains quite high compared to regular NIR instruments; thus hyperspectral imaging tends to only be used in applications which have a need for spatial information.

The use of infrared spectroscopy as a real-time, online (or "inline") process analytical technology is another principal area of interest. NIRS is commonly used in manufacturing

environments and processing plants for the online analysis of a range of food products, principally for the determination of proximate quality parameters such as moisture content, soluble solids and protein (Agbonkonkon et al., 2021; Torres et al., 2020). This real-time information can then be fed back into the manufacturing system, allowing various processing parameters to be adjusted accordingly in view of maintaining the optimal quality of the product. Online NIRS could potentially be extended to the quality assurance of bioactive compounds in functional food production systems, in addition to existing analytes already being monitored.

Finally, it is worth noting the importance of confirming the accuracy and reproducibility of infrared spectroscopy techniques using sufficiently large sample sizes and test sets which are independent to the calibration sets. Given that only a small fraction of the studies reviewed here used a fully independent test set for model validation purposes, it is likely that the reported accuracy is over-optimistic in many instances and not representative of the true accuracy which could be expected if applying the model for routine quality assurance purposes.

2.6 Summary

The technique of infrared spectroscopy has enjoyed considerable success in the food analysis industry over the past few decades. In recent years, an increasing number of studies are exploring the use of this technology for the analysis of bioactive compounds in food products, such as polyphenols, anthocyanins or carotenoids. While much reported work is still in the proof-of-concept or method development stage, infrared spectroscopy appears to show promise for the relative assessment – if not absolute quantification – of these bioactive analytes.

Chapter 3. Faba bean

This chapter uses information from three previously published papers:

Johnson, J.B., Collins, T., Skylas, D., Quail, K., Blanchard, C. and Naiker, M., 2020. Profiling the varietal antioxidative contents and macrochemical composition in Australian faba beans (*Vicia faba* L.). *Legume Science*, *2*(2), p.e28. DOI: 10.1002/leg3.28

Johnson, J.B., Walsh, K. and Naiker, M., 2020. Application of infrared spectroscopy for the prediction of nutritional content and quality assessment of faba bean (*Vicia faba* L.). *Legume Science*, *2*(3), p.e40. DOI: 10.1002/leg3.40

Johnson, J.B., Skylas, D.J., Mani, J.S., Xiang, J., Walsh, K.B. and Naiker, M., 2021. Phenolic Profiles of Ten Australian Faba Bean Varieties. *Molecules*, *26*(15), p.4642. DOI: 10.3390/molecules26154642

3.1 Introduction

Following on from the promising literature results discussed in Chapter 2, the first grain crop chosen for investigation was faba bean. This crop is well-known for its high levels of phenolic compounds and health-benefiting properties. Furthermore, Australia is the largest exporter of faba bean worldwide, making it of considerable economic significance. With increasing interest in the health benefits of this crop, there is potential for Australian growers/wholesalers to distinguish the high quality of Australian faba bean through marketing it with an emphasis on the levels of bioactive compounds present. However, this would necessitate (1) background information on the typical levels of phenolics and antioxidants found in Australian-grown faba bean, and (2) cost-effective methods of assessing the levels of these compounds in this crop.

This work aimed to provide some of this missing information through the analysis of a large number of faba bean samples (n=100) grown in controlled field trials. The samples investigated were comprised of ten different commercial faba bean varieties, grown in Pulse Breeding Australia (PBA) field trials across two growing sites (Charlick and Freeling, South Australia) and two seasons (2016 and 2017). This ensured that a wide range of genetic material and environmental variation was sampled.

This work had three major aspects:

- Firstly, the general phytochemical composition of the faba bean samples were assessed using benchtop spectrophotometric methods. These methods were used to measure the total phenolic content, total antioxidant capacity and total monomeric

anthocyanin content of the samples. This was conducted on all samples from 2016 and 2017 (n=100).

- To further investigate selected phenolic compounds present in the faba bean crop, high-performance liquid chromatography coupled with diode array detection (HPLC-DAD) was used to identify and quantify 10 phenolic compounds present. Due to the additional time-consuming sample concentration required, only the 60 samples from the 2017 season were analysed by HPLC-DAD.
- Finally, NIR and MIR spectra were collected from all faba bean flour samples (n=100) and analysed. This included qualitative analysis (identification of variety/growing site/season) and quantitative analysis (prediction of bioactive compound contents using PLSR models).

3.2 Background

Increased consumption of pulses world-wide is driven in part by growing consumer demand for new foods with enhanced nutrition and health benefits. Pulses such as faba bean (*Vicia faba* L.) have significant potential in the development of value-added foods and ingredients (López-Barrios et al., 2014; Vioque et al., 2012). Globally, faba bean is reported to be the third most important legume crop, with over 5.4 million tonnes harvested annually (Rahate et al., 2021).

Domestically, faba bean is grown in South Australia, Victoria, Western Australia, New South Wales and southern Queensland (Siddique et al., 2000). With annual production of around 300,000 tonnes, faba bean comprises 10-15% of the total pulse production in Australia (AEGIC, 2017). Australia is the third-largest producer worldwide and the largest exporter, responsible for one-third of international trade for this crop (AEGIC, 2017). Primary importers include the Middle East, particularly Egypt, and South East Asian countries (AEGIC, 2017).

Recent years have seen increasing interest in faba bean due to its nutritional content (Rahate et al., 2021; Sharan et al., 2021) and health-benefiting properties (Turco et al., 2016). In addition to containing high levels of protein and nearly all essential minerals (AEGIC, 2017), the reported health benefits of this crop include improving cardiovascular health (Siah et al., 2012), inhibiting xanthine oxidase activity (Spanou et al., 2012) and providing anti-obesity effects (Jakubczyk et al., 2019; Siah et al., 2012), anti-cancer activity (Siah et al., 2012) and anti-inflammatory activity (Boudjou et al., 2013). These beneficial biological activities are largely linked to the high levels of antioxidant and phenolic compounds found in faba bean (Boudjou et al., 2013; Siah et al., 2012; Turco et al., 2016). This has led to an interest in using faba bean or its isolates in functional food applications (Sharan et al., 2021).

However, it is important to note that genotypic variations can influence the phenolic and flavonoid biosynthetic pathways (Zanotto et al., 2020), resulting in significant variation in the phenolic contents of different grain/pulse varieties (Xiang et al., 2019c). Consequently, there has been recent interest in identifying faba bean varieties with high levels of phenolic content. For example, Valente et al. (2018) and Valente et al. (2019) profiled the phenolic content and antioxidant capacity of seven European faba bean varieties, finding that the levels of total and individual phenolic acids and flavonoids differed significantly between varieties. Similarly, Baginsky et al. (2013) found clear differences in the phenolic composition of 10 faba bean varieties grown in Chile, although it should be noted that this study was performed on immature seed material. Another earlier study highlighted the range in total phenolic contents and antioxidant activity among 13 Tunisian faba bean cultivars (Chaieb et al., 2011). However, despite Australia's international importance as a faba bean producer, there are few comparative studies reporting the phenolic contents of commercial faba bean varieties grown in this country. This is of particular significance as Australian pulse breeders have focused on selecting varieties for disease resistance and elevated yield (Siddique et al., 2000), without considering the impact these selections could have upon the phytochemical composition of the new varieties (Wrigley et al., 2019).

Nasar-Abbas et al. (2009) reported on the phenolic compounds found in one Australian faba bean variety, while Siah et al. (2014a) and Siah et al. (2012) investigated two and three varieties, respectively. In the largest study to date on the phytochemical composition of Australian faba bean, Siah et al. (2014b) examined the phenolic content and antioxidant activity in five Australian faba bean varieties.

In terms of non-invasive assessment, there have been a number of studies that have previously applied NIRS for the quality analysis of the faba bean crop (Johnson et al., 2020e), including the analysis of protein (El-Sherbeeny & Robertson, 1992; Williams et al., 1978), starch and oil (Wang et al., 2014a), tannins (De Haro et al., 1988) and total polyphenol content (Wang et al., 2014a). However, there have only been a limited number of studies investigating the prediction of bioactive compounds using this technique, and none investigating MIRS for this purpose.

Consequently, the first aspect of this work was to investigate and compare the total phenolic, antioxidant and anthocyanin content of a selection of Australian-grown faba bean samples. This allowed examination of the impacts of genotype, growing location and season upon the broad phytochemical composition of this crop. Subsequently, individual phenolic compounds were identified and quantified in a subset of these faba bean samples using HPLC-DAD.

Finally, infrared spectroscopy (NIRS and MIRS) was applied for the prediction of bioactive compounds in the faba bean matrix.

The results provide further insight into the nutritional and health-benefiting properties of common Australian faba bean cultivars, as well as providing valuable information on the extent of their genotypic variation present in terms of the phenolic acid and flavonoid biosynthesis pathways. Additionally, they highlight the prospect of using infrared spectroscopy for the quality assurance of faba bean.

3.3 Materials and methods

3.3.1 Faba bean samples

The faba bean seed material used in this study was sourced from a multi-environment field trial conducted as part of the Pulse Breeding Australia (PBA) Southern Node breeding program, via the Australian Export Grains Innovation Centre (AEGIC). The growing and harvest conditions are described in detail by Skylas et al. (2019). Briefly, 10 commercial faba bean varieties were grown at two different locations in South Australia: Charlick (35°19'43"S, 138°52'44"E) and Freeling (34°27'20"S, 138°47'17"E), over two consecutive growing seasons (2016 and 2017). These four environmental conditions are abbreviated here as 16Char, 17Char, 16Free and 17Free.

The 10 varieties included in these growing trials account for the bulk of domestic production of the faba bean crop (Pulse Australia, 2016b). Key features of each variety are provided in Table 3-1, including their origin, release date and disease resistance.

Two field replicates were analysed (each in duplicate) for each of the 2016 sites and three field replicates were analysed (each in duplicate) for each of the 2017 sites.

The seed samples were coarsely ground by AEGIC staff using a Van Gelder grinder with a 3 mm screen (J.P. Van Gelder and Co., Sydney), before being impact-milled to a fine flour using a Falling Number grinder with 0.8 mm screen (Falling Number AB, Stockholm). The moisture content of the samples was measured following AACC International Method 44-15.02 (AACC International, 1975). Prior to analysis, the flour samples were stored in the dark at 4°C.

Data supplied by AEGIC for each of the faba bean samples included their protein, total starch, amylose, amylopectin, vicine and convicine contents. The methods used for the determination of each of these analytes are described in detail by Skylas et al. (2019). Briefly, protein was measured by the Dumas combustion method (LECO TruMac N protein analyser; Saint Joseph, MI, USA), starch by AACC International Method 76-13.01, and amylose by AACC International Method 61-03.01. The amylopectin content was calculated as the difference between starch

and amylose contents. Vicine and convicine were measured by liquid chromatography-mass spectrometry (LC-MS); the method is provided in Skylas et al. (2019).

Table 3-1: Faba bean varieties included in this study.

Variety	Release	Seed size	Germplasm Origin	Disease resistance					
				Asc(f)	Asc(s)	CS	Rust	PSbMV	BLRV
Fiord	1980	Small	Greece	mS	mS	vS	S	S	-
Fiesta VF	1998	Medium	Spain	mS/mR	mR/R	S	S	S	-
Farah	2003	Medium	Spain	mR/R	mR/R	S	S	S	-
Nura	2005	Medium	Ecuador, Greece	mR/R	mR/R	S	mS	vS	-
Doza	2008	Medium	Ethiopia, Sudan	vS	vS	mS	mR/R	-	-
PBA Rana	2011	Large	Ecuador, Lebanon	R	R	mS	mS/mR	mR/R	-
PBA Warda	2012	Large	Ecuador, Greece	vS/S	vS/S	mS	mR/R	-	mT
PBA Samira	2014	Medium	Lebanon, UK, Spain, Ecuador, Greece	R	R	mS	mS	S	-
PBA Zahra	2015	Large	Morocco, Spain	R	-	mS	mS	S	-
PBA Nasma	2015	Large	China, Sudan, Italy	vS/S	vS/S	mS	mS	-	mT

Abbreviations: Asc = ascochyta blight; f = foliage; s = seed; CS = chocolate spot; PSbMV = pea seed-borne mosaic virus; BLRV = bean leafroll virus; R = resistant; S = susceptible; T = tolerant; m = moderately; v = very
A dash (-) indicates resistance is unknown
References: Pulse Australia (2016b), Skylas et al. (2019).

3.3.2 Reagents

All reagents used were of analytical grade. Methanol, hydrochloric acid and sodium carbonate were purchased from Chem Supply (Gillman, South Australia). All other reagents were purchased from Sigma-Aldrich Australia (Castle Hill, New South Wales). Unless otherwise specified, all dilutions and assay preparations were made using Milli-Q water (Merck; Bayswater, Victoria). All solutions were stored in the dark at 4°C until usage.

3.3.3 Extraction of polar phenolic compounds

Extracts were prepared in duplicate for each sample. Firstly, approximately 0.5 g of faba bean flour was combined with 8 mL of 90% v/v aqueous methanol and vortexed for 10 seconds (Ratek VM1 vortex mixer). Following this, the extracts were mixed for 60 minutes using an end-over-end shaker (Ratek RM4) operating at 50 rpm. After centrifugation at 1000 rcf for 10

minutes (Heraeus Multifuge; Thermo Fisher Scientific), the supernatant was collected. To extract any remaining phytochemicals, the extraction process was repeated with another 8 mL of 90% methanol added to the pellet, followed by end-over-end mixing for 20 minutes. The combined supernatants were finally made up to 20 mL volume using 90% methanol. Extracts were stored in the dark at 4°C until required for analysis.

Although previous researchers have investigated both free and bound phenolic fractions in other pulse crops (Xiang et al., 2019c), only the free phenolic fraction was investigated in the present study. The principal reason for this is that bound phenolics (i.e., those bound to cell wall structures) have reduced bioavailability (Manach et al., 2004); thus the free phenolic compounds are likely to be the primary contributors to any observed beneficial health effects.

3.3.4 Analysis of TPC

Total phenolic content (TPC) was determined through a modification of the Folin-Ciocalteu method developed by Singleton and Rossi (1965). Firstly, 2 mL of a 1:10 aqueous dilution of Folin-Ciocalteu reagent was combined with 400 µL of sample extract in a 10 mL centrifuge tube. The samples were incubated at room temperature in darkness for 10 minutes before 2 mL of 7.5% w/v aqueous sodium carbonate was added. The mixture was then vortexed for 10 seconds, incubated at 40°C for 30 minutes in a covered water bath, and vortexed for another 10 seconds. The absorbance at 760 nm was then measured using a UV-Vis spectrophotometer (Thermo Scientific Genesys 10S UV-Vis; Sydney, Australia) blanked with Milli-Q water. The total phenolic concentration was derived as a function of the equivalent absorbance of gallic acid in the range 20 to 120 mg L^{-1} (R^2 = 0.999; see Figure 3-1). Results were expressed as milligrams of gallic acid equivalents (GAE) per 100 g of oven dry sample weight (mg GAE/100 g).

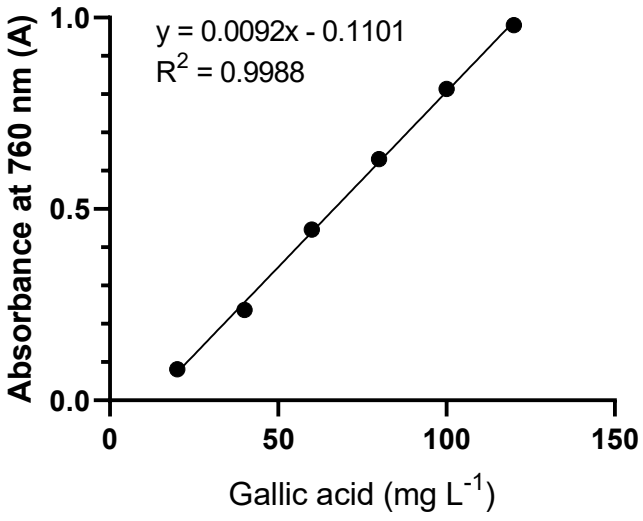

Figure 3-1: Calibration curve of gallic acid used in the Folin-Ciocalteu TPC assay.

3.3.5 Analysis of FRAP

To measure the total antioxidant capacity, the ferric reducing antioxidant potential (FRAP) assay developed by Benzie and Strain (1996) was performed on the samples. The FRAP reagent was prepared by combining 300 mM acetate buffer at pH 3.56, 20 mM aqueous ferric chloride and 10 mM TPTZ (made in 40 mM HCl) in the ratio 10:1:1. This solution was pre-equilibrated at a temperature of 37°C prior to use. The FRAP reagent, ferric chloride and TPTZ solutions were prepared fresh each day. To perform the assay, 3 mL of FRAP reagent, was combined with 100 µL of sample (pre-equilibrated at 37°C) and vortexed for 10 seconds. The samples were incubated in a covered water bath at 37°C for 4 minutes, vortexed for 10 seconds, and their absorbances read at 593 nm. The FRAP was derived as a function of the equivalent absorbance of Trolox (a water-soluble analog of vitamin E) in ethanol solution across the range of 25 to 175 mg L^{-1} (R^2 = 0.999; Figure 3-2). Results were expressed as milligrams of Trolox equivalents (TE) per 100 g of oven dry sample weight (mg TE/100 g).

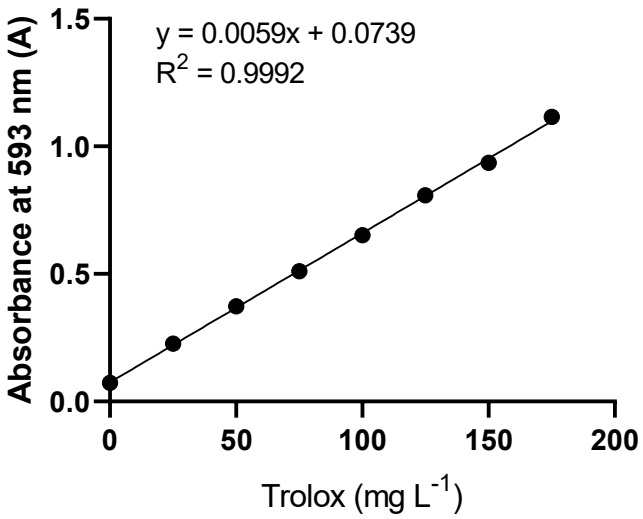

Figure 3-2: Calibration curve of Trolox used in the FRAP assay.

3.3.6 Analysis of TMAC

The total monomeric anthocyanin content (TMAC) was determined using a minor modification of the pH differential method described by Giusti and Wrolstad (2001). A pH 1 buffer solution was prepared from 0.025 M aqueous KCl, with the pH adjusted using concentrated (32%) HCl. Similarly, a pH 4.5 buffer solution was prepared using 0.4 M aqueous sodium acetate, adjusted to pH 4.5 with concentrated HCl.

In a cuvette, 400 µL of sample extract and 1.6 mL of pH 1 buffer were combined and mixed by inversion. After equilibration in darkness at room temperature for 15 minutes, its absorbance was read at 510 and 700 nm. This procedure was repeated on the same sample using the pH 4.5 buffer. The monomeric anthocyanin concentration was calculated using the formula shown in Equation 3.1 (Giusti & Wrolstad, 2001):

Anthocyanin content (mg cyd-3-glu L^{-1}) = (A × 449.38 × Dilution Factor × 1000) / (26900 × 1)

where A = (pH_1: $Absorbance_{510\ nm}$ – $Absorbance_{700\ nm}$) - ($pH_{4.5}$: $Absorbance_{510\ nm}$ – $Absorbance_{700\ nm}$) Equation 3.1

The molecular weight (449.38 g mol^{-1}) and molar extinction coefficient (26,900 M^{-1} cm^{-1}) of cyanidin-3-glucoside were used, as it is the most abundant anthocyanin bound in nature (Markakis, 1989). Results were expressed as milligrams of cyanidin-3-glucoside (cyd-3-glu) per 100 g of oven dried sample weight (mg cyd-3-glu/100 g).

3.3.7 Phenolic profiling by HPLC-DAD

As mentioned in the introduction (Section 3.1), phenolic profiling by high-performance liquid chromatography coupled with diode array detection (HPLC-DAD) was only conducted on 60 of the faba bean samples (all of the 2017 samples), due to the time-consuming rotary evaporating stage involved.

After conducting the benchtop phytochemical assays (Sections 3.3.4-3.3.6), approximately 18 mL of each methanol extract was concentrated using a rotary evaporator (Büchi Rotavapor R-114; Flawil, Switzerland) with an external vacuum pump (Edwards RV3 Rotary Vane Vacuum Pump; Burgess Hill, UK). The water bath (Büchi B-480) was limited to a temperature of 27°C to minimise enzymatic hydrolysis of any glycosidic bonds (Yabefa et al., 2014). After being evaporated to the point of dryness, the remaining solids were reconstituted in 1 mL of HPLC-grade methanol and syringe filtered (Livingstone 0.45 µm PTFE) into HPLC vials.

The concentrated extracts were analysed using an Agilent 1100 HPLC system (Waldbronn, Germany), comprising a G1313A autosampler, G1322A vacuum degasser, G1311A quaternary pump, G1316A thermostatted column compartment and G1315B diode array detector. The HPLC-DAD method used was developed by the author for the analysis of phenolic compounds in pulse matrices. A reversed-phase C_{18} column (Agilent Eclipse XDB-C18; 150 × 4.6 mm; 5 µm) and guard cartridge (Gemini C_{18} 4 × 2 mm) were used, with an injection volume of 5 µL and column temperature of 27 ± 0.8°C. The mobile phase comprised 0.01 M phosphoric acid (A) and methanol (B) at a flow rate of 1 mL/min, with the gradient beginning at 20% B and ramping linearly to reach 100% B at 20 min. The total run time was 25 min, with a post-run equilibration time of 7 min.

Compounds were putatively identified based on the comparison of their retention time and UV spectra with authentic standards run under the same conditions (Sigma Aldrich Australia). The purity of each peak was confirmed by examining the UV spectra at different time points throughout the peak. The concentrations of the identified compounds were calculated from the peak areas, using external calibration standards between 1-100 mg L^{-1}. The associated quality-of-analysis parameters are shown in Table 3-2. The limit of detection (LOD) was taken as the concentration at which the peak area was <2 mAU×s, while the limit of quantification (LOQ) was calculated as 3× the LOD.

Table 3-2: Quality-of-analysis parameters for the phenolic standards. All calibrations were performed at concentrations between 1–100 mg L^{-1}.

No.	Compound	Retention time (min)	Wavelength (nm)	Slope	LOD (mg L^{-1})	LOQ (mg L^{-1})	Calibration R^2
			Hydroxybenzoic acids				
1	Protocatechuic acid	3.94	250	13.5	0.1	0.4	1
4	*p*-hydroxybenzoic acid	5.78	250	25.9	0.1	0.2	1
5	Vanillic acid	6.26	250	12.3	0.1	0.4	1
6	Syringic acid	6.59	280	15.1	0.1	0.3	1
			Hydroxycinnamic acids				
3	Chlorogenic acid	5.26	320	14.1	0.1	0.4	1
7	*p*-coumaric acid	8.12	320	32.3	0.05	0.2	1
9	*trans*-ferulic acid	8.44	320	26.9	0.1	0.2	1
			Flavonoids				
2	Catechin	4.55	280	4.0	0.4	1.3	1
8	Vitexin	8.17	320	7.3	0.2	0.7	0.9999
10	Rutin	9.82	250	6.2	0.2	0.8	1

3.3.8 Near-infrared (NIR) spectroscopy

Near-infrared spectra were collected from the flour samples using an Antaris II FT-NIR Analyzer (Thermo Scientific; USA). The instrument was operated in reflectance mode, using the integrating sphere with a rotating sample cup (30 mm diameter). Spectra were collected between 1000-2500 nm (10,000-4,000 cm^{-1}), as the mean of 32 scans (resolution of 8 cm^{-1}). Spectra were collected in triplicate, repacking the sample cup with fresh flour each time. The

spectra were exported in *.csv format, with the mean of the triplicate spectra for each sample used in subsequent analysis.

3.3.9 Mid-infrared (MIR) spectroscopy

A Bruker Alpha FTIR (Fourier transform infrared) spectrophotometer (Bruker Optics Gmbh, Ettlingen, Germany) fitted with a platinum diamond ATR single reflection module was used to collect the MIR spectra. Homogenised faba bean flour was used to cover the reflection module and pressure applied to achieve uniform contact between the ATR interface and flour. Air was used as a reference background; the background measurement was performed every 10 samples. Cross-contamination of samples was minimised by cleaning and drying the platform with isopropyl alcohol and laboratory Kimwipes® between samples.

MIR spectra between 4000 and 400 cm^{-1} were recorded using the OPUS software version 7.5 (Bruker Optics Gmbh, Ettlingen, Germany) as the average of 24 scans at a resolution of 4 cm^{-1}. Five spectra were collected from each sample, repacking the instrument with fresh flour each time. The mean values from these replicate scans were used in subsequent analysis.

3.3.10 Statistical analysis of phenolic data

Statistical tests were performed on the phytochemical and phenolic data using IBM SPSS (v26) and R Studio running R 4.0.5 (R Core Team, 2020). Parametric testing was considered appropriate as the majority of data were approximately normally distributed, and the Central limit theorem could be applied due to the large sample size of the dataset (n = 60-100).

One-way ANOVAs were used for the analysis of the phytochemical constituents (i.e., FRAP, TPC, TMAC). A two-way ANOVA was performed on the HPLC-DAD dataset to assess the impact of variety and growing site on the content of various phenolic compounds.

Principal component analysis (PCA) was performed in the Unscrambler X software, version 10.5 (Camo ASA, Oslo, Norway). Where applicable, results are presented as mean ± 1 standard deviation.

3.3.11 Spectral data analysis

Qualitative exploratory analysis of the infrared spectral data was conducted using with the Unscrambler X software. The MIR spectra were pre-processed to the second derivative using a Savitzky-Golay algorithm at a polynomial number of 2 and a smoothing window of 41 points (Savitzky & Golay, 1964), following previous work on barley (Gordon et al., 2019) and mungbean (Johnson et al., 2019). Using a first or second derivative removes spectral variations in the baseline and slope (Savitzky & Golay, 1964), minimising differences due to non-compositional variables such as the pressure and contact with the reflection module.

Principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA) were performed in the Unscrambler X using the pre-processed MIR spectra.

Quantitative regression analysis of the infrared spectra was conducted in R Studio, using the spectrolab and prospectr packages. Spectra were pre-processed using the 1st and 2nd derivative method using a Savitzky-Golay algorithm (Savitzky & Golay, 1964) with varying numbers of smoothing points (5, 11, 15 or 21). Throughout the manuscript, these are referred to used abbreviated codes, e.g., 1d15 would be the Savitzky-Golay first derivative of the spectra, using a 15-point window (7 points on either side). In order to determine the optimum pre-processing method and number of components to use, the model performance of the calibration set was evaluated through full cross-validation (leave-one-out [LOO] method).

For model development, the 2017 samples (n=60) were used for the calibration set, while the 2016 samples (n=40) were used as an independent test set. A maximum of 10 components were considered for each PLSR model.

3.4 Results and discussion

3.4.1 Phytochemical profiles

The first stage of this work was to investigate the basic phytochemical composition of the faba bean samples using benchtop spectrophotometric-based methods, namely the analysis of total phenolic content, antioxidant capacity and total anthocyanin content. This information was used to gain a broad picture of the differences in phytochemical composition between the 10 faba bean varieties, and to examine any general trends between different growing seasons or locations.

3.4.1.1 Ferric reducing antioxidant potential

The average ferric reducing antioxidant potential determined for each of the faba bean varieties is shown in Table 3-3. There was a significant positive correlation between FRAP and the protein content of the samples (r_{100} = 0.346, P<0.001).

In order to determine the impact and interactions of genotype and environment on the FRAP, a three-way ANOVA was performed using variety, growing site and year as the independent variables. This revealed a significant interaction between variety and year (P<0.01), but not between variety and site. Variety had a highly significant impact on FRAP ($F_{10,60}$ = 632.267, P<0.001), while site and year had no significant impact (P>0.05 for both).

Table 3-3: Average ferric reducing antioxidant potential (FRAP; mg TE 100g^{-1} DW) of the ten faba bean varieties. Samples with the same letter in the last column were not statistically different at α = 0.05 according to post-hoc Tukey testing.

Variety	16Char	16Free	17Char	17Free	Mean ± SD
Fiord	255	223	261	231	243 ± 24 ^{d}
Fiesta VF	259	222	237	233	237 ± 23 ^{d}
Farah	253	250	249	231	245 ± 25 ^{d}
Nura	320	307	326	261	301 ± 33 ^{bc}
Doza	274	250	275	287	273 ± 27 ^{cd}
PBA Rana	456	471	542	610	531 ± 90 ^{a}
PBA Warda	265	244	269	259	260 ± 16 ^{cd}
PBA Samira	346	302	321	324	323 ± 25 ^{b}
PBA Zahra	284	264	299	262	278 ± 23 ^{bcd}
PBA Nasma	262	236	272	272	249 ± 27 ^{cd}
Mean ± SD	297 ± 62	277 ± 74	305 ± 91	297 ± 114	

While most varieties displayed similar FRAP levels, those for Nura and PBA Samira were significantly higher than the three lowest varieties (Fiesta VF, Fiord and Farah). PBA Rana had the highest FRAP values, being approximately double that of most of the other varieties. This highlighted the need for further investigation into the specific phenolic and antioxidant compounds that are elevated in this variety, in order to determine if it could potentially provide greater health benefits compared to other varieties.

3.4.1.2 Total phenolic content

The total phenolic contents followed a similar trend to the FRAP values (Table 3-4). There was a linear correlation between FRAP values and total phenolic contents (r_{100} = 0.917; P<0.001), as previously observed in faba bean (Chaieb et al., 2011) and other crops (Chen et al., 2018; Hung & Morita, 2008; Zhang et al., 2013; Žilić et al., 2012). There was also a significant positive correlation between protein content and total phenolics (r_{100} = 0.302, P<0.01) but no correlation between total phenolics and FRAP or moisture content (P>0.05).

A three-way ANOVA revealed a highly significant impact of variety on the TPC ($F_{10,60}$ = 72.646, P<0.001), with no interactions between variety, site or season (P>0.05 for all). As observed for FRAP, PBA Rana contained almost double the total phenolic content of most varieties, while PBA Samira contained less phenolics than PBA Rana but more than all other varieties.

The remaining eight varieties contained similar levels of phenolics (258-294 mg GAE $100g^{-1}$ DW). This is shown graphically in Figure 3-3.

Year also had a significant impact on the TPC ($P<0.05$), with samples from 2017 showing higher average TPCs compared to the 2016 samples.

Table 3-4: Average total phenolic content (TPC; expressed in mg GAE $100g^{-1}$ DW) of faba bean varieties. Samples with the same letter in the last column were not statistically different at α = 0.05 according to post-hoc Tukey testing.

Variety	16Char	16Free	17Char	17Free	Mean ± SD
Fiord	240	259	273	255	258 ± 22 [c]
Fiesta VF	263	252	259	286	266 ± 29 [c]
Farah	254	320	286	289	287 ± 38 [c]
Nura	269	292	278	243	268 ± 27 [c]
Doza	257	270	302	322	293 ± 33 [c]
PBA Rana	519	546	588	604	571 ± 59 [a]
PBA Warda	247	280	273	289	274 ± 21 [c]
PBA Samira	351	356	350	374	358 ± 32 [b]
PBA Zahra	288	290	330	264	294 ± 34 [c]
PBA Nasma	247	306	284	317	277 ± 39 [c]
Average	294 ± 86	317 ± 89	322 ± 96	324 ± 107	

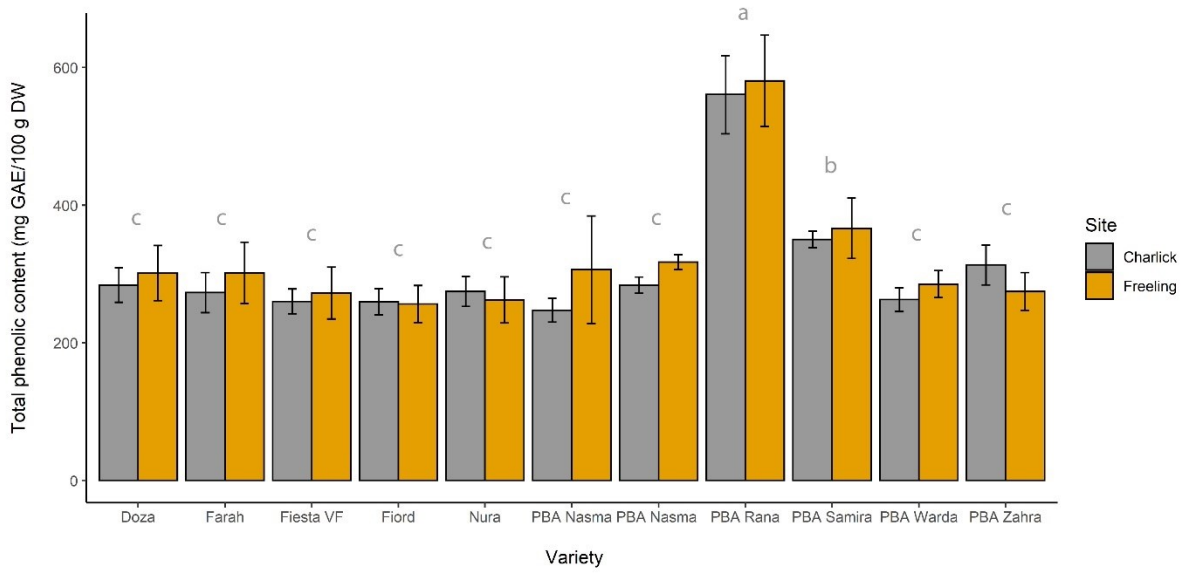

Figure 3-3: Total phenolic content of the 10 faba bean varieties at each of the growing sites (averaged across the 2016 and 2017 samples; n = 5 replicates for each bar). The letters (a–c) above each variety show the statistical significance of an ANOVA by variety averaged across both growing locations. Varieties with the same letter were not statistically different from one another at α = 0.05.

3.4.1.3 Total monomeric anthocyanin contents

The mean total monomeric anthocyanin contents of the faba bean samples are shown in Table 3-5. There was no significant correlation between TMAC and FRAP or TPC (P>0.05 for both), although it did show a negative correlation with moisture content (r_{100} = -0.202, P<0.05).

A three-way ANOVA showed a significant interaction between variety and year, as observed for FRAP. The TMAC was significantly impacted by all three variables – variety (P<0.001), site (P<0.05) and year (P<0.001), with the greatest impact from variety. The statistical differences are shown from the superscript annotations reported in Table 3-5. The variety Fiord contained the highest levels of anthocyanins, followed by PBA Warda. The anthocyanin content of Fiord was statistically higher than the five varieties containing the lowest anthocyanin levels, but not from the four remaining varieties.

Samples from 2016 displayed a significantly higher anthocyanin content than the 2017 samples, as can be seen from Table 3-5. Furthermore, closer examination of the impact of site revealed that most of the difference between sites occurred in 2017, with minimal difference observed in 2016 (Table 3-5).

Table 3-5: Average total monomeric anthocyanin content (TMAC; expressed in mg cyd-3-glu equivalents 100 g^{-1} DW) of the faba bean varieties. Samples with the same letter in the last column were not statistically different at α = 0.05 according to post-hoc Tukey testing.

Variety	16Char	16Free	17Char	17Free	Average
Fiord	18.9	18.6	19.2	18.0	18.6 ± 1.2 ^{ab}
Fiesta VF	19.8	20.3	16.5	12.4	16.7 ± 3.8 ^{abcd}
Farah	19.3	17.8	14.8	11.6	15.3 ± 3.7 ^{bcde}
Nura	21.1	16.5	15.2	17.0	17.2 ± 2.9 ^{abcd}
Doza	14.3	17.9	13.4	11.2	13.8 ± 3.6 ^{de}
PBA Rana	14.7	16.9	17.2	15.3	16.1 ± 2.6 ^{abcde}
PBA Warda	19.6	19.3	18.5	15.1	17.8 ± 2.8 ^{abc}
PBA Samira	12.2	10.6	12.8	14.4	12.7 ± 3.4 ^{e}
PBA Zahra	17.1	13.5	11.3	13.2	13.5 ± 2.9 ^{de}
PBA Nasma	21.7	20.2	15.8	12.1	21.0 ± 4.3 ^{cde}
Average	17.9 ± 3.7	17.2 ± 3.6	15.5 ± 3.1	14.0 ± 3.2	

3.4.1.4 PCA of moisture, antioxidant, phenolic and anthocyanin contents

In order to further explore the variation in the moisture, antioxidant, phenolics and anthocyanin contents, principal component analysis was conducted on the data for these four parameters. As these parameters varied considerably in terms of their absolute values, each datapoint was weighted by dividing by the overall standard deviation for that parameter.

The first two principal components (PCs) explained 78% of the total variation observed. Across PC1, broad separation was observed between PBA Rana and the remainder of the varieties (Figure 3-4). Examination of the loadings associated with this principal component indicated that PC1 scores were positively correlated with increased FRAP and total phenolics, confirming that levels of these compounds in PBA Rana were noticeably elevated compared to other varieties. PBA Samira also had higher scores along PC1 than the remainder of varieties, indicating that its FRAP and TPC levels were higher compared to other varieties, albeit not as distinct as those of PBA Rana.

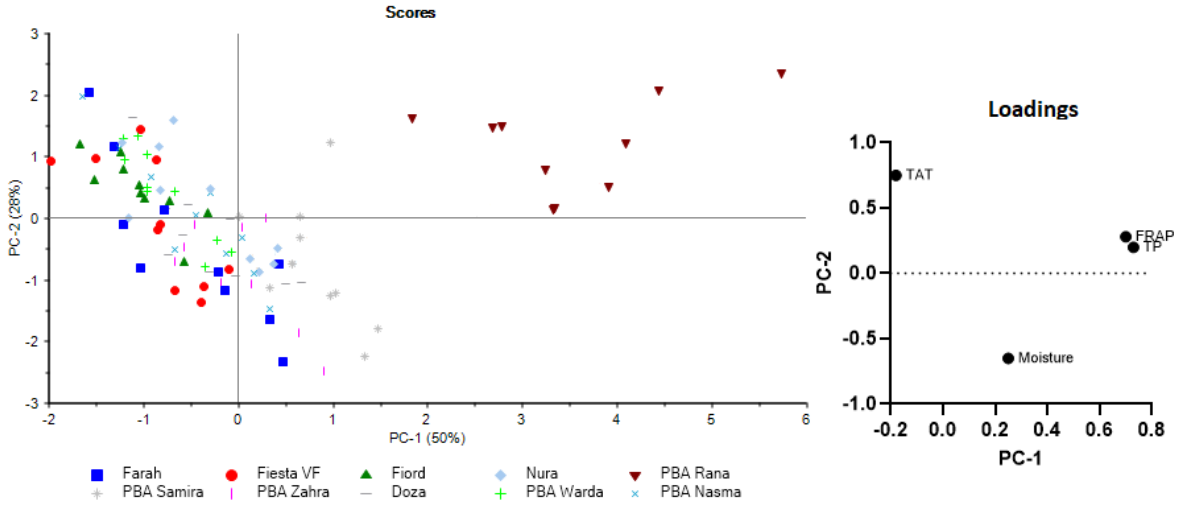

Figure 3-4: Principal component analysis of four chemical parameters (moisture, FRAP, TPC and TMAC), with the samples separated by variety. The loadings for the first two principal components of the PCA are also indicated.

Separation along the second PC was generally less clear. Both PBA Rana and Fiord were largely associated with positive scores along this axis, indicating above average anthocyanin levels and lower moisture contents. The negative PC2 scores observed for PBA Zahra indicated higher moisture contents and lower anthocyanin levels, as previously observed (Table 3-5). PBA Samira, which had the lowest mean anthocyanin levels of all faba bean varieties (Table 3-5), was also primarily associated with negative PC2 values.

To further visualise the general relationship between the chemical composition of the varieties obtained through PCA, a cluster analysis was performed on the mean moisture, antioxidant, phenolic and anthocyanin contents (Figure 3-5). This confirmed that the composition PBA Rana was highly distinct from the remaining varieties, while the composition of PBA Samira was moderately different. The hierarchical cluster analysis also suggested that based on their chemical composition, two general groups could be made of the remaining eight varieties, one comprising Nura, Doza and PBA Zahra, and the other comprising Fiord, Fiesta VF, PBA Nasma, PBA Warda and Farah.

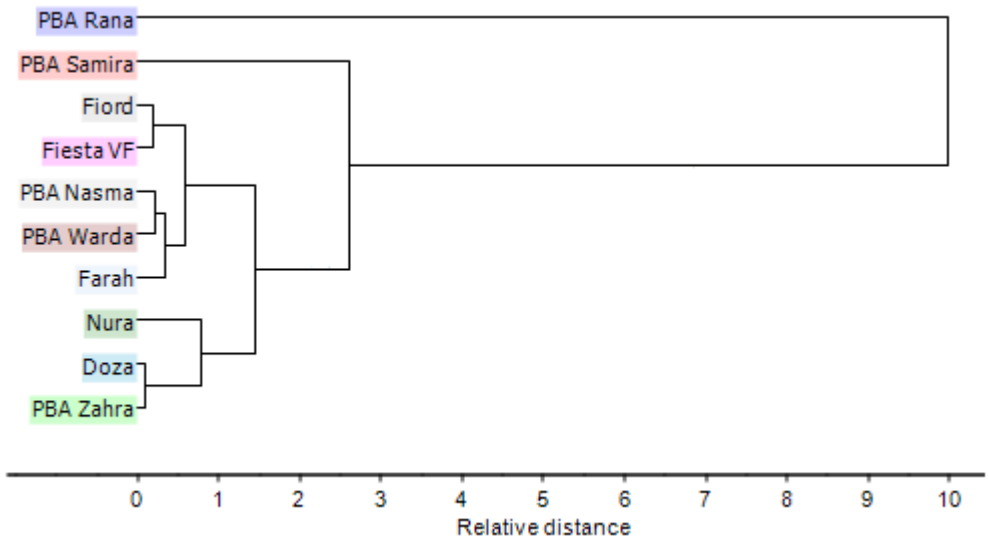

Figure 3-5: Hierarchical cluster analysis of the mean moisture, FRAP, TPC and TMAC of the ten faba bean varieties. The cluster analysis used Ward's method and the squared Euclidean distance.

3.4.1.5 General observations on the phytochemical composition

The ten varieties of Australian faba bean investigated showed considerable variation in their anthocyanin, phenolic and antioxidant contents. In particular, PBA Rana contained much higher levels of total phenolics and total antioxidants than all other varieties tested, indicating the need for further investigation into the specific phenolic compounds present in this variety. These results were supported by PCA and hierarchical cluster analysis, which demonstrated uniquely high levels of TPC and FRAP in PBA Rana, in addition to highlighting moderately elevated levels of these analytes in PBA Samira.

3.4.2 Phenolic profiles using HPLC-DAD

Following on from the benchtop phytochemical testing, HPLC-DAD was used to quantify selected phenolic compounds in a subset of the faba bean extracts (2017 samples only). This was considered appropriate as there were no significant differences in antioxidant capacity or total phenolic content between the 2016 and 2017 samples. The aim was to provide context data on the typical phenolic composition of Australian-grown faba bean, as well as to highlight potential differences in phenolic profiles between different commercial varieties of this crop. The results are presented and discussed throughout the remainder of this section.

3.4.2.1 Identification and quantification of phenolic compounds

A total of 10 phenolic compounds were putatively identified in the faba bean extracts based on their UV spectra and retention times (Figure 3-6). These comprised four hydroxybenzoic

acids, three hydroxycinnamic acids and three flavonoid-related compounds (Table 3-6). Additionally, the following compounds were all determined to be below the limit of detection in the extracts: gallic acid, gentisic acid, isovanillic acid, caffeic acid, sinapic acid, cinnamic acid and quercetin-3-glucoside.

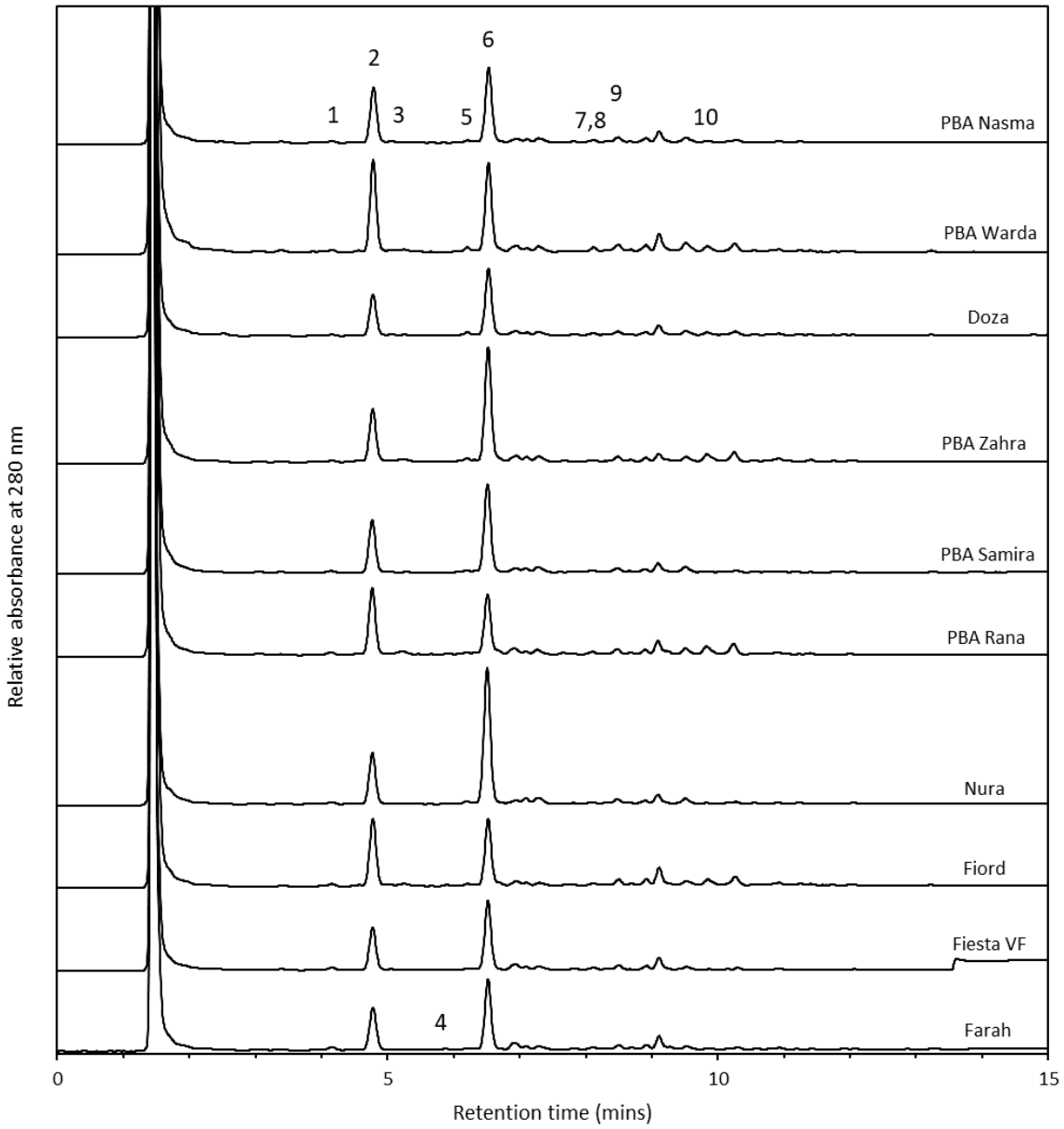

Figure 3-6: HPLC chromatograms of the phenolic compounds in the 10 faba bean varieties. The compounds indicated are (1) protocatechuic acid, (2) catechin, (3) chlorogenic acid, (4) p-hydroxybenzoic acid, (5) vanillic acid, (6) syringic acid, (7) p-coumaric acid, (8) vitexin, (9), trans-ferulic acid, (10) rutin.

Table 3-6: Mean phenolic acid and flavonoid contents in the 10 faba bean varieties (2017 samples only). Values given in µg/g (mean ± SD from 6 replicates, comprising 3 within-field triplicates from 2 field locations). The P value column indicates the significance between varieties, with results obtained from a two-way ANOVA between site × variety. Note that entries in the same row containing the same superscript letter (a–d) were not significantly different from one another at α = 0.05.

Compound	Doza	Farah	Fiesta VF	Fiord	Nura	PBA Nasma	PBA Rana	PBA Samira	PBA Warda	PBA Zahra	P value
Protocatechuic acid	1.88 ± 0.83 ^{b}	1.45 ± 0.61 ^{b}	1.44 ± 0.50 ^{b}	1.66 ± 0.77 ^{b}	1.29 ± 0.22 ^{b}	1.65 ± 0.60 ^{b}	2.93 ± 1.07 ^{a}	2.09 ± 0.77 ^{ab}	1.83 ± 0.56 ^{b}	1.81 ± 0.36 ^{b}	***
p-hydroxybenzoic acid	0.57 ± 0.06 ^{bcd}	0.52 ± 0.07 ^{cd}	0.52 ± 0.11 ^{cd}	0.64 ± 0.18 ^{bcd}	0.44 ± 0.08 ^{d}	0.62 ± 0.13 ^{bcd}	1.11 ± 0.21 ^{a}	0.73 ± 0.15 ^{bc}	0.61 ± 0.10 ^{bcd}	0.79 ± 0.11 ^{b}	***
Vanillic acid	2.46 ± 0.59 ^{ab}	1.87 ± 0.37 ^{b}	1.96 ± 0.36 ^{b}	1.88 ± 0.43 ^{b}	2.40 ± 0.24 ^{ab}	1.96 ± 0.29 ^{b}	2.24 ± 0.68 ^{ab}	2.81 ± 0.52 ^{a}	2.71 ± 0.40 ^{a}	2.76 ± 0.85 ^{a}	***
Syringic acid	77.6 ± 11.2 ^{c}	72.4 ± 7.9 ^{c}	80.5 ± 13.1 ^{c}	77.6 ± 11.8 ^{c}	149.8 ± 14.8 ^{a}	77.8 ± 9.5 ^{c}	72.5 ± 7.3 ^{c}	109.6 ± 21.6 ^{b}	80.9 ± 14.1 ^{c}	122.5 ± 14.4 ^{b}	***
Sum of hydroxybenzoic acids	82.5 ± 12.6 ^{c}	76.2 ± 8.5 ^{c}	84.4 ± 13.7 ^{c}	81.8 ± 12.7 ^{c}	153.9 ± 14.9 ^{a}	82.0 ± 10.3 ^{c}	78.8 ± 8.8 ^{c}	115.2 ± 22.8 ^{b}	86.0 ± 14.8 ^{c}	127.9 ± 15.4 ^{b}	***
Chlorogenic acid	0.89 ± 0.96	0.85 ± 0.52	1.27 ± 1.30	2.88 ± 2.56	0.78 ± 0.41	0.89 ± 0.44	3.02 ± 3.31	1.14 ± 0.73	1.70 ± 2.48	1.98 ± 3.41	NS
p-coumaric acid	1.21 ± 0.21 ^{bc}	1.64 ± 0.25 ^{ab}	1.86 ± 0.40 ^{a}	1.69 ± 0.28 ^{ab}	1.26 ± 0.16 ^{abc}	1.52 ± 0.37 ^{abc}	1.70 ± 0.27 ^{ab}	1.52 ± 0.38 ^{abc}	1.62 ± 0.54 ^{ab}	0.95 ± 0.17 ^{c}	***
trans-ferulic acid	1.27 ± 0.22 ^{b}	0.96 ± 0.21 ^{b}	1.11 ± 0.27 ^{b}	1.42 ± 0.42 ^{b}	1.36 ± 0.32 ^{b}	1.22 ± 0.35 ^{b}	2.99 ± 0.65 ^{a}	1.34 ± 0.25 ^{b}	1.26 ± 0.30 ^{b}	1.82 ± 0.18 ^{b}	***
Sum of hydroxycinnamic acids	3.37 ± 1.09 ^{b}	3.45 ± 0.76 ^{b}	4.24 ± 1.81 ^{ab}	5.99 ± 2.97 ^{ab}	3.40 ± 0.68 ^{b}	3.63 ± 0.85 ^{ab}	7.71 ± 2.83 ^{a}	4.00 ± 1.17 ^{ab}	4.58 ± 3.22 ^{ab}	4.11 ± 3.60 ^{ab}	*
Catechin	216 ± 64 ^{ab}	191 ± 37 ^{b}	215 ± 52 ^{ab}	245 ± 52 ^{ab}	232 ± 27 ^{ab}	207 ± 37 ^{b}	297 ± 53 ^{a}	240 ± 55 ^{ab}	258 ± 63 ^{ab}	220 ± 33 ^{ab}	*
Vitexin	0.88 ± 0.24 ^{b}	1.70 ± 1.82 ^{ab}	0.97 ± 0.39 ^{b}	1.52 ± 1.82 ^{ab}	0.58 ± 0.41 ^{b}	0.80 ± 0.28 ^{b}	3.50 ± 1.43 ^{a}	0.75 ± 0.43 ^{b}	1.21 ± 0.82 ^{b}	1.43 ± 1.52 ^{ab}	**
Rutin	5.55 ± 5.02	7.34 ± 5.11	7.66 ± 6.36	13.91 ± 11.81	4.04 ± 3.00	4.50 ± 2.30	15.87 ± 14.22	7.67 ± 4.09	10.48 ± 10.60	9.43 ± 16.29	NS
Sum of flavonoids	223 ± 61 ^{ab}	200 ± 37 ^{b}	223 ± 57 ^{ab}	261 ± 61 ^{ab}	237 ± 28 ^{ab}	212 ± 37 ^{b}	316 ± 45 ^{a}	248 ± 56 ^{ab}	269 ± 71 ^{ab}	231 ± 35 ^{ab}	*

NS = not significant (P > 0.05), * P < 0.05, ** P < 0.01, *** P < 0.001

The hydroxybenzoic acids found here (protocatechuic acid, *p*-hydroxybenzoic acid, vanillic acid and syringic acid) have all been previously reported from faba bean, as have two of the hydroxycinnamic acids (*p*-coumaric acid and *trans*-ferulic acid) (Choudhary et al., 2020; Liu et al., 2020b; Ryszard & Fereidoon, 2018). In addition, several of these phenolic acids have been found in faba bean pods (Valente et al., 2018). The concentrations of free *p*-coumaric and ferulic acids found here were similar to that reported by Liu et al. (2020b) in Canadian faba bean, although only one variety was included in that study. However, the concentration of syringic acid was much higher compared to previous studies (Liu et al., 2020b; Sosulski & Dabrowski, 1984). Although the reason for this difference is unclear, it is worth noting that levels of this compound varied significantly between the different genotypes studied (Table 3-6) and that previous research has indicated that the soil microbiota composition can have a significant impact on its concentration (Dong et al., 2016). Similarly, although chlorogenic acid does not appear to have been previously found in faba bean seed, it is produced in the roots of this plant (El-Akkad et al., 2002). The low concentrations and high levels of environmental variability may account for its absence in previous work.

The levels of catechin reported here (ranging from 191–297 mg/kg for different varieties) were at the lower range of concentrations reported by Baginsky et al. (2013) in the immature seed material of 10 Chilean faba bean varieties, with catechin contents varying between 85–978 mg/kg.

Although vitexin is more commonly known to occur in mungbean (Yao et al., 2011), it has been previously reported from faba bean using UHPLC-ESI-QTOF-MS-based metabolic profiling (Abu-Reidah et al., 2014), although it was not quantified by these authors. Similarly, although rutin (quercetin-3-rutinoside) does not appear to have been previously reported from faba bean seed, numerous other types of quercetin glycosides have been found in this matrix (Spanou et al., 2012; Valente et al., 2018). In addition, rutin has been reported from the flower tissue of several faba bean genotypes, indicating that the synthetic pathways for the production of this compound do occur in faba bean (Zanotto et al., 2020).

A two-way ANOVA, using variety and growing location as the independent variables, revealed that the content of all constituents, apart from chlorogenic acid and rutin, varied significantly with variety (Table 3-6). For most of these compounds, the highest concentrations were found in PBA Rana, which also contained the highest total phenolic content (Figure 3-6). However, Nura showed the highest levels of syringic acid and total hydroxybenzoic acids.

Similarly, the two-way ANOVA demonstrated that in the case of the 2017 growing season, the site had a significant impact on the content of protocatechuic acid, vanillic acid, chlorogenic acid, vitexin and rutin, as well as on the total amounts of hydroxybenzoic acids and

hydroxycinnamic acids (Table 3-7). For both hydroxybenzoic acids (protocatechuic acid and vanillic acid), samples grown at the Freeling site showed higher contents; while for chlorogenic acid and the flavonoids catechin and rutin, the Charlick samples showed higher concentrations.

Table 3-7: Impact of the growing site on phenolic acid and flavonoid contents. Values given in µg/g (mean ± SD from 3 replicates for each location). The P value column indicates the significance between sites, with results obtained from a two-way ANOVA between site × variety.

Compound	Charlick (*n* = 30)	Freeling (*n* = 30)	Site P Value	Variety × Site Interaction
Protocatechuic acid	1.43 ± 0.36	2.17 ± 0.87	***	NS
p-hydroxybenzoic acid	0.67 ± 0.19	0.64 ± 0.25	NS	NS
Vanillic acid	2.11 ± 0.41	2.50 ± 0.67	***	**
Syringic acid	89.3 ± 28.2	94.9 ± 27.9	NS	*
Sum of hydroxybenzoic acids	93.5 ± 28.3	100.2 ± 28.4	*	*
Chlorogenic acid	2.22 ± 2.57	0.86 ± 0.68	**	NS
p-coumaric acid	1.45 ± 0.42	1.55 ± 0.38	NS	NS
trans-ferulic acid	1.45 ± 0.47	1.38 ± 0.77	NS	**
Sum of hydroxycinnamic acids	5.11 ± 3.00	3.78 ± 1.39	*	NS
Catechin	224 ± 45	240 ± 60	NS	NS
Vitexin	1.67 ± 1.56	0.98 ± 0.86	*	NS
Rutin	12.21 ± 11.59	5.07 ± 3.54	**	NS
Sum of flavonoids	238 ± 52	246 ± 62	NS	NS

* P < 0.05, ** P < 0.01, *** P < 0.001

For the 2017 growing season, no significant effects of growing location were found for *p*-hydroxybenzoic acid, syringic acid, *p*-coumaric acid, *trans*-ferulic acid, catechin or the total amount of flavonoids. Significant interactions were found between the variety and growing site for several parameters, namely vanillic acid, syringic acid, *trans*-ferulic acid and the sum of the hydroxybenzoic acids. However, it should be noted that the present study investigated only one growing season; hence it remains for future work to generalise the results found here across a wider range of seasons and locations.

There appears to be limited previous literature investigating the impact of growing site and variety × growing site interaction on phenolic acid content in faba bean; however, Mpofu et al. (2006) found a significant impact of growing location on six phenolic acids in wheat. In contrast, Oomah et al. (2011) found minimal impact of growing location on the total phenolic content of 13 faba bean genotypes grown at two locations in Canada.

3.4.2.2 PCA of phenolic compounds

Overall, PBA Rana appeared to have the highest levels of most phenolic acids and flavonoids, possessing a distinct chemical profile compared to the other varieties. This observation was supported by the results of the principal component analysis performed on the normalised phenolic data, which revealed that most samples of PBA Rana were clustered toward the lower right of the scores plot, separated from the majority of other genotypes (Figure 3-7). Examination of the loadings plot revealed that this corresponded with higher concentrations of catechin and protocatechuic acid, and lower concentrations of syringic acid. In addition, this concurred with previous observations on the unique phytochemical profile of this genotype (Section 3.4.1.4). The unique phenolic acid profile of PBA Rana could be most easily visualised in its UV isoplot (Figure 3-8), where the plot showed the presence of a larger number of phenolic compounds and greater absorbance intensities for the compounds observed, particularly between elution times between 9-11 minutes.

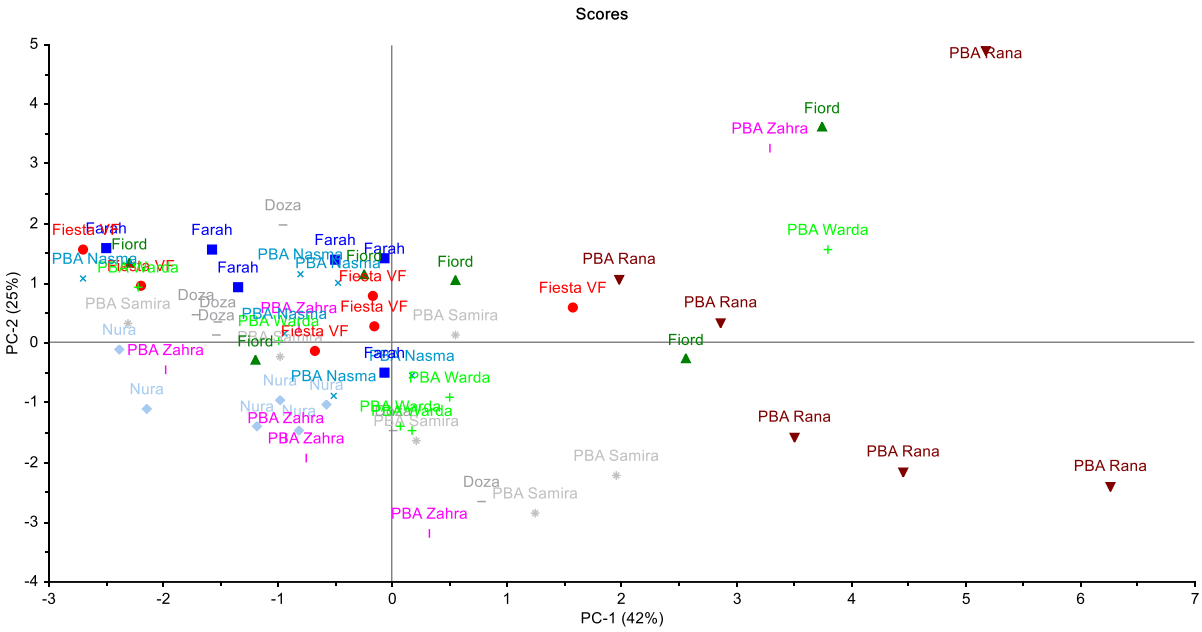

Figure 3-7: Scores plot showing the results of the principal component analysis performed on the normalised phenolic data. Each faba bean variety is indicated by a different symbol colour and shape.

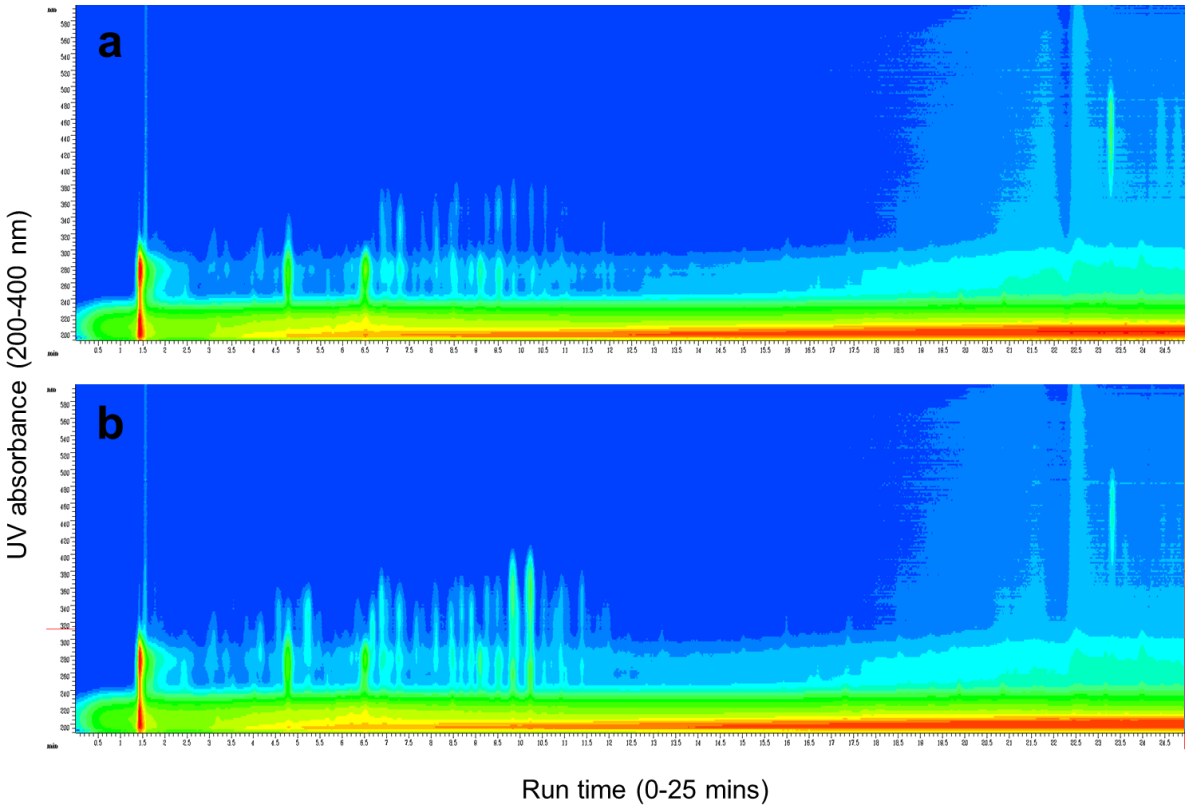

Figure 3-8: Isoplot showing the UV absorbance of various compounds eluting at various points throughout the HPLC run for PBA Samira (a) and PBA Rana (b). The x-axis shows the run time (from 0-25 minutes) and the y-axis shows the UV wavelength (from 200-400 nm). The colour of each pixel shows the relative absorbance (blue = low; red = high).

In contrast to the PCA results observed for PBA Rana, the variety Nura was clustered toward the lower left of the scores plot. Syringic acid was weighted on this region of the two PCs, corresponding with the higher concentrations of this compound found in Nura. The remaining varieties were more or less clustered around the centre of the scores plot, indicating a relatively similar phenolic composition between them.

3.4.2.3 Correlation analysis of phenolics

Correlation analysis was performed between the 10 phenolic compounds to ascertain if the concentrations of any specific compounds were closely linked to the concentrations of another compound. This may occur due to similar biosynthetic pathways between the compounds (Santos-Sánchez et al., 2019) or result from regulatory genes controlling multiple synthesis pathways. The correlation results demonstrated moderate to strong correlations between several compounds, most notably between rutin and chlorogenic acid (r_{60} = 0.979, $p < 0.001$), and between ferulic and *p*-hydroxybenzoic acid (r_{60} = 0.812, $P < 0.001$) (Figure 3-9). Rutin is

a quercetin glycoside, while chlorogenic acid is the ester of caffeic acid and quinic acid, hence these compounds are not closely structurally related. However, both can be synthesised through the phenylpropanoid pathway (Fraser & Chapple, 2011), suggesting that a regulatory gene may be responsible for the correlation between these compounds. Similarly, although ferulic acid is a hydroxycinnamic acid and *p*-hydroxybenzoic acid is a simple hydroxybenzoic acid, both can be produced through the shikimate biosynthesis pathway (Marchiosi et al., 2020).

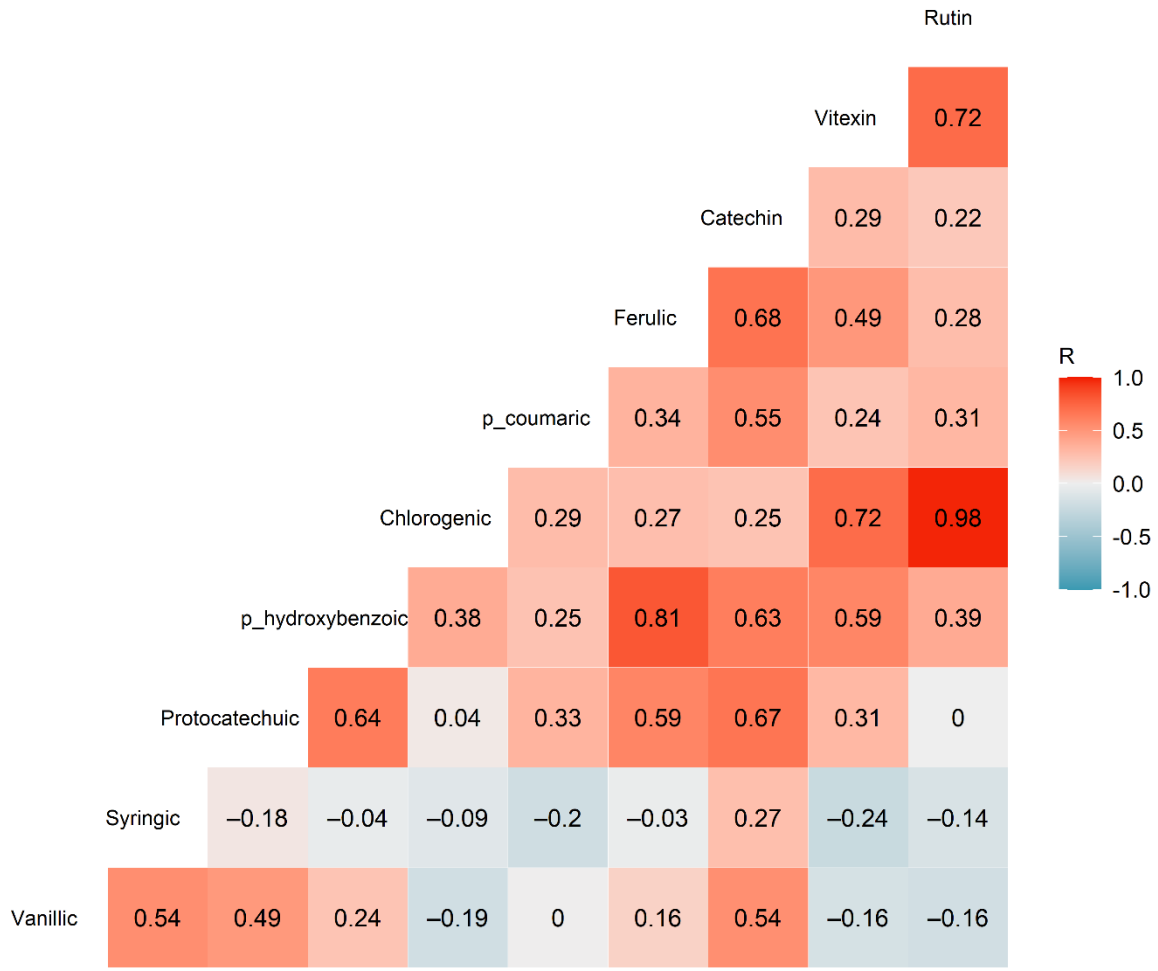

Figure 3-9: Correlogram showing the correlations between the various phenolic acids and flavonoids quantified in the faba bean samples (n = 60 samples). The numbers inside each square show the Pearson R correlation values.

3.4.2.4 Summary of phenolic profiling results

This work was the first study to profile the phenolic acid and flavonoid composition of >5 commercial Australian faba bean varieties. The most abundant compounds identified were catechin and syringic acid, with rutin, vitexin, protocatechuic, vanillic, *p*-hydroxybenzoic, chlorogenic, *p*-coumaric, and *trans*-ferulic acid all found in lower concentrations. The content

of most individual phenolics varied significantly with the variety, while growing location had a significant effect for around half of these compounds. Genotype × location interactions were only observed for vanillic, syringic, and *trans*-ferulic acids. Significant correlations were observed between a number of constituents, including between rutin and chlorogenic acid, and between ferulic and *p*-hydroxybenzoic acid. Notably, PBA Rana showed a distinct phenolic profile compared to the remaining nine varieties, supporting prior observations on the uniquely high levels of total phenolics and antioxidant capacity found in this variety (see Section 3.4.1.5).

3.4.3 Qualitative assessment using MIR spectroscopy

Infrared spectroscopy, particularly MIR spectroscopy, has shown considerable promise for the quality assurance of grain crops (Achten et al., 2019; Cozzolino, 2014b, 2016; Gordon et al., 2019; Pandiselvam et al., 2021), including product authentication, discrimination of origin and the detection of adulterants. Consequently, this section investigated the use of MIR spectroscopy for the qualitative analysis of faba bean flour, specifically, its potential for the discrimination of variety, growing location and season.

3.4.3.1 Qualitative examination of MIR spectra

The main spectral peaks observed in the MIR spectra were attributed to a range of constituent compounds, including water, compositional polysaccharides and protein (Table 3-8). There was a visible difference in the amplitude of the MIR absorbance between 2016 and 2017 samples, with the 2017 samples showing greater absorbance overall. However, the location of spectral peaks was virtually identical between the two years. There was little visible difference in the average spectra between the Charlick and Freeling sites, although a slightly larger peak at 2990-2950 cm^{-1} was observed in the Freeling samples. There was also little visible difference in the average spectra of the ten faba bean varieties, although varietal differences in the size of the 2990-2950 cm^{-1} peak (CH_2 vibration) were noted (Figure 3-10).

Table 3-8: Main spectral peaks observed in spectra obtained from the faba bean samples.

Range (cm^{-1})	Centre (cm^{-1})	Peak strength	Correlating bond(s)	Likely aetiological compound(s)	References
3600-3000	3260	S, broad	OH stretch	Water	(Gordon et al., 2019)
3000-2940	2980	M	Symmetric CH_3 stretch	Lipids, fatty acids	(Gordon et al., 2019; Li et al., 2013)
2980-2890	2920	M	Asymmetric CH_2 vibration	Lipids, fatty acids	(Gordon et al., 2019; Li et al., 2013)
1760-1720	1740	W	O-C=O stretch	Triglycerides	(Vlachos et al., 2006)
1700-1580	1640	S	N-H-C=O (amide bond)	Amide I in protein	(Bratu et al., 2007)
1580-1480	1540	M	N-H-C=O (amide bond)	Amide II in protein	(Bratu et al., 2007)
1480-1410	1440	W	CH_2, CH_3 bend	Lipids	(Yang et al., 2015)
1430-1350	1390	M-W	CH_3 bend	Lipids	(Pavia et al., 2001)
1350-1290	1330	W	C-O stretch	Polysaccharides	(Hoffmann & De Paola, 1984)
1270-1190	1240	M-W	C-O bond	Polysaccharides	(Gordon et al., 2019)
1170-1130	1150	M	C-O stretch	Polysaccharides	(Ambigaipalan et al., 2011)
1080-1060	1075	M	C-OH stretch	Polysaccharides	(Ambigaipalan et al., 2011)
1050-1030	1040	W	CH_2, C-OH bend	Starch/carbohydrates	(Ambigaipalan et al., 2011)
1050-950	995	v. S	CH_2, C-OH bend	Starch/carbohydrates	(Ambigaipalan et al., 2011)
950-910	930	W	C-OH bend	Starch	(Gordon et al., 2019)
870-820	850	M-W	C-O-C vibration	Amylopectin	(Bratu et al., 2007)
770-740	760	W	C-O-C ring vibration	Starch/carbohydrates	(Chen et al., 2008)

Abbreviations: v. = very, S = strong, M = medium, W = weak

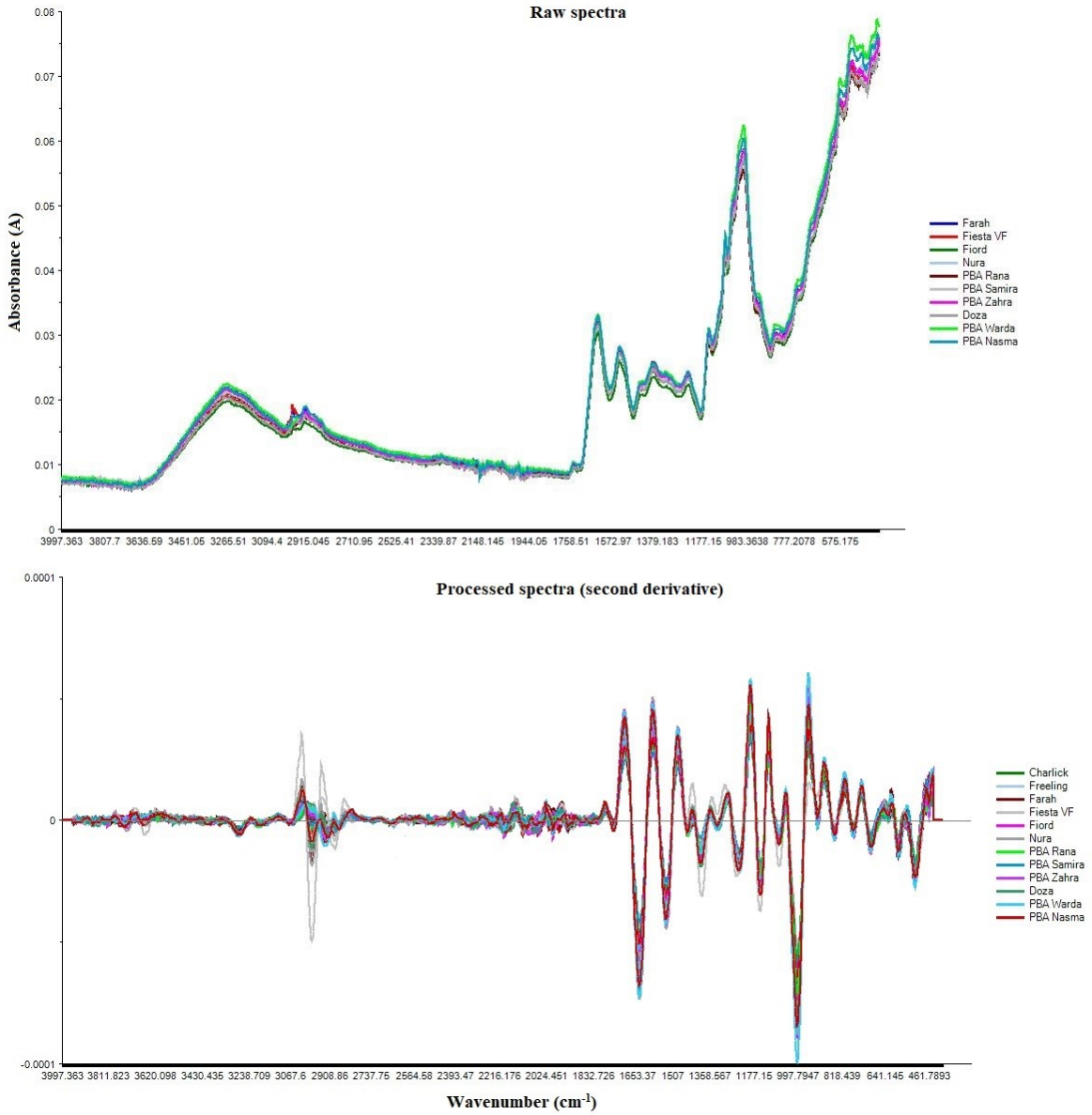

Figure 3-10: Average MIR spectra of the faba bean samples, before and after spectral processing through the use of the second derivative.

Prior to further analysis, the individual spectra were pre-processed to the second derivative (Savitzky-Golay algorithm; 41 smoothing points) to remove any differences in the absorbance amplitudes resulting from variation in the level of contact between the sample and the reflection module. This successfully removed any baseline amplitude variation, while amplifying the differences in peak positions, shapes and relative amplitude.

To further explore the spectral variation, principal component analysis was conducted on the second derivative of the MIR spectra. The first two principal components of the PCA explained 75% of the observed variation, while subsequent PCs explained <2% of the variation each. There was some incomplete separation by growing year (Figure 3-11A) and no clear

separation by site (Figure 3-11B). Most faba bean varieties were spread across the PCA scores plot (Figure 3-11C), again with no clear separation. However, some varieties, such as Fiesta VF, appeared to show a larger amount of intra-varietal spectral variation, being distributed much further across both PC1 and PC2, when compared to other varieties such as Fiord.

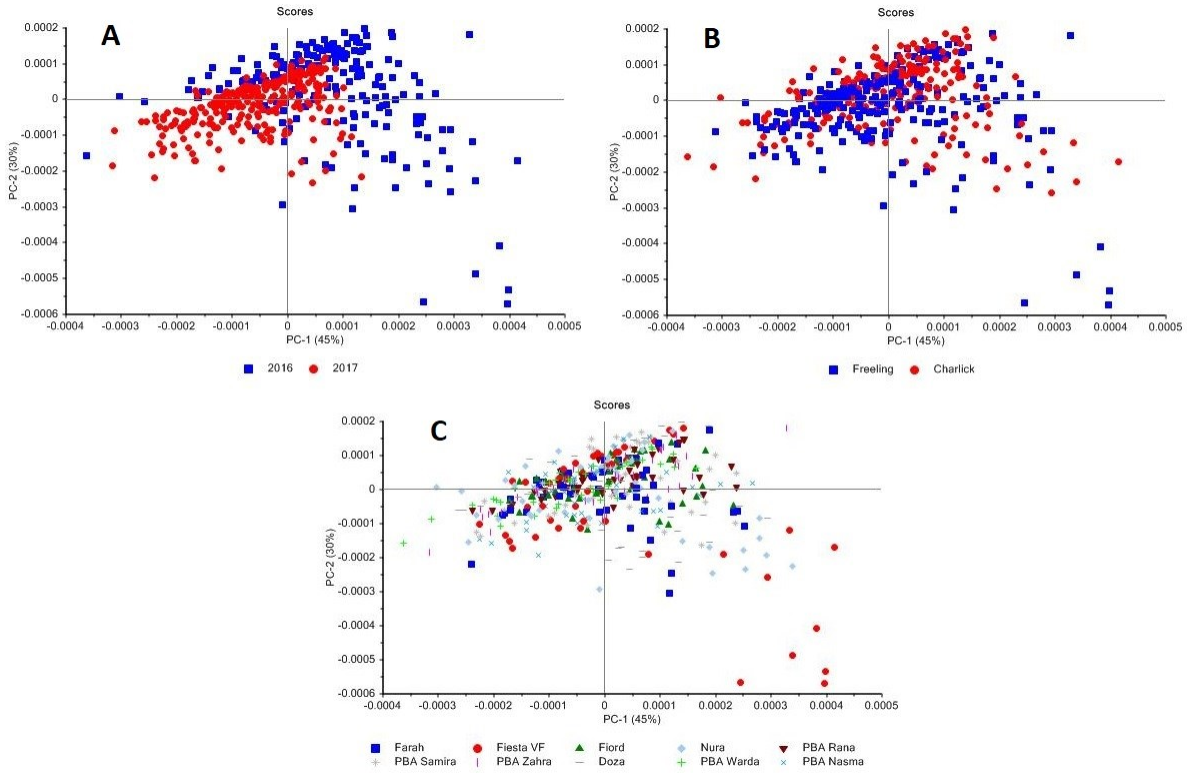

Figure 3-11: Principal component analysis of the second derivative of the MIR spectra, coloured by growing year (A), growing site (B) and variety (C).

Examination of the loadings plot revealed that the major differences between samples were due to bonds in the CH_2 and CH_3 stretch region (3000-2800 cm^{-1}), amide region (1700-1500 cm^{-1}) and C-OH bend region from starch/carbohydrates (1050-900 cm^{-1}) (Figure 3-12).

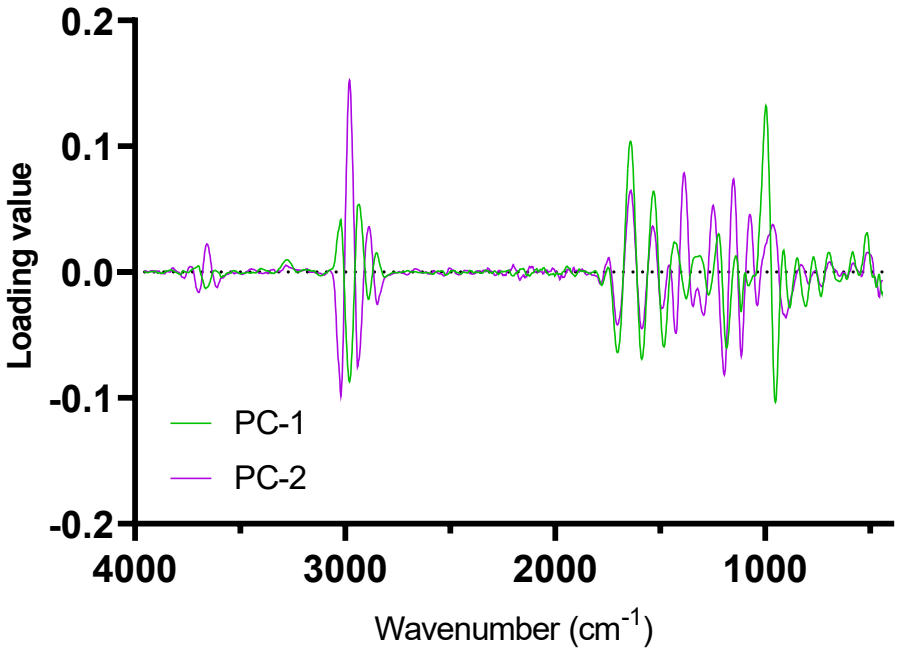

Figure 3-12: Loadings plot for the PCA performed on the MIR spectra.

3.4.3.2 PLS-DA of MIR spectra

Partial least squares discriminant analysis (PLS-DA) applied to MIR spectra has previously been highlighted as a powerful tool for the discrimination and authentication of grains of different origins (Gordon et al., 2019). Hence PLS-DA was performed on the second derivative of the faba bean spectra. The growing year was able to be correctly classified in 87.4% of the samples (Table 3-9), indicating that this technique may be suitable for some authentication purposes when applied to faba bean flour. However, the successful classification rate by growing site was much lower, with an average of 60.2% of samples correctly assigned to their growing site (Table 3-9). Further refinement of the PLS regression is required before this technique can be utilised for authentication of growing site in faba bean flour, including analysing samples from a wider range of growing locations and perhaps identifying and selecting the key model wavelengths responsible for site discrimination.

Overall, the MIR analysis was found to provide valuable qualitative information on the chemical composition of the faba bean varieties. The preliminary results presented here are quite promising, indicating that agronomic aspects such as the year of growth can be determined from the MIR spectra with moderately high accuracy. This could potentially be used for quality assurance/product authentication purposes. However, it should be noted that these results are for the calibration model only and were not tested on new populations. Nevertheless, with

the development of more sophisticated methods of data analysis, the usefulness of MIR spectroscopy for product authentication/discrimination is anticipated to only increase.

Table 3-9: PLS-DA classification of the faba bean samples by growing year and growing site.

	Actual growing year		Actual growing site	
Classified as	2016 (n=200)	2017 (n=300)	Charlick (n=250)	Freeling (n=250)
2016	87.0%	1.3%		
2017	1.0%	87.7%		
Ambiguous	12.0%	11.0%		
Charlick			55.5%	2.3%
Freeling			6.0%	63.3%
Ambiguous			38.5%	34.3%

Following on from the qualitative investigation of the infrared spectra, its use for the quantitative prediction of nutritional and bioactive analytes in the faba bean samples was investigated. These results are discussed in the following section.

3.4.4 Prediction of bioactive compounds using IR spectroscopy

The final aspect of the work on faba bean was to investigate the prospect of using infrared spectroscopy (NIRS or MIRS) for the quantification of bioactive compounds and other nutritional-related analytes in this matrix.

As previously mentioned in Section 3.2, Wang et al. (2014a) was the only study to date to investigate the use of NIRS for the prediction of total polyphenol content in faba bean. These authors reported good performance of the calibration model but did not validate their results using an independent test set. Furthermore, no quantitative use of MIRS has been reported in this crop. Consequently, this work aimed to investigate the use of these two techniques (NIRS and MIRS) for the prediction of various analytes (nutritional-related and bioactive compounds), using an independent test set for model validation.

3.4.4.1 Descriptive statistics

The descriptive statistics for the calibration and test sets are provided in Table 3-10. Although the samples were sourced from ten different varieties and two different growing sites, the level of variation in many of the analytes was not exceptionally high. For example, there was only

~1% coefficient of variation (CV) in the protein content for the calibration set (2017 samples). Other analytes such as starch, amylose and amylopectin also showed minimal variation, which could potentially lead to difficulty in creating predictive models. The amount of variation in the FRAP and TPC was moderately high, with CVs of 34 and 31%, respectively.

Table 3-10: Descriptive statistics for the parameters measured in this study, in the calibration and test sets.

	Calibration set – 2017 samples (n=60)		Test set – 2016 samples (n=40)	
Parameter	**Range**	**Mean ± SD**	**Range**	**Mean ± SD**
Moisture (%)	9.13-10.65	9.79 ± 0.32	9.13-10.42	9.68 ± 0.29
Protein (%)	26.46-30.23	27.78 ± 0.88	26.78-29.40	28.35 ± 0.54
FRAP (mg TE/100 g)	202-730	301 ± 102	191-488	287 ± 68
TPC (mg GAE/100 g)	223-691	323 ± 100	214-567	305 ± 87
TMAC (mg cyd-3-glu/100 g)	7.0-20.5	14.7 ± 3.2	8.7-23.0	17.5 ± 3.6
Starch (g/kg)	370-483	405 ± 20	372-412	395 ± 10
Amylose (g/kg)	135-195	166 ± 12	93-158	123 ± 16
Amylopectin (g/kg)	201-327	239 ± 23	230-304	272 ± 16
Vicine (mg/g)	4.38-7.32	5.63 ± 0.70	4.66-7.48	6.06 ± 0.70
Convicine (mg/g)	1.62-3.15	2.31 ± 0.36	2.04-3.18	2.52 ± 0.32
Total VC (mg/g)	6.19-9.22	7.93 ± 0.78	6.71-9.68	8.58 ± 0.77

Note: Total VC = total of vicine and convicine

3.4.4.2 NIR spectra

The raw and pre-processed NIR spectra are shown in Figure 3-13. Major peaks were located at 1200, 1470 and 1936 nm, corresponding to the CH second overtone from structural carbohydrates, OH second overtone from moisture, and OH first overtone, respectively (Manley, 2014). More minor peaks were observed between 1700-1840 nm (attributed to the first overtones of CH and CH_2), 2050-2200 nm (first overtones of the amide A/II and amide I/III bonds in protein) and 2270-2370 nm (combination bands of CH, CH_2 and CH_3 bonds) (Gergely & Salgó, 2007). Overall, the spectra were quite comparable to those reported by Wafula et al. (2020) from common bean (*Phaseolus vulgaris*). There was some minor variation in the spectral amplitude between samples (Figure 3-13A), which was successfully removed following application of the standard normal variate (SNV) algorithm (Figure 3-13B).

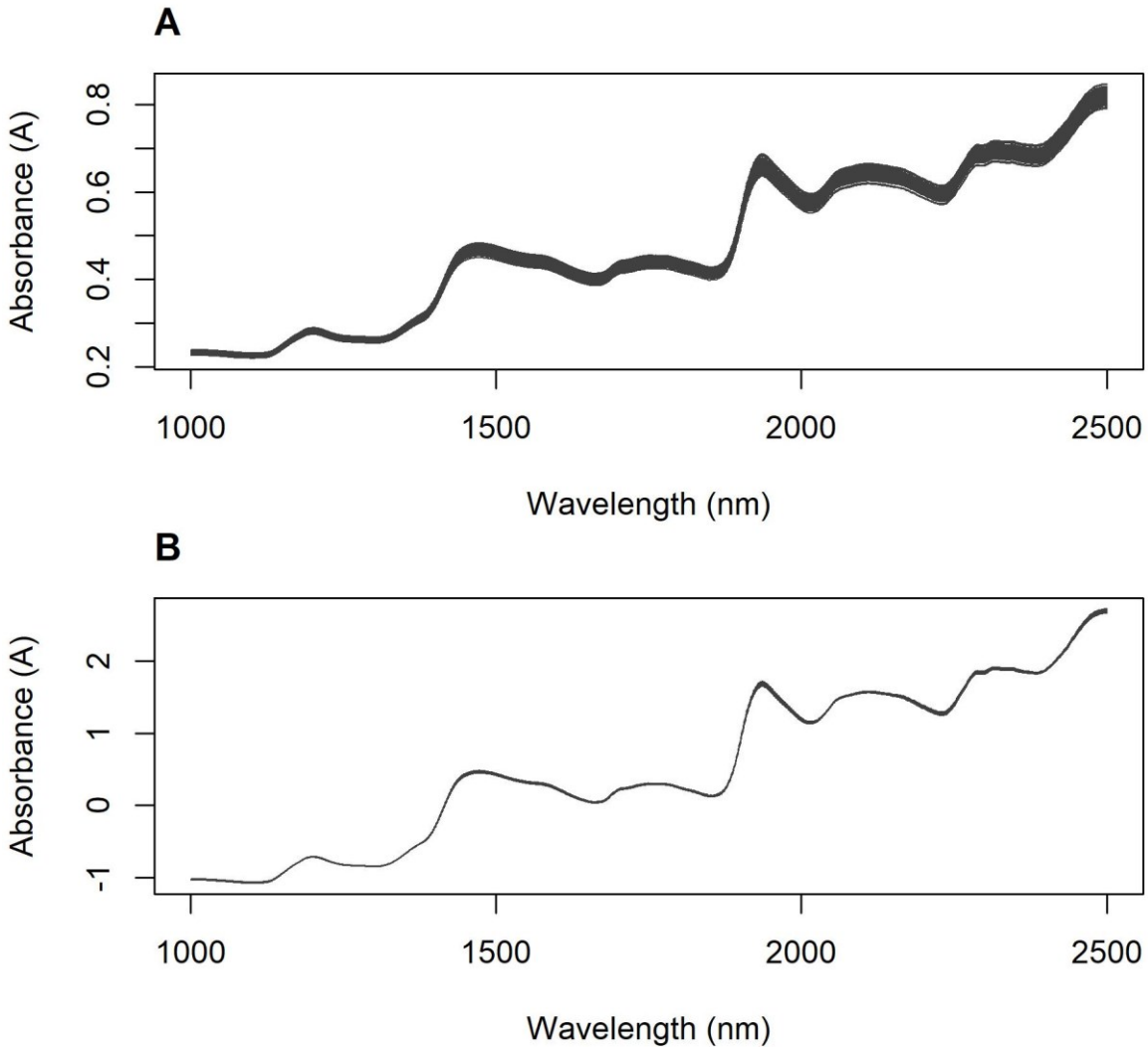

Figure 3-13: The raw absorbance NIR spectra (A) and SNV-processed spectra (B) of the faba bean flour samples.

3.4.4.3 NIR models

PLSR models were developed for each of the 11 analytes specified in Table 3-10, trialling 18 different pre-processing methods for each analyte. The following combinations of pre-processing methods were examined: none (raw spectra), SNV smoothed, 1d5 (1st derivative with 5 smoothing points), 1d11, 1d15, 1d21, 2d5 (2nd derivative with 5 smoothing points), 2d11, 2d15, 2d21, SNV + 1d5, SNV + 1d11, SNV + 1d15, SNV + 1d21, SNV + 2d5, SNV + 2d11, SNV + 2d15 and SNV + 2d21. For each analyte, the optimum model was selected from the R^2, RMSECV and ratio of performance to deviation (RPD) values from LOO cross-validation. The optimum number of components ("factors") for each model was identified from the location of the local minima on the model scree plot, which shows the RMSECV plotted against the number of components (see the example shown in Figure 3-14).

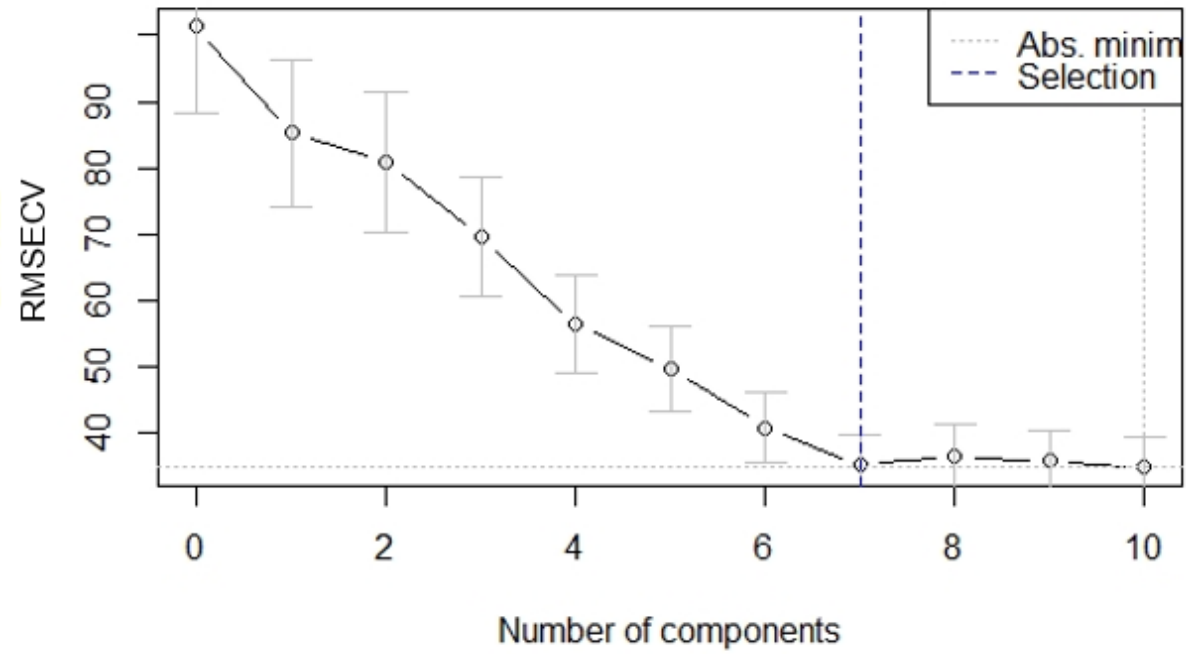

Figure 3-14: Scree plot showing the selection of the optimum number of components (7) for the PLSR model for the prediction of TP content, selected from the minimum RMSECV.

The optimised PLSR models for each analyte are shown in Table 3-11, along with their corresponding figures of merit. A summary of the results obtained on the independent test set are also included in this table, namely the R^2_{test} coefficient, root mean square error of prediction (RMSEP), and the bias, slope and intercept of the calibration model.

The best performing model was found for protein content, followed by TPC, FRAP and convicine. Certain other analytes (e.g., moisture, starch, amylopectin, vicine and total vicine/convicine) showed reasonable cross-validation statistics for the calibration, but much poorer results for the independent test set. Similarly, the models for the TMAC, amylose and amylopectin contents showed no predictive power.

Table 3-11: Optimum PLSR models found for the prediction of the specified analytes using NIR spectroscopy. The calibration set comprised the 2017 samples (n = 60) and the test set comprised the 2016 samples (n = 40).

Parameter	Spectral pre-processing^	Factors	R^2_{cv}	RMSECV	RPD	R^2_{test}	RMSEP	Bias	Slope	Intercept
Moisture (%)	2d11	6	0.65	0.19	1.69	0.19	0.52	-0.37	0.33	6.37
Protein (%)	1d5	7	0.94	0.21	4.15	0.86	0.35	-0.26	0.78	5.93
FRAP (mg TE/100 g)	SNV + 1d15	7	0.73	53	1.94	0.59	87	-61	0.53	102
TPC (mg GAE/100 g)	SNV + 1d15	7	0.88	35	2.86	0.66	76	-54	0.77	28
TMAC (mg cyd-3-glu/100 g)	2d15	1	0.01	3.12	1.01	0.12	4.96	3.6	0.99	3.7
Starch (g/kg)	2d15	6	0.63	12.1	1.66	0.12	20.8	-12.0	0.20	315
Amylose (g/kg)	2d11	2	0.39	9.14	1.29	0.04	34.9	-31.1	0.50	46.5
Amylopectin (g/kg)	2d15	6	0.55	15.3	1.50	0.05	37.1	20.4	-0.16	311
Vicine (mg/g)	SNV + 1d21	10	0.76	0.34	2.07	0.16	1.02	0.30	0.20	4.89
Convicine (mg/g)	SNV	10	0.78	0.16	2.17	0.48	0.41	0.34	0.73	0.93
Total VC (mg/g)	SNV + 1d21	9	0.70	0.42	1.84	0.15	2.93	0.67	0.26	6.5

^ As stated in the materials and methods (Section 3.3.11), the numerical codes refer to Savitzky-Golay derivative pre-processing. For example, 1d15 would be the Savitzky-Golay first derivative of the spectra, using a 15-point window (7 points on either side).

SNV = standard normal variate normalisation

Despite the low range of protein contents found among the calibration samples, the PLSR model for protein showed a high level of linearity (R^2) for both the calibration and test sets (Figure 3-15). The RMSEP of the test set was 0.35, indicating that the model could predict the protein content in independently sourced faba bean samples with ± 0.35% absolute error. As shown in Figure 3-15B, the model was slightly biased towards under-predicting protein contents (bias = -0.26), which appears to be due to the under-representation of high-protein content samples found in the calibration set (Figure 3-15A). Examination of the loadings plot for the protein prediction model (Figure 3-16) revealed the strongest influence at 1898 nm, apparently corresponding to the shoulder of the amide A/II region (Gergely & Salgó, 2007).

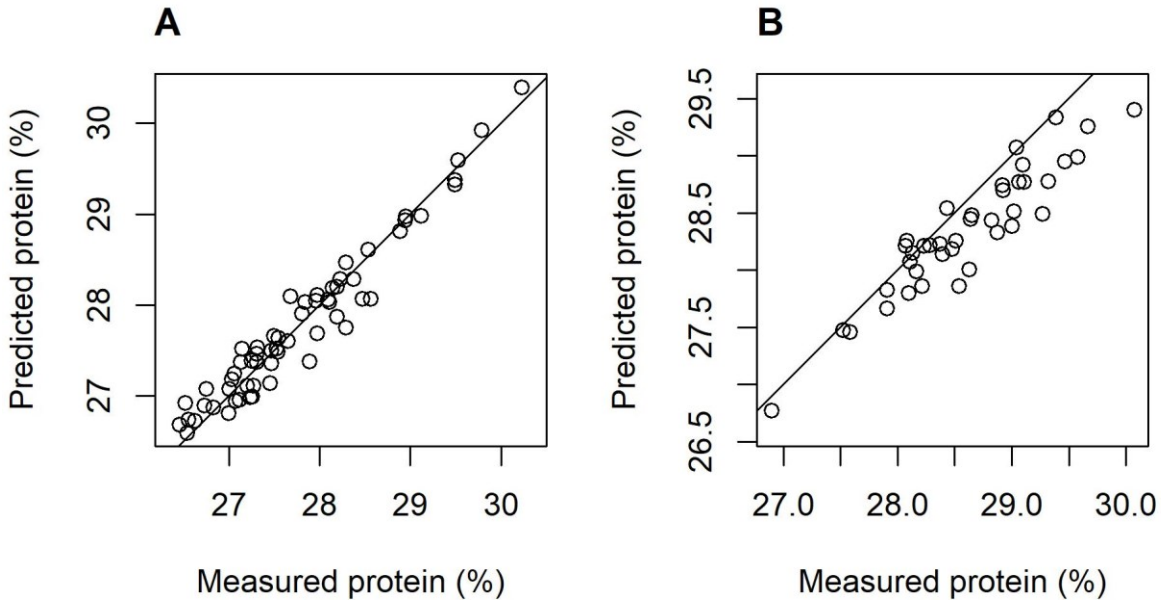

Figure 3-15: (A) Actual vs predicted protein contents for the calibration set (n=60). (B) Actual vs predicted protein contents for the test set (n=40).

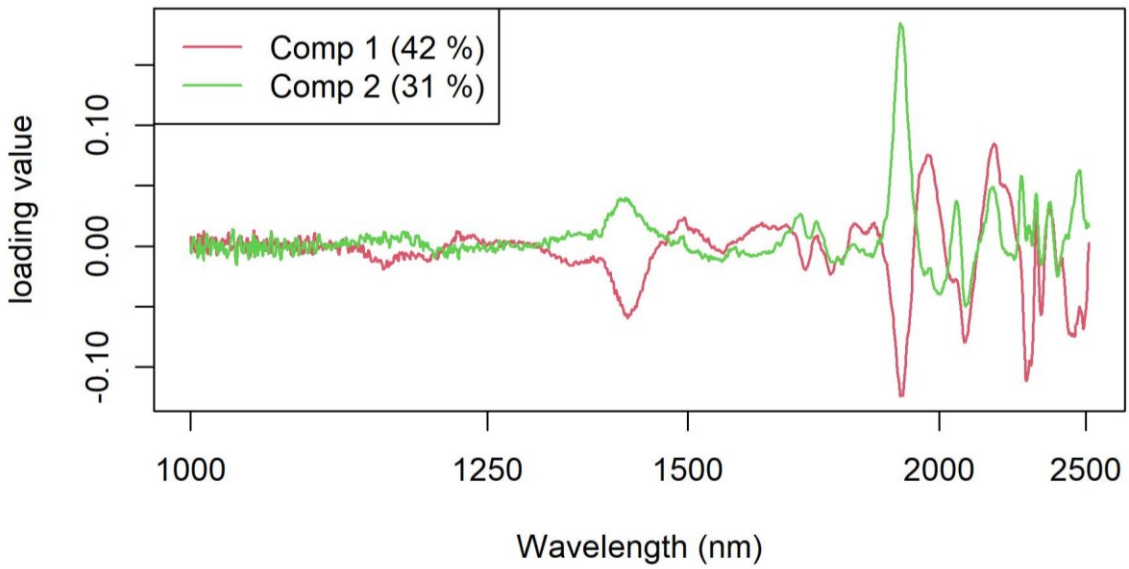

Figure 3-16: Loadings plot for the prediction of protein content in faba bean flour.

Compared to the protein model, the PLSR model for the prediction of TPC (Figure 3-17) did not perform as well, as anticipated for analytes that are present at lower concentrations. Although the cross-validation statistics showed a high linearity (R^2_{cv} of 0.88), the RMSEP (76 mg GAE/100 g) was almost twice that of the RMSECV (35 mg GAE/100 g), indicating relatively poorer performance of the model on independently sourced samples. This RMSEP value was also considerably higher than the standard laboratory error of the reference spectrophotometric method (mean SD of 8.3 mg GAE/100 g for n=100 samples analysed in duplicate). The TPC of most samples were under-predicted by the PLSR model, giving it an overall negative bias (-54 mg GAE/100 g). It is worthwhile noting that despite the broad range of genotypes and growing sites represented in the faba bean samples, there were relatively few samples containing moderate or high TPC levels, which is likely the source of bias in this model. Nevertheless, the model could still discriminate between samples with low and high TPC.

Interestingly, the loadings plot showed a strong influence of wavelengths around 1901 nm (Figure 3-18), which may be attributable to the amide bond region, as previously discussed for the protein model loadings. This may hint that the TPC model was using a secondary correlation between protein and TPC in order to estimate the TPC of the unknown samples.

Other influential wavelengths were 1960 and 2149 nm, which may correspond to the first overtones of OH stretch and CH stretch, respectively (Toledo-Martín et al., 2018). These wavelengths were found by previous authors to be important in the prediction of total phenolic

content in raspberries using NIRS (Toledo-Martín et al., 2018) and likely correspond to the OH of polyphenol groups and CH of aromatic rings, respectively.

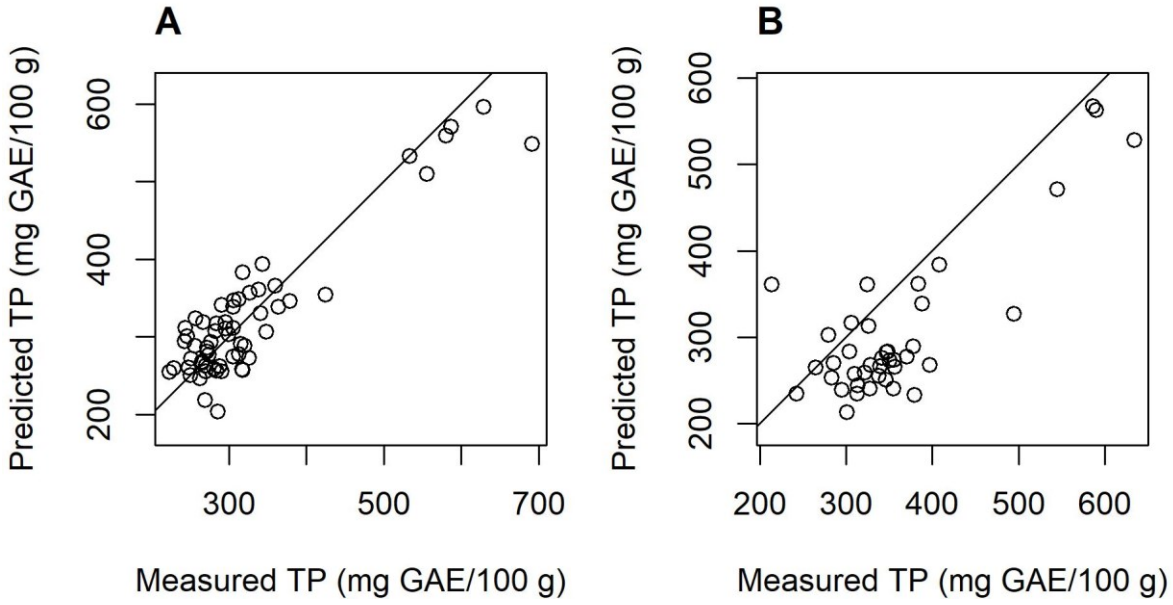

Figure 3-17: Actual vs predicted TP contents for the calibration set (A) and test set (B).

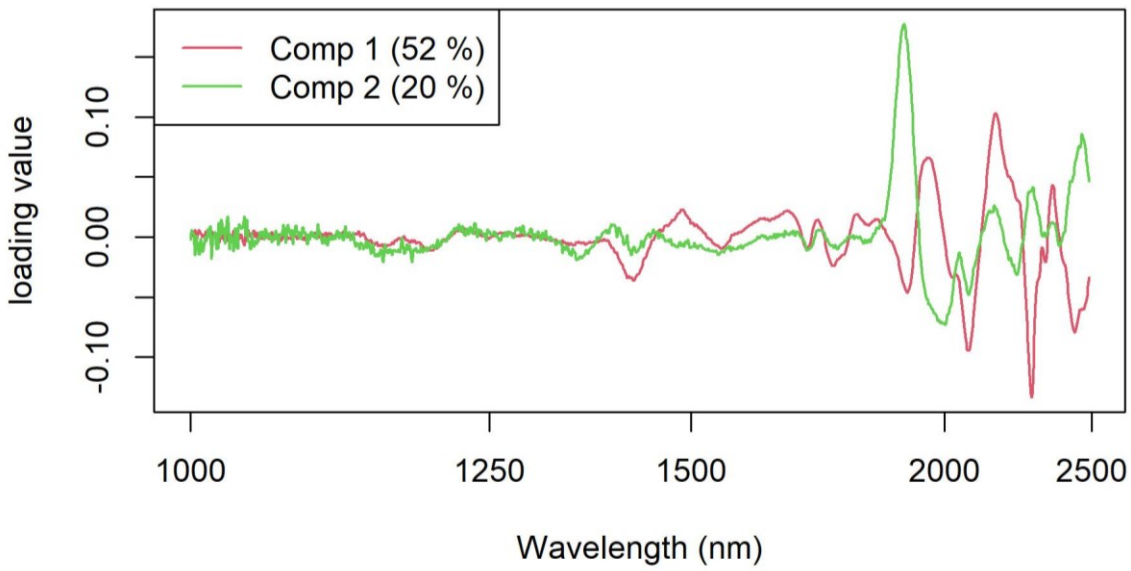

Figure 3-18: Loadings plot for the prediction of TP content in faba bean flour.

The calibration and test set results for the FRAP PLSR model (Figure 3-19) were quite similar to those found for the TPD model. Again, the model did not perform as well on the independent test set, with the results limited to discriminating between samples with high or low FRAP values.

The loadings plot was almost identical to that observed for the TPC model (compare Figures 3-18 and 3-20), with the predominating wavelengths being 1901, 1960 and 2149 nm. This indicates that the same analyte(s) were being measured by the TPC and FRAP models, which is a logical outcome if the phenolic compounds present in faba bean are primarily responsible for its antioxidant activity.

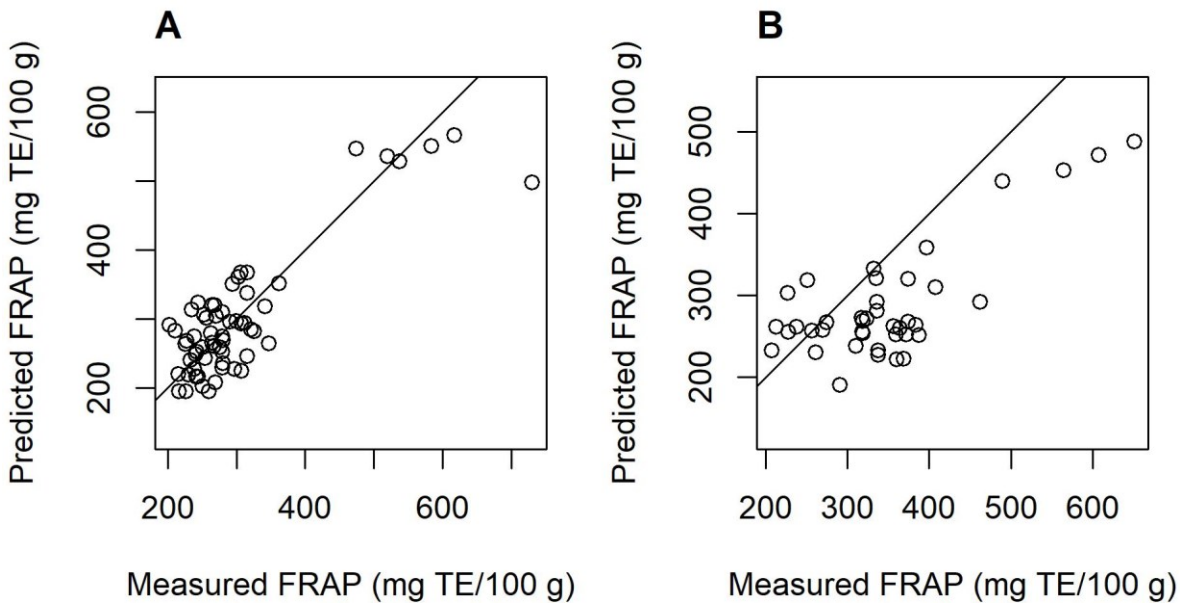

Figure 3-19: Actual vs predicted FRAP values for the calibration set (A) and test set (B).

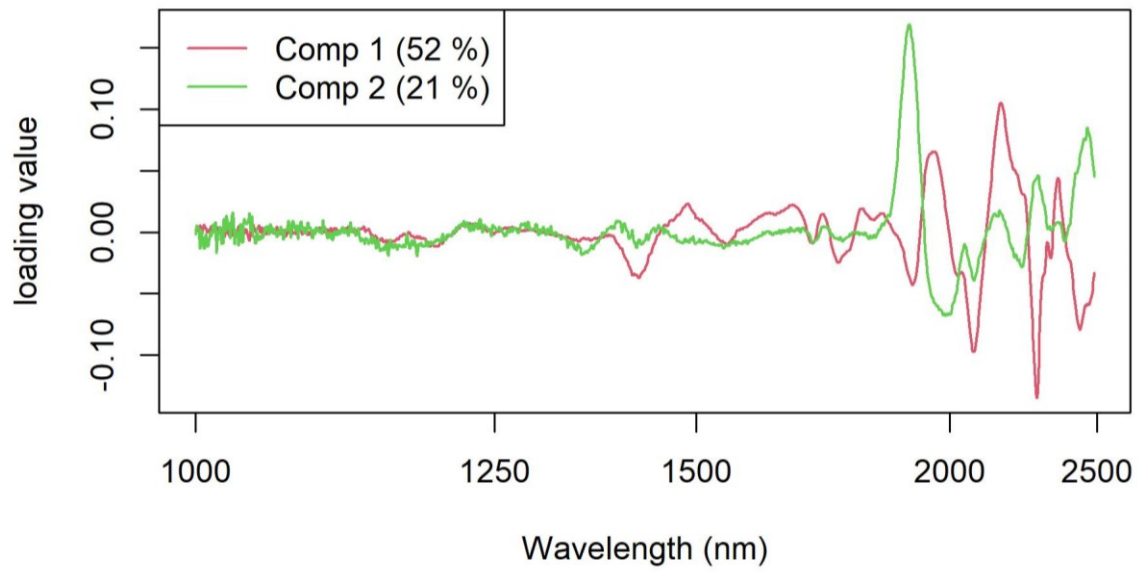

Figure 3-20: Loadings plot for the prediction of FRAP in faba bean flour.

3.4.4.4 MIR spectra

The raw and SNV-processed MIR spectra of the faba bean flour samples are shown in Figure 3-21. The spectra were more complex compared to the NIR region, with major peaks centred around 3250 cm^{-1} (attributable to OH stretch from moisture), 3000-2850 cm^{-1} (CH_2 and CH_3 stretch), 1640 cm^{-1} (C=O stretch of amides or other carbonyl-containing compounds), 1550-1200 cm^{-1} (various amide and phenol bonds) and 1000 cm^{-1} (aromatic rings in cellulose) (Abbas et al., 2017; Dufour, 2009; Karoui et al., 2010; Mecozzi & Sturchio, 2017). Compared to the NIR spectra, there was greater variation in the amplitude of the MIR spectra between different samples; however, most of this was removed through the SNV algorithm (Figure 3-21B). Although one sample appeared to be an outlier (Figure 3-21B), it was included in the PLSR model development in order to avoid artificially improving the reported model statistics (Williams et al., 2017).

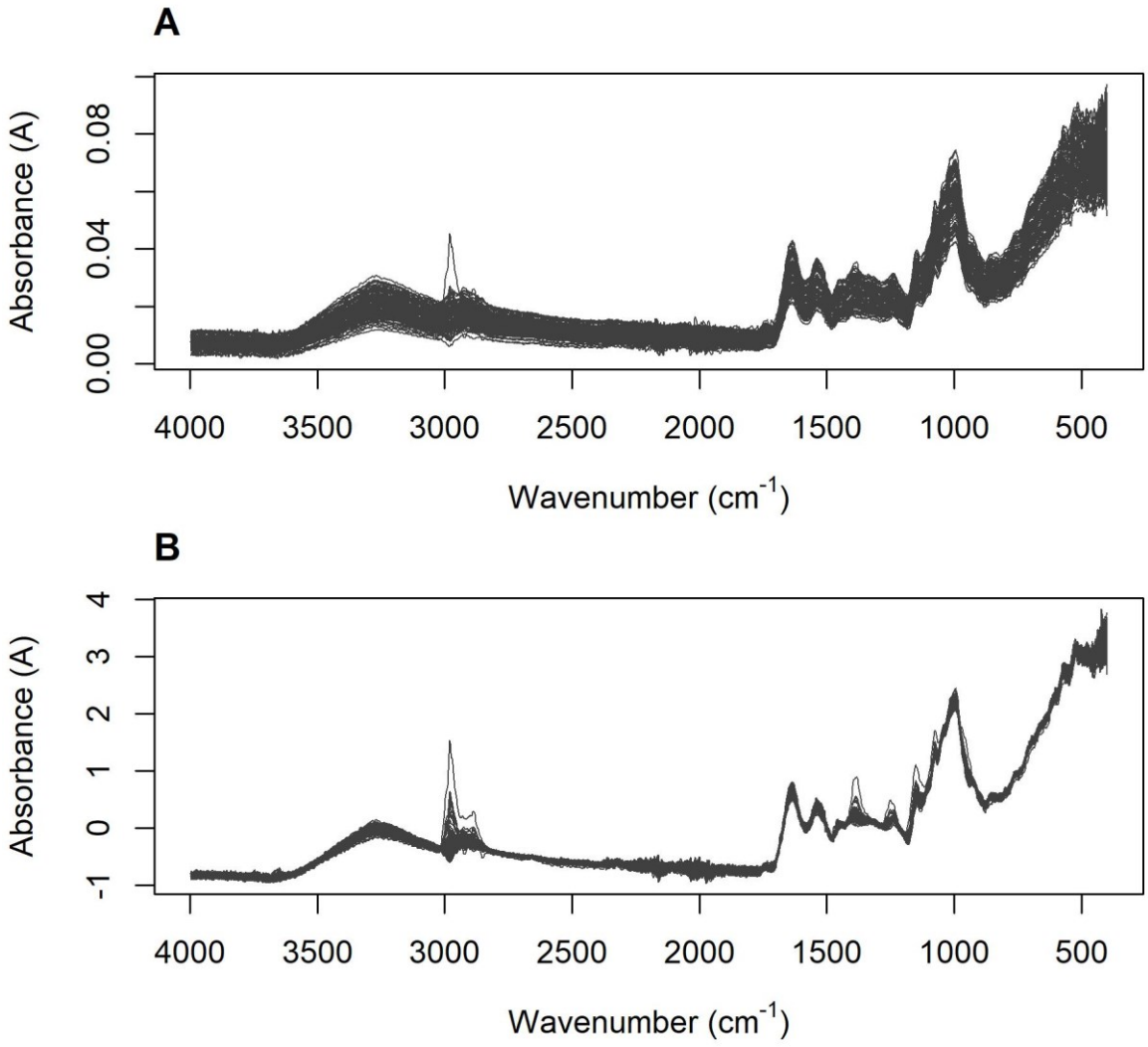

Figure 3-21: The raw MIR spectra (A) and SNV-processed spectra (B) of the faba bean flour samples.

3.4.4.5 MIR models

As detailed for the NIR spectra, PLSR models were created for each analyte, with the optimum pre-processing method determined through the LOO cross-validation statistics. The best-performing models for each analyte are detailed in Table 3-12. In contrast to the results found for the NIRS models, accurate calibration models were unable to be created from the MIR spectra (RPD < 1.5 for all). Similarly, the models had no predictive power when applied to the independent test set. For this reason, no calibration plots or loading plots are shown for the MIRS models.

Table 3-12: Optimum PLSR models found for the prediction of various analytes using MIR spectroscopy.

Parameter	Spectral pre-processing	Factors	R^2_{cv}	RMSECV	RPD	R^2_{test}	RMSEP	Bias	Slope	Intercept
Moisture (%)	None	4	0.47	0.23	1.38	0.09	0.46	-0.36	0.70	2.68
Protein (%)	SNV	2	0.17	0.80	1.11	0.11	1.20	0.90	0.23	22.00
FRAP (mg TE/100 g)	SNV + 2d21	2	0.26	87.1	1.18	0.00	79.0	2.10	0.09	260
TPC (mg GAE/100 g)	SNV + 1d21	5	0.37	79.2	1.27	0.04	112	-38.6	0.21	232
TMAC (mg cyd-3-glu/100 g)	SNV	1	0.10	3.0	1.06	0.00	4.45	1.57	0.09	16.10
Starch (g/kg)	SNV + 2d15	2	0.27	17.0	1.18	0.02	23.2	-15.5	-0.17	462
Amylose (g/kg)	SNV + 2d21	2	0.05	11.4	1.04	0.10	49.2	-46.7	0.57	26.3
Amylopectin (g/kg)	SNV + 2d11	8	0.45	17.0	1.36	0.01	37.7	27.2	-0.06	287
Vicine (mg/g)	None	1	-0.05	0.71	0.98	0.01	0.73	0.25	0.80	1.39
Convicine (mg/g)	SNV	6	0.14	0.33	1.09	0.29	0.49	0.39	0.55	1.34
Total VC (mg/g)	None	1	-0.06	0.79	0.98	0.00	0.89	0.47	0.10	7.74

3.4.4.6 General discussion on the IR models

The NIRS model for the prediction of protein content performed acceptably on the independent test set, confirming the suitability of this method for the rapid assessment of proximate analysis, as reported by numerous previous authors (El-Sherbeeny & Robertson, 1992; Wang et al., 2014a). The model performance on the independent test set (R^2_{test} = 0.86; RMSEP = 0.35%) was quite comparable to the (dependent) test set results reported by Wang et al. (2014a) (R^2_{test} of 0.94 and RMSECV of 0.33%). However, Wang et al. (2014a) used a much broader range of protein contents for calibration (23.8-33.1%) compared to those used here (26.5-30.2%), indicating that the model accuracy could be retained when using a narrower calibration range for this analyte. Examination of the model loadings plot confirmed that the selected wavelengths (principally 1898 nm) corresponded to the absorbance of amide bonds, which are found in protein. In other words, the model was looking in the correct region(s) of the NIR spectrum to be able to detect protein.

Although the NIRS models for TPC and FRAP did not perform as accurately as it did for the protein content, they showed some potential for the estimation of these parameters. Several of the key predictive wavelengths (1960 and 2149 nm) appeared to correspond to the first overtones of OH stretch and CH stretch of phenol-containing compounds, while the 1901 nm wavelength also showed a large contribution. This may indicate that at least part of the predictive power of the model was due to a potential secondary correlation between protein and TPC/FRAP levels.

There does not appear to be any previous work reporting the non-invasive prediction of antioxidant capacity in faba bean; however, Wang et al. (2014a) used NIRS to create prediction models for the total phenolic content in Chinese faba bean. The reported performance on the (dependent) test set was somewhat better than that reported here, with a R^2_{test} of 0.78 and RMSEP of 37 mg/100 g. However, this reported accuracy is likely somewhat over-optimistic, as the test set was not sourced independently from the calibration samples.

One of the challenges in creating accurate prediction models for TPC and FRAP from the present dataset may stem from the lack of samples containing intermediate TPC or FRAP values. This was despite the fact that this dataset included samples from ten different genotypes and two different growing locations, which would be anticipated to provide a wide range of phytochemical diversity. Consequently, the models developed here would appear to be only useful for screening purposes (e.g., classifying samples into high or low phenolic or FRAP contents) and not the absolute quantification of these analytes.

Although MIRS has previously been reported for the prediction of total phenolic content in common bean flour (Carbas et al., 2020), no accurate calibration models could be developed

from the MIR spectra in this study for any of the analytes investigated. Although several of the analytes showed moderate R^2 values in the calibration models, none showed acceptable performance when applied to the independent test set. This is likely due to the difficulty in applying a consistent level of pressure between the sample and the interface of the MIR instrument across all of the samples analysed. In turn, the amount of pressure applied has a strong influence on the signal amplitude and sensitivity. Consequently, the lack of predictive power observed here appears to be due to the lack of reproducibility in the MIR spectra. Although the main peak locations remain the same at different levels of pressure, the peak details can be obscured or altered in their relative heights at lower contact pressures (see the example shown in Figure 3-22). In future work, it might be possible to develop an automated device to apply a consistent amount of pressure between samples.

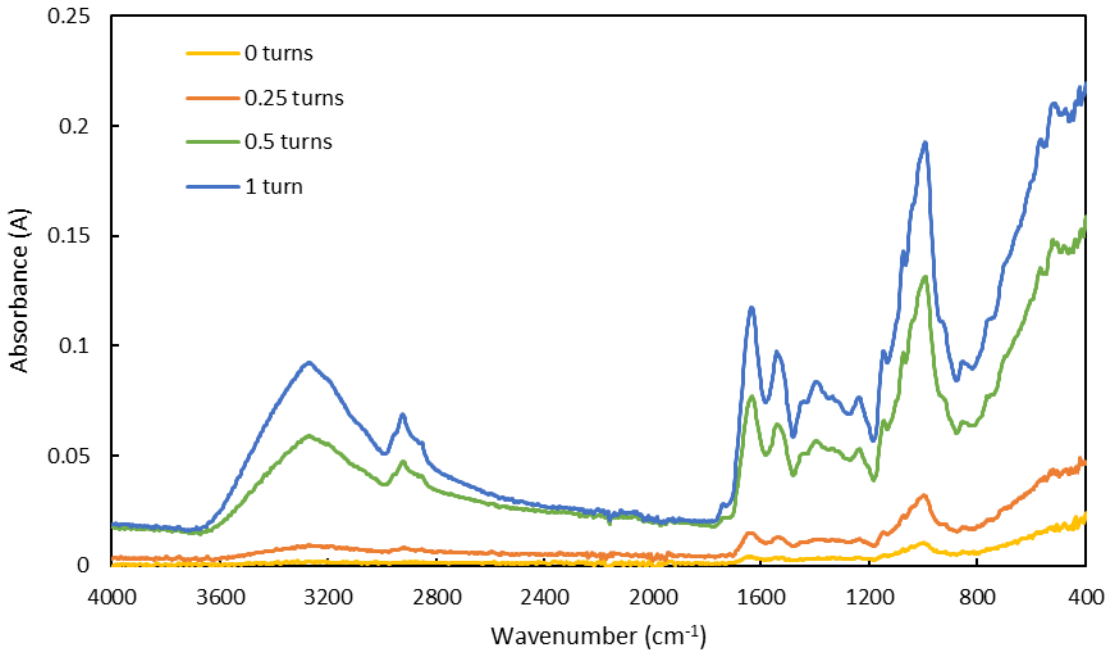

Figure 3-22: MIR spectra collected with increasing levels of pressure applied to the sample. Note that turns = the number of turns of the pressure collar following the first contact of the tip with the sample; higher values correspond to more applied pressure.

The NIR instrument used in this work does not suffer from the same drawback, as the sample was presented in a sample cup placed on top of the instrument, thus ensuring consistent presentation between different samples. Some authors have solved this issue with MIRS through the use of internal standards which are mixed with the sample (Sastre Toraño & Hattum, 2001). Although this is relatively minimal sample preparation, it would partially defeat the benefit of IR spectroscopy as a rapid analytical technique requiring no sample preparation.

Other authors have reported using ATR-MIR for the analysis of bioactive compounds without any special modifications to the instrument or sample matrices (Amanah et al., 2020; Carbas et al., 2020; Cozzolino et al., 2020; Kokalj Ladan et al., 2017; Uncu et al., 2019). However, it should be noted the majority of these studies did not confirm the model performance using independent test sets. Furthermore, most of studies using MIR spectroscopy for the analysis of bioactive compounds have been performed on liquid samples/extracts (see Table 2-3 in Section 2.5), which improves the reproducibility of sample presentation. On the other hand, phenolic concentrations are even lower in liquid extracts compared to the raw samples, meaning that the MIR analysis of grain extracts is unlikely to be a viable option.

3.5 Summary

The ten varieties of Australian faba bean investigated here displayed considerable variation in their anthocyanin, phenolic and antioxidant contents. In particular, the PBA Rana variety showed the highest level of total phenolics and the highest antioxidant capacity, followed by PBA Samira. Furthermore, PBA Rana contained significantly higher concentrations of most of the 10 individual phenolic compounds quantified in the faba bean samples.

In addition to being the first work to profile the phenolic acid and flavonoid composition in >5 commercial varieties of Australian faba bean, the impact of growing location on phenolic composition was also assessed. Location had a significant impact on the levels of five phenolic compounds (protocatechuic, vanillic and chlorogenic acids, vitexin and rutin), while genotype × location interactions were observed for three phenolic acids (vanillic, syringic and *trans*-ferulic acids).

Although most of the nutritional and anti-nutritional analytes included here could not be predicted using NIRS in this study (i.e., starch, amylose, amylopectin, vicine and convicine), good performance was found for the prediction of protein content, and acceptable performance for the estimation (i.e., high or low content) of TPC and FRAP values. In contrast, none of the MIR models showed acceptable results for any analyte. The results suggest that NIRS may be used for the rapid approximation of TPC and FRAP in faba bean. As NIRS is routinely used for the determination of protein in pulse crops for quality assurance purposes, application of such models for TPC/FRAP could potentially be utilised to extract further information from the spectra which are already collected for protein determination purposes.

Chapter 4. Wheat

4.1 Introduction

Following on from the promising results found for the use of infrared spectroscopy in faba bean (Section 3.7), the second grain crop chosen for investigation was wheat. In addition to its economic importance as the most widely grown grain crop in Australia, wheat contains lower phenolic levels compared to faba bean (Benayad et al., 2021). This makes it an ideal matrix to test the performance of infrared spectroscopy at low phenolic/antioxidant concentrations.

As with faba bean, there have been few studies reporting the use of infrared spectroscopy for the estimation of bioactive compounds in this crop, aside from a very recent study using NIRS to predict TPC in whole wheat flour (Tian et al., 2021c). Preliminary work by the author, presented at the 69th Australasian Grain Science Association Conference (Johnson et al., 2019), suggested that MIRS may show promise for the prediction of total phenolics and antioxidant capacity in wheat flour, although only a limited number of samples (n=17) were investigated in that study. Consequently, the work presented in this chapter aimed to investigate and compare the potential of NIRS and MIRS for the rapid prediction of total phenolic content and antioxidant capacity in Australian wheat flour using a larger sample size.

4.2 Background

Common or bread wheat (*Triticum aestivum* L.) is the second-most grown grain crop worldwide (after corn), with over 760 million tonnes harvested in 2020 (FAO, 2022). It is a vital crop for ensuring global food security, supplying approximately 20% of dietary calorie requirements across the globe (Shiferaw et al., 2013). Furthermore, wheat is the largest broadacre crop in Australia, with 14 million tonnes produced in the 2019-2020 season (Australian Bureau of Statistics, 2021).

Although the health benefits of wheat are generally considered to be lower than that of pulse crops (Curran, 2012; Rebello et al., 2014), it does contain a number of potentially health-benefitting compound classes, including polyphenols, carotenoids, vitamin E and phytosterols (Baublis et al., 2000; Dalton et al., 2012). Although these compounds are most abundant in whole grain products (Dalton et al., 2012), they are also found in refined wheat products, albeit at lower levels (Lu et al., 2014). The major polyphenol compounds previously identified from wheat include gallic acid, 4-hydroxybenzoic acid, vanillic acid protocatechuic acid, syringic acid, ferulic acid, caffeic acid, chlorogenic acid, sinapic acid, *p*-coumaric acid, cinnamic acid, rutin, quercetin, kaempferol, vitexin, isovitexin and resveratrol (Kaur et al., 2021; Tian et al.,

2021b). There is also increasing interest in developing wheat lines with elevated phenolic acid contents (Li et al., 2008; Tian et al., 2021a).

As with other grain crops, near-infrared spectroscopy has been widely applied for the quality assessment of wheat and wheat flour, including the prediction of protein (Ye et al., 2018), fibre (Stubbs et al., 2010), starch (Peng & Zhang, 2010) and other quality parameters important for bread-making (Williams, 2020). A recent study by Tian et al. (2021c) investigated the use of NIRS for the prediction of TPC in whole wheat flour, reporting a high level of linearity (R^2_{val} = 0.90) and precision (RMSEV = 7.1 mg GAE/100 g) on a dependent test set. However, the authors did not consider the prediction of antioxidant capacity in this study. Additionally – as with faba bean – there are few studies investigating the use of MIRS for the quality assessment of wheat.

Consequently, this work investigated and evaluated the performance of NIRS and MIRS for the prediction of phenolic content, anthocyanin content and antioxidant capacity in Australian-grown wheat.

4.3 Materials and methods

4.3.1 Wheat samples

A total of 65 wheat samples were sourced from the Australian Export Grains Innovation Centre (AEGIC). In order to provide maximum variability in their anticipated bioactive contents, these comprised samples sourced from a range of different wheat grades and production areas (n=17 samples), as well as samples from a pearling trial (n=48 samples from 3 different wheat grades). The samples from the pearling trial had varying levels of bran remaining on the kernel after the different pearling times (0-20 secs); thus, would be expected to display significantly different phenolic and antioxidant contents.

The sample details are provided in Appendix A. As commercial Australian wheat samples are sold by grade rather than variety, each sample likely comprised batches of wheat from a range of different varieties, rather than being from one pure wheat variety. Consequently, analysis by wheat variety was not performed here.

At AEGIC, the samples were impact milled to produce whole seed flour (Falling No. grinder, 0.8 mm screen), following the methods described for faba bean in Section 3.3.1. Moisture content was also determined by AACC International Method 44-15.02 and protein using a LECO TruMac N protein analyser, as described in Section 3.3.1. All results were expressed on an oven-dried weight basis.

4.3.2 Analysis of TPC, FRAP and TMAC

Following extraction of the polar phenolic compounds using 90% methanol (as described for faba bean in Section 3.3.3), the TPC, FRAP and TMAC were analysed following the methods in Sections 3.3.4, 3.3.5 and 3.3.6, respectively. Extractions and subsequent analyses were performed in duplicate. All results were expressed as mg of the corresponding standard per 100 g of sample.

Phenolic profiling by HPLC-DAD was not conducted on these samples as the phenolic composition of wheat flour is already well documented (Barros Santos et al., 2019; Liu et al., 2010; Lu et al., 2014; Mazzoncini et al., 2015; Mpofu et al., 2006; Tomé-Sánchez et al., 2020).

4.3.3 NIR and MIR spectroscopy

NIR and MIR spectra were collected from the wheat flour using the instruments and settings described in Sections 3.3.8 and 3.3.9. Five replicate MIR spectra and three replicate NIR spectra were collected from each flour sample, repacking the instrument with fresh flour each time. The mean of replicate scans was used in subsequent data analysis.

4.3.4 Spectral data analysis

Partial least squares regression analysis of the spectra was conducted in R Studio, using different pre-processing combinations of SNV and 1st or 2nd derivatives, as described in Section 3.3.11.

To create the calibration set, 50 of the samples (~75% of the total) were selected from the dataset using the Kennard-Stone algorithm, using the Mahalanobis distance across the first two 2 principal components. The remaining 15 samples (~25%) were used as a dependent test set to assess the performance of the optimised models. No outliers were removed from the dataset. The full range of wavelengths was used in the PLS regressions. The optimum pre-processing method and number of model components were determined through full LOO cross-validation performed on the calibration samples.

4.4 Results and discussion

4.4.1 Correlation analysis

To investigate the relationships that existed between the various phytochemical analytes, Pearson R linear correlation analysis was conducted on this data and a correlogram produced (Figure 4-1). The only significant positive correlation was between TPC and FRAP (r_{63} = 0.250, P<0.05). In addition, the FRAP was negatively associated with moisture content (r_{63} = -0.337, P<0.01), and the TPC was negatively correlated with protein content (r_{63} = -0.632, P<0.001).

This stood in contrast to the positive correlation between TPC and protein previously observed in faba bean (Section 3.4.1.2).

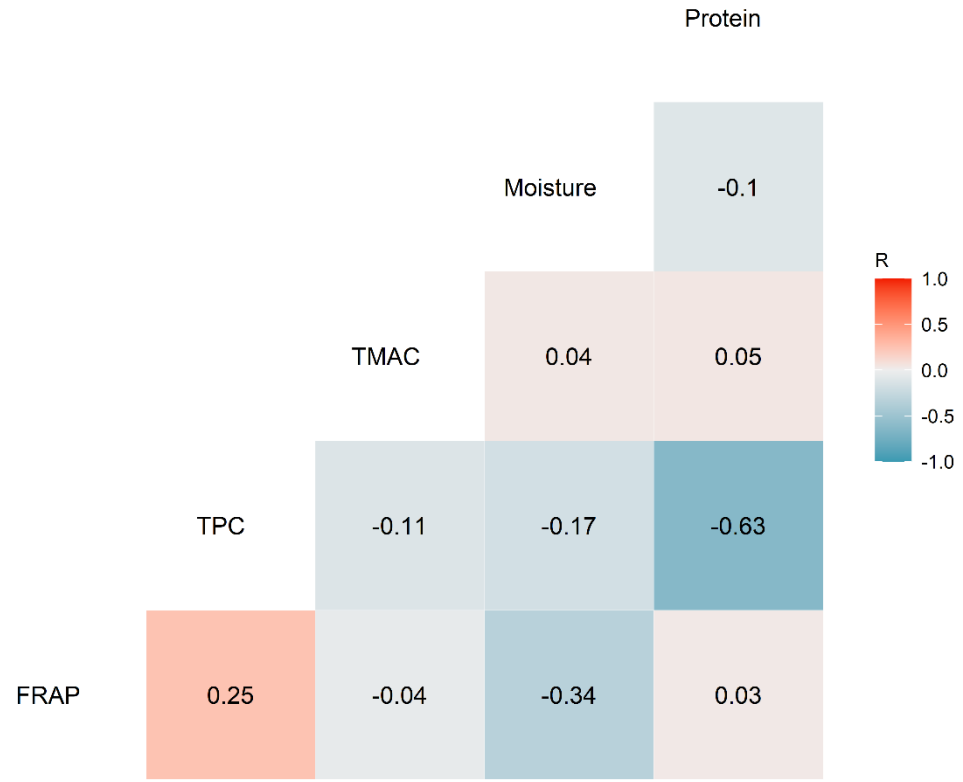

Figure 4-1: Correlogram showing the correlations between the phytochemical constituents and physical parameters of the wheat samples (n = 65). Correlations with R values above 0.24 or below -0.24 were statistically significant at α = 0.05.

4.4.2 Prediction of bioactive compounds using IR spectroscopy

4.4.2.1 Descriptive statistics

The descriptive statistics for the calibration and test sets are provided in Table 4-1. The range, mean and standard deviation were comparable between the calibration and test sets for most analytes, indicating that the Kennard-Stone algorithm did an adequate job of partitioning the spectra between these groups. The moisture and protein content showed a reasonably large level of variation between samples, with a moderate amount of variation also found in the FRAP content. However, the variation in the TPC was much lower, suggesting that it would be more difficult to create an accurate model for this analyte.

Table 4-1: Descriptive statistics for the parameters measured in the wheat flour samples, for both the calibration and test sets.

Parameter	Calibration set (n=50)		Test set (n=15)	
	Range	Mean ± SD	Range	Mean ± SD
Moisture (%)	10.1-14.5	11.90 ± 1.39	8.7-14.0	11.70 ± 1.30
Protein (%)	10.66-17.46	14.16 ± 1.98	11.43-16.84	13.97 ± 2.25
FRAP (mg TE/100 g)	14.4-64.0	32.8 ± 16.4	19.4-53.5	25.3 ± 8.5
TPC (mg GAE/100 g)	129.7-179.8	149.5 ± 11.3	139.0-167.6	151.9 ± 9.7
TMAC (mg cyd-3-glu/100 g)	0.0-10.0	2.5 ± 2.2	0.0-3.9	2.0 ± 1.1

4.4.2.2 NIR spectra

Figure 4-2 shows the mean raw and SNV-processed NIR spectra of all 65 wheat flour samples. The NIR spectra showed major peaks at 1467 nm and 1934 nm, corresponding to the second and first overtone of the OH bond (from moisture in the sample or other R-OH bonds), respectively (Ziegler et al., 2016a). Other smaller peaks included 1205 nm (CH second overtone from lipids and structural carbohydrates), 2107 nm (amide bond from protein), broad peaks between 1680-1850 nm (CH first overtone stretch) and 2260-2370 nm (CH stretching deformation of CH_2 and CH_3 bonds), and shoulders at 1360 nm (second overtone of CH_3 or ArOH) and 1580 nm (tentatively assigned to CH_3/ArCH first overtone from starch) (De Girolamo et al., 2019; Rodríguez et al., 2019).

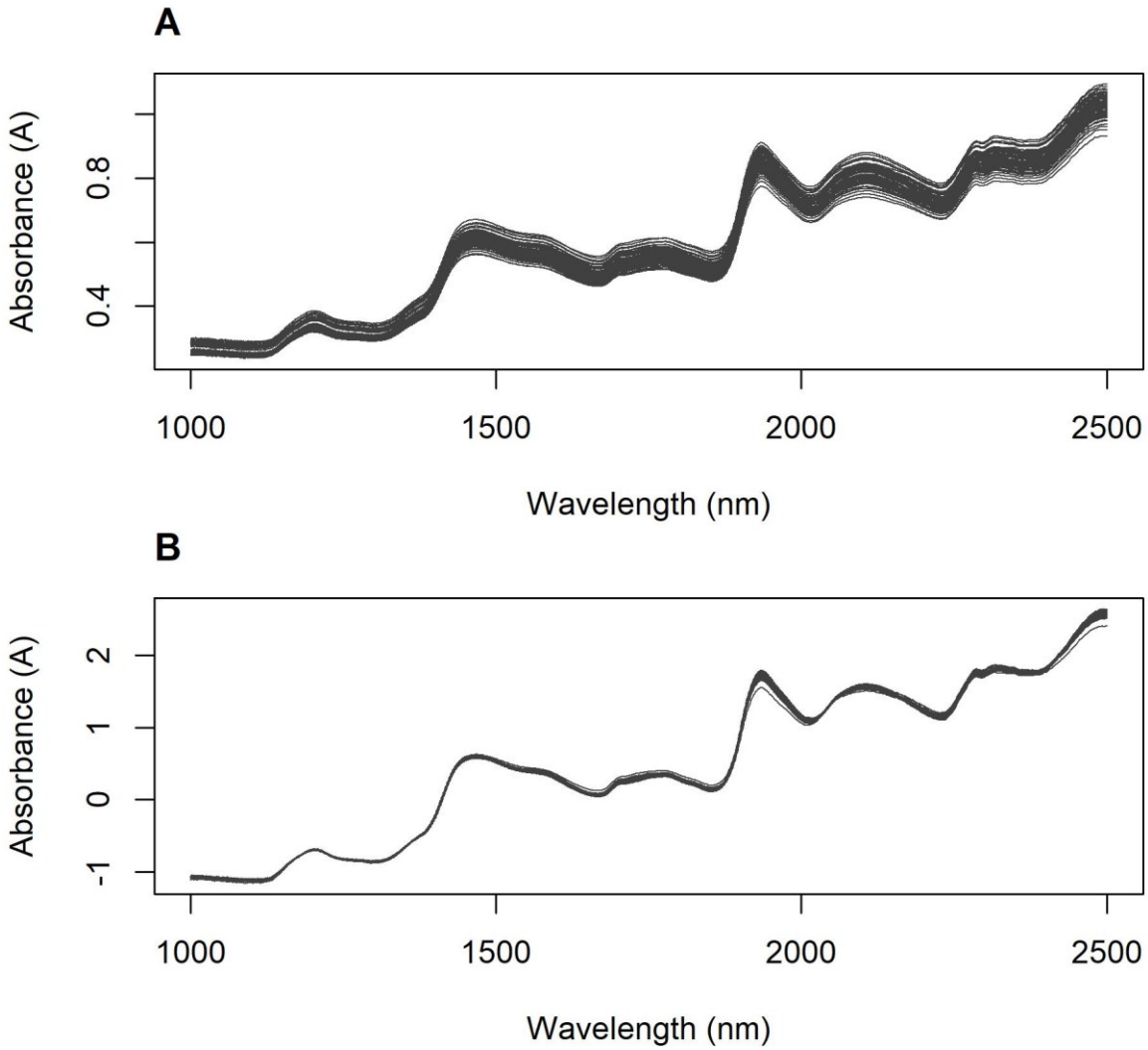

Figure 4-2: The absorbance raw NIR spectra (A) and SNV-processed spectra (B) of the wheat flour samples.

4.4.2.3 NIR models

The PLSR models with the optimised spectral pre-processing methods for each analyte are detailed in Table 4-2. The best-performing model was for protein content (Figure 4-3), which showed a R^2_{test} of 0.991 and RMSEP of 0.22%. Furthermore, the model showed minimal bias and the slope and intercept were extremely close to their ideal values of 1 and 0, respectively (Table 4-2). Examination of the loadings plot for the protein prediction model (Figure 4-4) showed predominant contributions in the N-H asymmetric stretch and amide II region (1928-1941 nm), as anticipated for a model measuring protein content (Kays et al., 2000). Other minor contributions were at 2226 nm (potentially corresponding to the amide I and III region) and 1732 nm (which may be due to CH first overtones in gluten, the major protein found in wheat) (Kays et al., 2000).

Table 4-2: Optimum PLSR models found for the prediction of the specified analytes using NIR spectroscopy, using separate calibration and dependent test sets.

Parameter	Spectral pre-processing	Factors	R^2_{cv}	RMSECV	RPD	R^2_{test}	RMSEP	Bias	Slope	Intercept
Moisture (%)	SNV + 1d15	6	0.84	0.56	2.49	0.76	0.62	-0.04	0.966	0.35
Protein (%)	SNV	5	0.974	0.32	6.25	0.991	0.22	-0.06	0.998	-0.03
FRAP (mg TE/100 g)	SNV	6	0.88	5.6	2.95	0.917	5.4	-0.64	0.620	9.2
TPC (mg GAE/100 g)	SNV + 1d21	5	0.61	7.0	1.62	0.83	3.9	-0.10	0.907	14.1
TMAC (mg cyd-3-glu/100 g)	1d11	4	0.05	2.1	1.03	0.00	1.9	-0.63	0.02	1.98

Table 4-3: Optimum PLSR models found for the prediction of various analytes using MIR spectroscopy, using separate calibration and dependent test sets.

Parameter	Spectral pre-processing	Factors	R^2_{cv}	RMSECV	RPD	R^2_{test}	RMSEP	Bias	Slope	Intercept
Moisture (%)	1d21	4	0.86	0.52	2.65	0.83	0.65	0.08	1.45	-5.10
Protein (%)	SNV + 2d15	6	0.92	0.55	3.65	0.93	0.62	-0.10	0.89	1.40
FRAP (mg TE/100 g)	SNV + 1d15	6	0.88	5.2	2.93	0.83	7.0	1.56	1.20	-4.21
TPC (mg GAE/100 g)	SNV	5	0.51	7.8	1.45	0.73	5.6	-1.53	0.76	33.8
TMAC (mg cyd-3-glu/100 g)	None	1	-0.04	2.1	0.99	0.02	1.6	0.54	-1.25	5.55

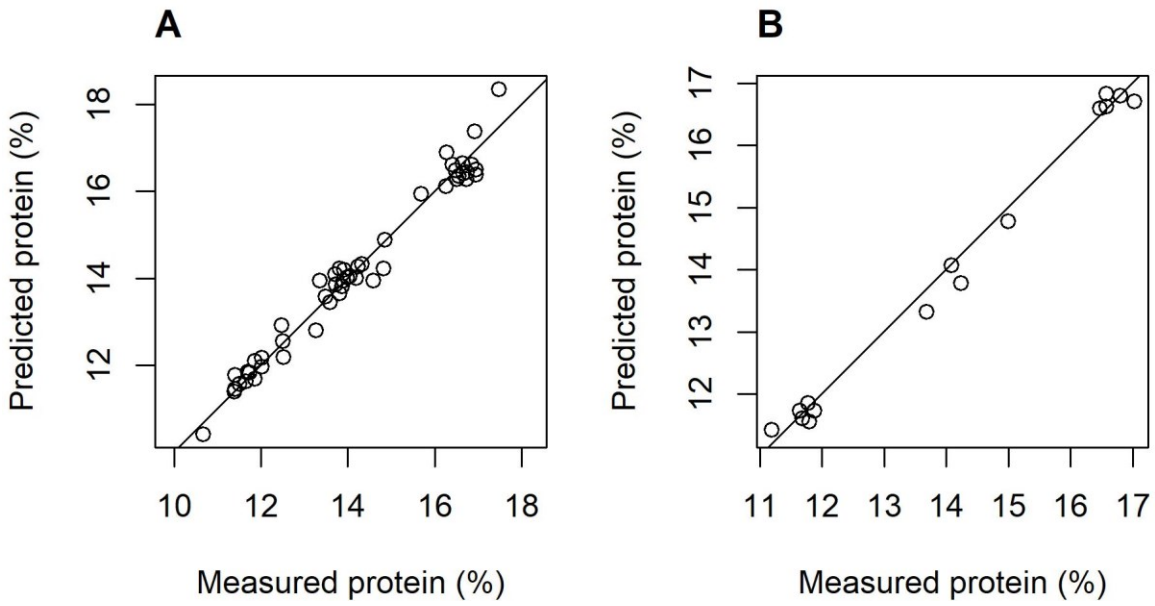

Figure 4-3: Actual vs NIRS-predicted protein values for the calibration (A) and test (B) samples.

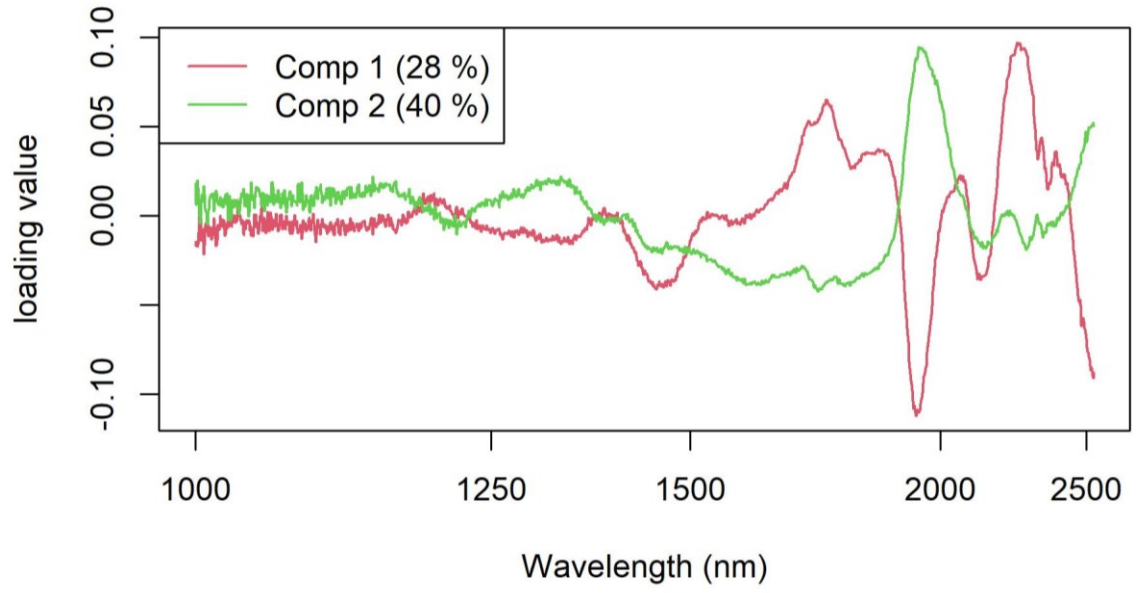

Figure 4-4: Loadings plot for the NIRS prediction of protein content in wheat flour.

The NIRS model for moisture content performed acceptably, with acceptable linearity, very low bias and a slope close to 1 (Table 4-2). However, the RPD was considerably lower than that for protein, indicating relatively higher error associated with the moisture model. The predominant loading for the moisture content model was at 1900, corresponding to the first overtone of the OH bond (Ziegler et al., 2016a), with more minor contributions at 1949 and 2156 nm.

In contrast to the results for protein and moisture, the model for TMAC showed no predictive power, likely due to the very low concentrations of this analyte.

The NIRS model for FRAP content also performed well, although it should be cautioned that there were no samples with intermediate FRAP values (Figure 4-5). For example, in the test set, only one sample had a high FRAP value, which was moderately underpredicted by the model. If this point was removed, then the R^2_{test} fell significantly to 0.46, although the RMSEP improved to 2.9 mg TE/100 g. Nevertheless, most of the other sample points fell very close to the regression line (Figure 4-5B). The main loadings for the FRAP model (Figure 4-6) showed a negative peak at 1934 nm (the N-H asymmetric stretch and amide II region) and a positive peak at 2102 nm (potentially resulting from the combination bands of ROH groups). Thus it appears that the antioxidant capacities of the samples were negatively correlated with their protein contents and positively associated with their levels of phenolic compounds (containing ROH groups).

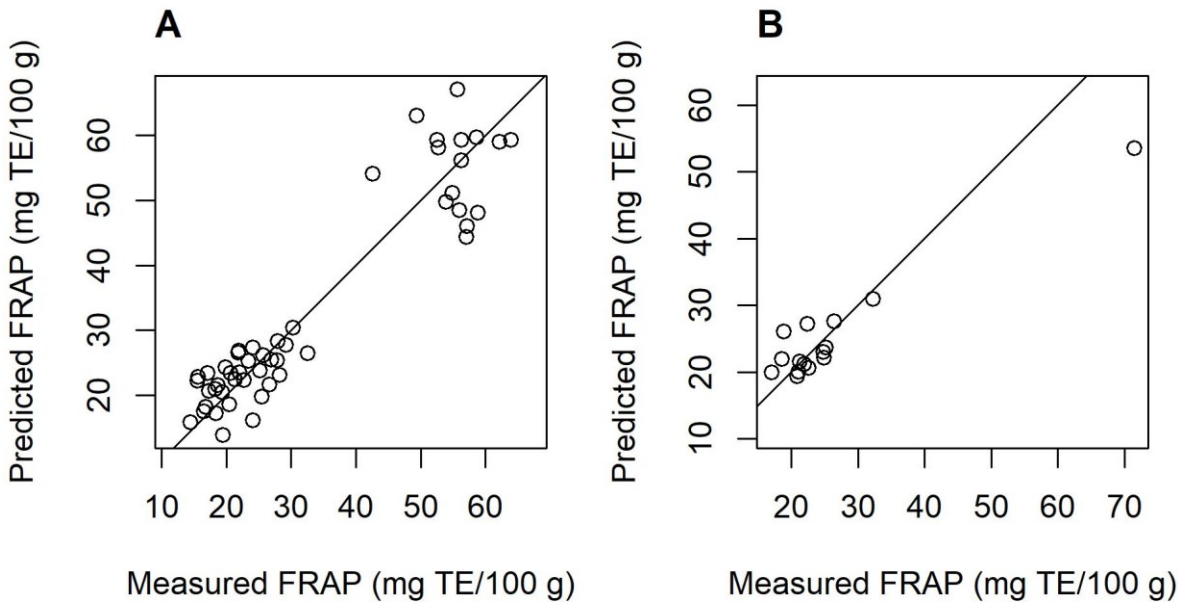

Figure 4-5: Actual vs NIRS-predicted FRAP values for the calibration (A) and test (B) samples.

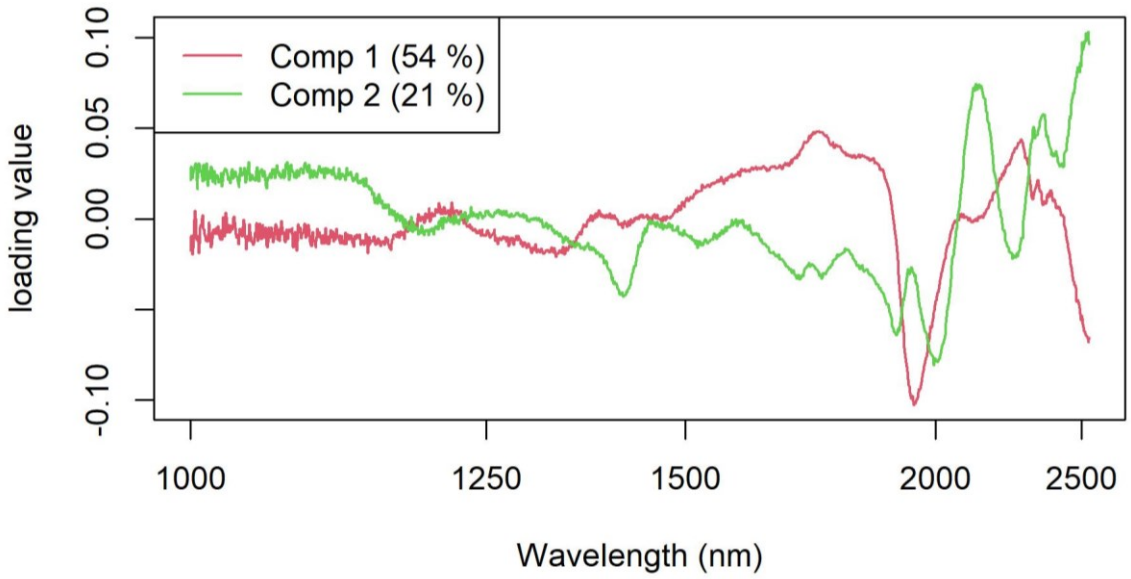

Figure 4-6: Loadings plot for the NIRS prediction of FRAP in wheat flour.

The TPC of the samples was more evenly distributed compared to the FRAP (for both the calibration and test sets), containing samples of a range of TPCs throughout the calibration range (Figure 4-7). Although the linearity was lower compared to the FRAP model (R^2_{test} = 0.73), the RMSEP was better at just 3.9 mg GAE/100 g. This was evidenced in most of the test set samples being very close to their true values across the entire calibration range (Figure 4-7B). The loadings plot for the TPC model (Figure 4-8) appeared visually different to the FRAP model due to the use of the first derivative of the spectra in the former. However, the main influential wavelengths were quite similar, being centred at approximately 1930 nm and 2110 nm. Again, these are likely to correspond to amide and ROH bonds, respectively.

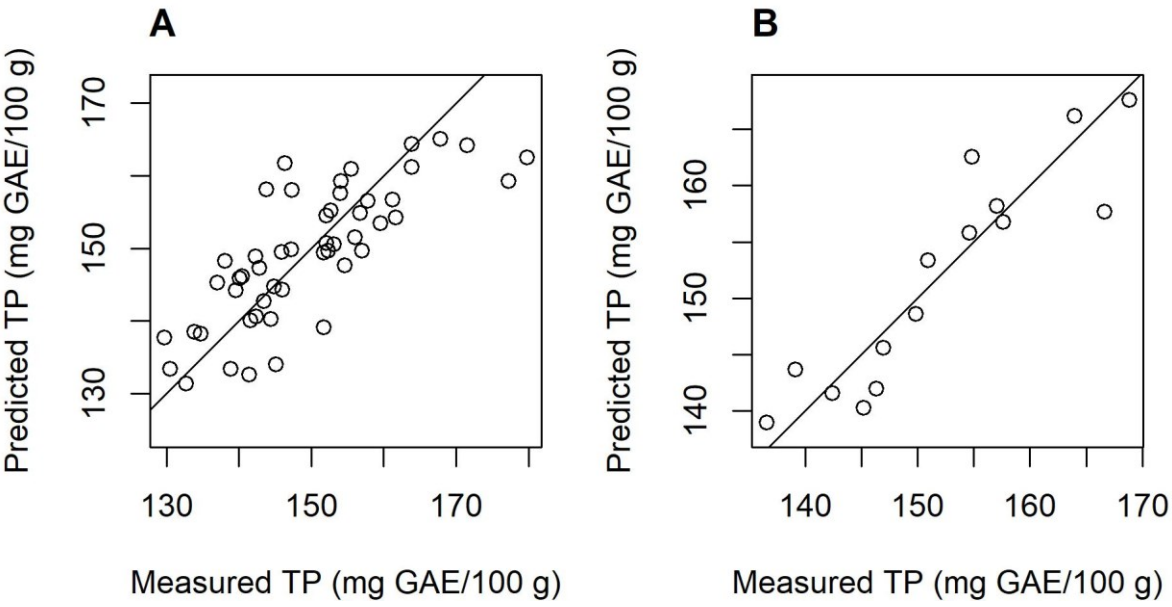

Figure 4-7: Actual vs NIRS-predicted TPC values for the calibration (A) and test (B) samples.

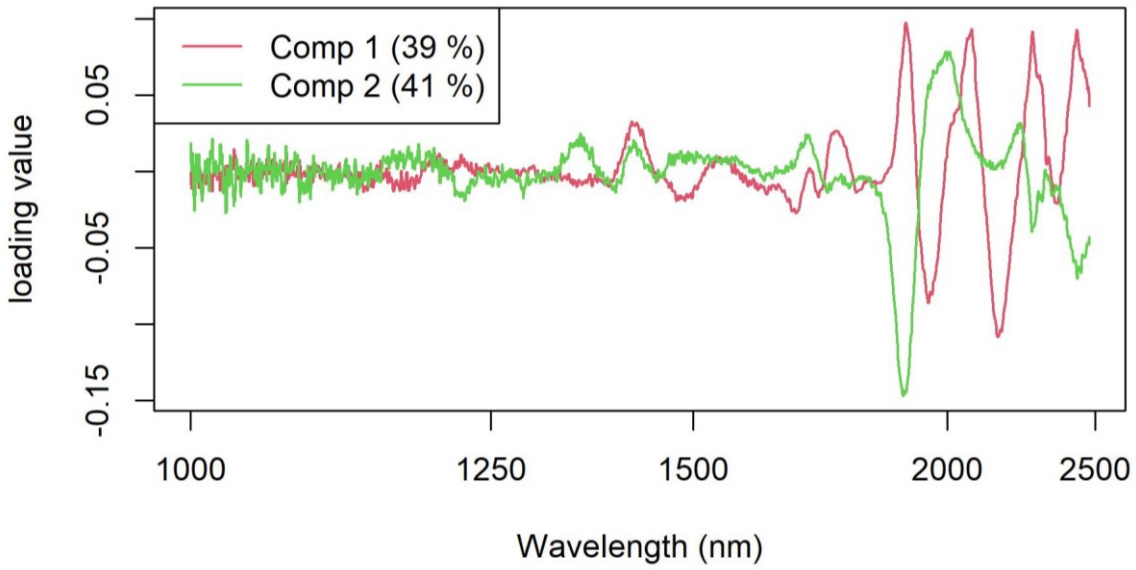

Figure 4-8: Loadings plot for the NIRS prediction of total phenolic content in wheat flour.

4.4.2.4 MIR spectra

The MIR spectra of the wheat samples are shown in Figure 4-9A. The spectra appeared broadly similar to that observed for faba bean (Section 3.4.4.2), indicating the presence of similar aetiological compounds/bonds. As can be seen from this figure, there was considerable variation in the amplitude of the MIR signal, resulting from inconsistencies in applying each of the samples to the ATR crystal with the same degree of pressure. For example, the spectra from one batch of samples showed a lower amplitude compared to another batch of samples analysed on the following day (Figure 4-9A). However, application of the SNV pre-processing algorithm removed most of this variation (Figure 4-9B).

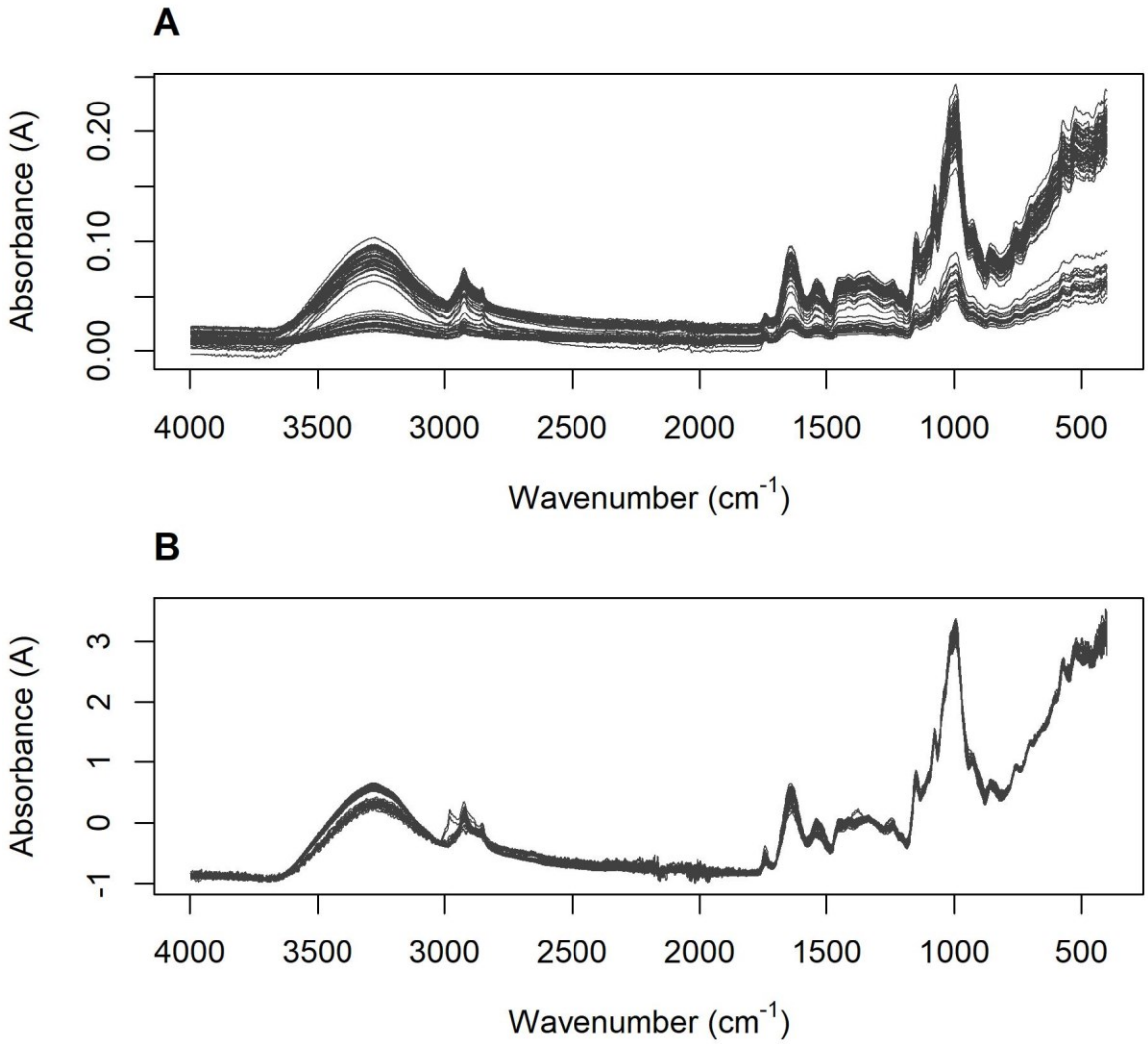

Figure 4-9: The raw MIR spectra (A) and SNV-processed spectra (B) of the wheat flour samples.

4.4.2.5 MIR models

Compared to the NIRS results, the MIRS models showed lower R^2_{test} and higher RMSEP values for all analytes (Table 4-3); although the magnitude of this difference was not very high for most analytes. For example, the RMSEP values for FRAP and TPC were only slightly higher, while the RMSEP for moisture content was almost identical between the two instrument types.

However, the MIRS results for the prediction of protein were less than half as accurate compared to NIRS (Figure 4-10). The loadings for the protein model (Figure 4-11) showed that the primary influences were from regions corresponding to CH stretch (~3050-2870 cm^{-1}), amide I and II bands (1760-1400 cm^{-1}) and aromatic groups from cellulose (~970 cm^{-1}) (Ji et al., 2020).

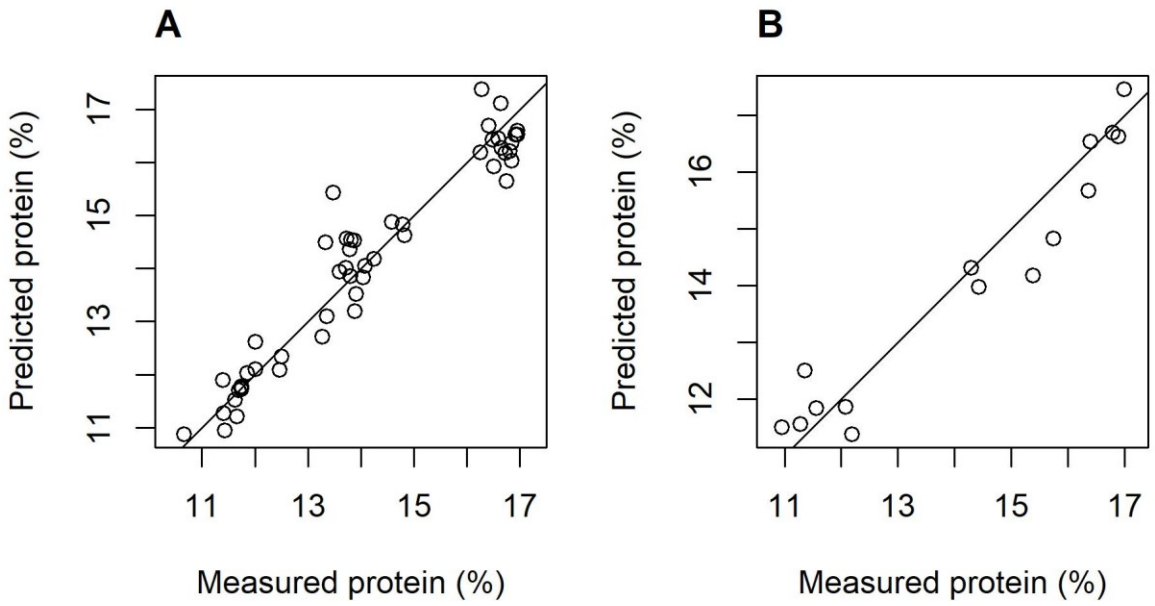

Figure 4-10: Actual vs MIRS-predicted protein values for the calibration (A) and test (B) samples.

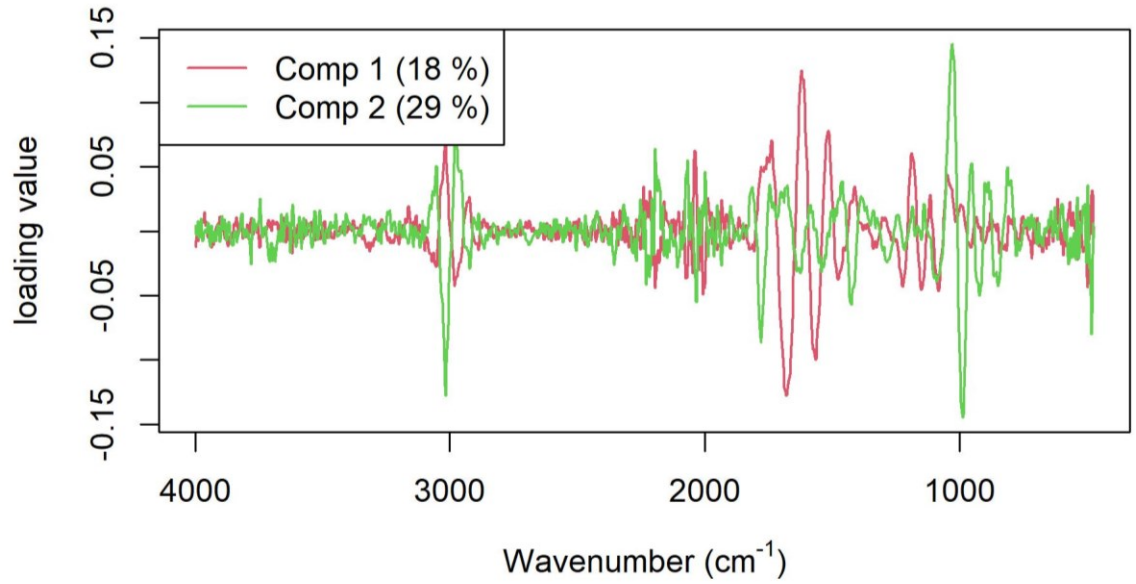

Figure 4-11: Loadings plot for the MIRS prediction of protein content in wheat flour.

Again, the included wheat samples displayed only low or high FRAP values (with no samples of intermediate values), while the TP contents of the samples were distributed roughly equally throughout the calibration range. Consequently, the prediction accuracy of the FRAP model (Figure 4-12) would likely be improved through inclusion of samples of intermediate values. The loadings for the FRAP model (Figure 4-13) were somewhat similar to those of the protein content model, but with the strongest influences from the CH stretch (~3000 cm^{-1}) and

cellulose bond (~1000 cm^{-1}) regions, with moderate influence from the carbonyl (~1700 cm^{-1}) and amide regions (1600-1400 cm^{-1}).

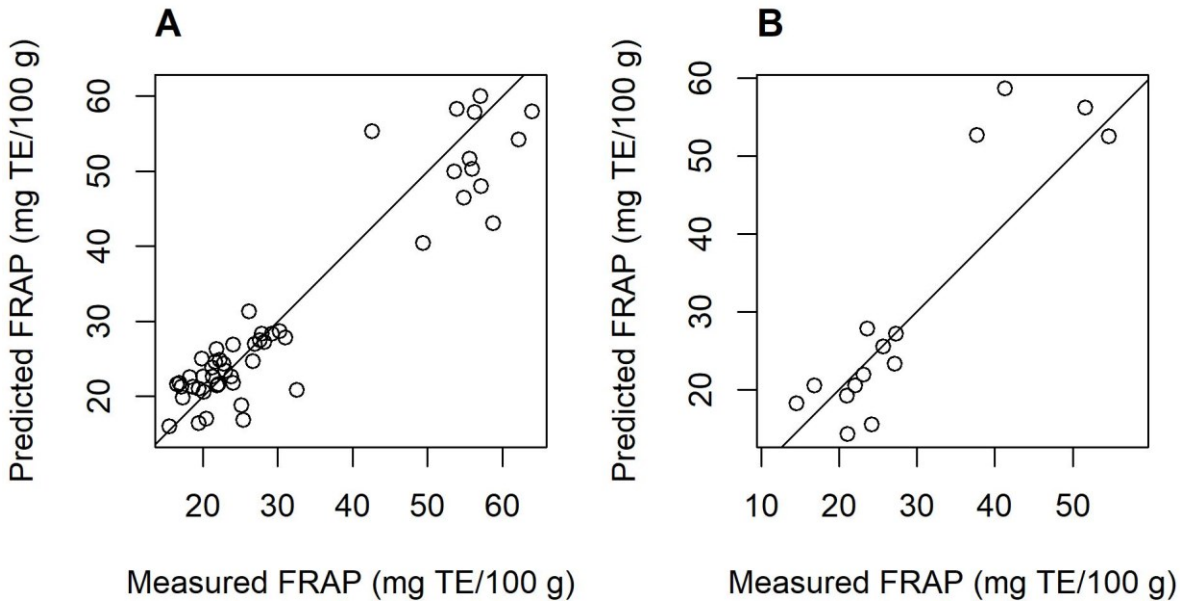

Figure 4-12: Actual vs MIRS-predicted FRAP values for the calibration (A) and test (B) samples.

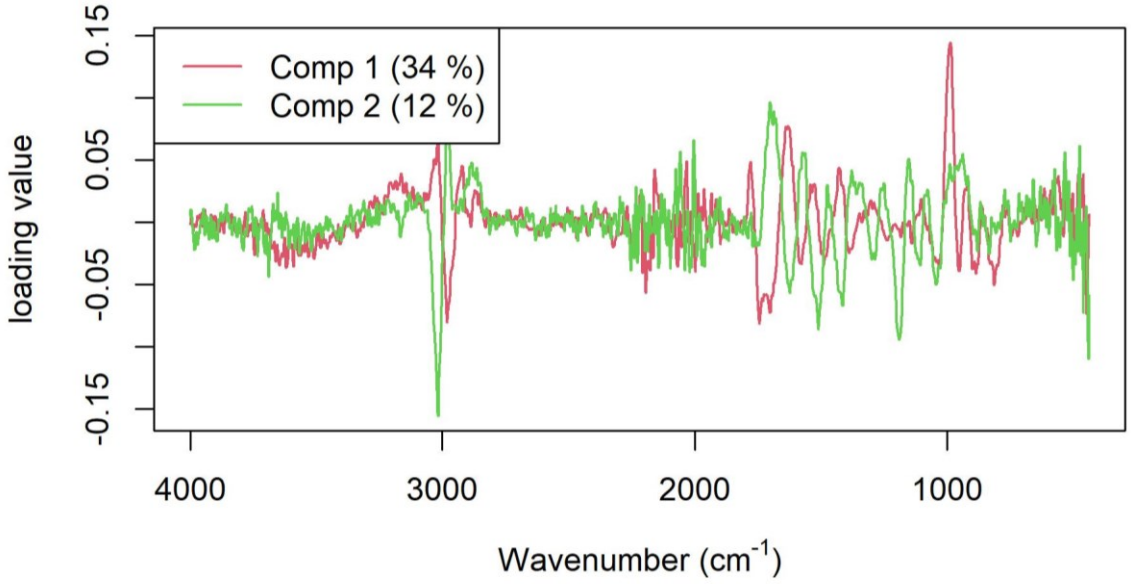

Figure 4-13: Loadings plot for the MIRS prediction of FRAP in wheat flour.

As observed for FRAP, the prediction accuracy of the MIRS model for TPC (Figure 4-14) was somewhat poorer than the NIRS model for this analyte (cf. Figure 4-7); however, MIRS still appeared suitable for the estimation (although not exact quantification) of TPC. In contrast to

the FRAP loadings, the loadings plot for the TPC prediction model showed influences from the OH (~3400 cm^{-1}), carbonyl (~1700 cm^{-1}) and cellulose (~1000 cm^{-1}) regions (Figure 4-15).

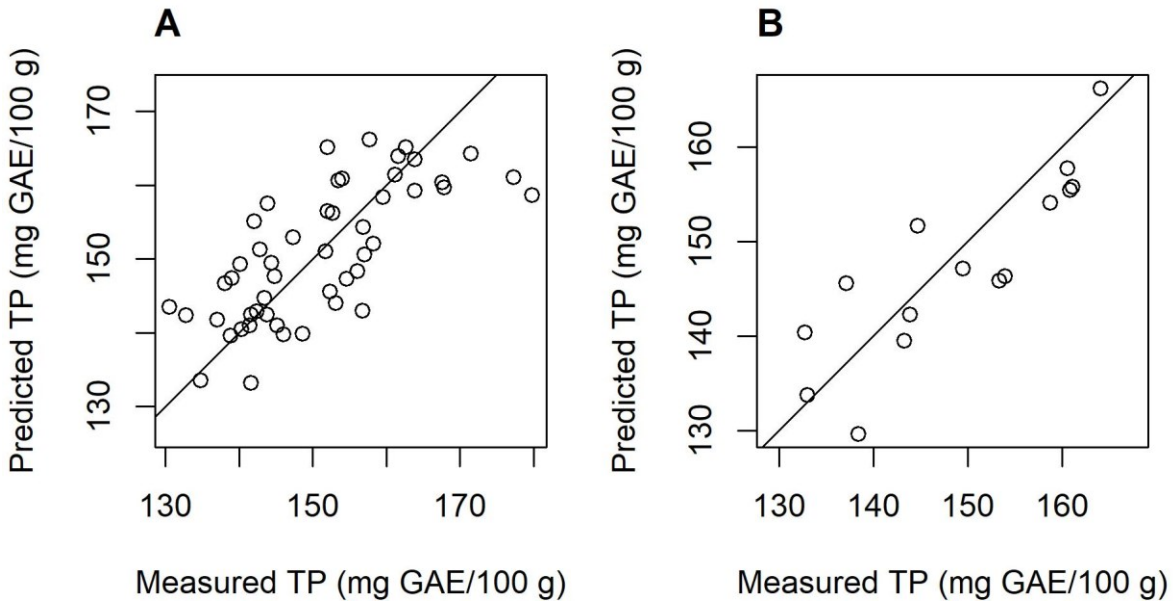

Figure 4-14: Actual vs MIRS-predicted TPC values for the calibration (A) and test (B) samples.

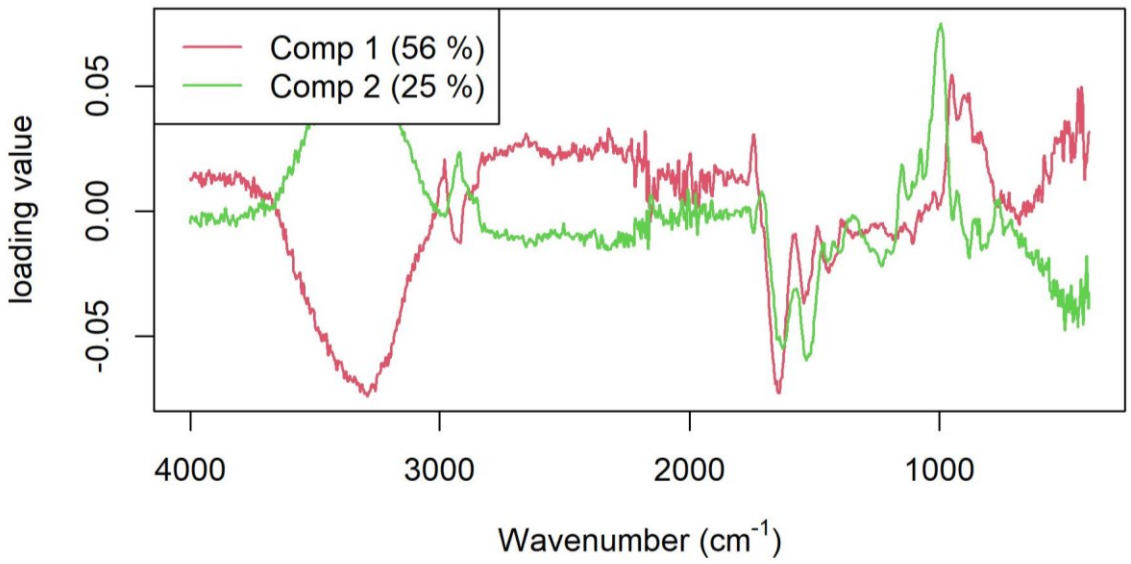

Figure 4-15: Loadings plot for the MIRS prediction of total phenolic content in wheat flour.

4.4.2.6 General discussion on the IR models

Overall, the NIRS results showed promise for the rapid, non-invasive prediction of most of the analytes investigated in the wheat samples, including proximate analytes (protein and

moisture) and the bioactive components (FRAP and TP). Although models for these proximate analytes have been previously reported by numerous authors, their success demonstrates that the NIRS instrumentation was functioning correctly and that there were no issues with sample degradation.

The best performing NIRS model was found for protein content (R^2_{test} = 0.991; RMSEP = 0.22%). This model showed ideal statistics – with minimal bias, a slope close to 1 and an intercept close to 0. Furthermore, the RPD of the test set (RPD_{test}) was 10.23, well above the limit of an excellent prediction model (RPD of 3) (Nicolaï et al., 2007). The model developed here was more accurate than NIRS models previously reported for Turkish wheat (R^2_{test} = 0.97, RMSEP = 0.38%; dependent test set) (Başlar & Ertugay, 2011) and triticale, a hybrid cross of wheat and rye (R^2_{test} = 0.96, RMSEP = 0.32%; test set from a different growing season) (Igne et al., 2007). It was not quite as accurate as the model reported by Ye et al. (2018) (R^2_{test} = 0.999, RMSEP = 0.05%; dependent test set), likely due to the smaller calibration sample size used in this study.

Although the model linearity was somewhat lower for the prediction of TPC (R^2_{test} = 0.83), the RMSEP for this analyte was exceptionally low at just 3.9 mg TE/100 g. This corresponds to a RPD_{test} of 2.49, indicating good prediction accuracy (Nicolaï et al., 2007). The only previous study found using NIRS for the prediction of TPC in wheat flour was performed by Tian et al. (2021c), who reported a higher R^2 value (R^2_{val} = 0.90), but lower precision (RMSEV = 7.1 mg GAE/100 g) on a dependent test set.

No previous studies were found using infrared spectroscopy for the prediction of antioxidant capacity in wheat flour, although several studies have attempted this in other crops such as maize (Redaelli et al., 2016), rice (Zhang et al., 2008) and gluten-free grains (Wiedemair et al., 2019). Although the FRAP prediction model developed here showed high linearity and a relatively low RMSEP, it should be cautioned that there were relatively few points with higher FRAP values included in the calibration and test sets. Consequently, this model had a slope that deviated from 1, indicating relative under-prediction of the highest-content sample. Nevertheless, the similar wavelength loadings between the TPC and FRAP models supports the observation that this model was looking at the “correct” wavelengths to measure compounds with antioxidant activity. While it is likely that analysis of a larger number of wheat samples containing a broader array of FRAP values would be required to create a robust NIRS model for the prediction of this analyte, the current model appeared suitable for rough estimation (not exact quantification) of FRAP values. Its RPD_{cal} was 2.95, although this fell to a RPD_{test} of 1.57 for the test set, indicating that it would be best suited to screening purposes only (Nicolaï et al., 2007).

Neither NIRS or MIRS were able to predict the total monomeric anthocyanin content, likely due to the very low concentration of this compound (mean content of just 2.5 mg cyd-3-glu/100 g in the calibration set).

In general, the MIRS models performed more poorly compared to the NIRS models in both the calibration and test sets (cf. Tables 4-2 and 4-3). The only exception to this trend was for the prediction of moisture content, which showed a higher R^2_{test} and similar RMSEP in the MIRS model.

In contrast to the results previously found in faba bean (Section 3.4.4.5), MIRS was able to predict the protein content of wheat with acceptable accuracy (R^2_{test} = 0.93; RMSEP = 0.62%). However, its accuracy was approximately three times lower compared to the use of NIRS for the same analyte (R^2_{test} = 0.991; RMSEP = 0.22%), reflecting the challenge in obtaining reproducible spectra when using ATR-MIR instrumentation.

Similarly, the linearity (R^2) and RMSEP were slightly poorer for the MIRS prediction of TPC and FRAP compared to their respective NIRS models. However, the slope of the FRAP MIRS model (1.20) was closer to the ideal value of 1. In contrast to the similar loadings seen in the NIRS models for TPC and FRAP prediction, the MIRS models for these analytes were noticeably different in their loadings. While both included the carbonyl region (~1700 cm^{-1}), the OH (~3400 cm^{-1}) and cellulose (~1000 cm^{-1}) regions showed greater influence in the TPC model, while the FRAP model included more influence from the CH stretch (~3000 cm^{-1}) and amide regions (1600-1400 cm^{-1}). This indicated that the FRAP model was relatively specific for the prediction of antioxidant compounds (potentially including non-phenolic antioxidant compounds such as tocopherols and or carotenoids), rather than just relying on the correlation between FRAP and TPC values to predict the former analyte.

4.5 Summary

The models developed here showed a high level of accuracy for the prediction of protein and considerable promise for the estimation of TPC and FRAP, particularly using the NIR spectra. However, the loadings plots for FRAP suggested that the MIRS model was more specific to antioxidant compounds, while the NIRS FRAP model loadings were very similar to the TPC model loadings.

As mentioned for faba bean (Section 3.4.4.6), NIR spectra are routinely collected from wheat samples for the prediction of protein content. Consequently, NIRS models for the estimation of TPC/FRAP – such as those demonstrated here – could be applied to these spectra to gain further information about the bioactive composition of the samples. This would provide an

additional layer of quality assurance information at no extra cost and minimal additional effort/setup.

It should be noted that the wheat samples investigated in this study were unable to be sourced from growing trials and consequently the models were validated using a dependent test set instead of an independent test set. Consequently, the reported accuracies may be over-optimistic. Future studies should incorporate a larger sample size and also confirm the model accuracy using an independent test set, ideally sourced from a different growing season to the samples comprising the calibration set.

Chapter 5. Mungbean

This chapter uses information from one previously published paper:

Johnson, J.B., Mani, J.S., Skylas, D., Walsh, K.B., Bhattarai, S.P. and Naiker, M., 2021. Phenolic profiles and nutritional quality of four new mungbean lines grown in northern Australia. *Legume Science*, *3*(2), p.e70. DOI: 10.1002/leg3.70

5.1 Introduction

The third crop chosen for investigation in this study was mungbean. Although it is not as widely grown in Australia as other grain crops such as wheat or faba bean, production has expanded rapidly in recent years. Furthermore, there is considerable opportunity to expand the production of this crop into the northern regions of Australia (Johnson et al., 2021d), as evidenced by recent field trials of this crop in northern Queensland and the Northern Territory (Surya Bhattarai, pers. comm.). In addition to this, there is considerable interest in the health-benefiting properties of mungbean, particularly in major international markets such as China.

Consequently, there is a need for context data on the typical levels of bioactive compounds in Australian-grown mungbean samples. Similarly, the development of rapid analytical techniques for the estimation of bioactive compound levels would be greatly beneficial for the quality assurance of mungbean samples, particularly those destined for the international export market.

For this study, a total of 100 mungbean samples (from 5 different varieties) were sourced from a controlled field trial conducted in Home Hill, north Queensland. Although samples could only be procured from one location and growing season, they did incorporate 20 within-field replicates of each variety. The first stage of work was the benchtop analysis of the phytochemical composition of all samples, in order to gain an understanding of the composition of different varieties and the level of within-field variation of the analytes.

Subsequently, the concentrations of selected phenolic compounds were measured in a subset of the mungbean samples (25 samples; comprising 5 field replicates of each variety) using HPLC-DAD. This aimed to provide more detailed insight into the specific phenolic profiles found in Australian-grown mungbean. This information could be useful for supporting specific health claims if the major phenolic compounds present in the samples have been clearly linked with certain health benefits via prior *in vitro* or *in vivo* studies. Basic proximate nutritional parameters (e.g., protein, ash) were also measured on this sample subset.

Finally, NIR and MIR spectra were collected from all 100 samples. As with faba bean and wheat, PLSR modelling was used to investigate the potential of IR spectroscopy for predicting the levels of bioactive compounds in mungbean.

5.2 Background

Mungbean (*Vigna radiata* L.) has traditionally been considered a relatively minor crop in the Australian pulse sector, with around 120,000 tonnes harvested annually (Chauhan & Williams, 2018). However, its popularity among producers has increased considerably in recent years, with the market value of Australian mungbean industry rapidly approaching $100 million p.a. (Australian Mungbean Association, 2017). Around 90% of the harvested crop is exported overseas, with the major importers including India, Vietnam, China and other south-east Asian countries. The Australian mungbean industry is considered to be a world leader in the development and adoption of industry-wide standards, in order to deliver the highest-quality produce possible (Australian Mungbean Association, 2017). This opens a huge potential to increase the production of mungbean in Australia, both for export purposes and to supply increasing demand in the domestic food market.

Mungbean is currently used in soups or dhal, milled to produce flour for various culinary uses, dried and roasted as snack products, or germinated and consumed as sprouts (Dahiya et al., 2015). The uses of mungbean continue to diversify, driven by opportunity for generating income at various points along the value chain. Hence an increasing number of mungbean-based value-added products are appearing on market shelves worldwide. Maintaining the high quality of both the mungbean seed and its value-added products are fundamental for the successful commercial future of the mungbean crop.

One of the major anticipated demands for this crop in the future is related to its high protein content, leading to the potential for meat replacement products (Alexeev et al., 2020). In addition to providing a valuable source of protein, mungbean also contain high levels of antioxidant and phenolic compounds (Ganesan & Xu, 2018; Hou et al., 2019). The major phenolic acids present in mungbean have been reported to be *trans*-ferulic acid, caffeic acid and coumaric acid, while the flavonoids present in the highest concentrations are isovitexin, vitexin and catechin (Hou et al., 2019; Meenu et al., 2016b). These and other bioactive compounds present in mungbean have been reported to possess a wide range of beneficial physiological activities, including scavenging free radicals (Tiwari et al., 2013), increasing insulin sensitivity (Yao et al., 2013), reducing plasma triglyceride levels (Tachibana et al., 2013), inhibiting the growth of pathogenic microbes (Hafidh et al., 2015), reducing hypertension (Hsu et al., 2011) and inflammation (Venkateshwarlu et al., 2016), and exerting anti-cancer effects (Hafidh et al., 2015). Increased research activities into the potential health

benefits of mungbean (e.g. Amaral et al., 2017; Hou et al., 2019; Hou et al., 2020) has led to increased consumer awareness and interest in this crop and its derived products as functional foods (Sehrawat et al., 2020), raising the profile and market value of this commodity.

Consequently, this study aimed to profile the bioactive composition of five mungbean varieties grown in Australia, using benchtop spectrophotometric methods, HPLC-DAD and infrared spectroscopy.

5.3 Materials and methods

5.3.1 Seed material

The mungbean seed material comprised a total of 100 samples, sourced from a variety-comparison field trial conducted near Home Hill, northern Queensland (19.8462°S, 147.2448°E). The samples comprised five different mungbean varieties: four newly developed varieties from AgriVentis Technology Ltd Australia (AVTMB 1-4), and one established commercial variety (Jade-AU). As samples were unable to be procured from more than one location or growing season, 20 within-field replicates of each variety were collected from different locations across the field at harvest maturity. The growing conditions, harvest details and physical seed quality – including yield, seed size and colour – are described in detail in Appendix B.

For each field replicate (n=20 per variety), approximately 20 g of seed material was ground to a fine flour (Breville Coffee & Spice Grinder; Botany, NSW) for subsequent extraction and analysis. The seed material was not dehulled before being processed.

Moisture content of the flour was determined according to AOAC Official Method 925.10. Briefly, flour samples (~3 g) were dried in a laboratory oven (Memmert 400; Buechenbach, Germany) at 105°C and the loss in mass quantified.

5.3.2 Analysis of FRAP, TPC, TMAC and CUPRAC

Polar phenolic compounds were extracted with 90% methanol, following the protocol described in Section 3.3.3, but using 1 g of flour and a final volume of 15 mL. Extractions and subsequent assays were performed in duplicate for each sample.

Ferric reducing antioxidant potential (FRAP), total polyphenolic content (TPC) and total monomeric anthocyanin content (TMAC) were determined from the extracts as described in Sections 3.3.4, 3.3.5 and 3.3.6. The results for FRAP were expressed in Trolox equivalents (TE), TPC results in gallic acid equivalents (GAE) and TMAC results in equivalents of cyanidin-3-glucoside (cyd-3-glu).

As an alternative measure of antioxidant capacity, the cupric reducing antioxidant potential (CUPRAC) assay was also performed on the samples. To perform the CUPRAC analysis, 1 mL of 10 mM aqueous copper (II) chloride, 1 mL of 1 M aqueous ammonium acetate, 1 mL of freshly prepared 7.5 mM neocuproine ethanol solution and 1 mL of Milli-Q water were combined with 100 µL of the sample extract. After vortexing for 30 seconds, the samples were incubated in a covered water bath at 50°C for 30 minutes. The resulting absorbances were measured at 450 nm using a UV-Vis spectrophotometer. As with the FRAP assay, the CUPRAC was derived as a function of the equivalent absorbance of Trolox standards in ethanol solution over the range of 50-600 mg L^{-1} (R^2 = 0.999; Figure 5-1). Results were expressed as milligrams of Trolox equivalents (TE) per 100 g of sample weight (mg TE/100 g).

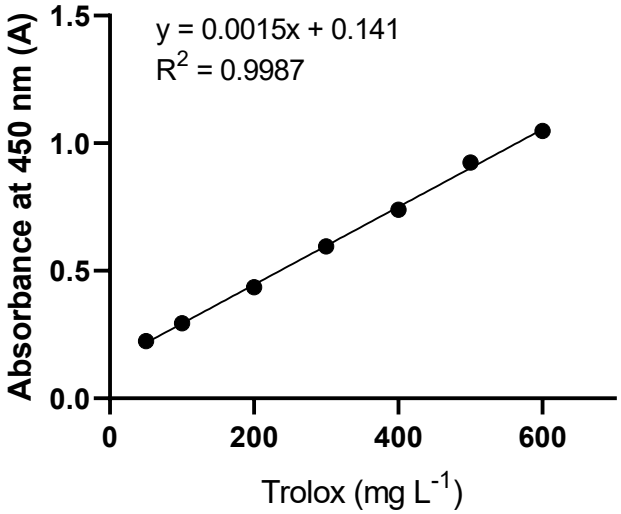

Figure 5-1: Calibration curve of Trolox used in the CUPRAC assay.

5.3.3 Analysis of basic proximate nutritional quality

As the results from the benchtop phytochemical analysis (Section 5.3.2) did not show a high level of variation between the within-field replicates, proximate nutritional analysis was only performed on a subset of the mungbean samples (5 within-field replicates for each variety, for a total of 25 samples).

Nitrogen and carbon content were determined on a LECO TruMac Series Carbon and Nitrogen Analyser (LECO, USA). Protein content was obtained by multiplying the nitrogen content by the standard conversion factor of 6.25 (Skylas et al., 2017). The ash content was determined as the proportion of mass remaining after combustion of the samples in a muffle furnace (ModuTemp; Midvale, WA) at 500°C for 8 h (Kalra, 1997).

5.3.4 Phenolic profiling by HPLC-DAD

To further characterise and investigate the individual phenolic compounds present in the different mungbean varieties, HPLC-DAD was used to conduct phenolic profiling on the subset of 25 samples described in Section 5.3.3.

Each methanol extract (10 mL) was concentrated (in duplicate) using a rotary evaporator with the water bath temperature kept at 27°C, before being reconstituted in 1 mL of methanol and syringe filtered (Livingstone 0.45 µm PTFE). The phenolic composition of these extracts was analysed using the Agilent 1100 HPLC system and methods described in Section 3.3.7.

Sixteen phenolic acids and flavonoids (Figure 5-2) were identified in the mungbean extracts based on comparison of their retention times and UV spectra to authentic standards (Sigma-Aldrich Australia). Additionally, isovitexin was tentatively identified in the extracts from its UV characteristics (λ_{max} and peak shapes) and relative retention time. This tentative identity was later confirmed using LC-MS (see Section 5.3.5), making a total of 17 identified compounds. All compounds were quantified using their respective standards (Table 5-1), except for isovitexin, for which no standard was available. This compound was quantified as equivalents of its structural isomer, vitexin (Figure 5-2).

Although hydroxybenzoic acids have been quantified at a detection wavelength of 280 nm by some authors (Xiang et al., 2019c), a wavelength of 210 nm was found to provide improved linearity and detection limits for most of these standards in this work (Table 5-1). This wavelength corresponds to the primary λ_{max} of most benzoic acids and has also been used by several previous researchers for their quantification (Chirinos et al., 2008; Pereira et al., 2010). The typical reproducibility (coefficient of variation) associated with the HPLC analysis, as measured by triplicate injections of quercetin, gallic acid and *p*-coumaric acid standards, was 2.5%, 0.73% and 0.25%, respectively.

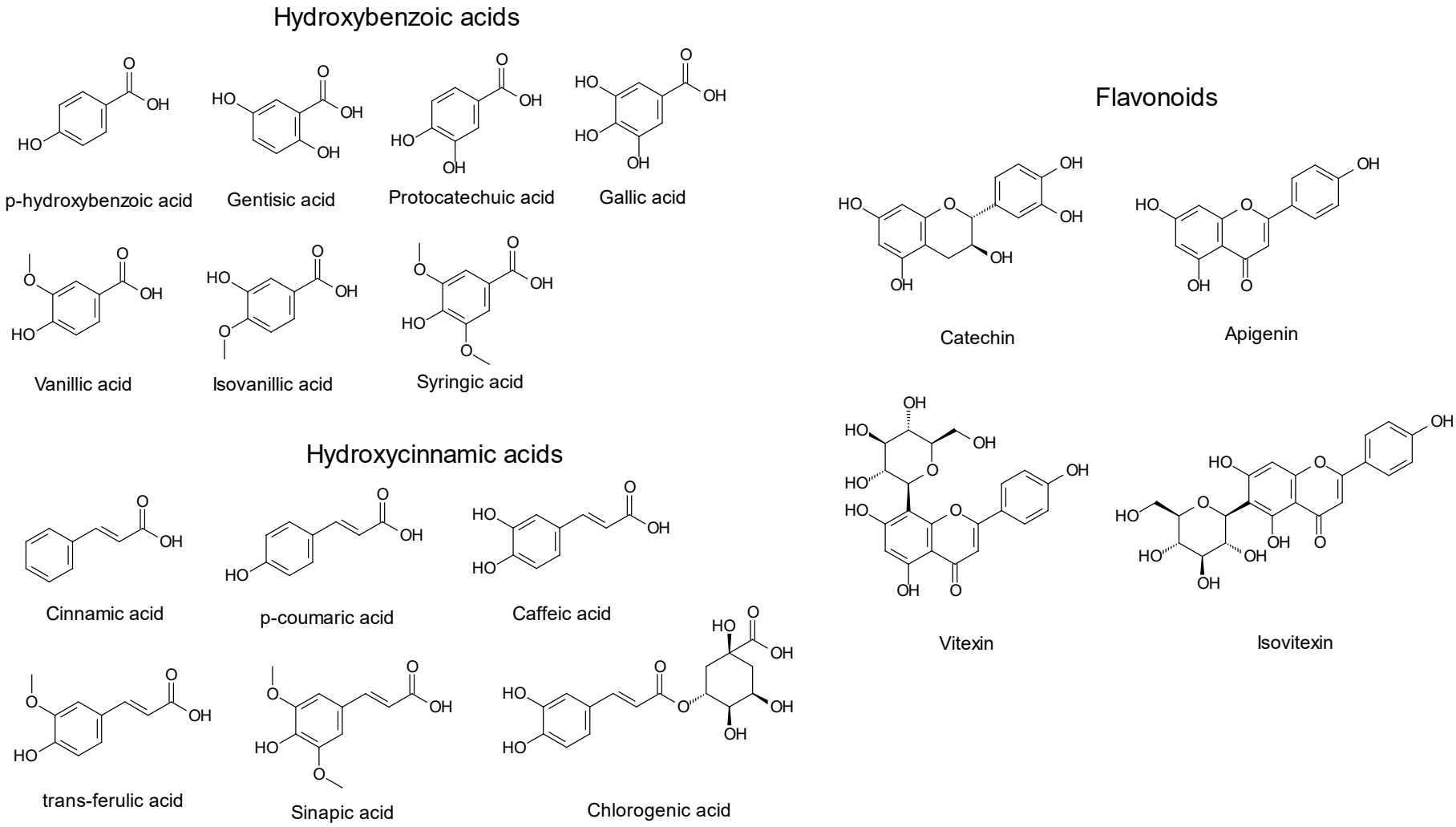

Figure 5-2: The structure of the phenolic acids and flavonoids analysed in the mungbean samples.

Table 5-1: Quality-of-analysis parameters associated with the phenolic acid and flavonoid standards. All standards were calibrated across the range of 1-100 mg L^{-1}.

No.	Compound	Retention time (min)	UV peaks (λ_{max}) (nm)	Quantification wavelength (nm)	Slope	Calibration R^2
Hydroxybenzoic acids						
1	Gallic acid	2.47	216, 272	210	40.1	0.9999
2	Protocatechuic acid	4.00	206, 218^{sh}, 261, 295	250	15.0	1
5	*p*-hydroxybenzoic acid	5.79	214^{sh}, 256	250	25.5	1
6	Gentisic acid	6.08	213, 238^{sh}, 330	210	39.7	1
7	Vanillic acid	6.36	206, 218, 262, 293	210	28.4	1
9	Isovanillic acid	6.58	206, 219, 260, 294	210	32.2	1
10	Syringic acid	6.68	218, 276	210	30.5	1
Hydroxycinnamic acids						
4	Chlorogenic acid	5.25	218, 246^{sh}, 303^{sh}, 327	320	13.2	0.9999
11	*p*-coumaric acid	8.25	211, 228, 300^{sh}, 310	320	32.3	1
8	Caffeic acid	6.43	218, 239, 298^{sh}, 325	320	25.0	0.9999
12	Sinapic acid	8.47	224^{sh}, 238, 325	320	23.6	1
13	*trans*-ferulic acid	8.54	218, 237, 298^{sh}, 324	320	25.5	0.999
17	Cinnamic acid	12.50	205, 218, 278	280	44.5	1
Flavonoids						
3	Catechin	4.86	204, 230^{sh}, 280	280	52.9	1
14	Vitexin	8.76	216, 269, 338	320	8.1	1
15	Isovitexin	9.34	215, 270, 338	320	(8.1)^	(1)^
16	Quercetin-3-glucoside	9.89	205, 257, 357	250	11.7	1

^quantified as equivalents of vitexin
sh = shoulder

5.3.5 LC-MS/MS identification of isovitexin

As mentioned in Section 5.3.4, isovitexin was tentatively identified in the mungbean extracts from the similarity of its UV spectrum to vitexin. Vitexin showed λ_{max} peaks at 2166, 269 and 338 nm, while the compound identified as isovitexin showed peaks at 215, 270 and 338 nm (Figure 5-3). Furthermore, the spectra showed a smoother dip between 280-310 nm compared to vitexin (compare Figures 5-3A and B), a characteristic feature of this compound (Jang et al., 2019). In addition, isovitexin showed a similar – albeit slightly longer – retention time than vitexin, consistent with the elution order found in previous literature analysing these two compounds using reversed-phase HPLC with a C_{18} column (Chen et al., 2013; Liu et al., 2009).

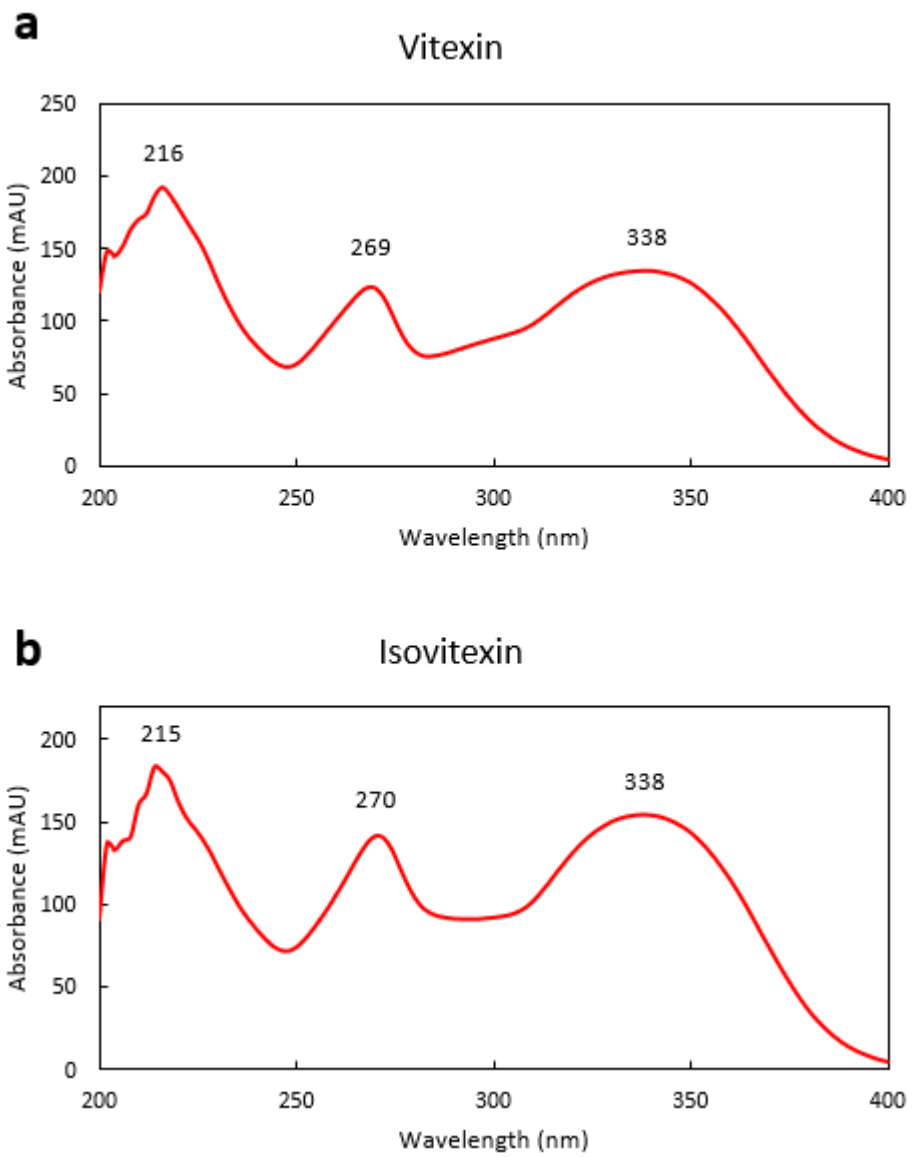

Figure 5-3: UV spectra of vitexin (A) and the compound tentatively identified as isovitexin (B), from the HPLC-DAD analysis.

To confirm the presence of isovitexin in the samples, liquid chromatography with tandem mass spectrometry (LC-MS/MS) analysis was performed on one of the concentrated methanol extracts used for HPLC-DAD analysis.

The LC-MS/MS analysis was performed on a Nexera X2 system coupled to a Shimadzu LCMS-8040 system, comprising a CBM-20A communications bus module, DGU-20A 5R degassing unit, LC-30AD pumps, SIL-30AC autosampler and CTO-20AC column oven. Chromatographic separation was achieved using a Raptor biphenyl column (100 mm × 2.1 mm, 2.7 µm), with an oven temperature of 40°C. Mobile phase A was water containing 5 mM ammonium formate and 0.1% formic acid, while mobile phase B comprised methanol containing 5 mM ammonium formate and 0.1% formic acid. The elution gradient began at 5% phase B, where it was held for 2 mins, before ramping to reach 85% phase B at 12 mins, where it was held for a further 2 mins. Between samples, the post-run equilibration time was 2 mins. The flow rate was 0.6 mL/min, with an injection volume of 5 µL.

The eluent from the LC was analysed using a Shimadzu LCMS-8040 model triple quadrupole mass spectrometer, equipped with an ESI source. The following ESI conditions were used: interface temperature of 350°C, DL temperature 250°C, heat block temperature 400°C. Nitrogen was used as both the nebulizing gas and drying gas, at flow rates of 3 L/min and 15 L/min, respectively. The interface voltage was set at 4.50 kV. Analysis was performed in positive and negative ionization modes, with Q3 scans collected in each mode. For the Q3 scans, all ions between *m/z* 100-700 were monitored.

In addition, product ion scans (PISs) were conducted using a precursor ion of *m/z* 431 in negative ionization mode and *m/z* 433 in positive mode, based on the expected values of isovitexin drawn from previous literature (Fu et al., 2008; Pereira et al., 2005). The collision energies for the PISs were set to 35 V and -35 V, respectively, with monitoring of all ions between *m/z* 50-450. The LC-MS/MS data were collected and analysed in the LabSolutions software (Shimadzu, Kyoto, Japan).

The identity of the vitexin peak was established by injecting an authentic standard of this compound (Sigma Aldrich, Australia).

5.3.6 NIR and MIR spectroscopy

NIR and MIR spectra were collected from all 100 samples of the mungbean flour using the instruments and settings described in Sections 3.3.8 and 3.3.9. Three replicate spectra were collected from each sample for both MIRS and NIRS, with the mean of the replicate scans used in subsequent data analysis.

5.3.7 Statistical analysis of phenolic data

Statistical tests were performed on the phytochemical and phenolic data using IBM SPSS (v26) and R Studio running R 4.0.5. Where applicable, results are presented as mean ± 1 standard deviation.

5.3.8 Spectral data analysis

Quantitative regression analysis of the infrared spectra was conducted in R Studio, as described in Section 3.3.11. Again, the model performance of the calibration set was evaluated through full LOO cross-validation. For model development, all of the AgriVentis samples (n=80) were used as a calibration set, while the Jade-AU samples (n=20) were used as an independent test set. A maximum of 10 components were considered for each PLSR model.

5.4 Results and discussion

5.4.1 Phytochemical and proximate composition

5.4.1.1 FRAP, CUPRAC, TPC and TMAC

The FRAP, CUPRAC, TPC and TMAC of the five mungbean varieties are provided in Table 5-2. The FRAP was exceptionally low in the mungbean samples (mean of 13.5-20.4 mg TE/100 g across the varieties) compared to the results found for faba bean. Furthermore, there was no significant difference in FRAP between the mungbean varieties (one-way ANOVA, $F_{4,95}$ = 1.79, P>0.05). However, the antioxidant capacity measured by the CUPRAC assay did vary significantly between varieties (one-way ANOVA, $F_{4,95}$ = 6.81, P<0.001). The varieties AVTMB 1-3 showed very similar CUPRAC levels, while that of AVTMB 4 and Jade-AU were significantly higher. The CUPRAC was at the higher end of the range of antioxidant capacities found by Johnson et al. (2020a) in five commercial mungbean samples from Australian and international growers.

The TPC showed a similar trend to CUPRAC, with significant variance between the mungbean varieties (one-way ANOVA, $F_{4,95}$ = 8.78, P<0.001). Again, AVTMB 4 showed a significantly higher TPC (104.7 ± 8.4 mg GAE/100 g) than all other varieties except Jade-AU (93.1 ± 14.5 mg GAE/100 g). The level of variation between samples was comparable to international studies on this crop (Kim et al., 2013; Shi et al., 2016; Zhang et al., 2013). Furthermore, the TPC was considerably higher than the range of TPCs (8-28 mg GAE/100 g) found in methanol extracts of 10 commercial Chinese mungbean samples (Zhang et al., 2013) and comparable to TPC values found by Yang et al. (2020) in nine mungbean cultivars from Sri Lanka (110-150 mg/100 g).

The mungbean varieties also displayed a significant difference in their TMAC (one-way ANOVA, $F_{4,95} = 28.2$, $P<0.001$), with AVTMB 4 showing very low anthocyanin levels, followed by Jade-AU. The three remaining AVTMB varieties all contained comparable TMAC values (12.4-15.5 mg cyd-3-glu/100 g).

Table 5-2: Phytochemical contents of the five mungbean varieties (n = 20 field replicates for each; results given as mean ± 1 SD). Varieties with the same superscript were not statistically different according to a post hoc Tukey test at α = 0.05.

Parameter	AVTMB 1	AVTMB 2	AVTMB 3	AVTMB 4	Jade-AU	P value
CUPRAC (mg TE/100 g)	500 ± 63^{b}	498 ± 78^{b}	504 ± 60^{b}	584 ± 83^{a}	567 ± 70^{a}	<0.001***
FRAP (mg TE/100 g)	20.4 ± 9.0	15.6 ± 9.7	13.5 ± 5.2	14.3 ± 5.0	16.8 ± 13.6	0.137 (NS)
TPC (mg GAE/100 g)	83.0 ± 14.3^{bc}	79.4 ± 11.4^{c}	90.5 ± 22.1^{bc}	104.7 ± 8.4^{a}	93.1 ± 14.5^{ab}	<0.001***
TMAC (mg cyd-3-glu/100 g)	14.4 ± 4.6^{a}	15.5 ± 3.6^{a}	12.4 ± 5.7^{a}	3.3 ± 4.5^{c}	7.3 ± 2.5^{b}	<0.001***

NS = not significant (P > 0.05), * P < 0.05, ** P < 0.01, *** P < 0.001

Table 5-3: Physical characteristics and nutritional parameters for the five mungbean varieties (n = 5 field replicates for each; results given as mean ± 1 SD). Varieties with the same superscript were not statistically different according to a post hoc Tukey test at α = 0.05.

Parameter	AVTMB 1	AVTMB 2	AVTMB 3	AVTMB 4	Jade-AU	P value
Moisture (%)	14.3 ± 2.1^{b,c}	10.1 ± 0.3^{a}	15.4 ± 1.8^{c}	16.0 ± 1.7^{c}	11.7 ± 0.9^{a,b}	<0.001***
Ash content (%)	4.26 ± 0.21^{a}	3.72 ± 0.13^{c}	3.90 ± 0.12^{b,c}	3.97 ± 0.12^{b,c}	4.15 ± 0.12^{a,b}	<0.001***
Protein (%)	28.01 ± 0.69^{a}	26.14 ± 0.05^{b}	27.34 ± 0.79^{a}	27.45 ± 0.65^{a}	27.32 ± 0.23^{a}	<0.001***
Carbon content (%)	47.33 ± 1.02^{a,b}	44.88 ± 0.14^{c}	47.56 ± 1.00^{b}	48.28 ± 0.88^{b}	46.00 ± 0.47^{a,c}	<0.001***
C:N ratio	10.56 ± 0.08^{a}	10.73 ± 0.06^{b}	10.88 ± 0.11^{b,c}	10.99 ± 0.09^{c}	10.52 ± 0.09^{a}	<0.001***

NS = not significant (P > 0.05), * P < 0.05, ** P < 0.01, *** P < 0.001

5.4.1.2 Proximate nutritional composition

As previously stated (Section 5.3.3), proximate nutritional analysis was only performed on 5 field replicates from each variety (n=25 samples in total), as the phytochemical results (Section 5.4.1.1) indicated relatively low within-field variation of their composition.

In terms of proximate composition (Table 5-3), there was a significant difference between varieties in moisture content at harvest (one-way ANOVA, $F_{4,20}$ = 14.452, P<0.001), ash content ($F_{4,20}$ = 10.799, P<0.001), protein content ($F_{4,20}$ = 7.358, P=0.001), carbon content ($F_{4,20}$ = 15.077, P<0.001) and the C:N ratio ($F_{4,20}$ = 27.743, P<0.001). The moisture content ranged between 10.1% in AVTMB 2 to 16.0% in AVTMB 4, with Jade-AU also possessing quite a low moisture content (11.7%). This is likely to be principally due to varietal differences rather than the genotypes being at different stages of maturity, given that all varieties displayed the onset of senescence directly prior to the application of defoliant and subsequent harvest of the crop.

The ash content was also the lowest in AVTMB 2 (3.72%) and highest in AVTMB 1 (4.26%), a similar range to that found in previous work on commercial varieties of Australian mungbean (Skylas et al., 2017). The protein content of AVTMB 2 (26.1%) was significantly lower than that of the remaining varieties, which ranged from 27.3-28.0%. Given that this variety (AVTMB 2) also had a significantly lower yield compared to the other varieties (Appendix B), its lower protein content may be indicative of increased susceptibility to environmental stressors in this particular variety. The C:N ratio was highest for AVTMB 3 and 4, and lowest for AVTMB 1 and Jade-AU.

5.4.1.3 Correlation analysis of phytochemical and proximate contents

As with faba bean and wheat, correlation analysis was performed on the phytochemical and proximate data for the 25 mungbean samples. The resultant correlogram is shown in Figure 5-4. There were a large number of significant correlations, most notably between TPC, CUPRAC and FRAP. Protein was positively correlated with the TPC and CUPRAC, but not FRAP. Positive correlations of varying strength were also found between protein, carbon, moisture and ash contents.

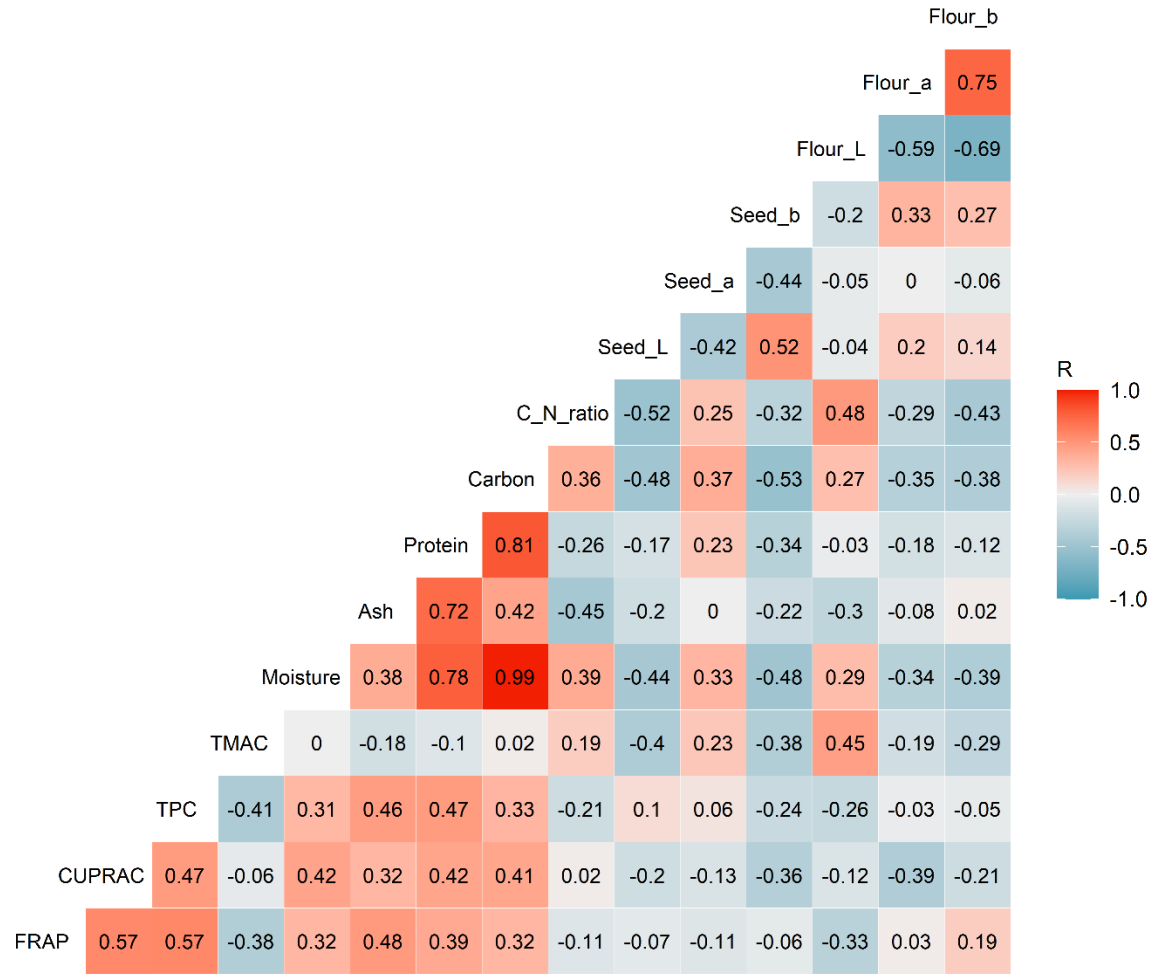

Figure 5-4: Correlogram showing the correlations between the phytochemical and proximate constituents of the mungbean samples (n = 25). Correlations with R values above 0.39 or below -0.39 were statistically significant at α = 0.05. Note that the "Seed" and "Flour" entries correspond to the CIE Lab colour of the mungbean whole seed and flour, respectively.

5.4.2 Identification of isovitexin using LC-MS/MS

Initial investigation of the *m/z* 431 ion chromatogram collected in negative ionization mode revealed the presence of two major peaks eluting at 6.66 and 6.97 minutes (labelled as peaks 1 and 2 in Figure 5-5). The precursor ions $[M - H]^-$ of both peaks were found to be *m/z* 430.6 (Table 5-4), agreeing closely with literature values of *m/z* 431 for both vitexin and isovitexin (Yan et al., 2013; Zhang et al., 2014). Furthermore, the positive precursor ions $[M + H]^+$ were both *m/z* 433.1, again agreeing with literature values (*m/z* 433.1) for the analysis of these compounds in positive ionization mode (Pereira et al., 2005).

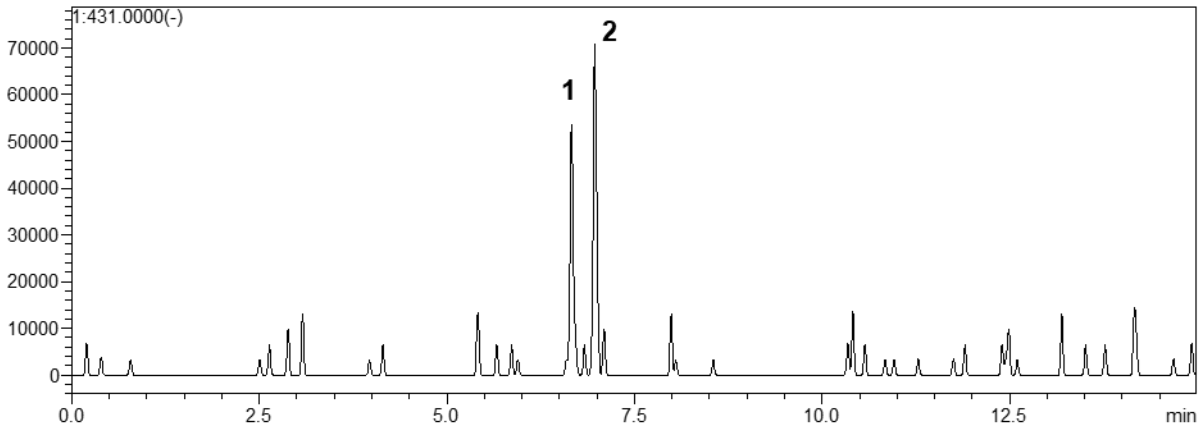

Figure 5-5: Ion chromatogram at 431 *m/z* showing the presence of vitexin (1) and isovitexin (2) in the mungbean extract. Analysis was conducted in negative ionization mode.

Table 5-4: Characteristics of vitexin and isovitexin analysed in the mungbean extract using LC-MS/MS.

Parameter	Peak 1 (vitexin)	Peak 2 (isovitexin)
Retention time (mins)	6.66	6.97
Peak area at *m/z* 431 (% of total)	159,664 (40.2%)	237,692 (59.8%)
Precursor ion $[M - H]^-$	430.6	430.6
Product ions (negative mode)	281.8, 310.0	281.9, 310.2
Precursor ion $[M + H]^+$	433.1	433.1
Product ions (positive mode)	282.8, 312.9	283.1, 312.8

The product ion scan in negative ionisation showed major ions at *m/z* 281.8 and 310.0 for peak 1, and at *m/z* 281.9 and 310.2 for peak 2 (Figure 5-6). Similarly, the major product ions observed in positive ionisation mode were *m/z* 282.8 and 312.9 (for peak 1) and *m/z* 283.1 and 312.8 (for peak 2), as shown in Figure 5-7. Again, this agreed with values from previous LC-MS studies of vitexin and isovitexin (Pereira et al., 2005; Yan et al., 2013). Other minor peaks (e.g., *m/z* 121.2 and 165.0 in positive mode) also agreed with literature observations (Pereira et al., 2005).

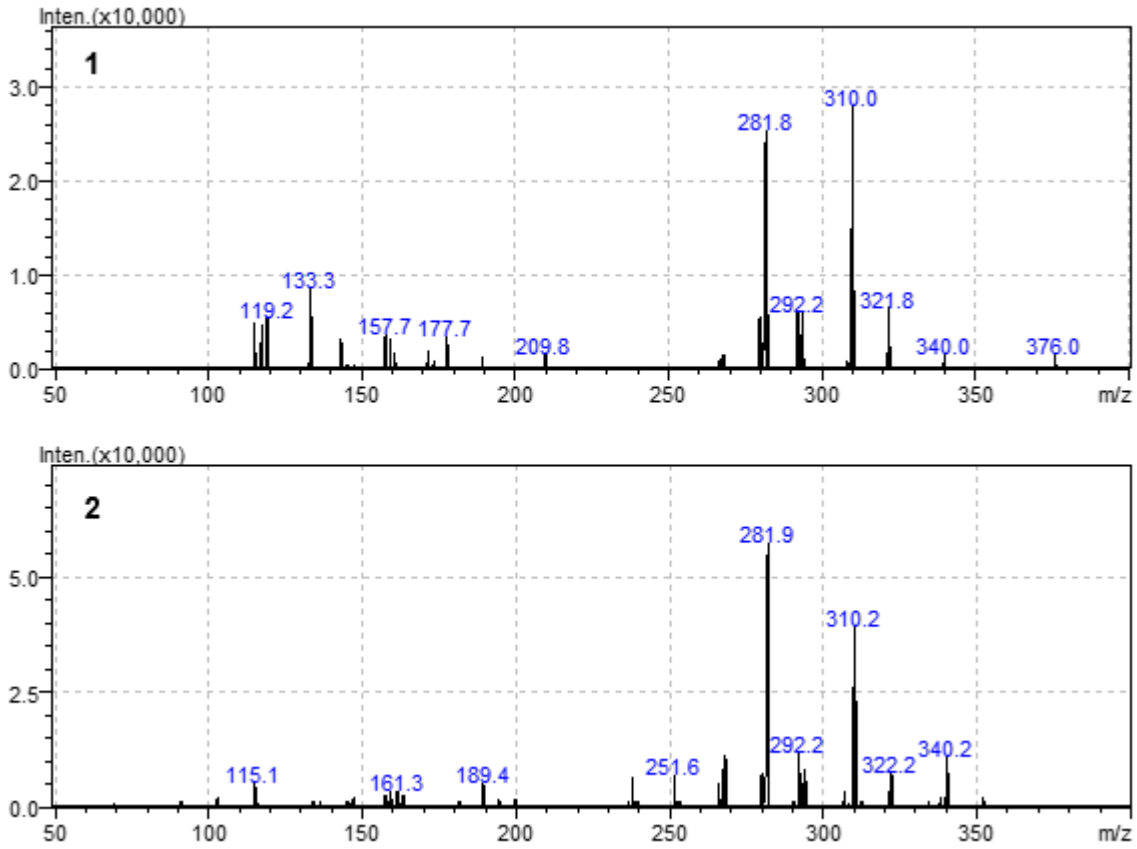

Figure 5-6: Product ion scans of peaks 1 and 2 from the mungbean extract, using negative ionization mode.

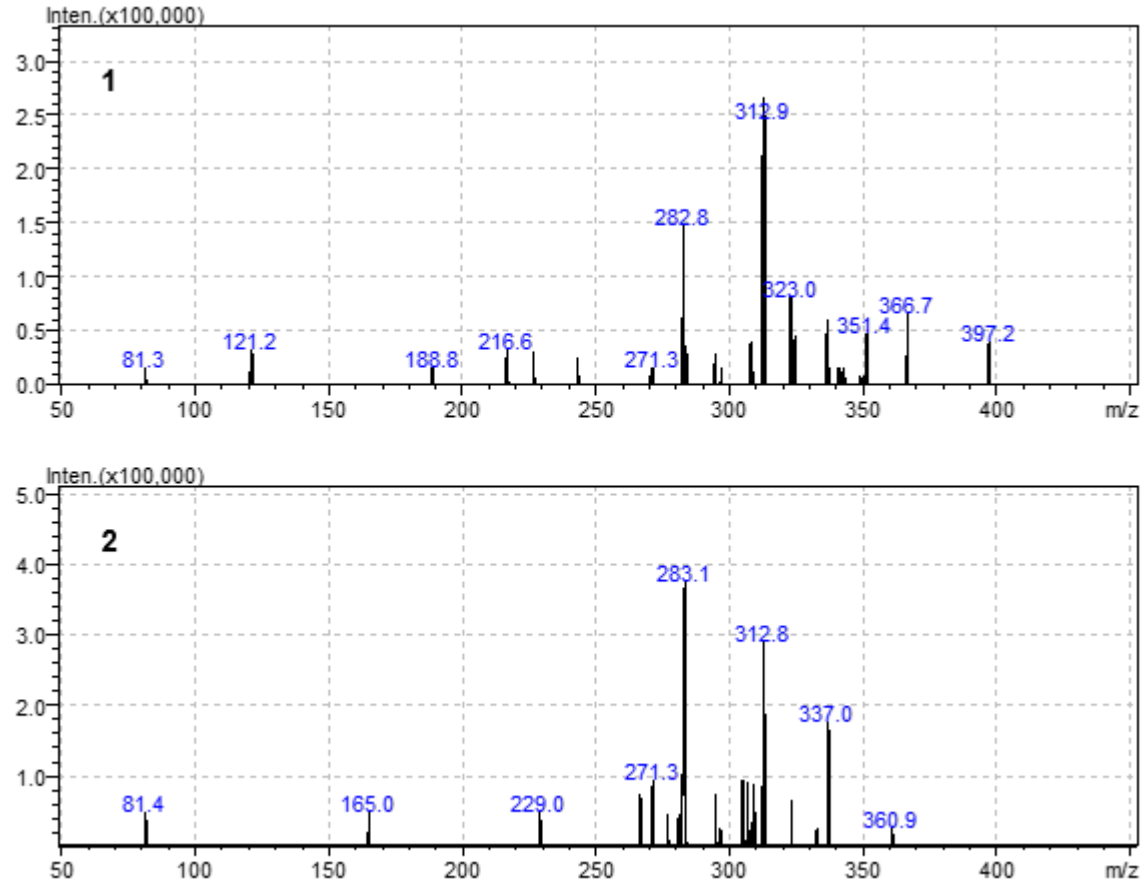

Figure 5-7: Product ion scans of peaks 1 and 2 from the mungbean extract, using positive ionization mode.

Finally, injection of an authentic standard of vitexin (5 ppm) revealed that it eluted at 6.67 mins; therefore, the identity of peak 1 was confirmed as vitexin. Given the high level of similarity observed in the mass spectra of peaks 1 and 2, as well as the similarity between the mass spectral features of peak 2 and literature values for isovitexin, the identity of peak 2 was assigned as isovitexin with a high level of confidence.

5.4.3 Phenolic profiles using HPLC-DAD

From the HPLC-DAD analysis of the mungbean sample extracts, 13 phenolic acids were identified (comprising seven hydroxybenzoic acids and six hydroxycinnamic acids), in addition to four flavonoids (Table 5-5). The elution order, relative abundance and UV spectral features of the compounds are illustrated in the HPLC chromatogram in Figure 5-8 and the UV isoplot in Figure 5-9.

The predominant hydroxybenzoic acid present was syringic acid, followed by vanillic and isovanillic acid in much lower concentrations. Protocatechuic acid, *p*-hydroxybenzoic acid, gentisic acid and gallic acids were also found in low concentrations. All hydroxycinnamic acids

were found in low concentrations, with chlorogenic, *p*-coumaric and caffeic acids found in the largest amounts, followed by cinnamic, sinapic and *trans*-ferulic acids. These trends broadly agreed with previous international studies profiling phenolic acids in mungbean (Hou et al., 2019; Meenu et al., 2016b; Yao et al., 2013). The predominant flavonoids were vitexin and isovitexin, followed closely by catechin. Quercetin-3-glucoside was found in low levels, while apigenin was below the limit of quantification in all varieties.

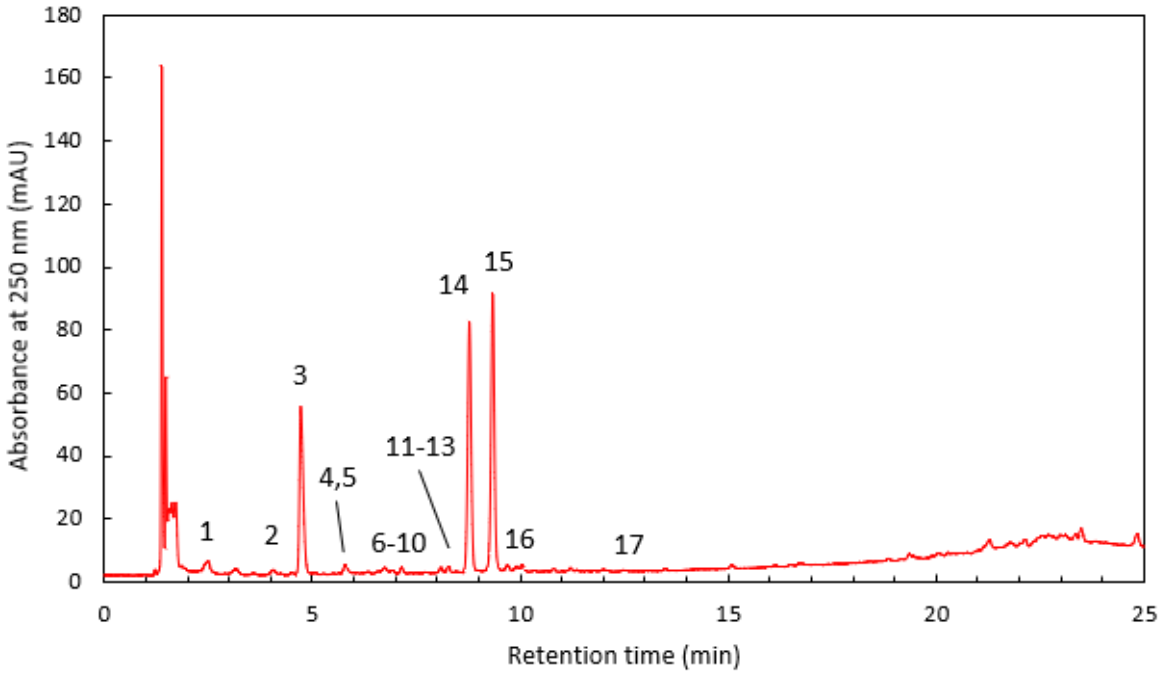

Figure 5-8: Annotated HPLC chromatogram from one of the mungbean samples (AVTMB 1). The numbered compounds correspond to those provided in Table 5-1.

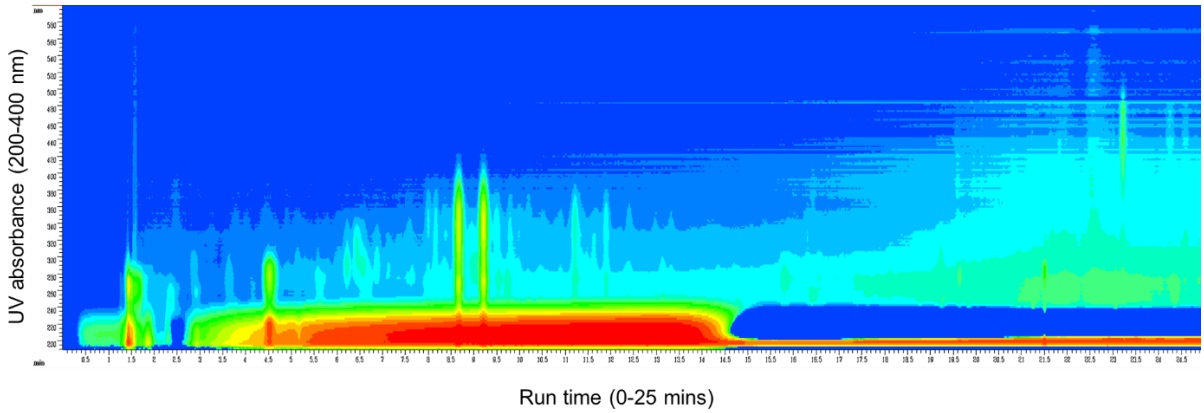

Figure 5-9: Isoplot showing the UV absorbance of various compounds eluting at various points throughout the HPLC run. The x-axis shows the run time (from 0-25 minutes) and the y-axis shows the UV wavelength (from 200-400 nm). The colour of each pixel shows the relative absorbance (blue = low; red = high).

Significant differences between the mungbean varieties were found for two hydroxybenzoic acids (*p*-hydroxybenzoic acid and vanillic acid), four hydroxycinnamic acids (caffeic acid, sinapic acid, *trans*-ferulic acid and cinnamic acid) and one flavonoid (vitexin). For *p*-hydroxybenzoic acid, and sinapic acid, the content found in the AVTMB 1 cultivar was significantly higher than that found in Jade-AU ($P<0.05$), with no significant differences found between the remaining cultivars and either AVTMB 1 or Jade-AU ($P>0.05$). For both caffeic acid and vitexin, the content found in AVTMB 4 was significantly higher than Jade-AU, while the vanillic acid content was higher in both AVTMB 3 and 4 compared to Jade-AU. Additionally, the cinnamic acid content of all varieties aside from AVTMB 1 were higher than Jade-AU. Although the one-way ANOVA for *trans*-ferulic acid indicated a significant difference between varieties ($P<0.05$), post-hoc Tukey testing was unable to determine which varieties were significantly different to one another. For this phenolic acid, AVTMB 2 had the highest concentration and AVTMB 1 the lowest. As all mungbean varieties were grown under the same environmental conditions, the differences in phenolic acid profiles between varieties can be attributed to their genetic differences, leading to differential expression of key enzymes in phenolic synthesis pathways, primarily the shikimic acid pathway (Santos-Sánchez et al., 2019). Although no genetic analysis was performed in this study, similar observations of upregulated gene expression have previously been made for varieties of other grain crops displaying increased content of specific phenolic acids (Laddomada et al., 2017; Ma et al., 2016).

Despite these differences in contents of individual phenolic acids and flavonoids, no significant differences were observed between varieties for the sum of hydroxybenzoic acids or flavonoids. AVTMB 4 did show a significantly higher content of total hydroxycinnamic acids compared to Jade-AU, but not to any other variety.

Table 5-5: Phenolic acid and flavonoid contents in the mungbean samples. Values given in µg/g (mean ± SD from 5 replicates for each variety).

Compound	AVTMB 1	AVTMB 2	AVTMB 3	AVTMB 4	Jade-AU	P value
Gallic acid	0.78 ± 0.04	0.70 ± 0.25	0.65 ± 0.15	0.66 ± 0.16	0.65 ± 0.27	*0.848*
Protocatechuic acid	1.18 ± 0.19	1.21 ± 0.20	1.30 ± 0.12	1.41± 0.41	1.42 ± 0.17	*0.338*
p-hydroxybenzoic acid	0.93 ± 0.29^{a}	0.83 ± 0.09^{a,b}	0.81 ± 0.06^{a,b}	0.81 ± 0.37^{a,b}	0.60 ± 0.19^{b}	*0.043**
Gentisic acid	0.75 ± 0.26	0.53 ± 0.09	0.68 ± 0.25	0.80 ± 0.25	0.45 ± 0.19	*0.058*
Vanillic acid	2.40 ± 0.12^{a,b}	2.37 ± 0.53^{a,b}	2.70 ± 0.17^{a}	3.00 ± 0.60^{a}	1.83 ± 0.42^{b}	*0.002***
Isovanillic acid	2.00 ± 0.47	2.18 ± 0.19	2.52 ± 0.45	2.78 ± 0.63	2.16 ± 0.65	*0.111*
Syringic acid	31.0 ± 2.7	30.8 ± 5.1	28.7 ± 5.1	38.9 ± 8.5	30.4 ± 5.0	*0.082*
Sum of hydroxybenzoic acids	38.3 ± 1.0	37.1 ± 4.4	39.2 ± 3.8	48.4 ± 9.9	38.0 ± 5.4	*0.103*
Chlorogenic acid	0.82 ± 0.26	0.71 ± 0.16	0.77 ± 0.07	0.97 ± 0.37	0.73 ± 0.19	*0.362*
p-coumaric acid	0.70 ± 0.12	0.78 ± 0.21	0.82 ± 0.14	0.81 ± 0.17	0.57 ± 0.12	*0.051*
Caffeic acid	0.84 ± 0.18^{a,b}	0.74 ± 0.10^{a,b}	0.85 ± 0.09^{a,b}	0.92 ± 0.22^{a}	0.61 ± 0.09^{b}	*0.015**
Sinapic acid	0.17 ± 0.04^{a}	0.15 ± 0.03^{a,b}	0.13 ± 0.04^{a,b}	0.17 ± 0.05^{a,b}	0.11 ± 0.01^{b}	*0.036**
trans-ferulic acid	0.11 ± 0.04^{a}	0.20 ± 0.08^{a}	0.19 ± 0.02^{a}	0.22 ± 0.08^{a}	0.14 ± 0.06^{a}	*0.046**
Cinnamic acid	0.17 ± 0.02^{a,b}	0.18 ± 0.03^{a}	0.19 ± 0.03^{a}	0.18 ± 0.03^{a}	0.13 ± 0.04^{b}	*0.017**
Sum of hydroxycinnamic acids	2.64 ± 0.30^{a,b}	2.65 ± 0.40^{a,b}	2.90 ± 0.23^{a}	3.26 ± 0.82^{a}	2.29 ± 0.45^{b}	*0.049**
Catechin	93.8 ± 7.0	93.4 ± 3.5	94.2 ± 2.9	105.2 ± 19.6	98.4 ± 18.7	*0.607*

Vitexin	117.9 ± 12.5[a,b]	132.0 ± 21.1[a,b]	140.4 ± 14.3[a,b]	148.9 ± 21.7[a]	114.9 ± 23.5[b]	*0.029**
Isovitexin	143.8 ± 28.9	152.6 ± 26.9	163.0 ± 15.4	173.5 ± 27.6	132.0 ± 26.9	*0.078*
Quercetin-3-glucoside	0.99 ± 0.27	0.88 ± 0.19	0.84 ± 0.17	0.89 ± 0.17	0.65 ± 0.12	*0.071*
Apigenin	<LOQ	<LOQ	<LOQ	<LOQ	<LOQ	*n/a*
Sum of flavonoids	349 ± 36	370 ± 38	402 ± 37	429 ± 68	346 ± 67	*0.103*

LOQ = limit of quantification
* P<0.05, ** P<0.01, *** P<0.001

The vitexin and isovitexin concentrations found here (115-149 µg/g and 132-174 µg/g, respectively) were comparable to or slightly higher than that found by Zhang et al. (2013) in methanol extracts from ten varieties of commercial mungbean in China (44-144 µg/g and 37-112 µg/g, respectively). The vitexin and isovitexin concentrations were highly correlated (R^2 = 0.969), with the isovitexin content consistently around 20% higher than vitexin content (Figure 5-10). This strong correlation is expected, as vitexin and isovitexin are structural isomers of apigenin glucoside. Vitexin is apigenin-8-C-glucoside, while isovitexin is apigenin-6-C-glucoside (Figure 5-2); hence both would be expected to form via a similar metabolic pathway (Abdullah et al., 2017). However, despite the structural similarity between these compounds, they differ in their biological activities, with vitexin displaying greater spasmolytic activity and inhibition of α-glucosidase compared to isovitexin (Choo et al., 2012; Ragone et al., 2007). However, isovitexin shows a higher antioxidant activity in many bioassay methods (He et al., 2016).

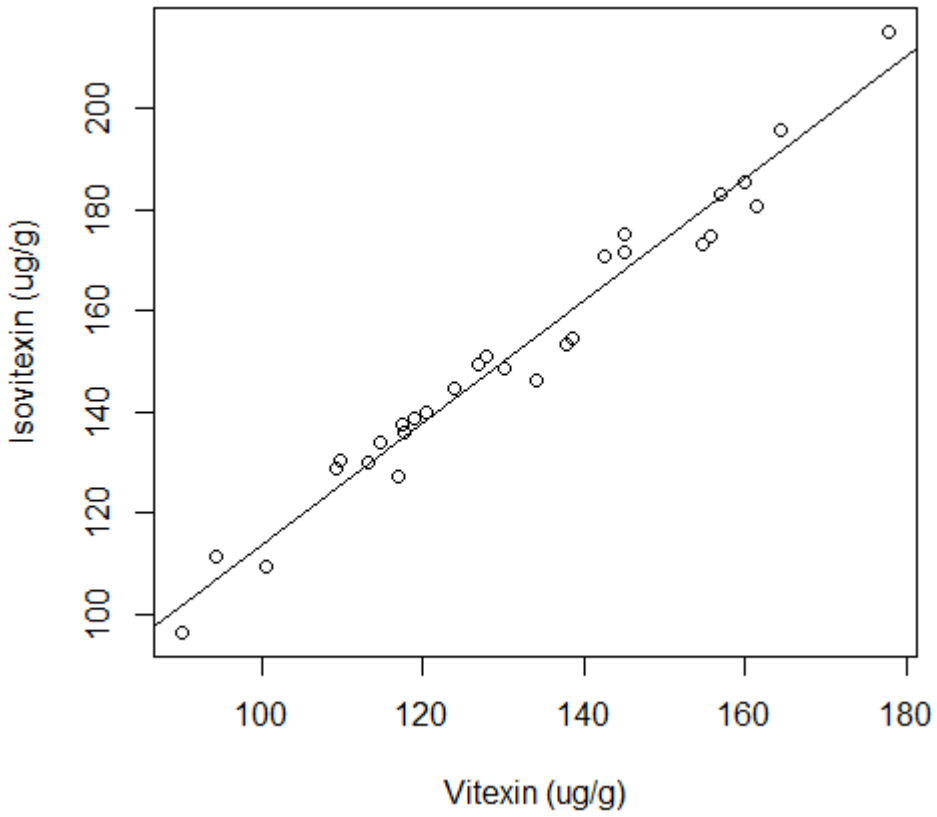

Figure 5-10: Relationship between vitexin and isovitexin content in the mungbean samples.

Notably, the free apigenin content was below the limit of quantification (<0.1 µg/g; Table 5-1), indicating that virtually all apigenin present in mungbean is in a glycosylated form, either as vitexin or isovitexin. Previous work has shown that vitexin and isovitexin contents decrease to

below detection limits when mungbean is sprouted (Yao et al., 2011), suggesting that these compounds may be hydrolysed into apigenin during germination. Furthermore, this agrees with separately conducted research that found apigenin concentrations increase during sprouting (Pająk et al., 2014). However, magnitude of reported increase in apigenin content (increasing from 0 to 1.9 µg/g) does not correspond directly to the decrease in combined vitexin/isovitexin content (decreasing from ~300 to 0 µg/g), suggesting that some intermediate form of these compounds or another side reaction may be at play during the germination process.

Correlation analysis revealed moderate to strong correlations between several phenolic acids and flavonoids (Figure 5-11), particularly between caffeic and coumaric acids, vanillic and coumaric acids, and vanillic and syringic acids. In contrast, the levels of gallic, protocatechuic or gentisic acid were not strongly correlated with any other compounds. Such correlations between individual phenolic and flavonoid constituents are likely related to their similar synthesis pathways (Santos-Sánchez et al., 2019), or upregulation of regulatory genes increasing expression of multiple unrelated synthesis pathways.

The TPC and antioxidant capacity were positively correlated with one another ($r = 0.47$-0.57; $P<0.05$), as observed in previous research (Johnson et al., 2020b; Xiang et al., 2019b), although the antioxidant capacity was not correlated with the total monomeric anthocyanin content (Table 5-6). However, the TPC was negatively correlated with the TMAC ($r = -0.41$, $P<0.05$), indicating that anthocyanins were not major contributors toward the overall phenolic content of this matrix. The sinapic acid content was also positively correlated with the FRAP and TPC ($r = 0.40$ and $r = 0.53$, respectively; $P<0.05$), although no other individual phenolic acids showed any significant correlation (Table 5-6).

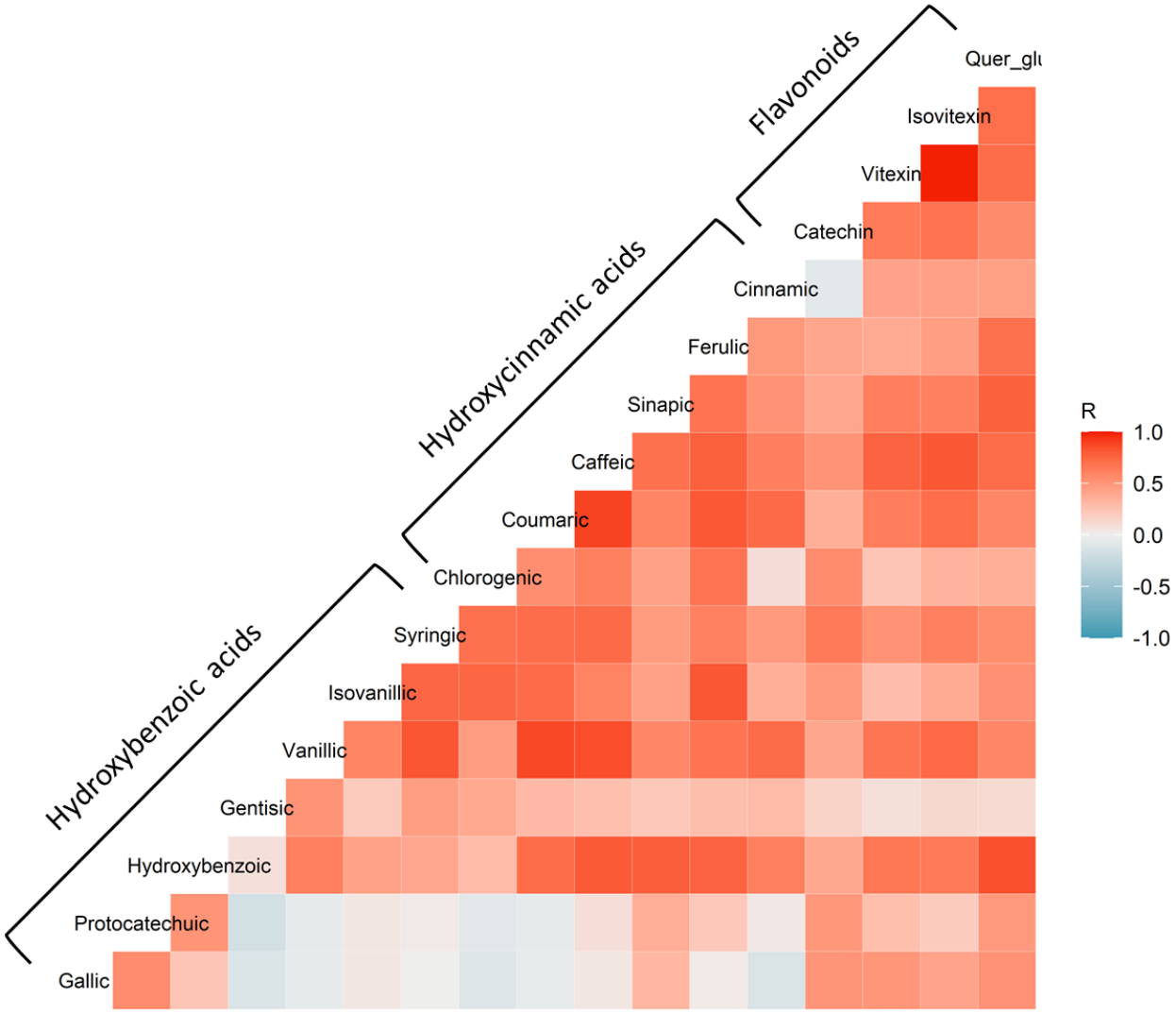

Figure 5-11: Correlations plot showing the relationships between the levels of different phenolic acid and flavonoid constituents.

Table 5-6: Correlation coefficients between the CUPRAC, FRAP, TPC, TMAC and individual phenolic acids in the mungbean samples (n=25).

Parameter	CUPRAC	FRAP	TPC	TMAC
CUPRAC	1	-	-	-
FRAP	0.57**	1	-	-
TPC	0.47*	0.57**	1	
TMAC	-0.06	-0.38	-0.41*	1
Gallic acid	-0.12	0.12	0.23	-0.19
Protocatechuic acid	-0.08	-0.49	-0.29	0.10
p-hydroxybenzoic acid	-0.03	0.14	0.23	-0.23
Gentisic acid	-0.14	-0.01	-0.25	-0.03
Vanillic acid	-0.12	-0.04	0.23	-0.11
Isovanillic acid	-0.18	-0.20	0.02	0.01
Syringic acid	-0.14	-0.07	0.12	-0.14
Chlorogenic acid	-0.05	-0.19	-0.20	0.01
Coumaric acid	-0.23	-0.09	0.01	0.16
Caffeic acid	0.14	-0.06	-0.04	0.07
Sinapic acid	0.21	0.40*	0.53**	-0.25
Ferulic acid	-0.09	-0.15	0.04	0.10
Cinnamic acid	-0.22	-0.06	-0.10	-0.17
Catechin	-0.15	-0.05	0.31	-0.08
Vitexin	-0.20	-0.07	0.29	0.00
Isovitexin	-0.23	-0.12	0.24	0.04
Quercetin-3-glucoside	0.14	0.33	0.22	-0.27

* P<0.05, ** P<0.01, *** P<0.001

5.4.4 Prediction of bioactive compounds using IR spectroscopy

The final stage of investigation into the mungbean samples was attempting the prediction of bioactive compounds using infrared spectroscopy. As was the case with the previous crops, PLSR was used to create prediction models for each analyte, using a variety of pre-processing techniques.

5.4.4.1 Descriptive statistics

The descriptive statistics for the calibration and test sets are provided in Table 5-7. The range of most analytes were comparable between the calibration and test sets, supporting this choice of the Jade-AU samples for use as an independent test set.

Table 5-7: Descriptive statistics for the parameters measured in this study, in the calibration and test sets.

Parameter	Calibration set (n=80)		Test set (n=20)	
	Range	Mean ± SD	Range	Mean ± SD
FRAP (mg TE/100 g)	10.0-55.3	16.0 ± 7.9	9.9-65.8	16.8 ± 13.6
CUPRAC (mg TE/100 g)	405.7-736.5	521.3 ± 79.1	499.2-776.3	567.2 ± 70.4
TPC (mg GAE/100 g)	65.7-130.0	89.4 ± 17.6	74.8-136.9	93.1 ± 14.5
TMAC (mg cyd-3-glu/100 g)	0.0-21.2	11.4 ± 6.6	2.7-11.6	7.3 ± 2.5

5.4.4.2 NIR spectra

The raw and SNV-processed NIR spectra are shown in Figure 5-12. These appeared very similar to those obtained for faba bean and chickpea, with major peaks at 1468 and 1935 nm (OH 2nd and 1st overtones), and other large peaks at 1198 nm (CH 2nd overtone), 1751 nm (CH 1st overtone), 2112 nm (amide combination band) and 2316 nm (CH deformation combination band).

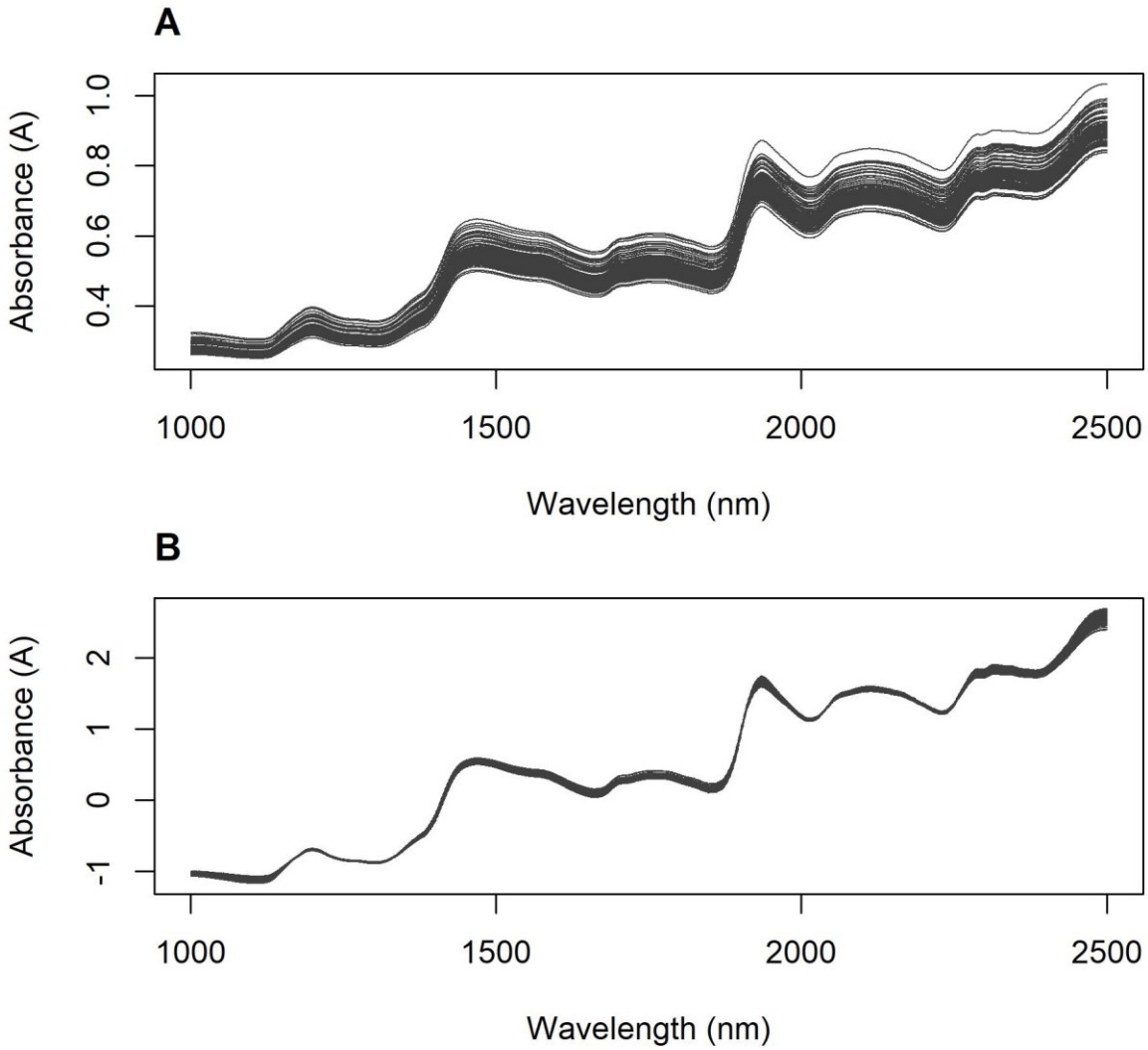

Figure 5-12:The raw absorbance NIR spectra (A) and SNV-processed spectra (B) of the mungbean samples.

5.4.4.3 NIR models

The optimum PLSR models found for each analyte are presented in Table 5-8. In contrast to the results found in faba bean (Section 3.4.4.3) and wheat (Section 4.4.2.3), none of the NIRS models showed good performance in the calibration and independent test sets. While the model for FRAP performed acceptably for the calibration samples (R^2_{cv} = 0.57, RMSECV = 5.1 mg TE/100 g), it was unable to predict this analyte in the test set (R^2_{test} = 0.20, RMSEP = 13.0 mg TE/100 g).

No other models for the investigated analytes (CUPRAC, TPC or TMAC) showed any predictive power, even for the calibration set (Table 5-8). This contrasted with the results of Meenu et al. (2016a), who used NIRS to predict the concentrations of 10 different phenolic acids and flavonoids in Chinese mungbean samples, with a high reported accuracy (R^2_{val} of up to 0.99 and RMSEs of <1.8%). However, it should be noted that these authors only used a

relatively small calibration set (42 samples) and did not report how many components were required for the PLSR models. Additionally, they used a dependent test set (18 samples) for model validation, rather than an independent test set. Consequently, it is likely that the results reported by these authors may be over-optimistic compared to the performance of the model when applied to an independent test set.

Nevertheless, the lack of predictive power observed even in the calibration set indicates that no spectral signal from the target analytes could be detected in the NIR region. This lack of predictive power may be due to several factors. Firstly, the lower concentrations of most analytes (e.g., compared to faba bean) would make it more difficult to detect using NIR wavelengths. However, this fact alone cannot account for the observation that accurate models could be created for the prediction of these analytes in wheat flour.

It is possible that matrix complexity could also play a role in the poor model performance. For example, other compounds present in varying concentrations may absorb in the same NIR regions as the compounds of interest (i.e., phenolic and antioxidant compounds), thus obscuring their signal.

Conversely, it is possible that variations in non-target constituents of the chickpea samples could be influencing model performance in another way. As discussed in Chapter 2, many NIRS applications targeting microconstituents actually rely on secondary or surrogate correlations with other macroconstituents which are more readily measured using NIRS (Walsh et al., 2020). If these correlations hold true for all samples, then the use of such secondary correlations is quite acceptable. However, these correlations may change between different sample populations or harvest years (Velasco et al., 1998), which might explain the poor performance in the independent test set.

Finally, and perhaps most likely, it is possible that the calibration set did not contain samples from a sufficiently wide range of origins and analyte concentrations. Specifically targeting samples to obtain a wider range of phenolic contents or antioxidant capacities may aid in this – perhaps through the inclusion of samples of different maturation stages or exposed to different environmental stressors.

Table 5-8: Optimum PLSR models found for the prediction of the specified analytes using NIR spectroscopy.

Parameter	Spectral pre-processing	Factors	R^2_{cv}	RMSECV	RPD	R^2_{test}	RMSEP	Bias	Slope	Intercept
FRAP (mg TE/100 g)	None	7	0.57	5.1	1.54	0.20	13.0	3.67	2.96	-22.1
CUPRAC (mg TE/100 g)	1d21	5	0.16	72.1	1.10	0.06	115.8	94.2	0.60	279.5
TPC (mg GAE/100 g)	SNV + 1d21	10	0.21	15.5	1.13	0.07	25.6	21.2	0.44	61.6
TMAC (mg cyd-3-glu/100 g)	SNV	7	0.29	5.6	1.19	0.00	11.9	-11.1	-0.05	8.3

Table 5-9: Optimum PLSR models found for the prediction of the specified analytes using MIR spectroscopy.

Parameter	Spectral pre-processing	Factors	R^2_{cv}	RMSECV	RPD	R^2_{test}	RMSEP	Bias	Slope	Intercept
FRAP (mg TE/100 g)	None	2	0.00	7.8	1.01	0.30	12.3	-0.21	3.70	-46.2
CUPRAC (mg TE/100 g)	None	1	-0.01	79.1	1.00	0.01	83.4	45.1	-0.97	1072
TPC (mg GAE/100 g)	None	0	-0.03	17.7	0.99	0.06	14.3	3.89	1.29	-21.5
TMAC (mg cyd-3-glu/100 g)	None	0	-0.03	6.7	0.99	0.00	4.8	-4.1	-0.13	8.8

5.4.4.4 MIR spectra

The MIR spectra (shown in Figure 5-13) appeared similar to the spectra from faba bean (Section 3.4.4.4), with major peaks centred around 3300 cm^{-1} (attributable to OH stretch from moisture), 3000-2850 cm^{-1} (CH_2 and CH_3 stretch), 1640 cm^{-1} (C=O stretch of amides or other carbonyl-containing compounds), 1550-1200 cm^{-1} (various amide and carbohydrate features) and 1000 cm^{-1} (pyranose rings in starch/cellulose) (Abbas et al., 2017; Dufour, 2009; Karoui et al., 2010; Mecozzi & Sturchio, 2017).

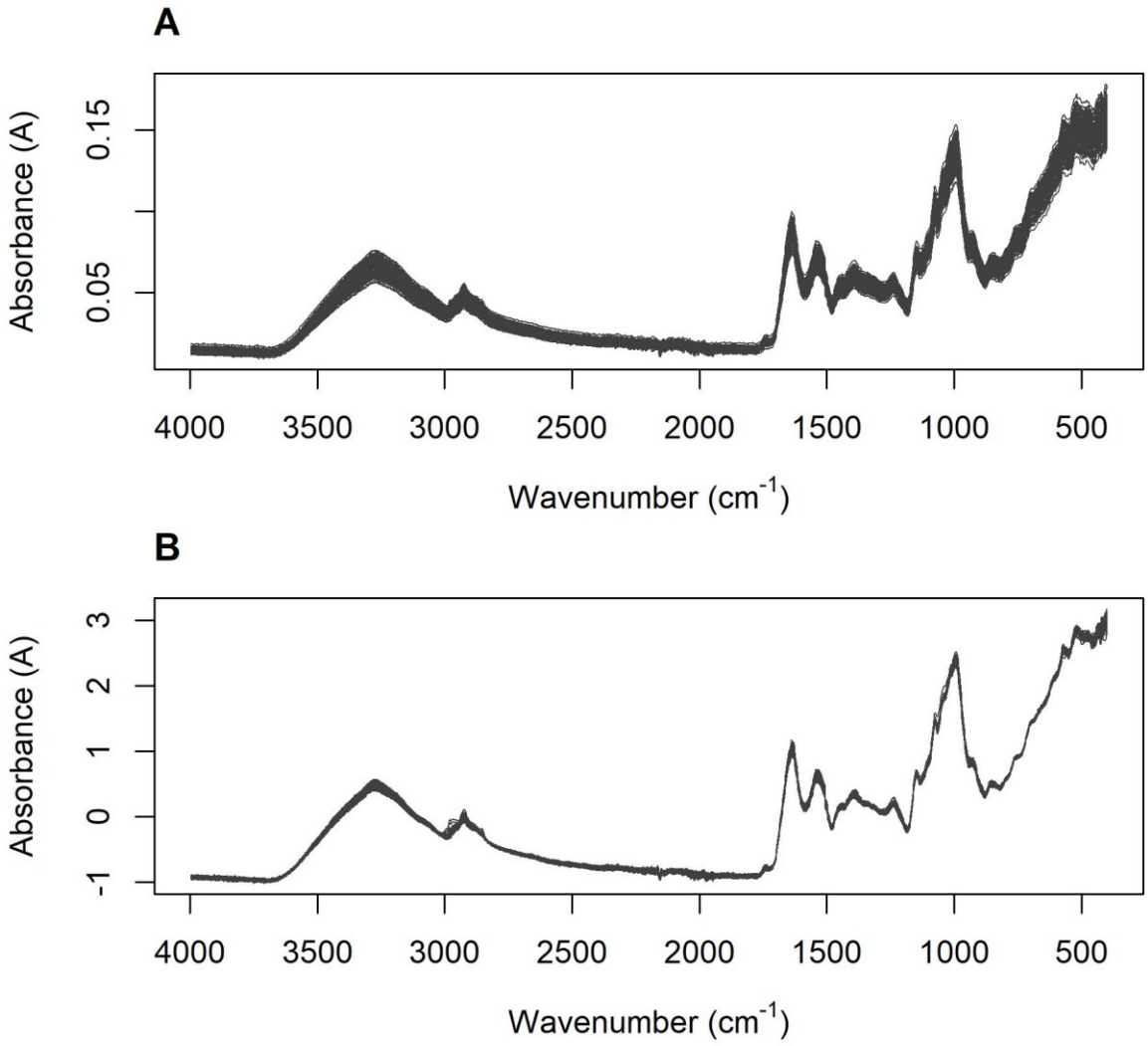

Figure 5-13:The raw MIR spectra (A) and SNV-processed spectra (B) of the mungbean flour samples.

5.4.4.5 MIR models

As with the NIRS models, none of the MIRS models were able to predict any of the bioactive analytes in the mungbean flour (Table 5-9). For this reason, no calibration plots or loadings are shown for these models.

The factors impacting on the NIRS models (see Section 5.4.4.3) are also likely to negatively influence the MIRS models. In particular, the relatively low range of analyte concentrations make the development of accurate models much more difficult. In addition, MIRS has the added challenge of trying to maintain a consistent level of pressure between the sample matrix and the ATR platform for each of the samples analysed. Although most of the visible impact of this variation is removed through use of SNV or derivative processing, it is possible that small variations between samples remain and consequently impact negatively on the model.

5.5 Summary

The five mungbean varieties investigated here showed significant variation in their CUPRAC, TPC and TMAC, although not in their FRAP. Additionally, the varietal differences in phenolic composition were supported by HPLC-DAD profiling of 17 individual phenolic acids and flavonoids in a subset of the mungbean samples. The greatest differences were seen for *p*-hydroxybenzoic, vanillic, caffeic, sinapic, ferulic and cinnamic acids, and the flavonoid glycoside vitexin.

In contrast to the previous results for faba bean and wheat, neither NIRS nor MIRS were able to predict any of the bioactive analytes (FRAP, CUPRAC, TPC or TMAC) measured in the mungbean samples.

Chapter 6. Chickpea

The introduction and discussion of this chapter uses information from a previously published paper:

Johnson, J.B., Walsh, K.B., Bhattarai, S.P. and Naiker, M., 2021. Partitioning of nutritional and bioactive compounds between the kernel, hull and husk of five new chickpea genotypes grown in Australia. *Future Foods*, *4*, p.100065. DOI: 10.1016/j.fufo.2021.100065

6.1 Introduction

The fourth and final crop selected for investigation in this study was chickpea. Although a relatively recent addition to the Australian cropping rotation, its popularity has grown rapidly in recent decades to position Australia as the largest global exporter of chickpea. As with mungbean, there has been increasing interest in the health benefits of this crop, with publications on this topic growing rapidly over the last decade. With Australian chickpea already internationally recognised for their high quality, there is potential to build on this reputation to market our produce with an emphasis on its health-benefiting constituents.

As with the previous crops investigated, the aim of this chapter was firstly to provide context data on the typical levels of phenolics and antioxidant compounds found in Australian-grown chickpea and secondly, to investigate the use of IR spectroscopy for the rapid, non-destructive prediction of these analytes. Only desi type chickpeas were investigated here, as kabuli type chickpeas comprise only around 5-10% of the total Australian production.

Due to time constraints in sample procurement, the chickpea samples were unable to be sourced from controlled trials growing the same varieties across the same locations in different years. However, they were sourced from a seedbank/collection containing archived seed from a large number of varieties, grown across several different locations and seasons. Consequently, they were considered to incorporate most of the variation that would typically be observed in the Australian desi chickpea crop. Furthermore, the inclusion of samples from a wide variety of growing conditions was anticipated to aid in the development of robust prediction models using infrared spectroscopy.

As with previous chapters, the basic phytochemical composition of the chickpea samples was characterised using benchtop spectrophotometric techniques. This was followed by more detailed examination of the phenolic composition using HPLC-DAD. Finally, the use of infrared spectroscopy (NIRS and MIRS) for the prediction of total phenolics, antioxidant capacity, and total anthocyanin content was investigated.

6.2 Background

Chickpea (*Cicer arietinum* L.) is one of the oldest known pulse crops and is widely grown across the world (Abbo et al., 2003b; Yadav & Chen, 2007). Globally, it is ranked as the second-most produced legume crop, with 14.8 million tonnes harvested in 2020 (FAO, 2022). Although this crop was not grown commercially in Australia until the 1970s (Pendergast et al., 2019; Siddique et al., 2000), Australia has grown to become the 8th largest producer and the largest exporter of chickpea. A total of 281,000 tonnes were harvested in 2020 (FAO, 2022), with over 95% of this being exported, primarily to the Indian subcontinent (Pulse Australia, 2016a). The current value of the Australian chickpea industry is estimated at AU $1.33 billion (Wood & Scott, 2021). Furthermore, Australian chickpeas are highly regarded on the international market for their quality (Wood & Scott, 2021).

There is considerable potential for Australian growers to expand the production of chickpea, particularly in northern Australia (Johnson et al., 2021c). Notably, data from the International Trade Centre estimates the current untapped demand for chickpea in international export markets to be worth over $400 million USD (KPMG, 2019).

One notable nutritional characteristic of chickpea is its high protein content (Clemente et al., 1999), making it an excellent replacement for meat in vegetarian diets. Furthermore, proteins and protein hydrolysates can be readily extracted from chickpea using wet or dry extraction methods (Boukid, 2021). These protein fractions can then be used in the production of artificial meat analogues and other protein-fortified products such as noodles, bread and cookies (Boukid, 2021).

In addition to this, chickpea has recently attracted interest due to its potential health-benefitting activity (de Camargo et al., 2019; Faridy et al., 2020; Kaur & Prasad, 2021; Wallace et al., 2016). Previous work has shown that chickpeas or compounds isolated from chickpea display a broad range of advantageous biological activities, including antioxidant activity (Domínguez-Arispuro et al., 2018), anti-cancer activity (Bochenek et al., 2019; Gupta & Bhagyawant, 2019; Gupta et al., 2018), hypocholesterolemic activity (Myint et al., 2017; Yust et al., 2012), hypoglycaemic activity (Akhtar et al., 2019; Ercan & El, 2016; Sreerama et al., 2012), anti-hypertensive activity (Mamilla & Mishra, 2017; Mokni Ghribi et al., 2015), and anti-inflammatory activity (Mahbub et al., 2021; Milán-Noris et al., 2018). The major compound classes believed to be responsible for these beneficial effects include polyphenols, carotenoids, tannins, sterols and peptides (Faridy et al., 2020; Wallace et al., 2016). International research has shown that the content of these phytochemicals – including phenolics and carotenoids – can vary significantly between different chickpea varieties (Bhagyawant et al., 2018; Heiras-Palazuelos et al., 2013; Quintero-Soto et al.,

2018; Rezaei et al., 2019; Serrano et al., 2017; Sharma et al., 2013), similar to that observed in other pulse species (Kim et al., 2013; Valente et al., 2018; Xiang et al., 2019a). Consequently, it is important to understand the variation in key phytochemical constituents between different varieties and growing conditions in the Australian setting.

Only desi type chickpeas were investigated in this study, as these are the predominant type grown in Australia. Desi chickpea comprises 90-95% of total production, with the remaining 5-10% being the kabuli type (Wood & Scott, 2021).

6.3 Materials and methods

6.3.1 Seed material

The 97 desi chickpea samples included in this study were sourced from archived samples stored at Agriculture Victoria Research (Horsham Victoria). The samples comprised 18 different varieties, grown in a range of field trials across four sites in Victoria and 3 growing seasons (2017, 2019 and 2020). The number of samples from each variety ranging from 1 to 20 (mean = 5 samples/variety; see Appendix D for more detail). When conducting varietal analysis, only varieties with ≥10 samples were included (n=5 varieties) to ensure a high level of statistical power. However, all samples were included in the analysis by year or location. The majority of samples (55) were grown under ambient conditions with no imposed treatments; however, 16 of the samples were from herbicide treatment trials and 25 samples were part of pathology trials. More details on the exact field conditions for each sample are provided in Appendix D.

6.3.2 Seed processing and analysis of physical characteristics

The hundred kernel weight of the whole seed was determined using an IC-VA seed counter (AIDEX Co, Japan), with measurements performed in triplicate for each sample. Following this, approximately 20 g of seed material from each sample was ground to a fine flour (Breville Coffee & Spice Grinder; Botany, NSW) prior to subsequent extraction and analysis. As with mungbean, the seed material was not dehulled before being processed.

The colour of the chickpea flour was quantified using a calibrated Konica Minolta chroma meter (CR-400), reported as CIE values of lightness (L*), yellow/blue (b*) and red/green colouration (a*). Again, measurements were performed in triplicate for each sample.

Finally, the moisture content of the flour was determined according to AOAC Official Method 925.10, as described in Section 5.3.1. All subsequent results were expressed on an oven-dry weight basis.

6.3.3 Analysis of FRAP, CUPRAC, TPC and TMAC

Polar phenolic compounds were extracted with 90% methanol, following the protocol described in Section 3.3.3, but using 1 g of flour and a final volume of 14 mL. Extractions and subsequent assays were performed in duplicate for each sample.

In order to reduce analysis times and costs, rapid analysis protocols using 96-well microplates were developed for each of the phytochemical assays (FRAP, CUPRAC, TPC and TMAC). In addition to reducing wastage of reagents and plasticware by a factor of 10 or more, this has the added benefit of reducing handling time for each sample and increasing throughput.

6.3.3.1 Microplate method for TPC

The TPC was determined using a 96-well microplate method developed by the author for the analysis of TPC in other sample extracts. Full details of the method development process are provided in Johnson et al. (2021b). Briefly, this involved investigation of the impacts of the measurement wavelength, incubation temperature and incubation time, as well as comparison of the reproducibility and correlation between the methods.

To perform the TPC microplate assay, 20 µL of sample extract was combined with 100 µL of 1:10 diluted Folin-Ciocalteu reagent in each well, followed by 10 mins incubation in darkness and addition of 100 µL of 7.5% aqueous sodium carbonate solution. After a further 10 mins incubation in darkness, the 96-well plate was shaken for 300 seconds (speed setting: mid) in a microplate reader (Bio-Rad iMark) and the absorbance measured at 750 nm. The results were expressed in gallic acid equivalents (GAE), based off the absorbances of a gallic acid calibration curve (20-120 mg L^{-1}; R^2 = 0.994). Each duplicate extract was analysed in duplicate wells, with the results averaged for each sample.

6.3.3.2 Microplate method for FRAP

The 96-well microplate method for FRAP was also previously published in Johnson et al. (2021c), although validation statistics were not reported there. The method development process, along with the relevant validation statistics, are presented in Appendix C.

To perform the assay, 10 µL of sample extract was combined in each well with 200 µL of FRAP reagent (prepared as described in Section 3.3.5) and shaken for 300 seconds (speed setting: mid), before the absorbance was measured at 593 nm. Results were expressed in Trolox equivalents (TE), based off the absorbances of Trolox standards between 50-300 mg L^{-1} (R^2 = 0.998).

6.3.3.3 Microplate method for CUPRAC

The 96-well microplate method for CUPRAC was also developed by the author, based on a 20× reduction in the proportion of reagents used in the full-scale assay (Section 5.3.2). The method development process, along with detailed validation statistics, are presented in Appendix C.

To perform the CUPRAC assay, 10 µL of sample extract was combined with 50 µL each of 10 mM aqueous copper (II) chloride, 1 M aqueous ammonium acetate, 7.5 mM neocuproine ethanol solution and Milli-Q water in a 96-well plate. After 30 mins incubation (in darkness), the plate was shaken for 60 seconds (speed setting: mid) and the absorbance measured at 450 nm. The results were expressed in Trolox equivalents (TE), based off the absorbances of Trolox standards between 100-1000 mg L^{-1} (R^2 = 0.999).

6.3.3.4 Microplate method for TMAC

The 96-well microplate method for TMAC was modified slightly from the protocol reported by Kwiatkowski et al. (2020). After combining 40 µL of sample extract in a 96-well plate with either 160 µL of pH 1 buffer (containing 0.025 M KCl) or 160 µL of pH 4.5 buffer (containing 0.4 M sodium acetate), the plate was shaken for 300 seconds (speed setting: mid). The absorbance of each well was measured at 510 nm and 700 nm, with the TMAC calculated from the difference in absorbance readings between the pH 1 and pH 4.5 extracts for each sample. Results were expressed in cyanidin-3-glucoside (cyd-3-glu) equivalents, following the formula given in Section 3.3.6, but using a path length of 0.6 cm instead of 1 cm.

6.3.4 Phenolic profiling by HPLC-DAD

Profiling of the phenolic compounds present in the chickpea extracts was conducted using HPLC-DAD. The analysis was conducted using the Agilent 1100 HPLC system and methods described in Section 3.3.7; however, the extracts were not concentrated prior to HPLC analysis. Analysis was conducted on all methanol extracts (n=194), with the results averaged for each sample. Results were expressed in mg/kg of the flour, on an oven-dried weight basis.

As only one of the major compounds matched the retention time and UV spectra of the authentic phenolic standards available (see the list in Section 5.3.4), the remaining compounds were identified through comparison of their UV spectra and relative retention times to previous literature profiling phenolic compounds from chickpea.

6.3.5 NIR and MIR spectroscopy

NIR and MIR spectra were collected from all 100 samples of the mungbean flour using the instruments and settings described in Sections 3.3.8 and 3.3.9. Three replicate spectra were collected from each sample for both MIRS and NIRS, with the mean of the replicate scans used in subsequent data analysis.

6.3.6 Statistical analysis of phenolic data

Statistical tests were performed on the phytochemical and phenolic data using R Studio running R 4.0.5. Where applicable, results are presented as mean ± 1 standard deviation.

6.3.7 Spectral data analysis

Quantitative regression analysis of the infrared spectra was conducted in R Studio, as described in Section 3.3.11. Again, the model performance of the calibration set was evaluated through full LOO cross-validation. For model development, the 2017 and 2019 samples (n=83) were used as a calibration set, while the 2020 samples (n=14) were used as an independent test set. A maximum of 10 components were considered for each PLSR model.

6.4 Results and discussion

6.4.1 Phytochemical profiles

6.4.1.1 Impact of variety, location and season

The first stage of investigation was the impact of variety, location and season on the phytochemical content of the chickpea samples. As the samples were not from a balanced genotype × environment × year trial with equal numbers of samples for each condition (see Appendix D), the impact of these variables was unable to be explored through a three-way ANOVA. However, each of these variables was investigated separately, thus averaging out the impacts of the other two variables (Tables 6-1 to 6-3). Consequently, while the interactions between these terms were unable to be investigated, their broad impacts on phytochemical composition and physical seed parameters could be observed.

Examination of these parameters by variety (Table 6-1) revealed a significant level of variation in the FRAP, TPC and TMAC between the major chickpea varieties, as well as in the seed size (HKW) and the yellow-blue colouration of their flour (CIE b value). The variety Howzat displayed the highest FRAP and TPC, while PBA Slasher showed the lowest concentrations of these analytes. However, this latter variety did contain the highest TMAC.

The FRAP of the chickpea samples was approximately 7-8 times lower than the results previously found for faba bean (Section 3.4.1.1), but higher than the values found for wheat

(Section 4.4.2.1) and mungbean (Section 5.4.1.1). However, the CUPRAC was around three times lower than that found for mungbean. Overall, the FRAP values were lower than the results reported by Johnson et al. (2021c) in the kernel flour of five new chickpea genotypes from Australia. No literature values were found for the CUPRAC analysis of chickpea.

The TPC of the chickpea extracts was around three times lower compared to faba bean, but comparable to the results observed in mungbean. Interestingly, the wheat samples were observed to contain a higher TPC than the chickpea samples, despite the latter showing higher FRAP values for most samples. The TPC of these chickpea samples were also comparable to those found by Johnson et al. (2021c). Furthermore, the TPC was comparable to values reported by Heiras-Palazuelos et al. (2013) for desi chickpea cultivars from Mexico, but lower than most values found by Segev et al. (2011) for Australian chickpeas.

Table 6-1: Impact of variety on the size, colour and phytochemical composition of desi chickpea. Note that only varieties with ≥10 samples were included. Varieties with the same superscript were not statistically different according to a post hoc Tukey test at α = 0.05.

Parameter	Howzat (n=10)	Kyabra (n=14)	PBA Slasher (n=11)	PBA Striker (n=20)	Sonali (n=10)	P value
HKW (g/100)	18.6 ± 1.7^{bc}	23.0 ± 1.9^{a}	18.8 ± 1.2^{bc}	20.2 ± 2.6^{b}	16.6 ± 0.8^{c}	<0.001***
Flour colour – L*	78.05 ± 1.41	80.66 ± 0.79	78.18 ± 6.30	79.28 ± 1.37	77.72 ± 1.10	0.066^{NS}
Flour colour – a*	1.93 ± 0.86	1.74 ± 0.39	1.78 ± 0.20	1.41 ± 0.60	1.51 ± 0.41	0.096^{NS}
Flour colour – b*	27.04 ± 1.65^{a}	26.09 ± 1.28^{ab}	24.96 ± 1.27^{b}	26.27 ± 1.52^{ab}	26.44 ± 0.30^{ab}	0.013*
Moisture (%)	9.21 ± 0.86	9.13 ± 0.83	9.24 ± 0.83	8.86 ± 0.72	8.39 ± 0.75	0.087^{NS}
FRAP (mg TE/100 g)	40.3 ± 16.2^{a}	29.5 ± 6.7^{ab}	24.9 ± 8.9^{b}	33.1 ± 10.4^{ab}	28.5 ± 13.3^{ab}	0.028*
CUPRAC (mg TE/100 g)	124 ± 20	129 ± 21	123 ± 17	132 ± 42	150 ± 26	0.232^{NS}
TPC (mg GAE/100 g)	93.7 ± 11.6^{a}	80.3 ± 14.1^{ab}	72.6 ± 8.8^{b}	91.1 ± 9.5^{a}	82.2 ± 13.7^{ab}	<0.001***
TMAC (mg cyd-3-glu/100 g)	5.8 ± 1.5^{ab}	5.0 ± 3.2^{ab}	7.2 ± 1.5^{a}	4.2 ± 2.0^{b}	4.5 ± 1.4^{b}	0.006**

NS – not significant (P > 0.05), * P < 0.05, ** P < 0.01, *** P < 0.001

All parameters differed significantly with the growing location (Table 6-2). The highest FRAP and TPC were found for the Curyo site, while the highest CUPRAC was at observed Horsham. Conversely, the highest TMAC was at Banyena.

Table 6-2: Impact of growing location on the size, colour and phytochemical composition of desi chickpea. Note that one location (Rupanyup) was excluded as it contained only 5 samples. Locations with the same superscript were not statistically different according to a post hoc Tukey test at α = 0.05.

Parameter	Banyena (n=25)	Curyo (n=18)	Horsham (n=49)	P value
HKW (g/100)	20.1 ± 2.6[a]	17.4 ± 1.2[b]	19.8 ± 3.1[a]	0.003**
Flour colour – L*	79.80 ± 1.06[a]	77.81 ± 1.17[b]	79.70 ± 1.46[a]	<0.001***
Flour colour – a*	1.67 ± 0.25[a]	1.26 ± 0.74[b]	1.98 ± 0.56[a]	<0.001***
Flour colour – b*	24.95 ± 0.68[c]	26.86 ± 1.04[b]	27.77 ± 0.98[a]	<0.001***
Moisture (%)	9.34 ± 0.51[a]	9.05 ± 0.48[a]	8.20 ± 0.96[b]	<0.001***
FRAP (mg TE/100 g)	27.0 ± 6.9[b]	38.3 ± 13.9[a]	34.9 ± 11.6[a]	0.002**
CUPRAC (mg TE/100 g)	114 ± 18[b]	133 ± 23[b]	157 ± 36[a]	<0.001***
TPC (mg GAE/100 g)	74.1 ± 6.9[b]	92.8 ± 13.9[a]	87.0 ± 11.6[a]	<0.001***
TMAC (mg cyd-3-glu/100 g)	6.6 ± 2.5[a]	5.0 ± 1.5[b]	3.5 ± 1.4[c]	<0.001***

NS – not significant (P > 0.05), * P < 0.05, ** P < 0.01, *** P < 0.001

Similarly, the growing season had a significant impact on all parameters measured (Table 6-3). Both FRAP and CUPRAC were higher in the 2020 samples, while the TPC was significantly lower in the 2017 samples. It is important to caution that as all samples were stored following harvest, there may have been some change in their composition over this period, particularly for the older samples. Although there does not appear to be any work documenting this specifically in chickpea, Nasar-Abbas et al. (2009) noted a minor reduction in the TPC of faba bean samples over the period of one year, with the loss accelerated under higher temperatures or exposure to light. However, Ziegler et al. (2016b) found contrasting results in soybean, with the free phenolic content increasing slightly over a storage period of one year.

In addition to possessing the highest TMAC, the oldest samples (2017) also tended to have a larger kernel size and higher moisture content. This latter parameter may be related to the absorption of moisture by the chickpea samples over the longer storage time.

Table 6-3: Impact of growing year on the size, colour and phytochemical composition of desi chickpea. Years with the same superscript were not statistically different according to a post hoc Tukey test at α = 0.05.

Parameter	2017 (n=30)	2019 (n=53)	2020 (n=14)	P value
HKW (g/100)	20.2 ± 2.4[a]	19.3 ± 2.8[a]	18.5 ± 3.1[a]	0.045*
Flour colour – L*	80.00 ± 1.14[a]	79.09 ± 1.71[b]	79.57 ± 1.12[ab]	0.0498*
Flour colour – a*	1.68 ± 0.27[b]	1.62 ± 0.65[b]	2.42 ± 0.40[a]	0.027*
Flour colour – b*	24.82 ± 0.74[b]	27.54 ± 1.12[a]	27.48 ± 0.88[a]	<0.001***
Moisture (%)	9.46 ± 0.56[a]	8.45 ± 0.97[b]	8.35 ± 0.81[b]	<0.001***
FRAP (mg TE/100 g)	26.9 ± 6.4[c]	33.8 ± 11.9[b]	43.7 ± 10.1[a]	<0.001***
CUPRAC (mg TE/100 g)	116 ± 18[c]	142 ± 32[b]	181 ± 28[a]	<0.001***
TPC (mg GAE/100 g)	76.4 ± 8.8[b]	88.7 ± 12.2[a]	88.3 ± 10.6[a]	<0.001***
TMAC (mg cyd-3-glu/100 g)	6.4 ± 2.4[a]	4.0 ± 1.6[b]	3.6 ± 1.0[b]	<0.001***

NS – not significant (P > 0.05), * P < 0.05, ** P < 0.01, *** P < 0.001

6.4.1.2 Correlation analysis

To further investigate the inter-relationships that may exist between the bioactive phytochemical constituents and the physical characteristics of the seed, Pearson R linear correlation analysis was performed. The results are summarised in the correlogram presented in Figure 6-1. Significant correlations were observed between TPC and FRAP, but not between TPC and CUPRAC or FRAP and CUPRAC. The CUPRAC was positively correlated with the b* colour, but negatively correlated with moisture content.

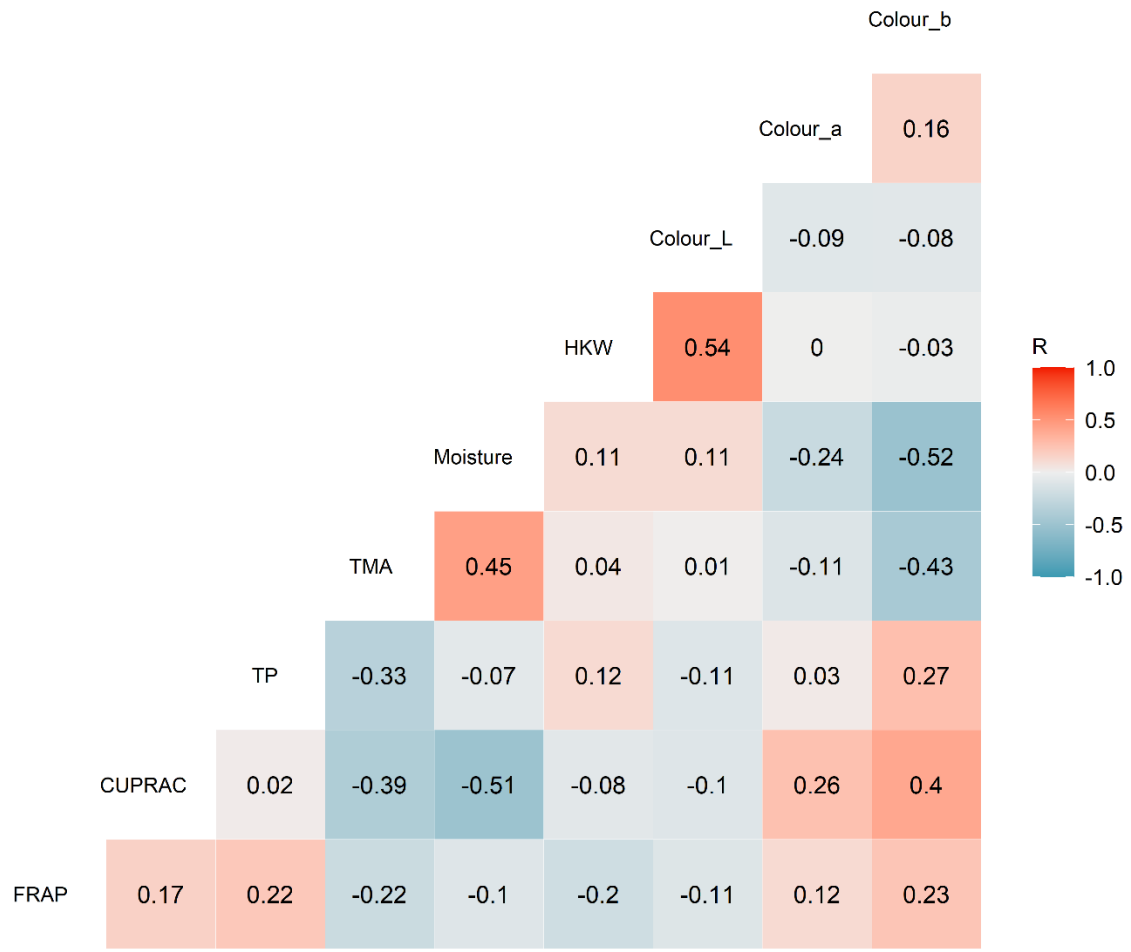

Figure 6-1: Correlogram showing the correlations between the phytochemical constituents and physical parameters of the chickpea seed (n = 97 samples). Correlations with R values above 0.21 or below -0.21 were statistically significant at α = 0.05.

6.4.2 Phenolic profiles using HPLC-DAD

Following on from the results of the benchtop phytochemical assays, phenolic profiling was performed on all of the chickpea extracts using HPLC-DAD. Analysis was performed in duplicate, with the results averaged for each sample. As mentioned in the methods (Section 6.3.4), the extracts were not concentrated prior to analysis.

A typical chromatogram obtained from the HPLC-DAD analysis is shown in Figure 6-2. Aside from the solvent front and glycosylated compounds emerging between 1-2 mins, two predominant compounds were observed at 2.93 and 4.68 mins, with another large peak occurring at 13.65 mins. In addition to this, a number of compounds were also present in lower concentrations. These appear as upward spikes on the UV isoplot (Figure 6-3) and can also

be seen on the 3D HPLC plot (Figure 6-4). Overall, a total of 12 compounds were found to occur in the majority of the chickpea extracts (Table 6-4).

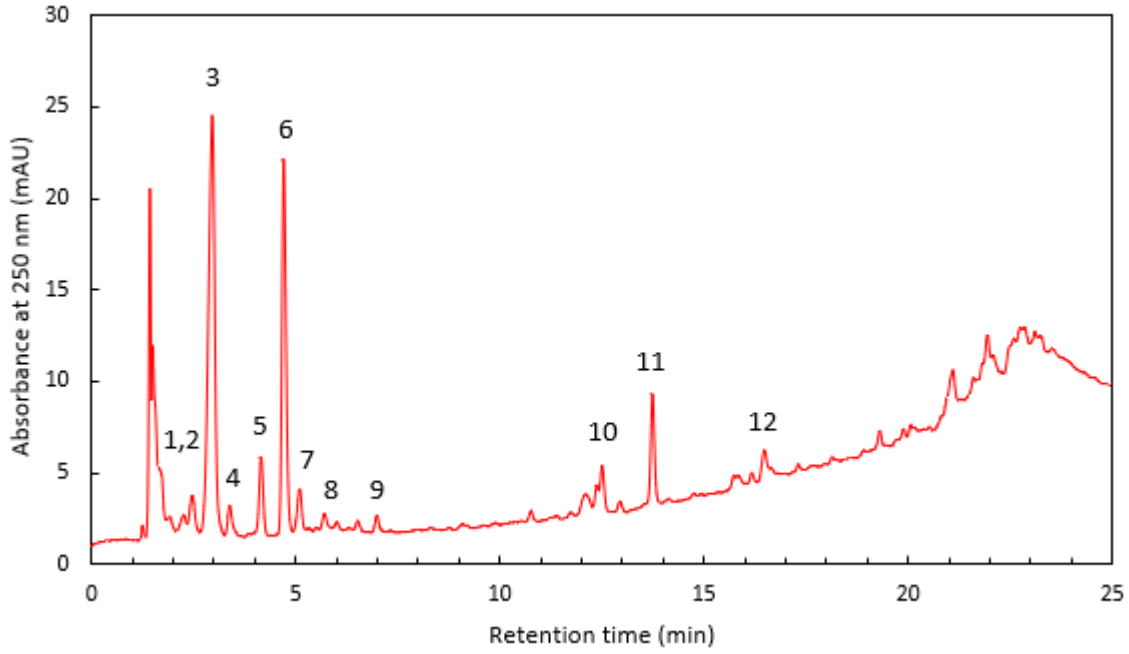

Figure 6-2: Annotated HPLC chromatogram from one of the chickpea samples, showing the absorbance trace at 250 nm.

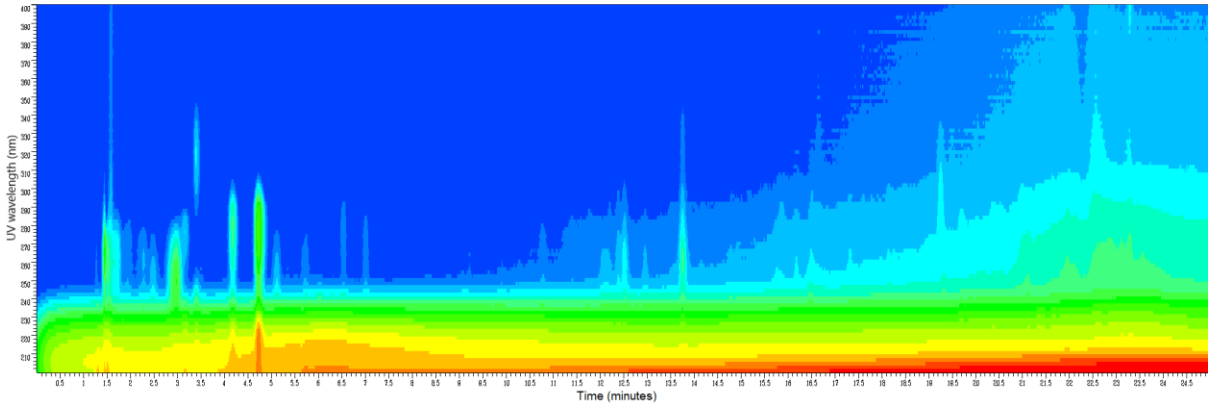

Figure 6-3: Isoplot showing the UV absorbance of various compounds eluting at various points throughout the HPLC run. The x-axis shows the run time (from 0-25 minutes) and the y-axis shows the UV wavelength (from 200-400 nm). The colour of each pixel shows the relative absorbance (blue = low; red = high).

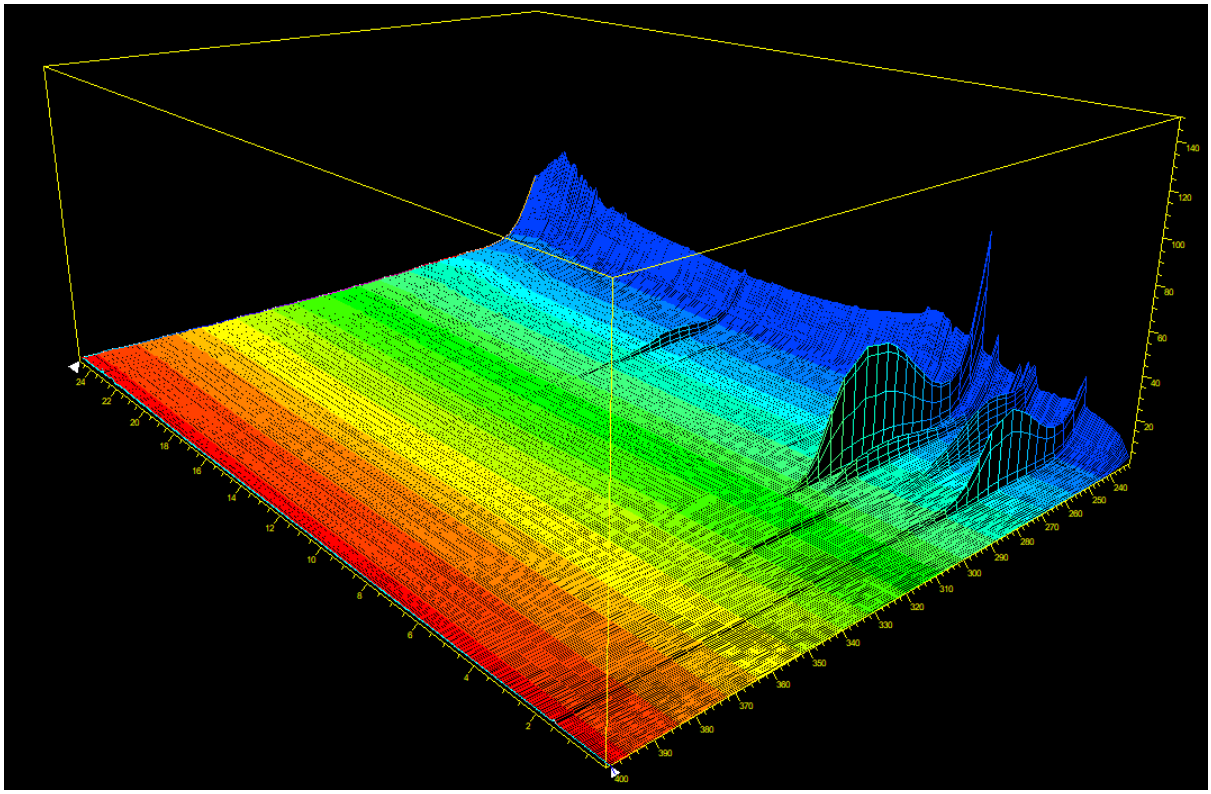

Figure 6-4: Three-dimensional plot showing the UV absorption (from 230 nm in blue to 400 nm in red) at different times throughout the HPLC run. Note that the time axis runs from 0 mins (on the right-hand side) to 25 mins (on the left-hand side).

As the majority of peaks did not match the retention times of the authentic phenolic standards available (see Section 5.3.4), the compounds were tentatively identified from their relative retention times and UV spectral features (Table 6-4) using relevant literature (Aguilera et al., 2011; Escarpa & González, 2000; Klejdus et al., 2007; Quintero-Soto et al., 2018; Wu et al., 2012). As these identities were only tentative, these compounds are referred to as Peaks 1-12 throughout the remainder of this section.

As shown in Table 6-5, there was a significant amount of variation found between the desi chickpea varieties for most of the phenolic compounds quantified. Howzat appeared to show the highest concentrations of most compounds, most notably for Peaks 2, 3 and 6. On the other hand, PBA Slasher typically contained lower levels of the phenolics.

Similar to the variation observed between genotypes, there was a significant difference observed between growing locations and year for most compounds. In nearly all cases, the 2017 samples showed the lowest concentrations of each compound. Likewise, the Banyena site showed the lowest concentrations for most compounds, while the Curyo site typically showed the highest concentrations.

Table 6-4: Tentative identifications of the compounds found in the chickpea methanol extracts using HPLC-DAD.

No.	Retention time (min)	UV peaks (λ_{max}) (nm)	Tentative identification	Quantification wavelength (nm)	Quantified as equivalents of
1	2.25	236, 263	Gallic acid	250	4-hydroxybenzoic acid
2	2.46	251	*p*-hydroxybenzoic acid hexoside	250	4-hydroxybenzoic acid
3	2.93	250	*p*-hydroxybenzoic acid pentoside?	250	4-hydroxybenzoic acid
4	3.33	211, 237, 318	Sinapic acid hexoside	320	Gentisic acid
5	4.13	219, 280	Catechin	280	Gallic acid
6	4.68	219, 280	Epicatechin	280	Gallic acid
7	5.05	251	4-hydroxybenzoic acid	250	4-hydroxybenzoic acid
8	5.64	212^{sh}, 254	Methoxybenzyl alcohol glycoside?	250	4-hydroxybenzoic acid
9	6.95	263	Genistein hexoside	250	4-hydroxybenzoic acid
10	12.43	254, 306^{sh}	Pseudobaptigenin	250	4-hydroxybenzoic acid
11	13.65	262	Biochanin A	250	4-hydroxybenzoic acid
12	16.42	255, 302^{sh}	Daidzein	250	4-hydroxybenzoic acid

sh = shoulder

Table 6-5: Individual phenolic contents of the chickpea samples (all values in mg/kg on a dry-weight basis), broken down by variety, growing location and year. Only varieties and locations with >9 samples each were included. The P value rows show the results of a one-way ANOVA performed for each variable. Entries with the same superscript were not statistically different according to a post hoc Tukey test at α = 0.05.

	Peak 1	Peak 2	Peak 3	Peak 4	Peak 5	Peak 6	Peak 7	Peak 8	Peak 9	Peak 10	Peak 11	Peak 12
Variety												
Howzat (n=10)	10.7 ± 5.7^{a}	22.2 ± 7.7^{a}	191.5 ± 29.6^{a}	11.4 ± 3.5^{a}	24.9 ± 5.6^{bc}	123.2 ± 19.5^{a}	17.9 ± 5.2^{a}	2.2 ± 3.0^{bc}	6.1 ± 1.6^{a}	17.1 ± 5.8^{a}	20.9 ± 8.0^{a}	19.1 ± 5.9^{ab}
Kyabra (n=14)	2.9 ± 4.8^{bc}	9.6 ± 2.7^{b}	36.7 ± 14.9^{d}	9.4 ± 2.2^{a}	27.7 ± 2.2^{ab}	94.9 ± 26.9^{b}	10.7 ± 7.2^{bc}	0.5 ± 1.8^{c}	1.4 ± 2.1^{c}	13.6 ± 6.4^{ab}	17.6 ± 8.8^{ab}	13.7 ± 6.3^{b}
PBA Slasher (n=11)	0.0 ± 0.0^{c}	8.8 ± 1.0^{b}	74.5 ± 16.8^{c}	6.0 ± 0.7^{b}	27.3 ± 1.6^{ab}	72.0 ± 25.3^{bc}	9.3 ± 7.4^{c}	4.3 ± 2.7^{ab}	3.7 ± 1.6^{b}	7.8 ± 3.2^{c}	9.9 ± 1.9^{c}	19.7 ± 13.1^{ab}
PBA Striker (n=20)	9.2 ± 3.9^{a}	12.1 ± 3.0^{b}	149.7 ± 31.1^{b}	6.6 ± 0.7^{b}	30.3 ± 6.0^{a}	95.7 ± 28.5^{b}	11.6 ± 5.6^{abc}	1.8 ± 2.7^{bc}	6.1 ± 1.0^{a}	11.5 ± 4.0^{bc}	14.7 ± 3.6^{bc}	18.5 ± 3.4^{ab}
Sonali (n=10)	7.1 ± 3.3^{ab}	8.5 ± 1.9^{b}	64.4 ± 15.8^{cd}	5.2 ± 0.4^{b}	20.3 ± 2.1^{c}	58.0 ± 19.4^{c}	17.7 ± 5.0^{ab}	6.6 ± 3.3^{a}	5.8 ± 0.8^{a}	13.5 ± 3.6^{abc}	16.2 ± 3.1^{abc}	23.3 ± 1.6^{a}
P value	*<0.001****	*<0.001****	*<0.001****	*0.161^{NS}*	*<0.001****	*<0.001****	*<0.001****	*0.041**	*<0.001****	*<0.001****	*<0.001****	*<0.001****
Year												
2017 (n=30)	1.6 ± 3.3^{b}	9.1 ± 2.1^{b}	76.7 ± 42.8^{a}	6.8 ± 1.1	28.5 ± 3.1	83.0 ± 33.6	7.5 ± 3.5^{c}	1.9 ± 2.6^{b}	3.0 ± 2.3^{b}	8.8 ± 1.9^{c}	11.0 ± 3.6^{c}	14.2 ± 8.9^{b}
2019 (n=53)	8.2 ± 4.4^{a}	12.0 ± 5.9^{ab}	105.5 ± 60.3^{a}	7.9 ± 3.0	28.2 ± 7.1	85.9 ± 30.8	18.0 ± 6.5^{b}	4.3 ± 5.6^{b}	5.5 ± 1.5^{a}	13.4 ± 5.0^{b}	18.0 ± 5.9^{a}	22.2 ± 2.2^{a}
2020 (n=14)	10.0 ± 7.5^{a}	13.7 ± 9.6^{a}	106.7 ± 45.3^{a}	7.5 ± 3.6	30.3 ± 6.0	88.4 ± 35.1	28.6 ± 7.2^{a}	9.0 ± 2.5^{a}	5.4 ± 0.7^{a}	18.8 ± 8.8^{a}	12.5 ± 5.7^{b}	23.6 ± 2.3^{a}

P value	*<0.001***	*0.008***	*0.020**	*0.136*^{NS}	*0.654*^{NS}	*0.590*^{NS}	*<0.001***	*<0.001***	*<0.001***	*<0.001***	*0.001***	*<0.001***
Location												
Banyena (n=25)	1.8 ± 3.5^{b}	8.8 ± 1.7^{b}	67.4 ± 36.4^{b}	6.9 ± 1.0	27.7 ± 2.3^{b}	70.0 ± 17.8^{b}	6.4 ± 1.7^{b}	1.7 ± 2.4^{b}	2.8 ± 2.4^{c}	8.6 ± 1.7^{b}	10.6 ± 1.4^{b}	14.7 ± 9.6^{b}
Curyo (n=18)	10.5 ± 3.6^{a}	16.4 ± 7.8^{a}	153.5 ± 60.4^{a}	7.7 ± 3.3	22.1 ± 4.3^{c}	96.7 ± 22.9^{a}	18.0 ± 3.4^{a}	4.5 ± 3.5^{ab}	6.7 ± 1.0^{a}	13.8 ± 5.0^{a}	19.3 ± 4.5^{a}	21.6 ± 1.8^{a}
Horsham (n=49)	7.9 ± 5.5^{a}	10.8 ± 5.8b	88.2 ± 45.0^{b}	7.9 ± 3.1	31.0 ± 6.1^{a}	82.7 ± 33.6^{ab}	21.0 ± 8.9^{a}	5.5 ± 6.0^{a}	5.0 ± 1.2^{b}	14.8 ± 6.7^{a}	16.0 ± 6.5^{a}	22.8 ± 2.4^{a}
P value	*<0.001***	*<0.001***	*<0.001***	*0.292*^{NS}	*<0.001***	*0.012**	*<0.001***	*0.006***	*<0.001***	*<0.001***	*<0.001***	*<0.001***

NS – not significant (P > 0.05), * P < 0.05, ** P < 0.01, *** P < 0.001

Finally, correlation analysis was performed on the HPLC-DAD data to determine if there were any significant correlations between any of the phenolic compounds and the phytochemical measurements (FRAP, CUPRAC, TPC, TMAC). This revealed significant positive correlations between a number of compounds and the FRAP, CUPRAC and TPC (Figure 6-5), the most notable of which was between Peak 6 and the TPC. Peak 7 was strongly correlated with the CUPRAC, while only correlations of moderate strength were found between FRAP and the phenolic compounds.

In addition, strong positive correlations were observed between several of the phenolic compounds (e.g., Peaks 1 & 2, 2 & 3, 7 & 8, 3 & 9). Overall, the correlation results were similar to those observed for faba bean (Section 3.4.2.3) and mungbean (Section 5.4.3), although there appeared to be fewer strong correlations in chickpea.

Figure 6-5: Correlogram showing the relationships between the 12 phenolic compounds measured by HPLC-DAD and the phytochemical constituents. Correlations with R values above 0.20 or below -0.20 were statistically significant at α = 0.05.

6.4.3 Prediction of bioactive compounds using IR spectroscopy

As with the preceding grain crops, infrared spectroscopy was applied for the prediction of bioactive analytes in the chickpea flour. The results are detailed throughout this section.

6.4.3.1 Descriptive statistics

The descriptive statistics for the calibration and test sets are provided in Table 6-6. The range and mean values were mostly comparable between the calibration and test sets, although the FRAP and CUPRAC were moderately higher in the latter.

Table 6-6: Descriptive statistics for the parameters measured in this study, in the calibration and test sets.

Parameter	Calibration set – 2017 and 2019 samples (n=83)		Test set – 2020 samples (n=14)	
	Range	Mean ± SD	Range	Mean ± SD
Moisture (%)	6.40-10.67	8.81 ± 0.97	7.20-9.84	8.35 ± 0.81
FRAP (mg TE/100 g)	9.6-56.7	28.3 ± 10.0	19.6-56.3	40.1 ± 9.3
CUPRAC (mg TE/100 g)	65.3-211.3	121.5 ± 28.6	137.9-224.7	165.5 ± 25.2
TPC (mg GAE/100 g)	53.8-106.1	76.8 ± 11.5	61.7-98.1	81.0 ± 10.0
TMAC (mg cyd-3-glu/100 g)	0.0-7.9	4.4 ± 2.0	1.7-4.8	3.3 ± 0.9

6.4.3.2 NIR spectra

The raw and pre-processed NIR spectra are shown in Figure 6-6. As observed in the NIR spectra of the other grain crops, the major peaks for the chickpea flour samples were located at 1470 and 1936 nm (OH 2nd and 1st overtones), with smaller peaks at 1202 nm (CH 2nd overtone), 1761 nm (CH 1st overtone), 2110 nm (amide 1st overtone) and 2311 nm (CH combination bands). However, it was noted that the spectra displayed a more prominent minor peak at 2346 nm (the right-most peak in the 2280-2360 nm region) compared to other crops such as faba bean and wheat.

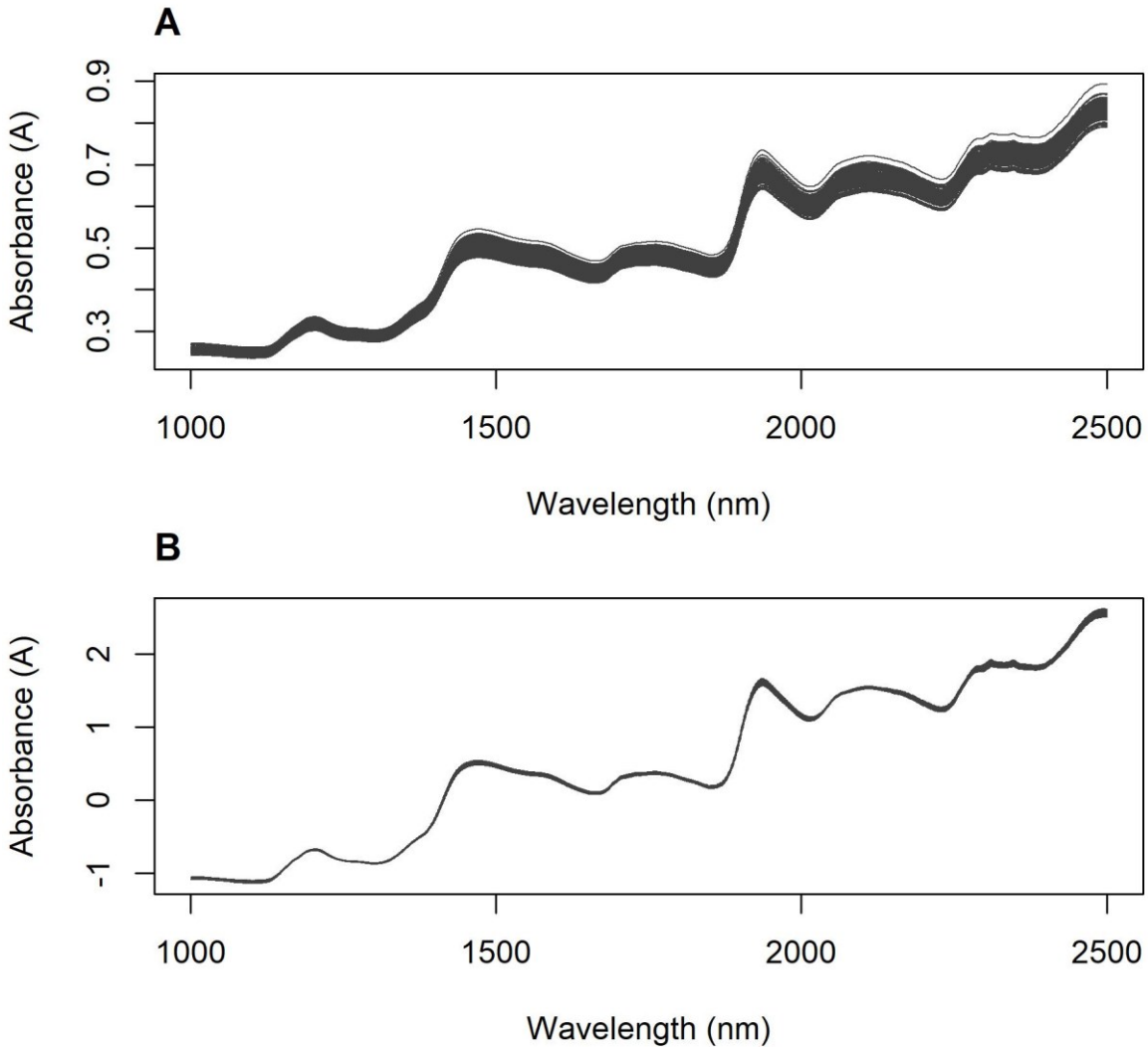

Figure 6-6: The raw absorbance NIR spectra (A) and SNV-processed spectra (B) of the chickpea flour samples.

6.4.3.3 NIR models

The optimum NIR models developed for each analyte are shown in Table 6-7. Although the model for FRAP performed acceptably in the calibration set, no predictive power was observed when it was applied to the independent test set (R^2 of 0). Similarly, none of the other analytes showed any predictive power in the independent test set. For this reason, no prediction plots or model loadings are shown.

As discussed for mungbean (Section 5.4.4.3), the failure of NIRS to predict these bioactive compounds is most likely related to the low concentrations and relatively low level of variation in the analytes, combined with other non-target matrix constituents obscuring the NIR signals that were successfully used for prediction in faba bean and wheat.

Table 6-7: Optimum PLSR models found for the prediction of the specified analytes using NIR spectroscopy.

Parameter	Spectral pre-processing	Factors	R^2_{cv}	RMSECV	RPD	R^2_{test}	RMSEP	Bias	Slope	Intercept
FRAP (mg TE/100 g)	SNV + 2d5	7	0.78	0.46	2.13	0.00	0.95	0.17	0.03	8.12
CUPRAC (mg TE/100 g)	SNV + 2d11	4	0.08	9.52	1.05	0.31	11.5	8.73	0.97	9.72
TPC (mg GAE/100 g)	SNV + 2d5	5	0.54	19.3	1.48	0.02	33.8	21.7	-0.54	242.4
TMAC (mg cyd-3-glu/100 g)	2d5	3	0.37	9.06	1.26	0.09	13.1	5.75	-0.68	131.7

Table 6-8: Optimum PLSR models found for the prediction of the specified analytes using MIR spectroscopy.

Parameter	Spectral pre-processing	Factors	R^2_{cv}	RMSECV	RPD	R^2_{test}	RMSEP	Bias	Slope	Intercept
FRAP (mg TE/100 g)	SNV + 2d5	3	0.59	0.62	1.58	0.07	1.13	-0.28	-0.36	11.4
CUPRAC (mg TE/100 g)	2d5	1	0.05	9.7	1.03	0.21	15.6	10.9	-1.13	72.9
TPC (mg GAE/100 g)	2d21	10	0.61	17.8	1.61	0.06	40.3	28.5	0.27	128.8
TMAC (mg cyd-3-glu/100 g)	1d5	6	0.48	8.2	1.40	0.00	11.7	0.83	-0.07	86.8

6.4.3.4 MIR spectra

The MIR spectra of the chickpea flour samples are shown in Figure 6-7. Again, the MIR spectra appeared broadly similar to that of the other grain crops, but with more prominent peaks at 2854 nm (attributed to CH_2 asymmetrical stretch) and 1744 nm (attributed to the C=O stretch of carbonyl groups). There was a moderate degree of variation in amplitude between the spectra, which was removed through use of the SNV algorithm (Figure 6-7) or derivative pre-processing.

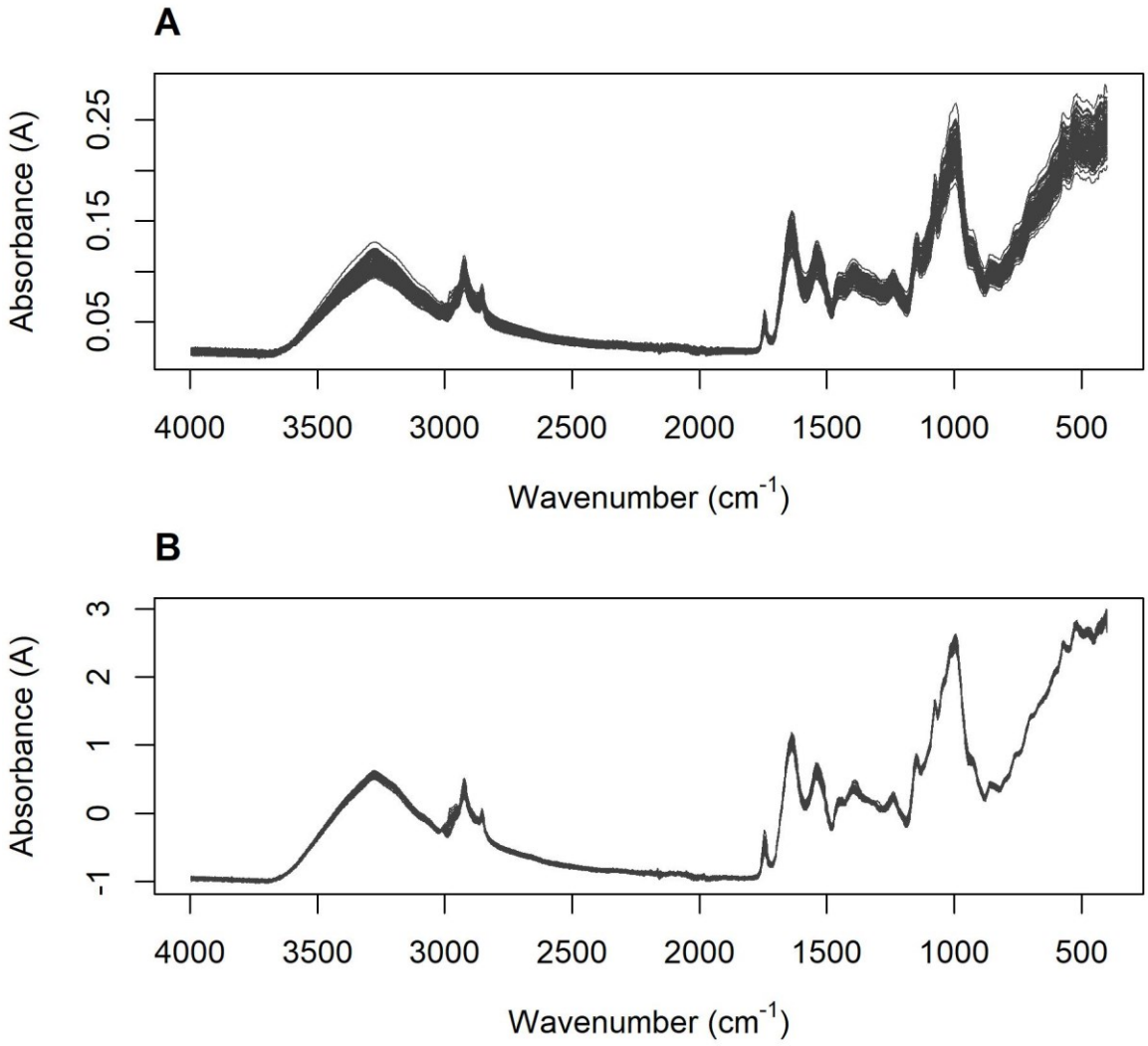

Figure 6-7: The raw MIR spectra (A) and SNV-processed spectra (B) of the chickpea flour samples.

6.4.3.5 MIR models

The optimum PLSR models developed from the MIR spectra of the chickpea samples are shown in Table 6-8. As with the NIR spectra, several models (i.e., FRAP, TPC) performed moderately well in the calibration stage, but were unable to provide any level of prediction in

the independent test set. Potential reasons for this prediction failure are likely to be the same as those discussed for the MIRS models in other crops (see Sections 3.4.4.6 and 5.4.4.5).

The only previous study using MIR spectroscopy for the assessment of bioactive constituents in chickpeas was performed by Kadiroğlu et al. (2018), who attempted the prediction of total phenolic content, iron chelating activity and free radical scavenging activity (a measure of antioxidant activity) in aqueous extracts from chickpeas, amongst other more proximate measures of quality (water soluble protein content, water binding capacity, oil binding capacity). The error found for the prediction of total phenolic content was the highest of all analytes investigated, but still appeared suitable for screening purposes (R^2_{CV} = 0.951, RMSECV = 5.03 mg/100 g GAE). However, it is important to note that this study used only 36 chickpea samples and did not examine the performance of the model on a test set (either dependent or independent), but only using internal cross-validation results. Given that the authors also used 8 latent variables for this regression, it is probable that model over-fitting in this study gave over-optimistic results.

6.5 Summary

The results demonstrated significant variation in the TPC, TMAC and antioxidant capacity (FRAP but not CUPRAC) of different desi chickpea cultivars grown in Victoria. Similarly, the growing location and year had a significant impact on the levels of these phytochemical constituents. The major phenolic compounds present were characterised using HPLC-DAD.

Similar to the results seen for mungbean, neither NIRS nor MIRS were able to predict any of the bioactive analytes (FRAP, CUPRAC, TPC or TMAC) quantified in the chickpea flour samples.

Chapter 7. General discussion and recommendations

7.1 General discussion

7.1.1 Typical phytochemical composition of Australian-grown grain crops

Despite its limitations, this work represents one of the largest conducted to date on the phytochemical and bioactive composition of Australian-grown grains. Consequently, it provides important context information on the typical levels of these compounds that can be expected in grain crops grown in Australia under comparable conditions. Furthermore, detailed HPLC-DAD phenolic profiles are reported for three of the less-studied crop types, allowing for future researchers to rapidly create targeted HPLC or LC-MS methods for the analysis of the major phenolic compounds in these crops.

Of the four grain crops investigated, the highest levels of phenolics and antioxidants were found in faba bean. This pulse also displayed the highest anthocyanin content. In the three remaining crops, wheat showed the highest TPC, followed by mungbean. However, chickpea showed the highest antioxidant activity (as measured by the FRAP assay), followed by wheat. In contrast, the CUPRAC was higher in mungbean compared to chickpea. This indicates that TPC is not necessarily an accurate indicator of antioxidant activity. The lack of clear correlation between TPC and antioxidant activity in some of the crops studied here stands in contrast to results reported by Li et al. (2016) in soybean, where it was found that the main quantitative trait loci regulating total antioxidants, phenolics, and flavonoids were found in one overlapping genomic region. This emphasises that correlations observed in one grain crop may not necessarily be extrapolated to other species. Furthermore, the results presented here demonstrate that the estimated antioxidant activity of a particular matrix varies depending on the assay method used. Consequently, this supports the recommendations of previous authors that at least two different assays should be used to measure antioxidant activity in a sample (Apak et al., 2013; Bartosińska et al., 2016).

In most cases, the levels of bioactive compounds found here were comparable to those previously reported in international studies. Although these results cannot be extrapolated to all pulse and grain crops, this suggests that for the four grain crops studied, Australian-grown produce should contain comparable or higher levels of health-benefiting compounds compared to other produce on the international market. This is in spite of the harsher growing conditions in this country (Page et al., 2018; Unkovich et al., 2020), principally poor soil quality and low rainfall. The maintenance of high quality in Australian grain crops may be attributable to a number of factors, including the selection of high-quality genotypes for breeding programs

(Aziz et al., 2018; Ryan, 2018; Sadras et al., 2021), the adoption of informed agronomic practices (Thomas et al., 2018; Wood et al., 2021b), such as crop rotations (Armstrong et al., 2019), targeted irrigation (Ara et al., 2021; Wood et al., 2021a) and targeted application of fertiliser (Page et al., 2018). These practices have been reported to provide increases in protein content (Armstrong et al., 2019) and consequently are likely to positively impact the contents of other phytochemical constituents.

Furthermore, it has been noted that both heat stress and drought stress generally increase the phenolic content and antioxidant capacity of grain crops (Britz & Kremer, 2002; Dwivedi et al., 2019; Laddomada et al., 2017; Rayee et al., 2018), although this impact is genotype-dependent (Liu et al., 2018a). Accordingly, the harsher Australian growing conditions may actually enhance the levels of these health-benefiting compounds compared to more mild climates such as Europe or Northern America.

7.1.2 Variation in phytochemical composition

In general, significant genotypic variation in phenolic content and antioxidant capacity was observed between different varieties of each grain crop. This was most notable in faba bean, where PBA Rana contained almost double the total phenolic content of the other nine varieties. A significant level of variation was also observed in the mungbean and chickpea varieties, albeit to a lesser magnitude.

This broadly agreed with the results of previous studies, which have highlighted a relatively high level of genotypic variation in the phenolic contents of most grain crops (Cai et al., 2015; MeiXiong et al., 2020; Oomah et al., 2011; Valente et al., 2019; Zhou et al., 2020). This is perhaps an unexpected finding in light of numerous authors warning of genetic bottlenecks and a lack of diversity existing among many grain crops (Abbo et al., 2003a; Furbank & Tester, 2011; Muñoz et al., 2017; Smýkal et al., 2015). However, this high degree of genetic diversity with regard to phytochemical constituents is likely to have arisen from the fact that grain breeders primarily select lines for their elevated yield, protein content and superior disease resistance (Iqbal et al., 2007; Michel et al., 2019; Miedaner et al., 2020; Siddique et al., 2000), with no regard for their phenolic content or antioxidant capacity. It is only very recently that selective breeding for bioactive compounds has been proposed in the grains sector (Gordeeva et al., 2020; Loskutov & Khlestkina, 2021).

Although genes controlling the production of phenolic compounds may be linked to some physical grain quality parameters, such as malting suitability (Cai et al., 2015), kernel size (Cai et al., 2016) and grain colour (Shao et al., 2011), they do not appear to be strongly linked to traits such as yield and protein content (Nigro et al., 2017; Rhodes et al., 2014). Consequently, there is unlikely to be any unidirectional selection pressure for high/low phenolic contents

imposed by previous and current breeding programs. Nevertheless, researchers are still working to identify the genes responsible for the quantitative production of many phenolic compounds found in grain crops (Czyczyło-Mysza et al., 2019; Han et al., 2020; Li et al., 2010; Sharma et al., 2020).

Compared to the influence of genotype, environmental factors (specifically growing location and season) had less of an impact on phenolic content and antioxidant capacity in faba bean and chickpea. However, the environmental influence was more obvious for chickpea, possibly due to the larger number of growing locations and seasons sampled. Similar results were found by Zrckova et al. (2018), who reported that genotype had the strongest impact on antioxidant capacity in wheat, while season had a stronger influence on TPC. Similarly, Sanghamitra et al. (2018) noted a high level of heritability in anthocyanin capacity and flavonoid content in pigmented rice. Interestingly, previous research on wheat has suggested lesser impacts of genotype × environment interactions under Australian conditions compared to international studies from North America and Europe (Williams et al., 2008). This contrasts with expected results, given the geographically diverse range of Australian growing locations studied; however, the reasons for this difference remain unclear (Williams et al., 2008). Nevertheless, this observation indicates the phytochemical composition of a given genotype should be more predictable when grown in Australia under known environmental conditions.

7.1.3 Performance of NIRS and MIRS models for the rapid prediction of nutritional and bioactive analytes

The application of NIRS and MIRS for the rapid prediction of nutritional and bioactive analytes showed mixed results. Excellent performance was found for the prediction of protein in the faba bean and wheat samples using NIRS; MIRS was also able to predict protein content in wheat, but not in faba bean. Furthermore, NIRS showed promise for the estimation (i.e., prediction of high or low content) of TPC and antioxidant capacity in faba bean and wheat, but not in mungbean or chickpea. Similar to the results for protein, MIRS could predict TPC and antioxidant capacity (FRAP) in wheat with reasonable accuracy, but was unable to predict TPC or antioxidant capacity in any of the three pulse crops. In the wheat samples, the loadings differed between the MIRS models for TPC and FRAP, whereas the loadings for the NIRS models of these analytes appeared very similar to one another. Consequently, the MIRS models appeared to be more specific for the analyte of interest, although the overall model performances were comparable between MIRS and NIRS in this matrix.

No predictive models were found for anthocyanin content in any crop. This was anticipated, given the very low concentration of this analyte. Furthermore, the only previous study attempting the prediction of anthocyanin content using infrared spectroscopy was Amanah et

al. (2020), who reported the prediction of anthocyanins in soybean using NIRS. However, the only model validation used by these authors was using spectra collected from the same samples included in calibration set – making the reported results extremely over-optimistic.

Nevertheless, NIRS appears to show promise for the prediction of the major bioactive analytes (TPC and antioxidant capacity) in some – but not all – grain crops. Given that NIR spectra are routinely collected from grain samples for the determination of protein content, application of models for the approximation of TPC and/or antioxidant capacity could be one way to gain further information about the bioactive composition of the samples. It is important to caution that this could only be the case if the wavelength range of the NIR spectra collected for protein determination included the key wavelengths involved in the TPC/FRAP prediction models (approximately 1800-2400 nm). However, if this was the case, there would be no extra cost or time required, aside from the initial investment of setting up the model and occasional calibration checks.

7.2 Recommendations

7.2.1 Recommendations for researchers

Based on the results found throughout this study, the following recommendations are made for researchers in this field:

- At least two different assays should be used to measure antioxidant activity in a sample, as measured antioxidant activities can vary significantly between matrices depending on the assay used.
- Future studies using infrared spectroscopy for the non-invasive prediction of bioactive (or other) compounds in grain crops should incorporate the use of an independent test set for model validation (i.e., samples grown in a different season or location to the samples used in the calibration set).
- More work is required to determine if NIRS/MIRS can predict TPC or antioxidant capacity in mungbean or chickpea flour, using a larger and/or more diverse calibration set.
- As all of the samples investigated in this study were ground to a flour before the collection of NIR/MIR spectra, it is unknown whether the levels of bioactive compounds could be predicted in samples of intact grain. Based on literature results it appears that this could be possible, although potentially with the trade-off of reduced accuracy.
- Future studies could also investigate whether non-linear regression techniques, such as neural networks or support vector regression, provide more accurate prediction compared to PLSR.

7.2.2 Recommendations for industry (buyers, wholesalers and processors)

Similarly, the following points are noted for industry stakeholders in the grains sector:

- Pulse crops with a higher bioactive content, such as faba bean, may potentially be marketed as functional foods products and attract a moderate price premium.
- NIRS may be able to predict the phenolic content/antioxidant capacity of some grain crops; hence it would bear investigating whether this can be conducted using instrumentation already in place for the measurement of protein content.
- Similarly, NIRS may be a suitable method for the routine quality assurance of functional food products by processors/manufacturers.

7.2.3 Recommendations for policy makers

Finally, there are several points of relevance for policy makers responsible for determining research directions/priorities:

- It would be worth investigating the feasibility of paying premiums to grain growers based on the overall nutritional properties of their crop – including the levels of bioactive compounds for functional food crop – rather than just the physical grain quality and protein content.
- Similarly, more detailed comparison between the levels of bioactive compounds found in Australian-grown and internationally grown grains would prove useful. If Australian grains can be marketed in international markets with an emphasis on their higher levels of bioactive compounds, then this could provide a significant competitive edge for Australian growers.

Chapter 8. Conclusions

The following conclusions are derived from the results presented throughout this thesis:

1. There is a significant level of genotypic variation in the phenolic content and antioxidant capacity of all the pulse crops studied (faba bean, mungbean and chickpea).
2. There is also a significant impact of environmental conditions (growing location and season) on the phenolic content and antioxidant capacity; however, this is generally less than genotypic effects.
3. Influences of genotype and environmental conditions on individual phenolic compounds often differ from their influence on the total phenolic content.
4. Infrared spectroscopy (both NIRS and MIRS) shows promise for the estimation of TPC and antioxidant capacity in some of the grain crops studied, with the matrix type strongly impacting on the relative performance of these techniques.
5. The prediction performance of NIRS for these bioactive analytes is typically higher than that of MIRS.
6. NIRS and MIRS were unable to determine anthocyanin content of any crop.
7. The optimal data pre-processing methods for creating prediction models varied between the compound classes and crop types; hence should be optimised for each application.
8. Preliminary results also suggest that infrared spectroscopy may show promise for the qualitative authentication of grain crops (e.g., growing location/variety).

Data availability

The main datasets associated with this thesis are available from Mendeley Data (DOI: 10.17632/8bb2d725sd.1). These include the sample information, mean NIR/MIR spectra for each sample, the chemical reference data, and HPLC phenolic results.

References

AACC International. (1975). Method 44-15.02. Moisture—air oven methods. Approved 30 October 1975; reapproved 3 November 1999. In *Approved Methods of Analysis*. St. Paul, MN, USA: AACC International.

Abbas, O., Compère, G., Larondelle, Y., Pompeu, D., Rogez, H., & Baeten, V. (2017). Phenolic compound explorer: A mid-infrared spectroscopy database. *Vibrational Spectroscopy, 92*, 111-118. https://doi.org/10.1016/j.vibspec.2017.05.008.

Abbo, S., Berger, J., & Turner, N. C. (2003a). Evolution of cultivated chickpea: four bottlenecks limit diversity and constrain adaptation. *Functional Plant Biology, 30*(10), 1081-1087. https://doi.org/10.1071/FP03084.

Abbo, S., Berger, J., & Turner, N. C. (2003b). Viewpoint: Evolution of cultivated chickpea: four bottlenecks limit diversity and constrain adaptation. *Functional Plant Biology, 30*(10), 1081-1087. https://doi.org/10.1071/FP03084.

Abdullah, F. I., Chua, L. S., & Rahmat, Z. (2017). Prediction of C-glycosylated apigenin (vitexin) biosynthesis in *Ficus deltoidea* based on plant proteins identified by LC-MS/MS. *Frontiers in Biology, 12*(6), 448-458. https://doi.org/10.1007/s11515-017-1472-0.

Abu-Reidah, I. M., del Mar Contreras, M., Arráez-Román, D., Fernández-Gutiérrez, A., & Segura-Carretero, A. (2014). UHPLC-ESI-QTOF-MS-based metabolic profiling of *Vicia faba* L. (Fabaceae) seeds as a key strategy for characterization in foodomics. *Electrophoresis, 35*(11), 1571-1581. https://doi.org/10.1002/elps.201300646.

Achten, E., Schütz, D., Fischer, M., Fauhl-Hassek, C., Riedl, J., & Horn, B. (2019). Classification of grain maize (*Zea mays* L.) from different geographical origins with FTIR spectroscopy—a suitable analytical tool for feed authentication? *Food Analytical Methods, 12*(10), 2172-2184. https://doi.org/10.1007/s12161-019-01558-9.

Adamson, D. (2013). Can Australia become the food bowl of Asia? Brisbane, Queensland: University of Queensland.

AEGIC. (2017). Australian pulses: quality, versatility, nutrition. Retrieved from: https://www.aegic.org.au. Accessed 6 February 2019.

Agbonkonkon, N., Wojciechowski, G., Abbott, D. A., Gaucher, S. P., Yim, D. R., Thompson, A. W., & Leavell, M. D. (2021). Faster, reduced cost calibration method development methods for the analysis of fermentation product using near-infrared spectroscopy (NIRS). *Journal of Industrial Microbiology and Biotechnology, 48*(5-6). https://doi.org/10.1093/jimb/kuab033.

Agudo, A., Cabrera, L., Amiano, P., Ardanaz, E., Barricarte, A., Berenguer, T., . . . González, C. A. (2007). Fruit and vegetable intakes, dietary antioxidant nutrients, and total mortality in Spanish adults: findings from the Spanish cohort of the European Prospective Investigation into Cancer and Nutrition (EPIC-Spain). *The American Journal of Clinical Nutrition, 85*(6), 1634-1642. https://doi.org/10.1093/ajcn/85.6.1634.

Aguilera, Y., Dueñas, M., Estrella, I., Hernández, T., Benitez, V., Esteban, R. M., & Martín-Cabrejas, M. A. (2011). Phenolic profile and antioxidant capacity of chickpeas (*Cicer arietinum* L.) as affected by a dehydration process. *Plant Foods for Human Nutrition, 66*(2), 187-195. https://doi.org/10.1007/s11130-011-0230-8.

Akhtar, H. M. S., Abdin, M., Hamed, Y. S., Wang, W., Chen, G., Chen, D., . . . Zeng, X. (2019). Physicochemical, functional, structural, thermal characterization and α-amylase inhibition of polysaccharides from chickpea (*Cicer arietinum* L.) hulls. *LWT-Food Science and Technology, 113*, 108265. https://doi.org/10.1016/j.lwt.2019.108265.

Alamar, P. D., Caramês, E. T. S., Poppi, R. J., & Pallone, J. A. L. (2016). Quality evaluation of frozen guava and yellow passion fruit pulps by NIR spectroscopy and chemometrics. *Food Research International, 85*, 209-214. https://doi.org/10.1016/j.foodres.2016.04.027.

Aleixandre-Tudo, J. L., Nieuwoudt, H., Aleixandre, J. L., & du Toit, W. (2018). Chemometric compositional analysis of phenolic compounds in fermenting samples and wines using different infrared spectroscopy techniques. *Talanta, 176*, 526-536. https://doi.org/10.1016/j.talanta.2017.08.065.

Alenazi, M. M., Shafiq, M., Alsadon, A. A., Alhelal, I. M., Alhamdan, A. M., Solieman, T. H. I., . . . Saad, M. A. O. (2020). Non-destructive assessment of flesh firmness and dietary antioxidants of greenhouse-grown tomato (*Solanum lycopersicum* L.) at different fruit maturity stages. *Saudi Journal of Biological Sciences, 27*(10), 2839-2846. https://doi.org/10.1016/j.sjbs.2020.07.004.

Alexeev, A., Alexeeva, T., Enaleva, L., Tupolskikh, T., & Shumskaia, N. (2020). Prospects for the use of protein-carbohydrate complex based on mung bean seeds in the functional meat products technology. *E3S Web Conf., 175*, 08004. https://doi.org/10.1051/e3sconf/202017508004.

Almeida, M., Torrance, K., & Datta, A. (2006). Measurement of optical properties of foods in near-and mid-infrared radiation. *International Journal of Food Properties, 9*(4), 651-664. https://doi.org/10.1080/10942910600853667.

Alves-Santos, A. M., Sugizaki, C. S. A., Lima, G. C., & Naves, M. M. V. (2020). Prebiotic effect of dietary polyphenols: A systematic review. *Journal of Functional Foods, 74*, 104169. https://doi.org/10.1016/j.jff.2020.104169.

Amanah, H. Z., Joshi, R., Masithoh, R. E., Choung, M.-G., Kim, K.-H., Kim, G., & Cho, B.-K. (2020). Nondestructive measurement of anthocyanin in intact soybean seed using Fourier Transform Near-Infrared (FT-NIR) and Fourier Transform Infrared (FT-IR) spectroscopy. *Infrared Physics & Technology, 111*, 103477. https://doi.org/10.1016/j.infrared.2020.103477.

Amaral, A. L., Ferreira, E. S., Silva, M. A., Neves, V. A., & Demonte, A. (2017). The Vicilin protein (*Vigna radiata* L.) of mung bean as a functional food: Evidence of “in vitro” hypocholesterolemic activity. *Nutrition & Food Science, 47*(6), 907-916. https://doi.org/10.1108/NFS-05-2017-0089.

Ambigaipalan, P., Hoover, R., Donner, E., Liu, Q., Jaiswal, S., Chibbar, R., . . . Seetharaman, K. (2011). Structure of faba bean, black bean and pinto bean starches at different levels of granule organization and their physicochemical properties. *Food Research International, 44*(9), 2962-2974. https://doi.org/10.1016/j.foodres.2011.07.006.

Anderson, N. T., Walsh, K. B., Flynn, J. R., & Walsh, J. P. (2021). Achieving robustness across season, location and cultivar for a NIRS model for intact mango fruit dry matter content. II. Local PLS and nonlinear models. *Postharvest Biology and Technology, 171*, 111358. https://doi.org/10.1016/j.postharvbio.2020.111358.

Apak, R., Gorinstein, S., Böhm, V., Schaich, K. M., Özyürek, M., & Güçlü, K. (2013). Methods of measurement and evaluation of natural antioxidant capacity/activity (IUPAC Technical Report). *Pure and Applied Chemistry, 85*(5), 957-998. https://doi.org/10.1351/pac-rep-12-07-15

Ara, I., Turner, L., Harrison, M. T., Monjardino, M., deVoil, P., & Rodriguez, D. (2021). Application, adoption and opportunities for improving decision support systems in irrigated agriculture: A review. *Agricultural Water Management, 257*, 107161. https://doi.org/10.1016/j.agwat.2021.107161.

Armstrong, R. D., Perris, R., Munn, M., Dunsford, K., Robertson, F., Hollaway, G. J., & Leary, G. J. O. (2019). Effects of long-term rotation and tillage practice on grain yield and protein of wheat and soil fertility on a vertosol in a medium-rainfall temperate environment. *Crop and Pasture Science, 70*(1), 1-15. https://doi.org/10.1071/CP17437.

Arslan, M., Xiaobo, Z., Shi, J., Elrasheid Tahir, H., Zareef, M., Rakha, A., & Bilal, M. (2020). In situ prediction of phenolic compounds in puff dried *Ziziphus jujuba* Mill. using hand-held spectral analytical system. *Food Chemistry, 331*, 127361. https://doi.org/10.1016/j.foodchem.2020.127361.

Ash, A., Gleeson, T., Hall, M., Higgins, A., Hopwood, G., MacLeod, N., . . . Wilson, P. (2017). Irrigated agricultural development in northern Australia: Value-chain challenges and opportunities. *Agricultural Systems, 155*, 116-125. https://doi.org/10.1016/j.agsy.2017.04.010.

Australian Bureau of Statistics. (2021). Agricultural commodities, Australia. Retrieved from: https://www.abs.gov.au/statistics/industry/agriculture/agricultural-commodities-australia/latest-release. Accessed 12 Jan 2022.

Australian Mungbean Association. (2017). Industry size and value. Retrieved from: http://www.mungbean.org.au/about-us.html. Accessed 12 Mar 2020.

Aziz, A., Mahmood, T., Mahmood, Z., Shazadi, K., Mujeeb-Kazi, A., & Rasheed, A. (2018). Genotypic variation and genotype × environment interaction for yield-related traits in synthetic hexaploid wheats under a range of optimal and heat-stressed environments. *Crop Science, 58*(1), 295-303. https://doi.org/10.2135/cropsci2017.01.0035.

Baginsky, C., Peña-Neira, Á., Cáceres, A., Hernández, T., Estrella, I., Morales, H., & Pertuzé, R. (2013). Phenolic compound composition in immature seeds of fava bean (*Vicia faba* L.) varieties cultivated in Chile. *Journal of Food Composition and Analysis, 31*(1), 1-6. https://doi.org/10.1016/j.jfca.2013.02.003.

Barnes, R. B., & Bonner, L. G. (1936). The early history and the methods of infrared spectroscopy. *American Journal of Physics, 4*(4), 181-189. https://doi.org/10.1119/1.1999112.

Barreiro-Hurlé, J., Colombo, S., & Cantos-Villar, E. (2008). Is there a market for functional wines? Consumer preferences and willingness to pay for resveratrol-enriched red wine. *Food Quality and Preference, 19*(4), 360-371. https://doi.org/10.1016/j.foodqual.2007.11.004.

Barros Santos, M. C., Ribeiro da Silva Lima, L., Ramos Nascimento, F., Pimenta do Nascimento, T., Cameron, L. C., & Simões Larraz Ferreira, M. (2019). Metabolomic approach for characterization of phenolic compounds in different wheat genotypes during grain development. *Food Research International, 124*, 118-128. https://doi.org/10.1016/j.foodres.2018.08.034.

Bartosińska, E., Buszewska-Forajta, M., & Siluk, D. (2016). GC–MS and LC–MS approaches for determination of tocopherols and tocotrienols in biological and food matrices. *Journal of Pharmaceutical and Biomedical Analysis, 127*, 156-169. https://doi.org/10.1016/j.jpba.2016.02.051

Başlar, M., & Ertugay, M. F. (2011). Determination of protein and gluten quality-related parameters of wheat flour using near-infrared reflectance spectroscopy (NIRS). *Turkish Journal of Agriculture and Forestry, 35*(2), 139-144.

Bastide, N., Dartois, L., Dyevre, V., Dossus, L., Fagherazzi, G., Serafini, M., & Boutron-Ruault, M.-C. (2017). Dietary antioxidant capacity and all-cause and cause-specific mortality in the E3N/EPIC cohort study. *European Journal of Nutrition, 56*(3), 1233-1243. https://doi.org/10.1007/s00394-016-1172-6.

Baublis, A. J., Lu, C., Clydesdale, F. M., & Decker, E. A. (2000). Potential of wheat-based breakfast cereals as a source of dietary antioxidants. *Journal of the American College of Nutrition, 19*(sup3), 308S-311S. https://doi.org/10.1080/07315724.2000.10718965.

Baysal, I., Ekizoglu, M., Ertas, A., Temiz, B., Agalar, H. G., Yabanoglu-Ciftci, S., . . . Turkmenoglu, F. P. (2021). Identification of phenolic compounds by LC-MS/MS and evaluation of bioactive properties of two edible halophytes: *Limonium effusum* and *L. sinuatum*. *Molecules, 26*(13), 4040. https://doi.org/10.3390/molecules26134040.

Beć, K. B., Grabska, J., & Huck, C. W. (2021). NIR spectroscopy of natural medicines supported by novel instrumentation and methods for data analysis and interpretation. *Journal of Pharmaceutical and Biomedical Analysis, 193*, 113686. https://doi.org/10.1016/j.jpba.2020.113686.

Beć, K. B., & Huck, C. W. (2019). Breakthrough potential in near-infrared spectroscopy: spectra simulation. A review of recent developments. *Frontiers in Chemistry, 7*(48). https://doi.org/10.3389/fchem.2019.00048.
Benayad, A., Taghouti, M., Benali, A., Aboussaleh, Y., & Benbrahim, N. (2021). Nutritional and technological assessment of durum wheat-faba bean enriched flours, and sensory quality of developed composite bread. *Saudi Journal of Biological Sciences, 28*(1), 635-642. https://doi.org/10.1016/j.sjbs.2020.10.053.
Benzie, I. F., & Strain, J. J. (1996). The ferric reducing ability of plasma (FRAP) as a measure of "antioxidant power": the FRAP assay. *Analytical Biochemistry, 239*(1), 70-76. https://doi.org/10.1006/abio.1996.0292
Bershtein, V. A., & Ryzhov, V. A. (1994). Far infrared spectroscopy of polymers. In *Polymer Analysis and Characterization* (pp. 43-121). Berlin, Heidelberg: Springer Berlin Heidelberg.
Betances-Salcedo, E., Revilla, I., Vivar-Quintana, A. M., & González-Martín, M. I. (2017). Flavonoid and antioxidant capacity of propolis prediction using near infrared spectroscopy. *Sensors, 17*(7), 1647. https://doi.org/10.3390/s17071647.
Bhagyawant, S. S., Gautam, A. K., Narvekar, D. T., Gupta, N., Bhadkaria, A., Srivastava, N., & Upadhyaya, H. D. (2018). Biochemical diversity evaluation in chickpea accessions employing mini-core collection. *Physiology and Molecular Biology of Plants, 24*(6), 1165-1183. https://doi.org/10.1007/s12298-018-0579-3.
Bhatt, M., Ayyalasomayajula, K., & Yalavarthy, P. (2016). Generalized Beer–Lambert model for near-infrared light propagation in thick biological tissues. *Journal of Biomedical Optics, 21*(7), 076012. https://doi.org/10.1117/1.JBO.21.7.076012.
Biesalski, H.-K., Dragsted, L. O., Elmadfa, I., Grossklaus, R., Müller, M., Schrenk, D., . . . Weber, P. (2009). Bioactive compounds: Definition and assessment of activity. *Nutrition, 25*(11), 1202-1205. https://doi.org/10.1016/j.nut.2009.04.023.
Bochenek, H. F., Santhakumar, A. B., Francis, N., Blanchard, C. L., & Chinkwo, K. A. (2019). Anti-cancer effects of chickpea extracts. In C. Walker & J. Panozzo (Eds.), *69th Australasian Grain Science Conference*. Melbourne, Victoria: AGSA.
Bokobza, L. (1998). Near infrared spectroscopy. *Journal of Near Infrared Spectroscopy, 6*(1), 3-17. https://doi.org/10.1255/jnirs.116.
BOM. (2020). Climate data online. Retrieved from: http://www.bom.gov.au/climate/data/. Accessed 29 Oct 2020.
Boudjou, S., Oomah, B. D., Zaidi, F., & Hosseinian, F. (2013). Phenolics content and antioxidant and anti-inflammatory activities of legume fractions. *Food Chemistry, 138*(2), 1543-1550. https://doi.org/10.1016/j.foodchem.2012.11.108.
Boukid, F. (2021). Chickpea (*Cicer arietinum* L.) protein as a prospective plant-based ingredient: a review. *International Journal of Food Science & Technology*. https://doi.org/10.1111/ijfs.15046.
Boukid, F., Folloni, S., Sforza, S., Vittadini, E., & Prandi, B. (2018). Current trends in ancient grains-based foodstuffs: Insights into nutritional aspects and technological applications. *Comprehensive Reviews in Food Science and Food Safety, 17*(1), 123-136. https://doi.org/10.1111/1541-4337.12315.
Bratu, I., Cotisel, M., Damian, G., & Mocanu, A. (2007). Secondary structure analysis of barley aleurone holoprotein by FTIR spectroscopy. *Journal of Optoelectronics and Advanced Materials, 9*(3), 672.
Britz, S. J., & Kremer, D. F. (2002). Warm temperatures or drought during seed maturation increase free α-tocopherol in seeds of soybean (*Glycine max* [L.] Merr.). *Journal of Agricultural and Food Chemistry, 50*(21), 6058-6063. https://doi.org/10.1021/jf0200016.
Bureau, S., Cozzolino, D., & Clark, C. J. (2019). Contributions of Fourier-transform mid infrared (FT-MIR) spectroscopy to the study of fruit and vegetables: A review. *Postharvest Biology and Technology, 148*, 1-14. https://doi.org/10.1016/j.postharvbio.2018.10.003.

Burks, C. S., Dowell, F. E., & Xie, F. (2000). Measuring fig quality using near-infrared spectroscopy. *Journal of Stored Products Research, 36*(3), 289-296. https://doi.org/10.1016/S0022-474X(99)00050-8.
Cai, S., Han, Z., Huang, Y., Chen, Z.-H., Zhang, G., & Dai, F. (2015). Genetic diversity of individual phenolic acids in barley and their correlation with barley malt quality. *Journal of Agricultural and Food Chemistry, 63*(31), 7051-7057. https://doi.org/10.1021/acs.jafc.5b02960.
Cai, S., Han, Z., Huang, Y., Hu, H., Dai, F., & Zhang, G. (2016). Identification of quantitative trait loci for the phenolic acid contents and their association with agronomic traits in Tibetan wild barley. *Journal of Agricultural and Food Chemistry, 64*(4), 980-987. https://doi.org/10.1021/acs.jafc.5b05441.
Câmara, J. S., Albuquerque, B. R., Aguiar, J., Corrêa, R. C. G., Gonçalves, J. L., Granato, D., . . . Ferreira, I. C. F. R. (2021). Food bioactive compounds and emerging techniques for their extraction: Polyphenols as a case study. *Foods, 10*(1), 37. https://doi.org/10.3390/foods10010037.
Campos-Vega, R., & Oomah, B. (2013). Chemistry and classification of phytochemicals. In B. Tiwari, N. Brunton & C. Brennan (Eds.), *Handbook of Plant Food Phytochemicals* (pp. 5-48).
Canal, C., & Ozen, B. (2017). Monitoring of wine process and prediction of its parameters with mid-infrared spectroscopy. *Journal of Food Process Engineering, 40*(1), e12280. https://doi.org/10.1111/jfpe.12280.
Caporaso, N., ElMasry, G., & Gou, P. (2021). Chapter 13 - Hyperspectral imaging techniques for noncontact sensing of food quality. In C. M. Galanakis (Ed.), *Innovative Food Analysis* (pp. 345-379): Academic Press.
Caporaso, N., Whitworth, M. B., & Fisk, I. D. (2018). Near-Infrared spectroscopy and hyperspectral imaging for non-destructive quality assessment of cereal grains. *Applied Spectroscopy Reviews, 53*(8), 667-687. https://doi.org/10.1080/05704928.2018.1425214.
Capuano, E., & van Ruth, S. M. (2016). Infrared spectroscopy: Applications. In B. Caballero, P. M. Finglas & F. Toldrá (Eds.), *Encyclopedia of Food and Health* (pp. 424-431). Oxford: Academic Press.
Caramês, E. T. S., Alamar, P. D., Poppi, R. J., & Pallone, J. A. L. (2017a). Quality control of cashew apple and guava nectar by near infrared spectroscopy. *Journal of Food Composition and Analysis, 56*, 41-46. https://doi.org/10.1016/j.jfca.2016.12.002.
Caramês, E. T. S., Alamar, P. D., Poppi, R. J., & Pallone, J. A. L. (2017b). Rapid assessment of total phenolic and anthocyanin contents in grape juice using infrared spectroscopy and multivariate calibration. *Food Analytical Methods, 10*(5), 1609-1615. https://doi.org/10.1007/s12161-016-0721-1.
Carbas, B., Machado, N., Oppolzer, D., Queiroz, M., Brites, C., Rosa, E. A. S., & Barros, A. I. R. N. A. (2020). Prediction of phytochemical composition, in vitro antioxidant activity and individual phenolic compounds of common beans using MIR and NIR spectroscopy. *Food and Bioprocess Technology, 13*(6), 962-977. https://doi.org/10.1007/s11947-020-02457-2.
Carvalho, D. G., Ranzan, L., Trierweiler, L. F., & Trierweiler, J. O. (2020). Determination of the concentration of total phenolic compounds in aged cachaça using two-dimensional fluorescence and mid-infrared spectroscopy. *Food Chemistry, 329*, 127142. https://doi.org/10.1016/j.foodchem.2020.127142.
Cassidy, L., Fernandez, F., Johnson, J. B., Naiker, M., Owoola, A. G., & Broszczak, D. A. (2020). Oxidative stress in Alzheimer's Disease: A review on emergent natural polyphenolic therapeutics. *Complementary Therapies in Medicine, 49*, 102294. https://doi.org/10.1016/j.ctim.2019.102294.
Cayuela, J. A., & García, J. F. (2017). Sorting olive oil based on alpha-tocopherol and total tocopherol content using near-infra-red spectroscopy (NIRS) analysis. *Journal of Food Engineering, 202*, 79-88. https://doi.org/10.1016/j.jfoodeng.2017.01.015.

Cayuela, J. A., & García, J. F. (2018). Nondestructive measurement of squalene in olive oil by near infrared spectroscopy. *LWT-Food Science and Technology, 88*, 103-108. https://doi.org/10.1016/j.lwt.2017.09.047.
Chaieb, N., González, J. L., López-Mesas, M., Bouslama, M., & Valiente, M. (2011). Polyphenols content and antioxidant capacity of thirteen faba bean (*Vicia faba* L.) genotypes cultivated in Tunisia. *Food Research International, 44*(4), 970-977. https://doi.org/10.1016/j.foodres.2011.02.026.
Chauhan, Y. S., & Williams, R. (2018). Physiological and agronomic strategies to increase mungbean yield in climatically variable environments of northern Australia. *Agronomy, 8*(6), 83. https://doi.org/10.3390/agronomy8060083.
Chen, L., Tan, G. J. T., Pang, X., Yuan, W., Lai, S., & Yang, H. (2018). Energy regulated nutritive and antioxidant properties during the germination and sprouting of broccoli sprouts (*Brassica oleracea* var. *italica*). *Journal of Agricultural and Food Chemistry, 66*(27), 6975-6985. https://doi.org/10.1021/acs.jafc.8b00466.
Chen, Y.-G., Li, P., Li, P., Yan, R., Zhang, X.-Q., Wang, Y., . . . Zhang, Q.-W. (2013). α-glucosidase inhibitory effect and simultaneous quantification of three major flavonoid glycosides in *Microctis folium*. *Molecules, 18*(4), 4221-4232. https://doi.org/10.3390/molecules18044221.
Chen, Y., Cao, X., Chang, P. R., & Huneault, M. A. (2008). Comparative study on the films of poly (vinyl alcohol)/pea starch nanocrystals and poly (vinyl alcohol)/native pea starch. *Carbohydrate Polymers, 73*(1), 8-17. https://doi.org/10.1016/j.carbpol.2007.10.015.
Chirinos, R., Campos, D., Costa, N., Arbizu, C., Pedreschi, R., & Larondelle, Y. (2008). Phenolic profiles of Andean mashua (*Tropaeolum tuberosum* Ruíz & Pavón) tubers: Identification by HPLC-DAD and evaluation of their antioxidant activity. *Food Chemistry, 106*(3), 1285-1298. https://doi.org/10.1016/j.foodchem.2007.07.024.
Choo, C. Y., Sulong, N. Y., Man, F., & Wong, T. W. (2012). Vitexin and isovitexin from the leaves of *Ficus deltoidea* with in-vivo α-glucosidase inhibition. *Journal of Ethnopharmacology, 142*(3), 776-781. https://doi.org/10.1016/j.jep.2012.05.062.
Choudhary, D. K., Chaturvedi, N., Singh, A., & Mishra, A. (2020). Characterization, inhibitory activity and mechanism of polyphenols from faba bean (gallic-acid and catechin) on α-glucosidase: insights from molecular docking and simulation study. *Preparative Biochemistry & Biotechnology, 50*(2), 123-132. https://doi.org/10.1080/10826068.2019.1679171.
Clark, C. J., McGlone, V. A., & Jordan, R. B. (2003). Detection of brownheart in 'Braeburn' apple by transmission NIR spectroscopy. *Postharvest Biology and Technology, 28*(1), 87-96. https://doi.org/10.1016/S0925-5214(02)00122-9.
Clemente, A., Vioque, J., Sánchez-Vioque, R., Pedroche, J., Bautista, J., & Millán, F. (1999). Protein quality of chickpea (*Cicer arietinum* L.) protein hydrolysates. *Food Chemistry, 67*(3), 269-274. https://doi.org/10.1016/S0308-8146(99)00130-2.
Cobaleda-Velasco, M., Almaraz-Abarca, N., Alanis-Bañuelos, R. E., Uribe-Soto, J. N., González-Valdez, L. S., Muñoz-Hernández, G., . . . Rojas-López, M. (2018). Rapid determination of phenolics, flavonoids, and antioxidant properties of *Physalis ixocarpa* Brot. ex Hornem. and *Physalis angulata* l. by infrared spectroscopy and partial least squares. *Analytical Letters, 51*(4), 523-536. https://doi.org/10.1080/00032719.2017.1331238.
Colarusso, L., Serafini, M., Lagerros, Y. T., Nyren, O., La Vecchia, C., Rossi, M., . . . Bellocco, R. (2017). Dietary antioxidant capacity and risk for stroke in a prospective cohort study of Swedish men and women. *Nutrition, 33*, 234-239. https://doi.org/10.1016/j.nut.2016.07.009.
Cooper, R. (2015). Re-discovering ancient wheat varieties as functional foods. *Journal of Traditional and Complementary Medicine, 5*(3), 138-143. https://doi.org/10.1016/j.jtcme.2015.02.004.
Cortés, V., Blasco, J., Aleixos, N., Cubero, S., & Talens, P. (2019). Monitoring strategies for quality control of agricultural products using visible and near-infrared spectroscopy: A

review. *Trends in Food Science & Technology, 85*, 138-148. https://doi.org/10.1016/j.tifs.2019.01.015.
Costa, C., Tsatsakis, A., Mamoulakis, C., Teodoro, M., Briguglio, G., Caruso, E., . . . Fenga, C. (2017). Current evidence on the effect of dietary polyphenols intake on chronic diseases. *Food and Chemical Toxicology, 110*, 286-299. https://doi.org/10.1016/j.fct.2017.10.023.
Cozzolino, D. (2014a). *Infrared spectroscopy: Theory, developments and applications*: Nova Science Publishers.
Cozzolino, D. (2014b). An overview of the use of infrared spectroscopy and chemometrics in authenticity and traceability of cereals. *Food Research International, 60*, 262-265. https://doi.org/10.1016/j.foodres.2013.08.034.
Cozzolino, D. (2015). Infrared spectroscopy as a versatile analytical tool for the quantitative determination of antioxidants in agricultural products, foods and plants. *Antioxidants, 4*(3), 482-497. https://doi.org/10.3390/antiox4030482.
Cozzolino, D. (2016). 16 - Authentication of cereals and cereal products. In G. Downey (Ed.), *Advances in Food Authenticity Testing* (pp. 441-457): Woodhead Publishing.
Cozzolino, D., Corbella, E., & Smyth, H. E. (2011). Quality control of honey using infrared spectroscopy: A review. *Applied Spectroscopy Reviews, 46*(7), 523-538. https://doi.org/10.1080/05704928.2011.587857.
Cozzolino, D., Phan, A. D. T., Netzel, M. E., Smyth, H., & Sultanbawa, Y. (2020). The use of vibrational spectroscopy to predict vitamin C in Kakadu plum powders (*Terminalia ferdinandiana* Exell, Combretaceae). *Journal of the Science of Food and Agriculture, 101*(8), 3208-3213. https://doi.org/10.1002/jsfa.10950.
Cozzolino, D., Roumeliotis, S., & Eglinton, J. (2014). Evaluation of the use of attenuated total reflectance mid infrared spectroscopy to determine fatty acids in intact seeds of barley (*Hordeum vulgare*). *LWT-Food Science and Technology, 56*(2), 478-483. https://doi.org/10.1016/j.lwt.2013.11.019
CRCNA. (2020). Northern Australian broadacre cropping situational analysis. Townsville, Queensland: CRC for Northern Australia.
Cunha Júnior, L. C., Teixeira, G. H. d. A., Nardini, V., & Walsh, K. B. (2016). Quality evaluation of intact açaí and juçara fruit by means of near infrared spectroscopy. *Postharvest Biology and Technology, 112*, 64-74. https://doi.org/10.1016/j.postharvbio.2015.10.001.
Curran, J. (2012). The nutritional value and health benefits of pulses in relation to obesity, diabetes, heart disease and cancer. *British Journal of Nutrition, 108*(S1), S1-S2. https://doi.org/10.1017/S0007114512003534.
Czyczyło-Mysza, I. M., Cyganek, K., Dziurka, K., Quarrie, S., Skrzypek, E., Marcińska, I., . . . Bocianowski, J. (2019). Genetic parameters and QTLs for total phenolic content and yield of wheat mapping population of CSDH lines under drought stress. *International Journal of Molecular Sciences, 20*(23), 6064. https://doi.org/10.3390/ijms20236064.
da Silva, A. P. G., Spricigo, P. C., Purgatto, E., de Alencar, S. M., & Jacomino, A. P. (2019). *Plinia trunciflora* and *Plinia cauliflora*: two species rich in bioactive compounds, terpenes, and minerals. *Journal of Food Measurement and Characterization, 13*(2), 921-931. https://doi.org/10.1007/s11694-018-0006-z.
Dahiya, P., Linnemann, A., Van Boekel, M., Khetarpaul, N., Grewal, R., & Nout, M. (2015). Mung bean: Technological and nutritional potential. *Critical Reviews in Food Science and Nutrition, 55*(5), 670-688. https://doi.org/10.1080/10408398.2012.671202
Dalton, S. M. C., Tapsell, L. C., & Probst, Y. (2012). Potential health benefits of whole grain wheat components. *Nutrition Today, 47*(4), 163-174. https://doi.org/10.1097/NT.0b013e31826069d0.
de Camargo, A. C., Favero, B. T., Morzelle, M. C., Franchin, M., Alvarez-Parrilla, E., de la Rosa, L. A., . . . Schwember, A. R. (2019). Is chickpea a potential substitute for soybean? Phenolic bioactives and potential health benefits. *International Journal of Molecular Sciences, 20*(11), 2644. https://doi.org/10.3390/ijms20112644.

De Girolamo, A., Cortese, M., Cervellieri, S., Lippolis, V., Pascale, M., Logrieco, A. F., & Suman, M. (2019). Tracing the geographical origin of durum wheat by FT-NIR spectroscopy. *Foods, 8*(10), 450. https://doi.org/10.3390/foods8100450.

De Haro, A., López-Medina, J., Cabrera, A., & Martín, A. (1988). Determination of tannin in the seeds of *Vicia faba* by NIR. In J. Huisman, T. van der Poel & I. Liener (Eds.), *Recent Advances of Research in Antinutritional Factors in Legume Seeds. Proceedings 1st International Workshop on Antinutritional Factors in Legume Seeds* (pp. 172-175). Wageningen: Wageningen Academic Publishers.

De Marchi, M., Penasa, M., Zidi, A., & Manuelian, C. L. (2018). Invited review: Use of infrared technologies for the assessment of dairy products—Applications and perspectives. *Journal of Dairy Science, 101*(12), 10589-10604. https://doi.org/10.3168/jds.2018-15202.

de Oliveira, G. A., Bureau, S., Renard, C. M.-G. C., Pereira-Netto, A. B., & de Castilhos, F. (2014). Comparison of NIRS approach for prediction of internal quality traits in three fruit species. *Food Chemistry, 143*, 223-230. https://doi.org/10.1016/j.foodchem.2013.07.122.

de Oliveira, I. R. N., Roque, J. V., Maia, M. P., Stringheta, P. C., & Teófilo, R. F. (2018). New strategy for determination of anthocyanins, polyphenols and antioxidant capacity of *Brassica oleracea* liquid extract using infrared spectroscopies and multivariate regression. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 194*, 172-180. https://doi.org/10.1016/j.saa.2018.01.006.

Del Rio, D., Agnoli, C., Pellegrini, N., Krogh, V., Brighenti, F., Mazzeo, T., . . . Panico, S. (2010). Total antioxidant capacity of the diet is associated with lower risk of ischemic stroke in a large Italian cohort. *The Journal of Nutrition, 141*(1), 118-123. https://doi.org/10.3945/jn.110.125120.

Detopoulou, P., Panagiotakos, D. B., Chrysohoou, C., Fragopoulou, E., Nomikos, T., Antonopoulou, S., . . . Stefanadis, C. (2010). Dietary antioxidant capacity and concentration of adiponectin in apparently healthy adults: the ATTICA study. *European Journal of Clinical Nutrition, 64*(2), 161-168. https://doi.org/10.1038/ejcn.2009.130.

Di Pasquale, J., Adinolfi, F., & Capitanio, F. (2011). Analysis of consumer attitudes and consumers' willingness to pay for functional foods. *International Journal on Food System Dynamics, 2*(2), 181-193. https://doi.org/10.18461/ijfsd.v2i2.227.

Diago, M. P., Fernández-Novales, J., Fernandes, A. M., Melo-Pinto, P., & Tardaguila, J. (2016). Use of visible and short-wave near-infrared hyperspectral imaging to fingerprint anthocyanins in intact grape berries. *Journal of Agricultural and Food Chemistry, 64*(40), 7658-7666. https://doi.org/10.1021/acs.jafc.6b01999.

Ding, X., Guo, Y., Ni, Y., & Kokot, S. (2016). A novel NIR spectroscopic method for rapid analyses of lycopene, total acid, sugar, phenols and antioxidant activity in dehydrated tomato samples. *Vibrational Spectroscopy, 82*, 1-9. https://doi.org/10.1016/j.vibspec.2015.10.004.

Domínguez-Arispuro, D. M., Cuevas-Rodríguez, E. O., Milán-Carrillo, J., León-López, L., Gutiérrez-Dorado, R., & Reyes-Moreno, C. (2018). Optimal germination condition impacts on the antioxidant activity and phenolic acids profile in pigmented desi chickpea (*Cicer arietinum* L.) seeds. *Journal of Food Science and Technology, 55*(2), 638-647. https://doi.org/10.1007/s13197-017-2973-1.

Dong, Y., Dong, K., Yang, Z. X., Zheng, Y., & Tang, L. (2016). [Microbial and physiological mechanisms for alleviating fusarium wilt of faba bean in intercropping system]. *Ying Yong Sheng Tai Xue Bao = The Journal of Applied Ecology, 27*(6), 1984-1992. https://doi.org/10.13287/j.1001-9332.201606.019.

Dotto, A. C., Dalmolin, R. S. D., ten Caten, A., & Grunwald, S. (2018). A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. *Geoderma, 314*, 262-274. https://doi.org/10.1016/j.geoderma.2017.11.006.

Dufour, E. (2009). Principles of infrared spectroscopy. In D.-W. Sun (Ed.), *Infrared Spectroscopy for Food Quality Analysis and Control* (pp. 1-27). Burlington, MA: Academic Press.
Dull, G. G., Birth, G. S., Smittle, D. A., & Leffler, R. G. (1989). Near infrared analysis of soluble solids in intact cantaloupe. *Journal of Food Science: an official publication of the Institute of Food Technologists (USA), 54*(2), 393-395.
Dwivedi, S. K., Basu, S., Kumar, S., Kumari, S., Kumar, A., Jha, S., . . . Kumar, G. (2019). Enhanced antioxidant enzyme activities in developing anther contributes to heat stress alleviation and sustains grain yield in wheat. *Functional Plant Biology, 46*(12), 1090-1102. https://doi.org/10.1071/FP19016.
Ed Nignpense, B., Chinkwo, K. A., Blanchard, C. L., & Santhakumar, A. B. (2020). Polyphenols: Modulators of platelet function and platelet microparticle generation? *International Journal of Molecular Sciences, 21*(1), 146. https://doi.org/10.3390/ijms21010146.
El-Akkad, S. S., Hassan, E. A., & Ali, M. E. (2002). Phenolic acid changes during *Orobanche* parasitism on faba bean and some other hosts. *Egyptian Journal of Biology, 4*, 37-44.
El-Sherbeeny, M. H., & Robertson, L. D. (1992). Protein content variation in a pure line faba bean (*Vicia faba*) collection. *Journal of the Science of Food and Agriculture, 58*(2), 193-196. https://doi.org/10.1002/jsfa.2740580206.
Ercan, P., & El, S. N. (2016). Inhibitory effects of chickpea and *Tribulus terrestris* on lipase, α-amylase and α-glucosidase. *Food Chemistry, 205*, 163-169. https://doi.org/10.1016/j.foodchem.2016.03.012.
Escarpa, A., & González, M. C. (2000). Optimization strategy and validation of one chromatographic method as approach to determine the phenolic compounds from different sources. *Journal of Chromatography A, 897*(1), 161-170. https://doi.org/10.1016/S0021-9673(00)00817-7.
Fanning, K., Edwards, D., Netzel, M., Stanley, R., Netzel, G., Russell, D., & Topp, B. (2013). Increasing anthocyanin content in Queen Garnet plum and correlations with in-field measures. In T. DeJong & C. DeBuse (Eds.), *X International Symposium on Plum and Prune Genetics, Breeding and Pomology* (Vol. 985, pp. 97-104). Davis, California: Acta Horticulturae.
FAO. (2022). FAOSTAT. Retrieved from: http://www.fao.org/faostat/en/#data/QC. Accessed 12 Jan 2022.
Faridy, J.-C. M., Stephanie, C.-G. M., Gabriela, M.-M. O., & Cristian, J.-M. (2020). Biological activities of chickpea in human health (*Cicer arietinum* L.). A review. *Plant Foods for Human Nutrition, 75*(2), 142-153. https://doi.org/10.1007/s11130-020-00814-2.
Feng, L., Wu, B., Zhu, S., He, Y., & Zhang, C. (2021). Application of visible/infrared spectroscopy and hyperspectral imaging with machine learning techniques for identifying food varieties and geographical origins. *Frontiers in Nutrition, 8*(332). https://doi.org/10.3389/fnut.2021.680357.
Ferrer-Gallego, R., Rodríguez-Pulido, F. J., Toci, A. T., & García-Estevez, I. (2020). Phenolic composition, quality and authenticity of grapes and wines by vibrational spectroscopy. *Food Reviews International*, 1-29. https://doi.org/10.1080/87559129.2020.1752231.
Food Standards Australia New Zealand. (2010). Proposal P 1004 – Primary production & processing standard for seed sprouts. Canberra, ACT: Food Standards Australia New Zealand.
Fraga, C. G., Oteiza, P. I., & Galleano, M. (2014). In vitro measurements and interpretation of total antioxidant capacity. *Biochimica et Biophysica Acta (BBA) - General Subjects, 1840*(2), 931-934. https://doi.org/10.1016/j.bbagen.2013.06.030.
Fragoso, S., Aceña, L., Guasch, J., Busto, O., & Mestres, M. (2011a). Application of FT-MIR spectroscopy for fast control of red grape phenolic ripening. *Journal of Agricultural and Food Chemistry, 59*(6), 2175-2183. https://doi.org/10.1021/jf104039g.
Fragoso, S., Aceña, L., Guasch, J., Mestres, M., & Busto, O. (2011b). Quantification of phenolic compounds during red winemaking using FT-MIR spectroscopy and PLS-

Regression. *Journal of Agricultural and Food Chemistry, 59*(20), 10795-10802. https://doi.org/10.1021/jf201973e.
Fraser, C. M., & Chapple, C. (2011). The penylpropanoid pathway in *Arabidopsis*. *The Arabidopsis Book, 2011*(9). https://doi.org/10.1199/tab.0152.
Fraser, D. G., Jordan, R. B., Künnemeyer, R., & McGlone, V. A. (2003). Light distribution inside mandarin fruit during internal quality assessment by NIR spectroscopy. *Postharvest Biology and Technology, 27*(2), 185-196. https://doi.org/10.1016/S0925-5214(02)00058-3.
Fu, Y., Zu, Y., Liu, W., Zhang, L., Tong, M., Efferth, T., . . . Chen, L. (2008). Determination of vitexin and isovitexin in pigeonpea using ultrasonic extraction followed by LC-MS. *Journal of Separation Science, 31*(2), 268-275. https://doi.org/10.1002/jssc.200700312.
Furbank, R. T., & Tester, M. (2011). Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Science, 16*(12), 635-644. https://doi.org/10.1016/j.tplants.2011.09.005.
Gabriëls, S. H. E. J., Mishra, P., Mensink, M. G. J., Spoelstra, P., & Woltering, E. J. (2020). Non-destructive measurement of internal browning in mangoes using visible and near-infrared spectroscopy supported by artificial neural network analysis. *Postharvest Biology and Technology, 166*, 111206. https://doi.org/10.1016/j.postharvbio.2020.111206.
Ganesan, K., & Xu, B. (2018). A critical review on phytochemical profile and health promoting effects of mung bean (*Vigna radiata*). *Food Science and Human Wellness, 7*(1), 11-33. https://doi.org/10.1016/j.fshw.2017.11.002.
Gardana, C., Scialpi, A., Fachechi, C., & Simonetti, P. (2018). Near-infrared spectroscopy and chemometrics for the routine detection of bilberry extract adulteration and quantitative determination of the anthocyanins. *Journal of Spectroscopy, 2018*, 4751247. https://doi.org/10.1155/2018/4751247.
Gautam, R., Vanga, S., Ariese, F., & Umapathy, S. (2015). Review of multidimensional data processing approaches for Raman and infrared spectroscopy. *EPJ Techniques and Instrumentation, 2*(1), 8. https://doi.org/10.1140/epjti/s40485-015-0018-6.
Gergely, S., & Salgó, A. (2007). Changes in protein content during wheat maturation—what is measured by near infrared spectroscopy? *Journal of Near Infrared Spectroscopy, 15*(1), 49-58. https://doi.org/10.1255/jnirs.687.
Giusti, M. M., & Wrolstad, R. E. (2001). Characterization and measurement of anthocyanins by UV-visible spectroscopy. *Current Protocols in Food Analytical Chemistry, 00*(1), F1.2.1-F1.2.13. https://doi.org/10.1002/0471142913.faf0102s00.
González Arbeláez, L. F., Ciocci Pardo, A., Fantinelli, J. C., Schinella, G. R., Mosca, S. M., & Ríos, J.-L. (2018). Cardioprotection and natural polyphenols: an update of clinical and experimental studies. *Food & Function, 9*(12), 6129-6145. https://doi.org/10.1039/C8FO01307A.
Gordeeva, E., Shamanin, V., Shoeva, O., Kukoeva, T., Morgounov, A., & Khlestkina, E. (2020). The strategy for marker-assisted breeding of anthocyanin-rich spring bread wheat (*Triticum aestivum* L.) cultivars in Western Siberia. *Agronomy, 10*(10), 1603. https://doi.org/10.3390/agronomy10101603.
Gordon, R., Chapman, J., Power, A., Chandra, S., Roberts, J., & Cozzolino, D. (2018). Unfrazzled by fizziness: Identification of beers using attenuated total reflectance mid-infrared spectroscopy and multivariate analysis. *Food Analytical Methods, 11*(9), 2360-2367. https://doi.org/10.1007/s12161-018-1225-y.
Gordon, R., Chapman, J., Power, A., Chandra, S., Roberts, J., & Cozzolino, D. (2019). Mid-infrared spectroscopy coupled with chemometrics to identify spectral variability in Australian barley samples from different production regions. *Journal of Cereal Science, 85*, 41-47. https://doi.org/10.1016/j.jcs.2018.11.004
Granato, D., Nunes, D. S., & Barba, F. J. (2017). An integrated strategy between food chemistry, biology, nutrition, pharmacology, and statistics in the development of

functional foods: A proposal. *Trends in Food Science & Technology, 62*, 13-22. https://doi.org/10.1016/j.tifs.2016.12.010.
Guaadaoui, A., Benaicha, S., Elmajdoub, N., Bellaoui, M., & Hamal, A. (2014). What is a bioactive compound? A combined definition for a preliminary consensus. *International Journal of Nutrition and Food Sciences, 3*(3), 174-179. https://doi.org/10.11648/j.ijnfs.20140303.16
Gunasekera, D., Tulloh, C., Ford, M., & Heyhoe, E. (2008). Climate change: Opportunities and challenges in Australian agriculture. *Proceedings of Faculty of Agriculture, Food and Natural Resources Annual Symposium* (Vol. 13). University of Sydney: Citeseer.
Gupta, N., & Bhagyawant, S. S. (2019). Enzymatic treatment improves ACE-I inhibiton and antiproliferative potential of chickpea. *Vegetos, 32*(3), 363-369. https://doi.org/10.1007/s42535-019-00031-6.
Gupta, N., Bisen, P. S., & Bhagyawant, S. S. (2018). Chickpea lectin inhibits human breast cancer cell proliferation and induces apoptosis through cell cycle arrest. *Protein and Peptide Letters, 25*(5), 492-499. https://doi.org/10.2174/0929866525666180406142900.
Hafidh, R. R., Abdulamir, A. S., Bakar, F. A., Sekawi, Z., Jahansheri, F., & Jalilian, F. A. (2015). Novel antiviral activity of mung bean sprouts against respiratory syncytial virus and herpes simplex virus− 1: an in vitro study on virally infected Vero and MRC-5 cell lines. *BMC Complementary and Alternative Medicine, 15*(1), 179. https://doi.org/10.1186/s12906-015-0688-2.
Hamzalıoğlu, A., & Gökmen, V. (2016). Chapter 18 - Interaction between bioactive carbonyl compounds and asparagine and impact on acrylamide. In V. Gökmen (Ed.), *Acrylamide in Food* (pp. 355-376): Academic Press.
Han, Y., Ling, S., Qi, Z., Shao, Z., & Chen, X. (2018). Application of far-infrared spectroscopy to the structural identification of protein materials. *Physical Chemistry Chemical Physics, 20*(17), 11643-11648. https://doi.org/10.1039/C8CP00802G.
Han, Z., Ahsan, M., Adil, M. F., Chen, X., Nazir, M. M., Shamsi, I. H., . . . Zhang, G. (2020). Identification of the gene network modules highly associated with the synthesis of phenolics compounds in barley by transcriptome and metabolome analysis. *Food Chemistry, 323*, 126862. https://doi.org/10.1016/j.foodchem.2020.126862.
He, M., Min, J.-W., Kong, W.-L., He, X.-H., Li, J.-X., & Peng, B.-W. (2016). A review on the pharmacological effects of vitexin and isovitexin. *Fitoterapia, 115*, 74-85. https://doi.org/10.1016/j.fitote.2016.09.011.
Heiras-Palazuelos, M. J., Ochoa-Lugo, M. I., Gutiérrez-Dorado, R., López-Valenzuela, J. A., Mora-Rochín, S., Milán-Carrillo, J., . . . Reyes-Moreno, C. (2013). Technological properties, antioxidant activity and total phenolic and flavonoid content of pigmented chickpea (*Cicer arietinum* L.) cultivars. *International Journal of Food Sciences and Nutrition, 64*(1), 69-76. https://doi.org/10.3109/09637486.2012.694854.
Henzell, T. (2007). *Australian agriculture: its history and challenges*: Csiro Publishing.
Hermsdorff, H. H. M., Puchau, B., Volp, A. C. P., Barbosa, K. B. F., Bressan, J., Zulet, M. Á., & Martínez, J. A. (2011). Dietary total antioxidant capacity is inversely related to central adiposity as well as to metabolic and oxidative stress markers in healthy young adults. *Nutrition & Metabolism, 8*(1), 59. https://doi.org/10.1186/1743-7075-8-59.
Hernández-Hernández, C., Fernández-Cabanás, V. M., Rodríguez-Gutiérrez, G., Bermúdez-Oria, A., & Morales-Sillero, A. (2021). Viability of near infrared spectroscopy for a rapid analysis of the bioactive compounds in intact cocoa bean husk. *Food Control, 120*, 107526. https://doi.org/10.1016/j.foodcont.2020.107526.
Herschel, W. (1800). XIII. Investigation of the powers of the prismatic colours to heat and illuminate objects; with remarks, that prove the different refrangibility of radiant heat. To which is added, an inquiry into the method of viewing the sun advantageously, with telescopes of large apertures and high magnifying powers. *Philosophical Transactions of the Royal Society of London, 90*, 255-283. https://doi.org/10.1098/rstl.1800.0014.

Hirogaki, M. (2013). Estimating consumers' willingness to pay for health food claims: A conjoint analysis. *International Journal of Innovation, Management and Technology, 4*(6), 541. https://doi.org/10.7763/IJIMT.2013.V4.458.
Hirri, A., Bassbasi, M., Souhassou, S., Kzaiber, F., & Oussama, A. (2016). Prediction of polyphenol fraction in virgin olive oil using mid-infrared attenuated total reflectance attenuated total reflectance accessory–mid-infrared coupled with partial least squares regression. *International Journal of Food Properties, 19*(7), 1504-1512. https://doi.org/10.1080/10942912.2015.1059854.
Hoffmann, F., & De Paola, R. (1984). Anomalous CO bond weakening of side-on-bonded carbon monoxide on a potassium-promoted Ru (001) surface. *Physical Review Letters, 52*(19), 1697. https://doi.org/10.1103/PhysRevLett.52.1697.
Hou, D., Yousaf, L., Xue, Y., Hu, J., Wu, J., Hu, X., . . . Shen, Q. (2019). Mung bean (*Vigna radiata* L.): Bioactive polyphenols, polysaccharides, peptides, and health benefits. *Nutrients, 11*(6), 1238. https://doi.org/10.3390/nu11061238.
Hou, D., Zhao, Q., Yousaf, L., Khan, J., Xue, Y., & Shen, Q. (2020). Consumption of mung bean (*Vigna radiata* L.) attenuates obesity, ameliorates lipid metabolic disorders and modifies the gut microbiota composition in mice fed a high-fat diet. *Journal of Functional Foods, 64*, 103687. https://doi.org/10.1016/j.jff.2019.103687.
Hsu, G.-S. W., Lu, Y.-F., Chang, S.-H., & Hsu, S.-Y. (2011). Antihypertensive effect of mung bean sprout extracts in spontaneously hypertensive rats. *Journal of Food Biochemistry, 35*(1), 278-288. https://doi.org/10.1111/j.1745-4514.2010.00381.x.
Hu, Y., Pan, Z. J., Liao, W., Li, J., Gruget, P., Kitts, D. D., & Lu, X. (2016). Determination of antioxidant capacity and phenolic content of chocolate by attenuated total reflectance-Fourier transformed-infrared spectroscopy. *Food Chemistry, 202*, 254-261. https://doi.org/10.1016/j.foodchem.2016.01.130.
Huang, H., Yu, H., Xu, H., & Ying, Y. (2008). Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review. *Journal of Food Engineering, 87*(3), 303-313. https://doi.org/10.1016/j.jfoodeng.2007.12.022.
Huang, Y., Xiao, D., Burton-Freeman, B. M., & Edirisinghe, I. (2016). Chemical changes of bioactive phytochemicals during thermal processing. In *Reference Module in Food Science*: Elsevier.
Hung, P. V., & Morita, N. (2008). Distribution of phenolic compounds in the graded flours milled from whole buckwheat grains and their antioxidant capacities. *Food Chemistry, 109*(2), 325-331. https://doi.org/10.1016/j.foodchem.2007.12.060.
Ignat, I., Volf, I., & Popa, V. I. (2011). A critical review of methods for characterisation of polyphenolic compounds in fruits and vegetables. *Food Chemistry, 126*(4), 1821-1835. https://doi.org/10.1016/j.foodchem.2010.12.026.
Igne, B., Gibson, L. R., Rippke, G. R., Schwarte, A., & Hurburgh Jr., C. R. (2007). Triticale moisture and protein content prediction by near-infrared spectroscopy (NIRS). *Cereal Chemistry, 84*(4), 328-330. https://doi.org/10.1094/CCHEM-84-4-0328.
Iqbal, M., Navabi, A., Salmon, D. F., Yang, R. C., & Spaner, D. (2007). Simultaneous selection for early maturity, increased grain yield and elevated grain protein content in spring wheat. *Plant Breeding, 126*(3), 244-250. https://doi.org/10.1111/j.1439-0523.2007.01346.x.
Jakubczyk, A., Karaś, M., Złotek, U., Szymanowska, U., Baraniak, B., & Bochnak, J. (2019). Peptides obtained from fermented faba bean seeds (*Vicia faba*) as potential inhibitors of an enzyme involved in the pathogenesis of metabolic syndrome. *LWT-Food Science and Technology, 105*, 306-313. https://doi.org/10.1016/j.lwt.2019.02.009.
Jang, D., Jung, Y. S., Kim, M.-S., Oh, S. E., Nam, T. G., & Kim, D.-O. (2019). Developing and validating a method for separating flavonoid isomers in common buckwheat sprouts using HPLC-PDA. *Foods, 8*(11), 549. https://doi.org/10.3390/foods8110549.
Jara-Palacios, M. J., Rodríguez-Pulido, F. J., Hernanz, D., Escudero-Gilete, M. L., & Heredia, F. J. (2016). Determination of phenolic substances of seeds, skins and stems from white grape marc by near-infrared hyperspectral imaging. *Australian*

Journal of Grape and Wine Research, 22(1), 11-15. https://doi.org/10.1111/ajgw.12165.
Jha, S. N. (2010). Near Infrared Spectroscopy. In S. N. Jha (Ed.), *Nondestructive evaluation of food quality: Theory and practice* (pp. 141-212). Berlin, Germany: Springer Berlin Heidelberg.
Ji, Y., Yang, X., Ji, Z., Zhu, L., Ma, N., Chen, D., . . . Cao, Y. (2020). DFT-calculated IR spectrum amide I, II, and III band contributions of N-methylacetamide fine components. *ACS Omega, 5*(15), 8572-8578. https://doi.org/10.1021/acsomega.9b04421.
Johnson, J., Collins, T., Power, A., Chandra, S., Portman, D., Blanchard, C., & Naiker, M. (2020a). Antioxidative properties and macrochemical composition of five commercial mungbean varieties in Australia. *Legume Science, 2*(1), e27. https://doi.org/10.1002/leg3.27.
Johnson, J., Collins, T., Skylas, D., & Naiker, M. (2019). ATR-MIR: A valuable tool for the rapid assessment of biochemically active compounds in grains. In C. Walker & J. Panozzo (Eds.), *69th Australasian Grain Science Conference* (pp. 73-79). Melbourne, Australia: AGSA.
Johnson, J., Collins, T., Skylas, D., Quail, K., Blanchard, C., & Naiker, M. (2020b). Profiling the varietal antioxidative content and macrochemical composition in Australian faba beans (*Vicia faba* L.). *Legume Science, 2*(2), e28. https://doi.org/10.1002/leg3.28.
Johnson, J., Collins, T., Walsh, K., & Naiker, M. (2020c). Solvent extractions and spectrophotometric protocols for measuring the total anthocyanin, phenols and antioxidant content in plums. *Chemical Papers, 74*(12), 4481-4492. https://doi.org/10.1007/s11696-020-01261-8.
Johnson, J., Mani, J., Ashwath, N., & Naiker, M. (2020d). Potential for Fourier transform infrared (FTIR) spectroscopy toward predicting antioxidant and phenolic contents in powdered plant matrices. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 118228. https://doi.org/10.1016/j.saa.2020.118228.
Johnson, J. B. (2020). An overview of near-infrared spectroscopy (NIRS) for the detection of insect pests in stored grains. *Journal of Stored Products Research, 86*, 101558. https://doi.org/10.1016/j.jspr.2019.101558.
Johnson, J. B., Broszczak, D. A., Mani, J. S., Anesi, J., & Naiker, M. (2021a). A cut above the rest: oxidative stress in chronic wounds and the potential role of polyphenols as therapeutics. *Journal of Pharmacy and Pharmacology*. https://doi.org/10.1093/jpp/rgab038.
Johnson, J. B., Mani, J. S., & Naiker, M. (2021b). Development and validation of a 96-well microplate assay for the measurement of total phenolic content in ginger extracts. *Food Analytical Methods*, 1-8. https://doi.org/10.1007/s12161-021-02127-9.
Johnson, J. B., & Naiker, M. (2019). Seeing red: A review of the use of near-infrared spectroscopy (NIRS) in entomology. *Applied Spectroscopy Reviews*, 1-30. https://doi.org/10.1080/05704928.2019.1685532.
Johnson, J. B., Walsh, K., & Naiker, M. (2020e). Application of infrared spectroscopy for the prediction of nutritional content and quality assessment of faba bean (*Vicia faba* L.). *Legume Science, 2*(3), e40. https://doi.org/10.1002/leg3.40.
Johnson, J. B., Walsh, K. B., Bhattarai, S. P., & Naiker, M. (2021c). Partitioning of nutritional and bioactive compounds between the kernel, hull and husk of five new chickpea genotypes grown in Australia. *Future Foods*, 100065. https://doi.org/10.1016/j.fufo.2021.100065.
Johnson, J. B., Walsh, K. B., Mani, J. S., Bhattarai, S., & Naiker, M. (2020f). More than protein: The potential for rapid assessment of bioactive compounds in Australian crops. *Developing Northern Australia Conference*. Rockhampton, Australia.
Johnson, J. B., Walsh, K. B., Mani, J. S., Hoyos, E., Bhattarai, S., & Naiker, M. (2021d). Healthy, wealthy and wise: how can the North capitalise on the emerging functional food market? *Developing Northern Australia Conference*. Online.

Kadiroğlu, P., Aydemir, L. Y., & Akcakaya, F. G. (2018). Prediction of functional properties of registered chickpea samples using FT-IR spectroscopy and chemometrics. *LWT-Food Science and Technology, 93*, 463-469. https://doi.org/10.1016/j.lwt.2018.03.080.
Kalra, E. K. (2003). Nutraceutical - definition and introduction. *AAPS PharmSci, 5*(3), 27-28. https://doi.org/10.1208/ps050325.
Kalra, Y. (1997). *Handbook of reference methods for plant analysis*. Boca Raton, Florida: CRC Press.
Karoui, R., Downey, G., & Blecker, C. (2010). Mid-infrared spectroscopy coupled with chemometrics: A tool for the analysis of intact food systems and the exploration of their molecular structure−quality relationships − A review. *Chemical Reviews, 110*(10), 6144-6168. https://doi.org/10.1021/cr100090k.
Kaur, N., Singh, B., Kaur, A., Yadav, M. P., Singh, N., Ahlawat, A. K., & Singh, A. M. (2021). Effect of growing conditions on proximate, mineral, amino acid, phenolic composition and antioxidant properties of wheatgrass from different wheat (*Triticum aestivum* L.) varieties. *Food Chemistry, 341*, 128201. https://doi.org/10.1016/j.foodchem.2020.128201.
Kaur, R., & Prasad, K. (2021). Technological, processing and nutritional aspects of chickpea (*Cicer arietinum*) - A review. *Trends in Food Science & Technology, 109*, 448-463. https://doi.org/10.1016/j.tifs.2021.01.044.
Kawano, S. (2008). Sampling and sample presentation. In H. W. Siesler, Y. Ozaki, S. Kawata & H. M. Heise (Eds.), *Near-Infrared Spectroscopy: Principles, Instruments, Applications*. Weinheim, Germany: John Wiley & Sons.
Kays, S. E., Barton, F. E., & Windham, W. R. (2000). Predicting protein content by near infrared reflectance spectroscopy in diverse cereal food products. *Journal of Near Infrared Spectroscopy, 8*(1), 35-43. https://doi.org/10.1255/jnirs.262.
Khan, A., Munir, M. T., Yu, W., & Young, B. R. (2020). A review towards hyperspectral imaging for real-time quality control of food products with an illustrative case study of milk powder production. *Food and Bioprocess Technology, 13*(5), 739-752. https://doi.org/10.1007/s11947-020-02433-w.
Kiani, S., van Ruth, S. M., Minaei, S., & Ghasemi-Varnamkhasti, M. (2018). Hyperspectral imaging, a non-destructive technique in medicinal and aromatic plant products industry: Current status and potential future applications. *Computers and Electronics in Agriculture, 152*, 9-18. https://doi.org/10.1016/j.compag.2018.06.025.
Kim, J.-K., Kim, E.-H., Lee, O.-K., Park, S.-Y., Lee, B., Kim, S.-H., . . . Chung, I.-M. (2013). Variation and correlation analysis of phenolic compounds in mungbean (*Vigna radiata* L.) varieties. *Food Chemistry, 141*(3), 2988-2997. https://doi.org/10.1016/j.foodchem.2013.05.060.
Klejdus, B., Vacek, J., Benešová, L., Kopecký, J., Lapčík, O., & Kubáň, V. (2007). Rapid-resolution HPLC with spectrometric detection for the determination and identification of isoflavones in soy preparations and plant extracts. *Analytical and Bioanalytical Chemistry, 389*(7), 2277-2285. https://doi.org/10.1007/s00216-007-1606-3.
Kobayashi, S., Murakami, K., Sasaki, S., Uenishi, K., Yamasaki, M., Hayabuchi, H., . . . Sugiyamama, Y. (2012). Dietary total antioxidant capacity from different assays in relation to serum C-reactive protein among young Japanese women. *Nutrition Journal, 11*(1), 91. https://doi.org/10.1186/1475-2891-11-91.
Koch, W. (2019). Dietary polyphenols—important non-nutrients in the prevention of chronic noncommunicable diseases. A systematic review. *Nutrients, 11*(5), 1039. https://doi.org/10.3390/nu11051039.
Kokalj Ladan, M., Straus, J., Tavčar Benković, E., & Kreft, S. (2017). FT-IR-based method for rutin, quercetin and quercitrin quantification in different buckwheat (*Fagopyrum*) species. *Scientific Reports, 7*(1), 7226. https://doi.org/10.1038/s41598-017-07665-z.
KPMG. (2019). North Queensland market and agricultural supply chain study. Townsville, Queensland: Townsville Enterprise Limited.

Kukula-Koch, W., Koch, W., Czernicka, L., Głowniak, K., Asakawa, Y., Umeyama, A., . . . Kuzuhara, T. (2018). MAO-A inhibitory potential of terpene constituents from ginger rhizomes—A bioactivity guided fractionation. *Molecules, 23*(6), 1301.

Kumar, S., Singh, R., & Dhanani, T. (2017). Rapid estimation of bioactive phytochemicals in vegetables and fruits using near infrared reflectance spectroscopy. In E. M. Yahia (Ed.), *Fruit and Vegetable Phytochemicals* (pp. 781-802).

Kwiatkowski, M., Kravchuk, O., Skouroumounis, G. K., & Taylor, D. K. (2020). Response surface parallel optimization of extraction of total phenolics from separate white and red grape skin mixtures with microwave-assisted and conventional thermal methods. *Journal of Cleaner Production, 251*, 119563. https://doi.org/10.1016/j.jclepro.2019.119563.

Laddomada, B., Durante, M., Mangini, G., D'Amico, L., Lenucci, M. S., Simeone, R., . . . Blanco, A. (2017). Genetic variation for phenolic acids concentration and composition in a tetraploid wheat (*Triticum turgidum* L.) collection. *Genetic Resources and Crop Evolution, 64*(3), 587-597. https://doi.org/10.1007/s10722-016-0386-z.

Le, B. T. (2020). Application of deep learning and near infrared spectroscopy in cereal analysis. *Vibrational Spectroscopy, 106*, 103009. https://doi.org/10.1016/j.vibspec.2019.103009.

Lee, L. C., Liong, C.-Y., & Jemain, A. A. (2017). A contemporary review on Data Preprocessing (DP) practice strategy in ATR-FTIR spectrum. *Chemometrics and Intelligent Laboratory Systems, 163*, 64-75. https://doi.org/10.1016/j.chemolab.2017.02.008.

Li, E., Hasjim, J., Singh, V., Tizzotti, M., Godwin, I. D., & Gilbert, R. G. (2013). Insights into sorghum starch biosynthesis from structure changes induced by different growth temperatures. *Cereal Chemistry, 90*(3), 223-230. https://doi.org/10.1094/CCHEM-09-12-0113-R.

Li, L., Shewry, P. R., & Ward, J. L. (2008). Phenolic acids in wheat varieties in the HEALTHGRAIN diversity screen. *Journal of Agricultural and Food Chemistry, 56*(21), 9732-9739. https://doi.org/10.1021/jf801069s.

Li, M.-W., Muñoz, N. B., Wong, C.-F., Wong, F.-L., Wong, K.-S., Wong, J. W.-H., . . . Lam, H.-M. (2016). QTLs regulating the contents of antioxidants, phenolics, and flavonoids in soybean seeds share a common genomic region. *Frontiers in Plant Science, 7*. https://doi.org/10.3389/fpls.2016.00854.

Li, X., Park, N. I., Xu, H., Woo, S.-H., Park, C. H., & Park, S. U. (2010). Differential epression of flavonoid biosynthesis genes and accumulation of phenolic compounds in common buckwheat (*Fagopyrum esculentum*). *Journal of Agricultural and Food Chemistry, 58*(23), 12176-12181. https://doi.org/10.1021/jf103310g.

Liu, D., Zeng, X.-A., & Sun, D.-W. (2013). NIR spectroscopy and imaging techniques for evaluation of fish quality—A review. *Applied Spectroscopy Reviews, 48*(8), 609-628. https://doi.org/10.1080/05704928.2013.775579.

Liu, H., Bruce, D. R., Sissons, M., Able, A. J., & Able, J. A. (2018a). Genotype-dependent changes in the phenolic content of durum under water-deficit stress. *Cereal Chemistry, 95*(1), 59-78. https://doi.org/10.1002/cche.10007.

Liu, M., Zhu, K., Yao, Y., Chen, Y., Guo, H., Ren, G., . . . Li, J. (2020a). Antioxidant, anti-inflammatory, and antitumor activities of phenolic compounds from white, red, and black *Chenopodium quinoa* seed. *Cereal Chemistry, 97*(3), 703-713. https://doi.org/10.1002/cche.10286.

Liu, Q., Qiu, Y., & Beta, T. (2010). Comparison of antioxidant activities of different colored wheat grains and analysis of phenolic compounds. *Journal of Agricultural and Food Chemistry, 58*(16), 9235-9241. https://doi.org/10.1021/jf101700s.

Liu, W., Fu, Y., Zu, Y., Kong, Y., Zhang, L., Zu, B., & Efferth, T. (2009). Negative-pressure cavitation extraction for the determination of flavonoids in pigeon pea leaves by liquid chromatography–tandem mass spectrometry. *Journal of Chromatography A, 1216*(18), 3841-3850. https://doi.org/10.1016/j.chroma.2009.02.073.

Liu, Y., Ragaee, S., Marcone, M. F., & Abdel-Aal, E.-S. M. (2020b). Composition of phenolic acids and antioxidant properties of selected pulses cooked with different heating conditions. *Foods, 9*(7), 908. https://doi.org/10.3390/foods9070908.
Liu, Y., Xu, M., Wu, H., Jing, L., Gong, B., Gou, M., . . . Li, W. (2018b). The compositional, physicochemical and functional properties of germinated mung bean flour and its addition on quality of wheat flour noodle. *Journal of Food Science and Technology, 55*(12), 5142-5152. https://doi.org/10.1007/s13197-018-3460-z.
López-Maestresalas, A., Pérez, C., Tierno, R., Arazuri, S., De Galarreta, J. I. R., & Jarén, C. (2017). Prediction of main potato compounds by NIRS. *Chemical Engineering Transactions, 58*, 385-390. https://doi.org/10.3303/CET1758065.
López‐Barrios, L., Gutiérrez‐Uribe, J. A., & Serna‐Saldívar, S. O. (2014). Bioactive peptides and hydrolysates from pulses and their potential use as functional ingredients. *Journal of Food Science, 79*(3), R273-R283. https://doi.org/10.1111/1750-3841.12365.
Loskutov, I. G., & Khlestkina, E. K. (2021). Wheat, barley, and oat breeding for health benefit components in grain. *Plants, 10*(1), 86. https://doi.org/10.3390/plants10010086.
Lu, X., & Rasco, B. A. (2012). Determination of antioxidant content and antioxidant activity in foods using infrared spectroscopy and chemometrics: A review. *Critical Reviews in Food Science and Nutrition, 52*(10), 853-875. https://doi.org/10.1080/10408398.2010.511322.
Lu, Y., Luthria, D., Fuerst, E. P., Kiszonas, A. M., Yu, L., & Morris, C. F. (2014). Effect of processing on phenolic composition of dough and bread fractions made from refined and whole wheat flour of three wheat varieties. *Journal of Agricultural and Food Chemistry, 62*(43), 10431-10436. https://doi.org/10.1021/jf501941r.
Lu, Y., Saeys, W., Kim, M., Peng, Y., & Lu, R. (2020). Hyperspectral imaging technology for quality and safety evaluation of horticultural products: A review and celebration of the past 20-year progress. *Postharvest Biology and Technology, 170*, 111318. https://doi.org/10.1016/j.postharvbio.2020.111318.
Ludwig, B., Murugan, R., Parama, V. R. R., & Vohland, M. (2019). Accuracy of estimating soil properties with mid-infrared spectroscopy: Implications of different chemometric approaches and software packages related to calibration sample size. *Soil Science Society of America Journal, 83*(5), 1542-1552. https://doi.org/10.2136/sssaj2018.11.0413.
Lyon, N. (2019). Drought drives mungbean production to 20-year low. Retrieved from: https://www.graincentral.com/cropping/pulses/drought-drives-mungbean-production-to-20-year-low. Accessed 9 Oct 2020.
Ma, D., Li, Y., Zhang, J., Wang, C., Qin, H., Ding, H., . . . Guo, T. (2016). Accumulation of phenolic compounds and expression profiles of phenolic acid biosynthesis-related genes in developing grains of white, purple, and red wheat. *Frontiers in Plant Science, 7*(528). https://doi.org/10.3389/fpls.2016.00528.
Macavilca, E. A., & Condezo-Hoyos, L. (2020). Assessment of total antioxidant capacity of altiplano colored quinoa (*Chenopodium quinoa* willd) by visible and near-infrared diffuse reflectance spectroscopy and chemometrics. *LWT-Food Science and Technology, 134*, 110182. https://doi.org/10.1016/j.lwt.2020.110182.
Mahbub, R., Francis, N., Blanchard, C., & Santhakumar, A. (2021). The anti-inflammatory and antioxidant properties of chickpea hull phenolic extracts. *Food Bioscience, 40*, 100850. https://doi.org/10.1016/j.fbio.2020.100850.
Mahesar, S. A., Lucarini, M., Durazzo, A., Santini, A., Lampe, A. I., & Kiefer, J. (2019). Application of infrared spectroscopy for functional compounds evaluation in olive oil: A current snapshot. *Journal of Spectroscopy, 2019*, 5319024. https://doi.org/10.1155/2019/5319024.
Mallet, A., Tsenkova, R., Muncan, J., Charnier, C., Latrille, É., Bendoula, R., . . . Roger, J.-M. (2021). Relating near-infrared light path-length modifications to the water content of scattering media in near-infrared spectroscopy: Toward a new Bouguer–Beer–

Lambert law. *Analytical Chemistry, 93*(17), 6817-6823. https://doi.org/10.1021/acs.analchem.1c00811.
Mamilla, R. K., & Mishra, V. K. (2017). Effect of germination on antioxidant and ACE inhibitory activities of legumes. *LWT-Food Science and Technology, 75*, 51-58. https://doi.org/10.1016/j.lwt.2016.08.036.
Mamouei, M., Budidha, K., Baishya, N., Qassem, M., & Kyriacou, P. A. (2021). An empirical investigation of deviations from the Beer–Lambert law in optical estimation of lactate. *Scientific Reports, 11*(1), 13734. https://doi.org/10.1038/s41598-021-92850-4.
Manach, C., Scalbert, A., Morand, C., Rémésy, C., & Jiménez, L. (2004). Polyphenols: food sources and bioavailability. *The American Journal of Clinical Nutrition, 79*(5), 727-747. https://doi.org/10.1093/ajcn/79.5.727.
Mani, J. S., Johnson, J. B., Hosking, H., Ashwath, N., Walsh, K. B., Neilsen, P. M., . . . Naiker, M. (2021). Antioxidative and therapeutic potential of selected Australian plants: A review. *Journal of Ethnopharmacology, 268*, 113580. https://doi.org/10.1016/j.jep.2020.113580.
Manley, M. (2014). Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chemical Society Reviews, 43*(24), 8200-8214. https://doi.org/10.1039/C4CS00062E.
Marchiosi, R., dos Santos, W. D., Constantin, R. P., de Lima, R. B., Soares, A. R., Finger-Teixeira, A., . . . Ferrarese-Filho, O. (2020). Biosynthesis and metabolic actions of simple phenolic acids in plants. *Phytochemistry Reviews, 19*(4), 865-906. https://doi.org/10.1007/s11101-020-09689-2.
Markakis, P. C. (1989). Food colorants: Anthocyanins. *Critical Reviews in Food Science and Nutrition, 28*(4), 273-314. https://doi.org/10.1080/10408398909527503.
Markosyan, A., McCluskey, J. J., & Wahl, T. I. (2009). Consumer response to information about a functional food product: Apples enriched with antioxidants. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, 57*(3), 325-341. https://doi.org/10.1111/j.1744-7976.2009.01154.x.
Martín-Tornero, E., de Jorge Páscoa, R. N. M., Espinosa-Mansilla, A., Martín-Merás, I. D., & Lopes, J. A. (2020). Comparative quantification of chlorophyll and polyphenol levels in grapevine leaves sampled from different geographical locations. *Scientific Reports, 10*(1), 6246. https://doi.org/10.1038/s41598-020-63407-8.
Martínez-Sandoval, J. R., Nogales-Bueno, J., Rodríguez-Pulido, F. J., Hernández-Hierro, J. M., Segovia-Quintero, M. A., Martínez-Rosas, M. E., & Heredia, F. J. (2016). Screening of anthocyanins in single red grapes using a non-destructive method based on the near infrared hyperspectral technology and chemometrics. *Journal of the Science of Food and Agriculture, 96*(5), 1643-1647. https://doi.org/10.1002/jsfa.7266.
Mayerhöfer, T. G., Pahlow, S., & Popp, J. (2020). The Bouguer-Beer-Lambert law: Shining sight on the obscure. *ChemPhysChem, 21*(18), 2029-2046. https://doi.org/10.1002/cphc.202000464.
Mazzoncini, M., Antichi, D., Silvestri, N., Ciantelli, G., & Sgherri, C. (2015). Organically vs conventionally grown winter wheat: Effects on grain yield, technological quality, and on phenolic composition and antioxidant properties of bran and refined flour. *Food Chemistry, 175*, 445-451. https://doi.org/10.1016/j.foodchem.2014.11.138.
McClure, W. F. (2003). 204 years of near infrared technology: 1800–2003. *Journal of Near Infrared Spectroscopy, 11*(6), 487-518.
McGoverin, C. M., Weeranantanaphan, J., Downey, G., & Manley, M. (2010). Review: The application of near infrared spectroscopy to the measurement of bioactive compounds in food commodities. *Journal of Near Infrared Spectroscopy, 18*(2), 87-111. https://doi.org/10.1255/jnirs.874.
McKenzie, F. C., & Williams, J. (2015). Sustainable food production: constraints, challenges and choices by 2050. *Food Security, 7*(2), 221-233. https://doi.org/10.1007/s12571-015-0441-1.

Mecozzi, M., & Sturchio, E. (2017). Computer assisted examination of infrared and near infrared spectra to assess structural and molecular changes in biological samples exposed to pollutants: A case of study. *Journal of Imaging, 3*(1), 11. https://doi.org/10.3390/jimaging3010011.

Meenu, M., Kamboj, U., Sharma, A., Guha, P., & Mishra, S. (2016a). Green method for determination of phenolic compounds in mung bean (*Vigna radiata* L.) based on near-infrared spectroscopy and chemometrics. *International Journal of Food Science & Technology, 51*(12), 2520-2527. https://doi.org/10.1111/ijfs.13232.

Meenu, M., Sharma, A., Guha, P., & Mishra, S. (2016b). A rapid high-performance liquid chromatography photodiode array detection method to determine phenolic compounds in mung bean (*Vigna radiata* L.). *International Journal of Food Properties, 19*(10), 2223-2237. https://doi.org/10.1080/10942912.2015.1121396.

Mehmood, T., & Ahmed, B. (2016). The diversity in the applications of partial least squares: an overview. *Journal of Chemometrics, 30*(1), 4-17. https://doi.org/10.1002/cem.2762.

MeiXiong, MengliZhao, Zhen-XiangLu, ParthibaBalasubramanian, & BrianBeres. (2020). Genotypic variation for phenolic compounds in developing and whole seeds, and storage conditions influence visual seed quality of yellow dry bean genotypes. *Canadian Journal of Plant Science, 100*(3), 284-295. https://doi.org/10.1139/cjps-2019-0153.

Michel, S., Löschenberger, F., Ametz, C., Pachler, B., Sparry, E., & Bürstmayr, H. (2019). Simultaneous selection for grain yield and protein content in genomics-assisted wheat breeding. *Theoretical and Applied Genetics, 132*(6), 1745-1760. https://doi.org/10.1007/s00122-019-03312-5.

Miedaner, T., Boeven, A. L. G.-C., Gaikpa, D. S., Kistner, M. B., & Grote, C. P. (2020). Genomics-assisted breeding for quantitative disease resistances in small-grain cereals and maize. *International Journal of Molecular Sciences, 21*(24), 9717. https://doi.org/10.3390/ijms21249717.

Milán-Noris, A. K., Gutiérrez-Uribe, J. A., Santacruz, A., Serna-Saldívar, S. O., & Martínez-Villaluenga, C. (2018). Peptides and isoflavones in gastrointestinal digests contribute to the anti-inflammatory potential of cooked or germinated desi and kabuli chickpea (*Cicer arietinum* L.). *Food Chemistry, 268*, 66-76. https://doi.org/10.1016/j.foodchem.2018.06.068.

Mishra, P., Biancolillo, A., Roger, J. M., Marini, F., & Rutledge, D. N. (2020). New data preprocessing trends based on ensemble of multiple preprocessing techniques. *TrAC Trends in Analytical Chemistry, 132*, 116045. https://doi.org/10.1016/j.trac.2020.116045.

Miškolci, S. (2014). Consumer preferences and willingness to pay for the health aspects of food. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 59*(4), 167-176.

Mokni Ghribi, A., Sila, A., Maklouf Gafsi, I., Blecker, C., Danthine, S., Attia, H., . . . Besbes, S. (2015). Structural, functional, and ACE inhibitory properties of water-soluble polysaccharides from chickpea flours. *International Journal of Biological Macromolecules, 75*, 276-282. https://doi.org/10.1016/j.ijbiomac.2015.01.037.

Mora-Ruiz, M. E., Reboredo-Rodríguez, P., Salvador, M. D., González-Barreiro, C., Cancho-Grande, B., Simal-Gándara, J., & Fregapane, G. (2017). Assessment of polar phenolic compounds of virgin olive oil by NIR and mid-IR spectroscopy and their impact on quality. *European Journal of Lipid Science and Technology, 119*(1), 1600099. https://doi.org/10.1002/ejlt.201600099.

Mpofu, A., Sapirstein, H. D., & Beta, T. (2006). Genotype and environmental variation in phenolic content, phenolic acid composition, and antioxidant activity of hard spring wheat. *Journal of Agricultural and Food Chemistry, 54*(4), 1265-1270. https://doi.org/10.1021/jf052683d.

Muñoz, N., Liu, A., Kan, L., Li, M.-W., & Lam, H.-M. (2017). Potential uses of wild germplasms of grain legumes for crop improvement. *International Journal of Molecular Sciences, 18*(2), 328. https://doi.org/10.3390/ijms18020328.
Myint, H., Kishi, H., Koike, S., & Kobayashi, Y. (2017). Effect of chickpea husk dietary supplementation on blood and cecal parameters in rats. *Animal Science Journal, 88*(2), 372-378. https://doi.org/10.1111/asj.12651.
Nasar-Abbas, S., Siddique, K., Plummer, J., White, P., Harris, D., Dods, K., & D'antuono, M. (2009). Faba bean (*Vicia faba* L.) seeds darken rapidly and phenolic content falls when stored at higher temperature, moisture and light intensity. *LWT-Food Science and Technology, 42*(10), 1703-1711. https://doi.org/10.1016/j.lwt.2009.05.013.
Ncama, K., Tesfay, S. Z., Fawole, O. A., Opara, U. L., & Magwaza, L. S. (2018). Non-destructive prediction of 'Marsh' grapefruit susceptibility to postharvest rind pitting disorder using reflectance Vis/NIR spectroscopy. *Scientia Horticulturae, 231*, 265-271. https://doi.org/10.1016/j.scienta.2017.12.028.
Netzel, M., Fanning, K., Netzel, G., Zabaras, D., Karagianis, G., Treloar, T., . . . Stanley, R. (2012). Urinary excretion of antioxidants in healthy humans following Queen Garnet plum juice ingestion: a new plum variety rich in antioxidant compounds. *Journal of Food Biochemistry, 36*(2), 159-170. https://doi.org/10.1111/j.1745-4514.2010.00522.x.
Netzel, M., Netzel, G., Tian, Q., Schwartz, S., & Konczak, I. (2006). Sources of antioxidant activity in australian native fruits. Identification and quantification of anthocyanins. *Journal of Agricultural and Food Chemistry, 54*(26), 9820-9826. https://doi.org/10.1021/jf0622735.
Ni, W., Nørgaard, L., & Mørup, M. (2014). Non-linear calibration models for near infrared spectroscopy. *Analytica Chimica Acta, 813*, 1-14. https://doi.org/10.1016/j.aca.2013.12.002.
Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46*(2), 99-118. https://doi.org/10.1016/j.postharvbio.2007.06.024.
Nigro, D., Laddomada, B., Mita, G., Blanco, E., Colasuonno, P., Simeone, R., . . . Blanco, A. (2017). Genome-wide association mapping of phenolic acids in tetraploid wheats. *Journal of Cereal Science, 75*, 25-34. https://doi.org/10.1016/j.jcs.2017.01.022.
Nogales-Bueno, J., Baca-Bocanegra, B., Romero-Molina, L., Martínez-López, A., Rato, A. E., Heredia, F. J., . . . González-Miret, M. L. (2020). Control of the extractable content of bioactive compounds in coffee beans by near infrared hyperspectral imaging. *LWT-Food Science and Technology, 134*, 110201. https://doi.org/10.1016/j.lwt.2020.110201.
Norris, K. (1965). Direct spectrophotometric determination of moisture content of grain and seeds. In A. Wexter (Ed.), *Proceedings of the 1963 International Symposium on Humidity and Moisture* (Vol. 4, pp. 19-25): Reinhold Publishing Corporation, New York.
Oomah, B. D., Luc, G., Leprelle, C., Drover, J. C. G., Harrison, J. E., & Olson, M. (2011). Phenolics, phytic acid, and phytase in Canadian-hrown low-tannin faba bean (*Vicia faba* L.) genotypes. *Journal of Agricultural and Food Chemistry, 59*(8), 3763-3771. https://doi.org/10.1021/jf200338b.
Orman, B. A., & Schumann, R. A. (1991). Comparison of near-infrared spectroscopy calibration methods for the prediction of protein, oil, and starch in maize grain. *Journal of Agricultural and Food Chemistry, 39*(5), 883-886. https://doi.org/10.1021/jf00005a015.
Ozaki, Y., Huck, C., Tsuchikawa, S., & Engelsen, S. B. (2021). *Near-infrared spectroscopy: Theory, spectral analysis, instrumentation, and applications*. Singapore: Springer.
Page, K. L., Dalal, R. C., Wehr, J. B., Dang, Y. P., Kopittke, P. M., Kirchhof, G., . . . Menzies, N. W. (2018). Management of the major chemical soil constraints affecting yields in

the grain growing region of Queensland and New South Wales, Australia – a review. *Soil Research, 56*(8), 765-779. https://doi.org/10.1071/SR18233.
Pająk, P., Socha, R., Gałkowska, D., Rożnowski, J., & Fortuna, T. (2014). Phenolic profile and antioxidant activity in selected seeds and sprouts. *Food Chemistry, 143*, 300-306. https://doi.org/10.1016/j.foodchem.2013.07.064.
Pallone, J. A. L., Caramês, E. T. d. S., & Alamar, P. D. (2018). Green analytical chemistry applied in food analysis: alternative techniques. *Current Opinion in Food Science, 22*, 115-121. https://doi.org/10.1016/j.cofs.2018.01.009.
Pandiselvam, R., Sruthi, N. U., Kumar, A., Kothakota, A., Thirumdas, R., Ramesh, S. V., & Cozzolino, D. (2021). Recent applications of vibrational spectroscopic techniques in the grain industry. *Food Reviews International*, 1-31. https://doi.org/10.1080/87559129.2021.1904253.
Park, J.-R., Kang, H.-H., Cho, J.-K., Moon, K.-D., & Kim, Y.-J. (2020). Feasibility of rapid piperine quantification in whole and black pepper using near infrared spectroscopy and chemometrics. *Journal of Food Science, 85*(10), 3094-3101. https://doi.org/10.1111/1750-3841.15428.
Pasquini, C. (2003). Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society, 14*(2), 198-219. https://doi.org/10.1590/S0103-50532003000200006.
Pavia, D. L., Lampman, G. M., & Kriz, G. S. (2001). *Introduction to spectroscopy: A guide for student of organic chemistry*. (3rd ed.). London: Brooks/Cole.
Pellegrini, N., Vitaglione, P., Granato, D., & Fogliano, V. (2020). Twenty-five years of total antioxidant capacity measurement of foods and biological fluids: merits and limitations. *Journal of the Science of Food and Agriculture, 100*(14), 5064-5078. https://doi.org/10.1002/jsfa.9550.
Pendergast, L., Bhattarai, S. P., & Midmore, D. J. (2019). Evaluation of aerated subsurface drip irrigation on yield, dry weight partitioning and water use efficiency of a broad-acre chickpea (*Cicer arietinum*, L.) in a vertosol. *Agricultural Water Management, 217*, 38-46. https://doi.org/10.1016/j.agwat.2019.02.022.
Peng, J., & Zhang, Z.-m. (2010). Rapid determination of starch and amylose content in whole wheat seeds by near infrared reflectance spectroscopy (NIRS)[J]. *Journal of Triticeae Crops, 30*(2), 276-279.
Pereira, C. A. M., Yariwake, J. H., & McCullagh, M. (2005). Distinction of the C-glycosylflavone isomer pairs orientin/isoorientin and vitexin/isovitexin using HPLC-MS exact mass measurement and in-source CID. *Phytochemical Analysis, 16*(5), 295-301. https://doi.org/10.1002/pca.820.
Pereira, V., Câmara, J. S., Cacho, J., & Marques, J. C. (2010). HPLC-DAD methodology for the quantification of organic acids, furans and polyphenols by direct injection of wine samples. *Journal of Separation Science, 33*(9), 1204-1215. https://doi.org/10.1002/jssc.200900784.
Petheram, C., Gallant, J., Stone, P., Wilson, P., & Read, A. (2018). Rapid assessment of potential for development of large dams and irrigation across continental areas: application to northern Australia. *The Rangeland Journal, 40*(4), 431-449. https://doi.org/10.1071/RJ18012.
Pompella, A., Sies, H., Wacker, R., Brouns, F., Grune, T., Biesalski, H. K., & Frank, J. (2014). The use of total antioxidant capacity as surrogate marker for food quality and its effect on health is to be discouraged. *Nutrition, 30*(7), 791-793. https://doi.org/10.1016/j.nut.2013.12.002.
Power, A., & Cozzolino, D. (2020). How fishy is your fish? Authentication, provenance and traceability in fish and seafood by means of vibrational spectroscopy. *Applied Sciences, 10*(12), 4150. https://doi.org/10.3390/app10124150.
Pulse Australia. (2016a). Chickpea production: northern region. Retrieved from: http://www.pulseaus.com.au/growing-pulses/bmp/chickpea/northern-guide. Accessed 19 Mar 2021.

Pulse Australia. (2016b). Faba bean production: Southern and Western region. Retrieved from: http://www.pulseaus.com.au/growing-pulses/bmp/faba-and-broad-bean/southern-guide. Accessed Aug 13 2019.
Queji, M. D., Wosiacki, G., Cordeiro, G. A., Peralta‐Zamora, P. G., & Nagata, N. (2010). Determination of simple sugars, malic acid and total phenolic compounds in apple pomace by infrared spectroscopy and PLSR. *International Journal of Food Science & Technology, 45*(3), 602-609. https://doi.org/10.1111/j.1365-2621.2010.02173.x.
Quintero-Soto, M. F., Saracho-Peña, A. G., Chavez-Ontiveros, J., Garzon-Tiznado, J. A., Pineda-Hidalgo, K. V., Delgado-Vargas, F., & Lopez-Valenzuela, J. A. (2018). Phenolic profiles and their contribution to the antioxidant activity of selected chickpea genotypes from Mexico and ICRISAT collections. *Plant Foods for Human Nutrition, 73*(2), 122-129. https://doi.org/10.1007/s11130-018-0661-6.
R Core Team. (2020). R: A language and environment for statistical computing. (version 4.0.5 ed.). Vienna, Austria: R Foundation for Statistical Computing.
Rady, A. M., Sugiharto, S., & Adedeji, A. A. (2018). Evaluation of carrot quality using visible near infrared spectroscopy and multivariate analysis. *Journal of Food Research, 7*(4), 80-93. https://doi.org/10.5539/jfr.v7n4p80
Ragone, M. I., Sella, M., Conforti, P., Volonté, M. G., & Consolini, A. E. (2007). The spasmolytic effect of *Aloysia citriodora*, Palau (South American cedrón) is partially due to its vitexin but not isovitexin on rat duodenums. *Journal of Ethnopharmacology, 113*(2), 258-266. https://doi.org/10.1016/j.jep.2007.06.003.
Rahate, K. A., Madhumita, M., & Prabhakar, P. K. (2021). Nutritional composition, anti-nutritional factors, pretreatments-cum-processing impact and food formulation potential of faba bean (*Vicia faba* L.): A comprehensive review. *LWT-Food Science and Technology, 138*, 110796. https://doi.org/10.1016/j.lwt.2020.110796.
Rajalakshmi, G., & Gopal, A. (2020). Performance evaluation of preprocessing techniques for near-infrared spectroscopy signals. *Microprocessors and Microsystems*, 103372. https://doi.org/10.1016/j.micpro.2020.103372.
Rasines-Perea, Z., & Teissedre, P.-L. (2017). Grape polyphenols' effects in human cardiovascular diseases and diabetes. *Molecules, 22*(1), 68. https://doi.org/10.3390/molecules22010068.
Rasouli, H., Farzaei, M. H., & Khodarahmi, R. (2017). Polyphenols and their benefits: A review. *International Journal of Food Properties, 20*(sup2), 1700-1741. https://doi.org/10.1080/10942912.2017.1354017.
Rautiainen, S., Larsson, S., Virtamo, J., & Wolk, A. (2012a). Total antioxidant capacity of diet and risk of stroke. *Stroke, 43*(2), 335-340. https://doi.org/10.1161/STROKEAHA.111.635557.
Rautiainen, S., Levitan, E. B., Mittleman, M. A., & Wolk, A. (2013). Total antioxidant capacity of diet and risk of heart failure: A population-based prospective cohort of women. *The American Journal of Medicine, 126*(6), 494-500. https://doi.org/10.1016/j.amjmed.2013.01.006.
Rautiainen, S., Levitan, E. B., Orsini, N., Åkesson, A., Morgenstern, R., Mittleman, M. A., & Wolk, A. (2012b). Total antioxidant capacity from diet and risk of myocardial infarction: A prospective cohort of women. *The American Journal of Medicine, 125*(10), 974-980. https://doi.org/10.1016/j.amjmed.2012.03.008.
Rayee, R., Tran, H.-D., Xuan, T. D., & Khanh, T. D. (2018). Imposed water deficit after anthesis for the improvement of macronutrients, quality, phytochemicals, and antioxidants in rice grain. *Sustainability, 10*(12), 4843. https://doi.org/10.3390/su10124843.
Rebello, C. J., Greenway, F. L., & Finley, J. W. (2014). Whole grains and pulses: A comparison of the nutritional and health benefits. *Journal of Agricultural and Food Chemistry, 62*(29), 7029-7049. https://doi.org/10.1021/jf500932z.
Redaelli, R., Alfieri, M., & Cabassi, G. (2016). Development of a NIRS calibration for total antioxidant capacity in maize germplasm. *Talanta, 154*, 164-168. https://doi.org/10.1016/j.talanta.2016.03.048.

Renfroe, W., & Kays, S. (1985). Nondestructive spectrophotometric determination of dry matter in onions. *Journal of the American Society for Horticultural Science, 110*(2), 297-303.
Rezaei, M. K., Deokar, A. A., Arganosa, G., Roorkiwal, M., Pandey, S. K., Warkentin, T. D., . . . Tar′ an, B. (2019). Mapping quantitative trait loci for carotenoid concentration in three f2 populations of chickpea. *The Plant Genome, 12*(3), 190067. https://doi.org/10.3835/plantgenome2019.07.0067.
Rhodes, D. H., Hoffmann, L., Rooney, W. L., Ramu, P., Morris, G. P., & Kresovich, S. (2014). Genome-wide association study of grain polyphenol concentrations in global sorghum [*Sorghum bicolor* (L.) Moench] germplasm. *Journal of Agricultural and Food Chemistry, 62*(45), 10916-10927. https://doi.org/10.1021/jf503651t.
Rinnan, Å. (2014). Pre-processing in vibrational spectroscopy – when, why and how. *Analytical Methods, 6*(18), 7124-7129. https://doi.org/10.1039/C3AY42270D.
Ristic, R., Cozzolino, D., Jeffery, D. W., Gambetta, J. M., & Bastian, S. E. P. (2016). Prediction of phenolic composition of Shiraz wines using attenuated total reflectance mid-infrared (ATR-MIR) spectroscopy. *American Journal of Enology and Viticulture, 67*(4), 460-465. https://doi.org/10.5344/ajev.2016.16030.
Rodriguez-Otero, J. L., Hermida, M., & Cepeda, A. (1995). Determination of fat, protein, and total solids in cheese by near-infrared reflectance spectroscopy. *Journal of AOAC International, 78*(3), 802-806. https://doi.org/10.1093/jaoac/78.3.802.
Rodríguez-Pulido, F. J., Gil-Vicente, M., Gordillo, B., Heredia, F. J., & González-Miret, M. L. (2017). Measurement of ripening of raspberries (*Rubus idaeus* L) by near infrared and colorimetric imaging techniques. *Journal of Food Science and Technology, 54*(9), 2797-2803. https://doi.org/10.1007/s13197-017-2716-3.
Rodríguez, S. D., López-Fernández, M. P., Maldonado, S., & Buera, M. P. (2019). Evidence on the discrimination of quinoa grains with a combination of FT-MIR and FT-NIR spectroscopy. *Journal of Food Science and Technology, 56*(10), 4457-4464. https://doi.org/10.1007/s13197-019-03948-7.
Ryan, P. R. (2018). Assessing the role of genetics for improving the yield of Australia's major grain crops on acid soils. *Crop and Pasture Science, 69*(3), 242-264. https://doi.org/10.1071/CP17310.
Ryszard, A., & Fereidoon, S. (2018). Antioxidant activity of faba bean extract and fractions thereof. *Journal of Food Bioactives, 2*(2), 112-118. https://doi.org/10.31665/JFB.2018.2146.
Sadras, V. O., Rosewarne, G. M., & Lake, L. (2021). Australian lentil breeding between 1988 and 2019 has delivered greater yield gain under stress than under high-yield conditions. *Frontiers in Plant Science, 12*. https://doi.org/10.3389/fpls.2021.674327.
Sanches-Silva, A., Testai, L., Nabavi, S. F., Battino, M., Pandima Devi, K., Tejada, S., . . . Farzaei, M. H. (2020). Therapeutic potential of polyphenols in cardiovascular diseases: Regulation of mTOR signaling pathway. *Pharmacological Research, 152*, 104626. https://doi.org/10.1016/j.phrs.2019.104626.
Sanghamitra, P., Sah, R. P., Bagchi, T. B., Sharma, S. G., Kumar, A., Munda, S., & Sahu, R. K. (2018). Evaluation of variability and environmental stability of grain quality and agronomic parameters of pigmented rice (*O. sativa* L.). *Journal of Food Science and Technology, 55*(3), 879-890. https://doi.org/10.1007/s13197-017-2978-9.
Santana, M. C. d., Ferreira, M. M. C., & Pallone, J. A. L. (2020). Control of ascorbic acid in fortified powdered soft drinks using near-infrared spectroscopy (NIRS) and multivariate analysis. *Journal of Food Science and Technology, 57*(4), 1233-1241. https://doi.org/10.1007/s13197-019-04154-1.
Santhakumar, A. B., Kundur, A. R., Fanning, K., Netzel, M., Stanley, R., & Singh, I. (2015). Consumption of anthocyanin-rich Queen Garnet plum juice reduces platelet activation related thrombogenesis in healthy volunteers. *Journal of Functional Foods, 12*, 11-22. https://doi.org/10.1016/j.jff.2014.10.026.
Santos-Sánchez, N. F., Salas-Coronado, R., Hernández-Carlos, B., & Villanueva-Cañongo, C. (2019). Shikimic acid pathway in biosynthesis of phenolic compounds. In M. Soto-

Hernández, R. García-Mateos & M. Palma-Tenango (Eds.), *Plant Physiological Aspects of Phenolic Compounds*. London, UK: IntechOpen.
Santos, W. N. L. d., da Silva Sauthier, M. C., dos Santos, A. M. P., de Andrade Santana, D., Almeida Azevedo, R. S., & da Cruz Caldas, J. (2017). Simultaneous determination of 13 phenolic bioactive compounds in guava (*Psidium guajava* L.) by HPLC-PAD with evaluation using PCA and Neural Network Analysis (NNA). *Microchemical Journal, 133*, 583-592. https://doi.org/10.1016/j.microc.2017.04.029.
Sastre Toraño, J., & Hattum, S. (2001). Quantitative analysis of active compounds in pharmaceutical preparations by use of attenuated total-reflection Fourier transform mid-infrared spectrophotometry and the internal standard method. *Fresenius' Journal of Analytical Chemistry, 371*(4), 532-535. https://doi.org/10.1007/s002160101031.
Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*(8), 1627-1639. https://doi.org/10.1021/ac60214a047.
Schoot, M., Kapper, C., van Kollenburg, G. H., Postma, G. J., van Kessel, G., Buydens, L. M. C., & Jansen, J. J. (2020). Investigating the need for preprocessing of near-infrared spectroscopic data as a function of sample size. *Chemometrics and Intelligent Laboratory Systems, 204*, 104105. https://doi.org/10.1016/j.chemolab.2020.104105.
Segev, A., Badani, H., Galili, L., Hovav, R., Kapulnik, Y., Shomer, I., & Galili, S. (2011). Total phenolic content and antioxidant activity of chickpea (*Cicer arietinum* L.) as affected by soaking and cooking conditions. *Food and Nutrition Sciences, 2*(7), 724-730. https://doi.org/10.4236/fns.2011.27099.
Sehrawat, N., Yadav, M., Kumar, S., Upadhyay, S. K., Singh, M., & Sharma, A. K. (2020). Review on health promoting biological activities of mungbean: A potent functional food of medicinal importance. *Plant Archives, 20*(2), 2969-2975.
Sen, I., Ozturk, B., Tokatli, F., & Ozen, B. (2016). Combination of visible and mid-infrared spectra for the prediction of chemical parameters of wines. *Talanta, 161*, 130-137. https://doi.org/10.1016/j.talanta.2016.08.057.
Serrano, C., Carbas, B., Castanho, A., Soares, A., Patto, M. C. V., & Brites, C. (2017). Characterisation of nutritional quality traits of a chickpea (*Cicer arietinum*) germplasm collection exploited in chickpea breeding in Europe. *Crop and Pasture Science, 68*(11), 1031-1040. https://doi.org/10.1071/CP17129.
Shahidi, F., & Ambigaipalan, P. (2015). Phenolics and polyphenolics in foods, beverages and spices: Antioxidant activity and health effects – A review. *Journal of Functional Foods, 18*, 820-897. https://doi.org/10.1016/j.jff.2015.06.018.
Shan, J., Wang, X., Han, S., & Kondo, N. (2017). Application of curve fitting and wavelength selection methods for determination of chlorogenic acid concentration in coffee aqueous solution by VIS/NIR spectroscopy. *Food Analytical Methods, 10*(4), 999-1006. https://doi.org/10.1007/s12161-016-0650-z.
Shao, Y., Jin, L., Zhang, G., Lu, Y., Shen, Y., & Bao, J. (2011). Association mapping of grain color, phenolic content, flavonoid content and antioxidant capacity in dehulled rice. *Theoretical and Applied Genetics, 122*(5), 1005-1016. https://doi.org/10.1007/s00122-010-1505-4.
Sharabiani, V. R., Nazarloo, A. S., & Taghinezhad, E. (2019). Prediction of protein content of winter wheat by canopy of near infrared spectroscopy (NIRS), using partial least squares regression (PLSR) and artificial neural network (ANN) models. *Yüzüncü Yıl Üniversitesi Tarım Bilimleri Dergisi, 29*(1), 43-51. https://doi.org/10.29133/yyutbd.447926.
Sharan, S., Zanghelini, G., Zotzel, J., Bonerz, D., Aschoff, J., Saint-Eve, A., & Maillard, M.-N. (2021). Fava bean (*Vicia faba* L.) for food applications: From seed to ingredient processing and its effect on functional properties, antinutritional factors, flavor, and color. *Comprehensive Reviews in Food Science and Food Safety, 20*(1), 401-428. https://doi.org/10.1111/1541-4337.12687.

Sharma, M., Rahim, M. S., Kumar, P., Mishra, A., Sharma, H., & Roy, J. (2020). Large-scale identification and characterization of phenolic compounds and their marker–trait association in wheat. *Euphytica, 216*(8), 127. https://doi.org/10.1007/s10681-020-02659-x.

Sharma, S., Yadav, N., Singh, A., & Kumar, R. (2013). Nutritional and antinutritional profile of newly developed chickpea (*Cicer arietinum* L) varieties. *International Food Research Journal, 20*(2), 805-810.

Shi, Z., Yao, Y., Zhu, Y., & Ren, G. (2016). Nutritional composition and antioxidant activity of twenty mung bean cultivars in China. *The Crop Journal, 4*(5), 398-406. https://doi.org/10.1016/j.cj.2016.06.011.

Shiferaw, B., Smale, M., Braun, H.-J., Duveiller, E., Reynolds, M., & Muricho, G. (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security, 5*(3), 291-317. https://doi.org/10.1007/s12571-013-0263-y.

Siah, S., Konczak, I., Wood, J. A., Agboola, S., & Blanchard, C. L. (2014a). Effects of roasting on phenolic composition and in vitro antioxidant capacity of Australian grown faba beans (*Vicia faba* L.). *Plant Foods for Human Nutrition, 69*(1), 85-91. https://doi.org/10.1007/s11130-013-0400-y.

Siah, S., Wood, J. A., Agboola, S., Konczak, I., & Blanchard, C. L. (2014b). Effects of soaking, boiling and autoclaving on the phenolic contents and antioxidant activities of faba beans (*Vicia faba* L.) differing in seed coat colours. *Food Chemistry, 142*, 461-468. https://doi.org/10.1007/s11130-013-0400-y.

Siah, S. D., Konczak, I., Agboola, S., Wood, J. A., & Blanchard, C. L. (2012). In vitro investigations of the potential health benefits of Australian-grown faba beans (*Vicia faba* L.): chemopreventative capacity and inhibitory effects on the angiotensin-converting enzyme, α-glucosidase and lipase. *British Journal of Nutrition, 108*(S1), S123-S134. https://doi.org/10.1017/S0007114512000803.

Siddique, K., Brinsmead, R., Knight, R., Knights, E., Paull, J., & Rose, I. (2000). Adaptation of chickpea (*Cicer arietinum* L.) and faba bean (*Vicia faba* L.) to Australia. In R. Knight (Ed.), *Linking Research and Marketing Opportunities for Pulses in the 21st Century* (pp. 289-303). Berlin/Heidelberg, Germany: Springer.

Singh, B., Singh, J. P., Shevkani, K., Singh, N., & Kaur, A. (2017). Bioactive constituents in pulses and their health benefits. *Journal of Food Science and Technology, 54*(4), 858-870. https://doi.org/10.1007/s13197-016-2391-9.

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

Skylas, D. J., Blanchard, C. L., & Quail, K. J. (2017). Variation in nutritional composition of Australian mungbean varieties. *Journal of Agricultural Science, 9*(5), 45-53. https://doi.org/10.5539/jas.v9n5p45.

Skylas, D. J., Paull, J. G., Hughes, D. G., Gogel, B., Long, H., Williams, B., . . . Quail, K. J. (2019). Nutritional and anti-nutritional seed-quality traits of faba bean (*Vicia faba*) grown in South Australia. *Crop and Pasture Science, 70*(5), 463-472. https://doi.org/10.1071/CP19017.

Smýkal, P., Coyne, C. J., Ambrose, M. J., Maxted, N., Schaefer, H., Blair, M. W., . . . Varshney, R. K. (2015). Legume crops phylogeny and genetic diversity for science and breeding. *Critical Reviews in Plant Sciences, 34*(1-3), 43-104. https://doi.org/10.1080/07352689.2014.897904.

Sørensen, L. K. (2002). True accuracy of near infrared spectroscopy and its dependence on precision of reference data. *Journal of Near Infrared Spectroscopy, 10*(1), 15-25.

Sosulski, F. W., & Dabrowski, K. J. (1984). Composition of free and hydrolyzable phenolic acids in the flours and hulls of ten legume species. *Journal of Agricultural and Food Chemistry, 32*(1), 131-133. https://doi.org/10.1021/jf00121a033.

Spanou, C., Veskoukis, A. S., Kerasioti, T., Kontou, M., Angelis, A., Aligiannis, N., . . . Kouretas, D. (2012). Flavonoid glycosides isolated from unique legume plant extracts

as novel inhibitors of xanthine oxidase. *PLOS One, 7*(3), e32214. https://doi.org/10.1371/journal.pone.0032214.
Sreerama, Y. N., Sashikala, V. B., & Pratape, V. M. (2012). Phenolic compounds in cowpea and horse gram flours in comparison to chickpea flour: Evaluation of their antioxidant and enzyme inhibitory properties associated with hyperglycemia and hypertension. *Food Chemistry, 133*(1), 156-162. https://doi.org/10.1016/j.foodchem.2012.01.011.
Stubbs, T. L., Kennedy, A. C., & Fortuna, A.-M. (2010). Using NIRS to predict fiber and nutrient content of dryland cereal cultivars. *Journal of Agricultural and Food Chemistry, 58*(1), 398-403. https://doi.org/10.1021/jf9025844.
Sunoj, S., Igathinathane, C., & Visvanathan, R. (2016). Nondestructive determination of cocoa bean quality using FT-NIR spectroscopy. *Computers and Electronics in Agriculture, 124*, 234-242. https://doi.org/10.1016/j.compag.2016.04.012.
Susi, H., & Ard, J. S. (1973). Determination of crystallinity of α-D-lactose by far infrared spectroscopy. *Journal of Association of Official Analytical Chemists, 56*(1), 177-180. https://doi.org/10.1093/jaoac/56.1.177.
Tachibana, N., Wanezaki, S., Nagata, M., Motoyama, T., Kohno, M., & Kitagawa, S. (2013). Intake of mung bean protein isolate reduces plasma triglyceride level in rats. *Functional Foods in Health and Disease, 3*(9), 365-376. https://doi.org/10.31989/ffhd.v3i9.39.
Tahir, H. E., Xiaobo, Z., Tinting, S., Jiyong, S., & Mariod, A. A. (2016). Near-infrared (NIR) spectroscopy for rapid measurement of antioxidant properties and discrimination of Sudanese honeys from different botanical origin. *Food Analytical Methods, 9*(9), 2631-2641. https://doi.org/10.1007/s12161-016-0453-2.
Tahir, H. E., Xiaobo, Z., Zhihua, L., Jiyong, S., Zhai, X., Wang, S., & Mariod, A. A. (2017). Rapid prediction of phenolic compounds and antioxidant activity of Sudanese honey using Raman and Fourier transform infrared (FT-IR) spectroscopy. *Food Chemistry, 226*, 202-211. https://doi.org/10.1016/j.foodchem.2017.01.024.
Temiz, H. T., & Ulaş, B. (2021). A review of recent studies employing hyperspectral imaging for the determination of food adulteration. *Photochem, 1*(2), 125-146. https://doi.org/10.3390/photochem1020008.
Terhoeven-Urselmans, T., Schmidt, H., Georg Joergensen, R., & Ludwig, B. (2008). Usefulness of near-infrared spectroscopy to determine biological and chemical soil properties: Importance of sample pre-treatment. *Soil Biology and Biochemistry, 40*(5), 1178-1188. https://doi.org/10.1016/j.soilbio.2007.12.011.
Thomas, D. T., Moore, A. D., Bell, L. W., & Webb, N. P. (2018). Ground cover, erosion risk and production implications of targeted management practices in Australian mixed farming systems: Lessons from the Grain and Graze program. *Agricultural Systems, 162*, 123-135. https://doi.org/10.1016/j.agsy.2018.02.001.
Thomas, N. C. (1991). The early history of spectroscopy. *Journal of Chemical Education, 68*(8), 631. https://doi.org/10.1021/ed068p631.
Tian, W., Chen, G., Gui, Y., Zhang, G., & Li, Y. (2021a). Rapid quantification of total phenolics and ferulic acid in whole wheat using UV–Vis spectrophotometry. *Food Control, 123*, 107691. https://doi.org/10.1016/j.foodcont.2020.107691.
Tian, W., Chen, G., Tilley, M., & Li, Y. (2021b). Changes in phenolic profiles and antioxidant activities during the whole wheat bread-making process. *Food Chemistry, 345*, 128851. https://doi.org/10.1016/j.foodchem.2020.128851.
Tian, W., Chen, G., Zhang, G., Wang, D., Tilley, M., & Li, Y. (2021c). Rapid determination of total phenolic content of whole wheat flour using near-infrared spectroscopy and chemometrics. *Food Chemistry, 344*, 128633. https://doi.org/10.1016/j.foodchem.2020.128633.
Tilahun, S., Park, D. S., Seo, M. H., Hwang, I. G., Kim, S. H., Choi, H. R., & Jeong, C. S. (2018). Prediction of lycopene and β-carotene in tomatoes by portable chroma-meter and VIS/NIR spectra. *Postharvest Biology and Technology, 136*, 50-56. https://doi.org/10.1016/j.postharvbio.2017.10.007.

Tiwari, A., Sahana, C., Zehra, A., Madhusudana, K., Kumar, D., & Agawane, S. (2013). Mitigation of starch-induced postprandial glycemic spikes in rats by antioxidants-rich extract of *Cicer arietinum* Linn. seeds and sprouts. *Journal of Pharmacy and Bioallied Sciences, 5*(4), 270-276. https://doi.org/10.4103/0975-7406.120077.
Toledo-Martín, E. M., García-García, M. d. C., Font, R., Moreno-Rojas, J. M., Salinas-Navarro, M., Gómez, P., & Del Río-Celestino, M. (2018). Quantification of total phenolic and carotenoid content in blackberries (*Rubus fructicosus* L.) using near infrared spectroscopy (NIRS) and multivariate Analysis. *Molecules, 23*(12), 3191. https://doi.org/10.3390/molecules23123191.
Tomé-Sánchez, I., Martín-Diana, A. B., Peñas, E., Bautista-Expósito, S., Frias, J., Rico, D., . . . Martinez-Villaluenga, C. (2020). Soluble phenolic composition tailored by germination conditions accompany antioxidant and anti-inflammatory properties of wheat. *Antioxidants, 9*(5), 426. https://doi.org/10.3390/antiox9050426.
Torres, I., Sánchez, M.-T., Garrido-Varo, A., & Pérez-Marín, D. (2020). *New generation NIRS sensors for quality and safety assurance in summer squashes along the food supply chain*: SPIE.
Trapani, S., Migliorini, M., Cecchi, L., Giovenzana, V., Beghi, R., Canuti, V., . . . Zanoni, B. (2017). Feasibility of filter-based NIR spectroscopy for the routine measurement of olive oil fruit ripening indices. *European Journal of Lipid Science and Technology, 119*(6), 1600239. https://doi.org/10.1002/ejlt.201600239.
Tschannerl, J., Ren, J., Jack, F., Krause, J., Zhao, H., Huang, W., & Marshall, S. (2019). Potential of UV and SWIR hyperspectral imaging for determination of levels of phenolic flavour compounds in peated barley malt. *Food Chemistry, 270*, 105-112. https://doi.org/10.1016/j.foodchem.2018.07.089.
Turco, I., Ferretti, G., & Bacchetti, T. (2016). Review of the health benefits of Faba bean (*Vicia faba* L.) polyphenols. *Journal of Food & Nutrition Research, 55*(4), 283-293.
Tzanova, M., Atanassova, S., Atanasov, V., & Grozeva, N. (2020). Content of polyphenolic compounds and antioxidant potential of some Bulgarian red grape varieties and red wines, determined by HPLC, UV, and NIR spectroscopy. *Agriculture, 10*(6), 193. https://doi.org/10.3390/agriculture10060193.
Uncu, O., Ozen, B., & Tokatli, F. (2019). Use of FTIR and UV–visible spectroscopy in determination of chemical characteristics of olive oils. *Talanta, 201*, 65-73. https://doi.org/10.1016/j.talanta.2019.03.116.
Unkovich, M., McBeath, T., Llewellyn, R., Hall, J., Gupta, V. V., & Macdonald, L. M. (2020). Challenges and opportunities for grain farming on sandy soils of semi-arid south and south-eastern Australia. *Soil Research, 58*(4), 323-334. https://doi.org/10.1071/SR19161.
Urala, N., & Lähteenmäki, L. (2007). Consumers' changing attitudes towards functional foods. *Food Quality and Preference, 18*(1), 1-12. https://doi.org/10.1016/j.foodqual.2005.06.007.
Valente, I. M., Cabrita, A. R. J., Malushi, N., Oliveira, H. M., Papa, L., Rodrigues, J. A., . . . Maia, M. R. G. (2019). Unravelling the phytonutrients and antioxidant properties of European *Vicia faba* L. seeds. *Food Research International, 116*, 888-896. https://doi.org/10.1016/j.foodres.2018.09.025.
Valente, I. M., Maia, M. R. G., Malushi, N., Oliveira, H. M., Papa, L., Rodrigues, J. A., . . . Cabrita, A. R. J. (2018). Profiling of phenolic compounds and antioxidant properties of European varieties and cultivars of *Vicia faba* L. pods. *Phytochemistry, 152*, 223-229. https://doi.org/10.1016/j.phytochem.2018.05.011.
Velasco, L., Schierholt, A., & Becker, H. C. (1998). Performance of near-infrared reflectance spectroscopy (NIRS) in routine analysis of C18 unsaturated fatty acids in intact rapeseed. *Lipid / Fett, 100*(2), 44-48. https://doi.org/10.1002/(SICI)1521-4133(199802)100:2<44::AID-LIPI44>3.0.CO;2-G.
Venkateshwarlu, E., Reddy, K. P., & Dilip, D. (2016). Potential of *Vigna radiata* (L.) sprouts in the management of inflammation and arthritis in rats: Possible biochemical alterations. *Indian Journal of Experimental Biology, 54*(1), 37-43.

Viegas, T. R., Mata, A. L. M. L., Duarte, M. M. L., & Lima, K. M. G. (2016). Determination of quality attributes in wax jambu fruit using NIRS and PLS. *Food Chemistry, 190*, 1-4. https://doi.org/10.1016/j.foodchem.2015.05.063.
Vioque, J., Alaiz, M., & Girón-Calle, J. (2012). Nutritional and functional properties of *Vicia faba* protein isolates and related fractions. *Food Chemistry, 132*(1), 67-72. https://doi.org/10.1016/j.foodchem.2011.10.033.
Vlachos, N., Skopelitis, Y., Psaroudaki, M., Konstantinidou, V., Chatzilazarou, A., & Tegou, E. (2006). Applications of Fourier transform-infrared spectroscopy to edible oils. *Analytica Chimica Acta, 573*, 459-465. https://doi.org/10.1016/j.aca.2006.05.034.
Wafula, E. N., Wainaina, I. N., Buvé, C., Nguyen, N.-D.-T., Kinyanjui, P. K., Saeys, W., . . . Hendrickx, M. (2020). Application of near-infrared spectroscopy to predict the cooking times of aged common beans (*Phaseolus vulgaris* L.). *Journal of Food Engineering, 284*, 110056. https://doi.org/10.1016/j.jfoodeng.2020.110056.
Wallace, T. C., Murray, R., & Zelman, K. M. (2016). The nutritional value and health benefits of chickpeas and hummus. *Nutrients, 8*(12), 766. https://doi.org/10.3390/nu8120766.
Walsh, K. B., Blasco, J., Zude-Sasse, M., & Sun, X. (2020). Visible-NIR 'point' spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use. *Postharvest Biology and Technology, 168*, 111246. https://doi.org/10.1016/j.postharvbio.2020.111246.
Wang, B., Sun, J., Xia, L., Liu, J., Wang, Z., Li, P., . . . Sun, X. (2021a). The applications of hyperspectral imaging technology for agricultural products quality analysis: a review. *Food Reviews International*, 1-20. https://doi.org/10.1080/87559129.2021.1929297.
Wang, J., Kalt, W., & Sporns, P. (2000). Comparison between HPLC and MALDI-TOF MS analysis of anthocyanins in highbush blueberries. *Journal of Agricultural and Food Chemistry, 48*(8), 3330-3335. https://doi.org/10.1021/jf000101g.
Wang, J., Liu, H., & Ren, G. (2014a). Near-infrared spectroscopy (NIRS) evaluation and regional analysis of Chinese faba bean (*Vicia faba* L.). *The Crop Journal, 2*(1), 28-37. https://doi.org/10.1016/j.cj.2013.10.001.
Wang, Y.-J., Li, T.-H., Li, L.-Q., Ning, J.-M., & Zhang, Z.-Z. (2021b). Evaluating taste-related attributes of black tea by micro-NIRS. *Journal of Food Engineering, 290*, 110181. https://doi.org/10.1016/j.jfoodeng.2020.110181.
Wang, Y., Yang, M., Lee, S.-G., Davis, C. G., Koo, S. I., Fernandez, M. L., . . . Chun, O. K. (2014b). Diets high in total antioxidant capacity improve risk biomarkers of cardiovascular disease: a 9-month observational study among overweight/obese postmenopausal women. *European Journal of Nutrition, 53*(6), 1363-1369. https://doi.org/10.1007/s00394-013-0637-0.
Weeranantanaphan, J., Downey, G., Allen, P., & Sun, D.-W. (2011). A review of near infrared spectroscopy in muscle food analysis: 2005–2010. *Journal of Near Infrared Spectroscopy, 19*(2), 61-104.
Wiedemair, V., & Huck, C. W. (2018). Evaluation of the performance of three hand-held near-infrared spectrometer through investigation of total antioxidant capacity in gluten-free grains. *Talanta, 189*, 233-240. https://doi.org/10.1016/j.talanta.2018.06.056.
Wiedemair, V., Ramoner, R., & Huck, C. W. (2019). Investigations into the total antioxidant capacities of cultivars of gluten-free grains using near-infrared spectroscopy. *Food Control, 95*, 189-195. https://doi.org/10.1016/j.foodcont.2018.07.045.
Williams, P., Dardenne, P., & Flinn, P. (2017). Tutorial: Items to be included in a report on a near infrared spectroscopy project. *Journal of Near Infrared Spectroscopy, 25*(2), 85-90. https://doi.org/10.1177/0967033517702395.
Williams, P., & Norris, K. (1987). *Near-infrared technology in the agricultural and food industries*: American Association of Cereal Chemists, Inc.
Williams, P. C. (2020). Application of chemometrics to prediction of some wheat quality factors by near-infrared spectroscopy. *Cereal Chemistry, 97*(5), 958-966. https://doi.org/10.1002/cche.10318.

Williams, P. C., Stevenson, S. G., Starkey, P. M., & Hawtin, G. C. (1978). The application of near infrared reflectance spectroscopy to protein-testing in pulse breeding programmes. *Journal of the Science of Food and Agriculture, 29*, 285-292. https://doi.org/10.1002/jsfa.2740290315.

Williams, R. M., O'Brien, L., Eagles, H. A., Solah, V. A., & Jayasena, V. (2008). The influences of genotype, environment, and genotype × environment interaction on wheat quality. *Australian Journal of Agricultural Research, 59*(2), 95-111. https://doi.org/10.1071/AR07185.

Wilson, R. H., & Tapp, H. S. (1999). Mid-infrared spectroscopy for food analysis: recent new applications and relevant developments in sample presentation methods. *TrAC Trends in Analytical Chemistry, 18*(2), 85-93. https://doi.org/10.1016/S0165-9936(98)00107-1.

Wood, J. A., & Scott, J. F. (2021). Economic impacts of chickpea grain classification: how 'seed quality is Queen' must be considered alongside 'yield is King' to provide a princely income for farmers. *Crop and Pasture Science, 72*(2), 136-145. https://doi.org/10.1071/CP20282.

Wood, R. M., Dunn, B. W., Waters, D. L. E., Blanchard, C. L., Mawson, A. J., & Oli, P. (2021a). Effect of agronomic management on rice grain quality Part III: Australian water-saving irrigation practices. *Cereal Chemistry, 98*(2), 249-262. https://doi.org/10.1002/cche.10340.

Wood, R. M., Waters, D. L. E., Mawson, A. J., Blanchard, C. L., Dunn, B. W., & Oli, P. (2021b). Effect of agronomic management on rice grain quality Part I: A review of Australian practices. *Cereal Chemistry, 98*(2), 222-233. https://doi.org/10.1002/cche.10343.

Wrigley, C., Matakovsky, E., Melnik, V., Pascual, L., & Romanov, G. A. (2019). Adressing global wheat issues one-grain-at-a-time, based on gliadin alleles. In C. Walker & J. Panozzo (Eds.), *69th Australasian Grain Science Conference* (pp. 9-15). Melbourne, Australia: AGSA.

Wu, Z., Song, L., Feng, S., Liu, Y., He, G., Yioe, Y., . . . Huang, D. (2012). Germination dramatically increases isoflavonoid content and diversity in chickpea (*Cicer arietinum* L.) seeds. *Journal of Agricultural and Food Chemistry, 60*(35), 8606-8615. https://doi.org/10.1021/jf3021514.

Xiang, J., Apea-Bah, F. B., Ndolo, V. U., Katundu, M. C., & Beta, T. (2019a). Profile of phenolic compounds and antioxidant activity of finger millet varieties. *Food Chemistry, 275*, 361-368. https://doi.org/10.1016/j.foodchem.2018.09.120.

Xiang, J., Li, W., Ndolo, V. U., & Beta, T. (2019b). A comparative study of the phenolic compounds and in vitro antioxidant capacity of finger millets from different growing regions in Malawi. *Journal of Cereal Science, 87*, 143-149. https://doi.org/10.1016/j.jcs.2019.03.016.

Xiang, J., Zhang, M., Apea-Bah, F. B., & Beta, T. (2019c). Hydroxycinnamic acid amide (HCAA) derivatives, flavonoid C-glycosides, phenolic acids and antioxidant properties of foxtail millet. *Food Chemistry, 295*, 214-223. https://doi.org/10.1016/j.foodchem.2019.05.058.

Xiao, H., Li, A., Li, M., Sun, Y., Tu, K., Wang, S., & Pan, L. (2018). Quality assessment and discrimination of intact white and red grapes from *Vitis vinifera* L. at five ripening stages by visible and near-infrared spectroscopy. *Scientia Horticulturae, 233*, 99-107. https://doi.org/10.1016/j.scienta.2018.01.041.

Yabefa, J., Ocholi, Y., & Odubo, G. (2014). Effect of substrate concentration and cell loading on the hydrolysis of β (1-4) glycosidic bond in orange mesocarp (*Citrus sinensis*) by *Trichoderma reesei* for glucose production. *Asian Journal of Plant Science and Research, 4*(4), 21-24.

Yadav, S. S., & Chen, W. (2007). *Chickpea breeding and management*. Oxfordshire, UK: CABI.

Yamaguchi, K. K. d. L., Pereira, L. F. R., Lamarão, C. V., Lima, E. S., & da Veiga-Junior, V. F. (2015). Amazon acai: Chemistry and biological activities: A review. *Food Chemistry, 179*, 137-151. https://doi.org/10.1016/j.foodchem.2015.01.055.
Yan, C., Liu, H., & Lin, L. (2013). Simultaneous determination of vitexin and isovitexin in rat plasma after oral administration of *Santalum album* L. leaves extract by liquid chromatography tandem mass spectrometry. *Biomedical Chromatography, 27*(2), 228-232. https://doi.org/10.1002/bmc.2780.
Yan, H., Li, P.-H., Zhou, G.-S., Wang, Y.-J., Bao, B.-H., Wu, Q.-N., & Huang, S.-L. (2021). Rapid and practical qualitative and quantitative evaluation of non-fumigated ginger and sulfur-fumigated ginger via Fourier-transform infrared spectroscopy and chemometric methods. *Food Chemistry, 341*, 128241. https://doi.org/10.1016/j.foodchem.2020.128241.
Yang, Q.-Q., Ge, Y.-Y., Gunaratne, A., Kong, K.-W., Li, H.-B., Gul, K., . . . Gan, R.-Y. (2020). Phenolic profiles, antioxidant activities, and antiproliferative activities of different mung bean (*Vigna radiata*) varieties from Sri Lanka. *Food Bioscience, 37*, 100705. https://doi.org/10.1016/j.fbio.2020.100705.
Yang, W., Kenny, J. M., & Puglia, D. (2015). Structure and properties of biodegradable wheat gluten bionanocomposites containing lignin nanoparticles. *Industrial Crops and Products, 74*, 348-356. https://doi.org/10.1016/j.indcrop.2015.05.032.
Yang, Y., Xu-zhen, C., & Gui-xing, R. (2011). Application of near-infrared reflectance spectroscopy to the evaluation of C-chiro-inositol, vitexin, and isovitexin contents in mung bean. *Agricultural Sciences in China, 10*(12), 1986-1991. https://doi.org/10.1016/S1671-2927(11)60200-9.
Yao, Y., Cheng, X.-z., & Ren, G.-x. (2011). Contents of D-chiro-inositol, vitexin, and isovitexin in various varieties of mung bean and its products. *Agricultural Sciences in China, 10*(11), 1710-1715. https://doi.org/10.1016/S1671-2927(11)60169-7.
Yao, Y., Yang, X., Tian, J., Liu, C., Cheng, X., & Ren, G. (2013). Antioxidant and antidiabetic activities of black mung bean (*Vigna radiata* L.). *Journal of Agricultural and Food Chemistry, 61*(34), 8104-8109. https://doi.org/10.1021/jf401812z.
Ye, D., Sun, L., Zou, B., Zhang, Q., Tan, W., & Che, W. (2018). Non-destructive prediction of protein content in wheat using NIRS. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 189*, 463-472. https://doi.org/10.1016/j.saa.2017.08.055.
Yust, M. d. M., Millán-Linares, M. d. C., Alcaide-Hidalgo, J. M., Millán, F., & Pedroche, J. (2012). Hypocholesterolaemic and antioxidant activities of chickpea (*Cicer arietinum* L.) protein hydrolysates. *Journal of the Science of Food and Agriculture, 92*(9), 1994-2001. https://doi.org/10.1002/jsfa.5573.
Zanotto, S., Khazaei, H., Elessawy, F. M., Vandenberg, A., & Purves, R. W. (2020). Do faba bean genotypes carrying different zero-tannin genes (zt1 and zt2) differ in phenolic profiles? *Journal of Agricultural and Food Chemistry, 68*(28), 7530-7540. https://doi.org/10.1021/acs.jafc.9b07866.
Zareef, M., Chen, Q., Ouyang, Q., Arslan, M., Hassan, M. M., Ahmad, W., . . . Ancheng, W. (2019). Rapid screening of phenolic compounds in congou black tea (*Camellia sinensis*) during in vitro fermentation process using portable spectral analytical system coupled chemometrics. *Journal of Food Processing and Preservation, 43*(7), e13996. https://doi.org/10.1111/jfpp.13996.
Zhang, C., Shen, Y., Chen, J., Xiao, P., & Bao, J. (2008). Nondestructive prediction of total phenolics, flavonoid contents, and antioxidant capacity of rice grain using near-infrared spectroscopy. *Journal of Agricultural and Food Chemistry, 56*(18), 8268-8272. https://doi.org/10.1021/jf801830z.
Zhang, J., Chu, C.-J., Li, X.-L., Yao, S., Yan, B., Ren, H.-L., . . . Zhao, Z.-Z. (2014). Isolation and identification of antioxidant compounds in *Vaccinium bracteatum* Thunb. by UHPLC-Q-TOF LC/MS and their kidney damage protection. *Journal of Functional Foods, 11*, 62-70. https://doi.org/10.1016/j.jff.2014.09.005.
Zhang, N., Liu, X., Jin, X., Li, C., Wu, X., Yang, S., . . . Yanne, P. (2017). Determination of total iron-reactive phenolics, anthocyanins and tannins in wine grapes of skins and

seeds based on near-infrared hyperspectral imaging. *Food Chemistry, 237*, 811-817. https://doi.org/10.1016/j.foodchem.2017.06.007.
Zhang, X., Shang, P., Qin, F., Zhou, Q., Gao, B., Huang, H., . . . Yu, L. L. (2013). Chemical composition and antioxidative and anti-inflammatory properties of ten commercial mung bean samples. *LWT-Food Science and Technology, 54*(1), 171-178. https://doi.org/10.1016/j.lwt.2013.05.034.
Zhang, Z.-L., Zhou, M.-L., Tang, Y., Li, F.-L., Tang, Y.-X., Shao, J.-R., . . . Wu, Y.-M. (2012). Bioactive compounds in functional buckwheat food. *Food Research International, 49*(1), 389-395. https://doi.org/10.1016/j.foodres.2012.07.035.
Zhou, B., Jin, Z., Schwarz, P., & Li, Y. (2020). Impact of genotype, environment, and malting conditions on the antioxidant activity and phenolic content in US malting barley. *Fermentation, 6*(2), 48. https://doi.org/10.3390/fermentation6020048.
Ziegler, J. U., Leitenberger, M., Longin, C. F. H., Würschum, T., Carle, R., & Schweiggert, R. M. (2016a). Near-infrared reflectance spectroscopy for the rapid discrimination of kernels and flours of different wheat species. *Journal of Food Composition and Analysis, 51*, 30-36. https://doi.org/10.1016/j.jfca.2016.06.005.
Ziegler, V., Vanier, N. L., Ferreira, C. D., Paraginski, R. T., Monks, J. L. F., & Elias, M. C. (2016b). Changes in the bioactive compounds content of soybean as a function of grain moisture content and temperature during long-term storage. *Journal of Food Science, 81*(3), H762-H768. https://doi.org/10.1111/1750-3841.13222.
Žilić, S., Serpen, A., Akıllıoğlu, G., Janković, M., & Gökmen, V. (2012). Distributions of phenolic compounds, yellow pigments and oxidative enzymes in wheat grains and their relation to antioxidant capacity of bran and debranned flour. *Journal of Cereal Science, 56*(3), 652-658. https://doi.org/10.1016/j.jcs.2012.07.014.
Zrckova, M., Capouchova, I., Eliášová, M., Paznocht, L., Pazderů, K., Dvořák, P., . . . Štěrba, Z. (2018). The effect of genotype, weather conditions and cropping system on antioxidant activity and content of selected antioxidant compounds in wheat with coloured grain. *Plant, Soil and Environment, 64*(11), 530-538. https://doi.org/10.17221/430/2018-PSE.

Appendix A – Wheat sample details

The sample details (grade and growing location) for the 65 wheat samples sourced from Australian Export Grains Innovation Centre (AEGIC) are provided in Table A1. In Australia, wheat grades are primarily based on the protein content and rheological properties of the wheat.

Table A1: Sample details for the 65 wheat samples.

Sample code	Grade/description^	Location	Notes
W1	H2	NSW	
W2	APH2	NSW	
W3	APW1	NSW	
W4	H2	Qld	
W5	APH2	Qld	
W6	H2	Qld	
W7	APH1	Qld	
W8	ASW1	NSW	
W9	Sunco[#]	NSW	
W10	H2	Qld	
W11	APH2	NSW	
W12	APH1	NSW	
W13	APW1	WA	
W14	ASW1	WA	
W15	APW1	NSW	
W16	H2	NSW	
W17	APH (mix of grades 1+2)	NSW	
W18	APH1	Qld	Un-pearled sample
W19	APH1	Qld	Pearling time: 4 secs
W20	APH1	Qld	Pearling time: 4 secs
W21	APH1	Qld	Pearling time: 4 secs
W22	APH1	Qld	Pearling time: 8 secs
W23	APH1	Qld	Pearling time: 8 secs
W24	APH1	Qld	Pearling time: 8 secs
W25	APH1	Qld	Pearling time: 12 secs
W26	APH1	Qld	Pearling time: 12 secs
W27	APH1	Qld	Pearling time: 12 secs
W28	APH1	Qld	Pearling time: 16 secs
W29	APH1	Qld	Pearling time: 16 secs
W30	APH1	Qld	Pearling time: 16 secs
W31	APH1	Qld	Pearling time: 20 secs
W32	APH1	Qld	Pearling time: 20 secs
W33	APH1	Qld	Pearling time: 20 secs
W34	ASW1	NSW	Un-pearled sample
W35	ASW1	NSW	Pearling time: 4 secs
W36	ASW1	NSW	Pearling time: 4 secs
W37	ASW1	NSW	Pearling time: 4 secs
W38	ASW1	NSW	Pearling time: 8 secs
W39	ASW1	NSW	Pearling time: 8 secs

W40	ASW1	NSW	Pearling time: 8 secs
W41	ASW1	NSW	Pearling time: 12 secs
W42	ASW1	NSW	Pearling time: 12 secs
W43	ASW1	NSW	Pearling time: 12 secs
W44	ASW1	NSW	Pearling time: 16 secs
W45	ASW1	NSW	Pearling time: 16 secs
W46	ASW1	NSW	Pearling time: 16 secs
W47	ASW1	NSW	Pearling time: 20 secs
W48	ASW1	NSW	Pearling time: 20 secs
W49	ASW1	NSW	Pearling time: 20 secs
W50	Grist (from Allied Mills)	Unknown	Un-pearled sample
W51	Grist (from Allied Mills)	Unknown	Pearling time: 4 secs
W52	Grist (from Allied Mills)	Unknown	Pearling time: 4 secs
W53	Grist (from Allied Mills)	Unknown	Pearling time: 4 secs
W54	Grist (from Allied Mills)	Unknown	Pearling time: 8 secs
W55	Grist (from Allied Mills)	Unknown	Pearling time: 8 secs
W56	Grist (from Allied Mills)	Unknown	Pearling time: 8 secs
W57	Grist (from Allied Mills)	Unknown	Pearling time: 12 secs
W58	Grist (from Allied Mills)	Unknown	Pearling time: 12 secs
W59	Grist (from Allied Mills)	Unknown	Pearling time: 12 secs
W60	Grist (from Allied Mills)	Unknown	Pearling time: 16 secs
W61	Grist (from Allied Mills)	Unknown	Pearling time: 16 secs
W62	Grist (from Allied Mills)	Unknown	Pearling time: 16 secs
W63	Grist (from Allied Mills)	Unknown	Pearling time: 20 secs
W64	Grist (from Allied Mills)	Unknown	Pearling time: 20 secs
W65	Grist (from Allied Mills)	Unknown	Pearling time: 20 secs

^ APH1 = Australian Prime Hard 1; APH2 = Australian Prime Hard 2; APW1 = Australian Premium White 1; ASW1 = Australian Standard White 1; H2 = Hard 2.

Note that Sunco is a pure wheat variety, not a grade of wheat

Appendix B – Mungbean field trials

This appendix provides details on the field trial results for the mungbean samples investigated in Chapter 5.

B1. Seed material

The mungbean seed material comprised four new varieties from AgriVentis Technology Ltd Australia (AVTMB 1-4), which had previously been grown in small-scale field trials in 2017 (Rockhampton, Queensland) and 2018 (Biloela, Queensland). A well-established commercial variety (Jade-AU) was used for comparative purposes. Marketed by the Australian Mungbean Association, Jade-AU is a large-seeded, shiny green variety typically grown in the region between central Queensland and northern New South Wales. Compared to other large-seeded varieties such as Crystal, it displays improved yield and increased resistance to powdery mildew.

B2. Growing conditions and harvest

The mungbean varieties were grown on a commercial farm in northern Queensland as a sugarcane break crop. The farm site was 25 km SW of Home Hill, adjacent to the Burdekin River (19.8462°S, 147.2448°E) (Figure B1), with a sandy clay loam soil type (type 6Umb in the DPI classification system). This location is just north of the potential range currently deemed suitable for mungbean cropping (Figure B1), with the aim being to demonstrate the versatility of these new varieties for extending this crop to the northern regions of Queensland.

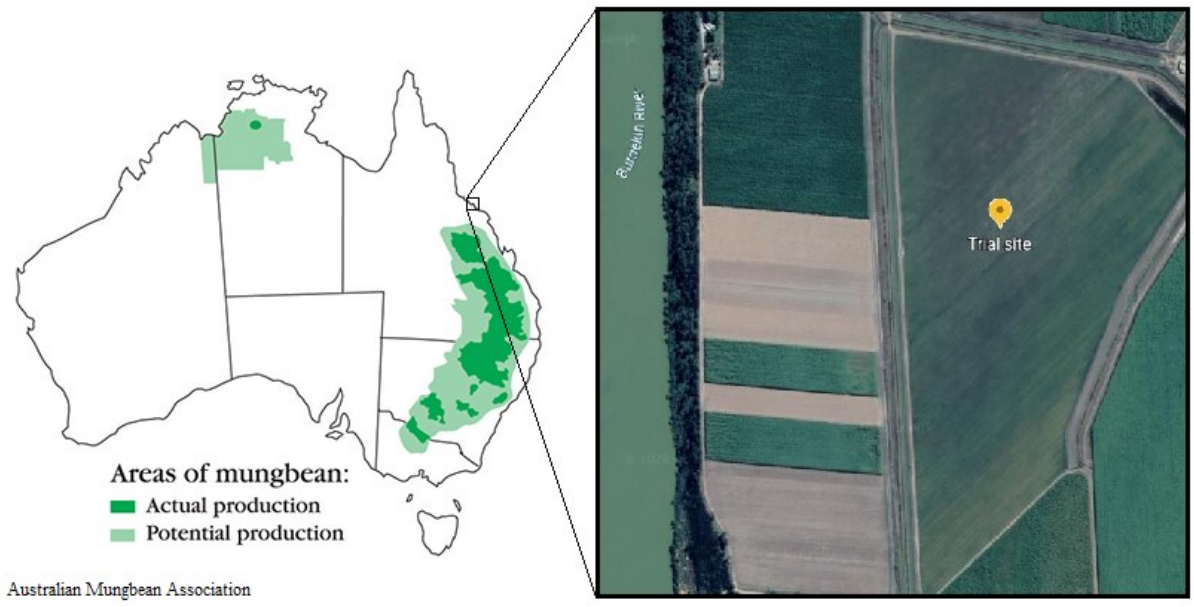

Figure B1: Location of the growing trial at Home Hill (19.8462°S, 147.2448°E). Map of mungbean production areas provided by the Australian Mungbean Association. Used with permission.

Between 6-18 ha of each new variety (1 ha for Jade-AU) was sown into seed beds (prepared with 100 kg/ha diammonium phosphate) between 25-27 August 2019 as a spring season crop. Pre-emergent and post-emergent herbicides were utilised for weed control, alongside two sprays of micro-nutrients. Furrow irrigation commenced after the soil water deficit reached the refill point. This period of the year is generally dry (Figure B2), with just 7 mm of cumulative rainfall received throughout the growing trial (BOM, 2020). The daily minimum and maximum temperatures ranged between 7.7 and 38.7°C, respectively, with a mean maximum temperature of 30.3°C (BOM, 2020).

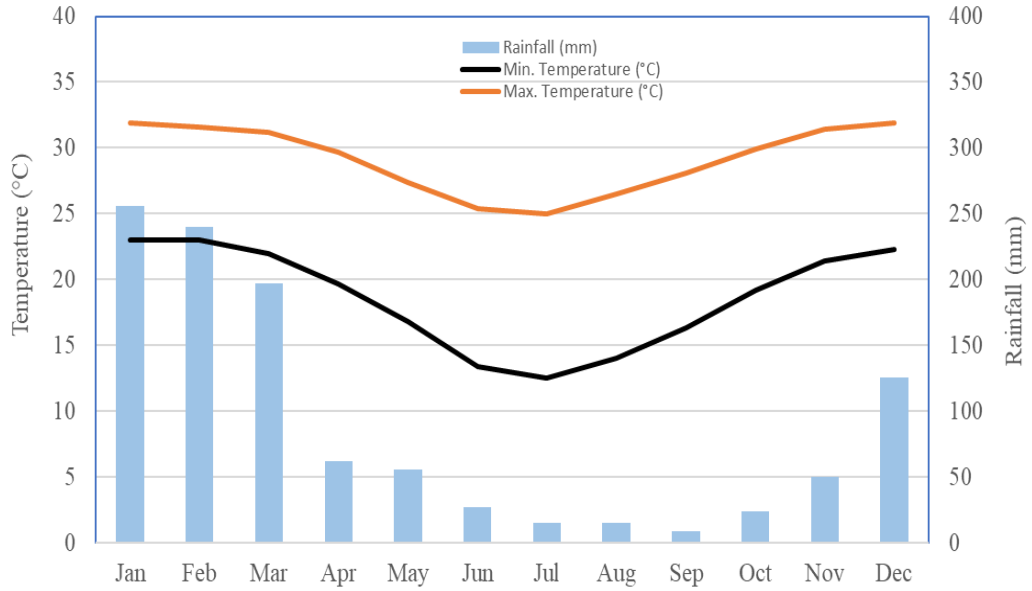

Figure B2: Climate graph showing the mean temperature and rainfall at the mungbean trial site (data sourced from the Australian Bureau of Meteorology).

Upon commencement of senescence, the crop was chemically defoliated (200 g/L diquat dibromide monohydrate) on 25th November and harvested on 27th November 2019 (93 days after sowing). Twenty within-field replicates were collected for each variety from different, randomly selected locations throughout the field.

B3. Measurement of seed quality

Thousand kernel weight (TKW) was determined using an IC-VA seed counter (AIDEX Co, Japan) (n=3 replicates for each variety).

The colour of the intact mungbean seed material and subsequent flour was quantified using a Konica Minolta chroma meter (CR-400), reported as CIE values of brightness (L*), yellowness (b*) and red/green colouration (a*). Measures were made in triplicate for each sample and the results averaged.

Seed sprouting tests were conducted by the Agricultural Testing Laboratory for Seed and Grain (AgEtal; Toowoomba, Qld), following standard methods for the Primary Production and Processing Standard for Seed Sprouts (FSANZ 2010). Resultant sprouts were classified as acceptable or unacceptable for human consumption, and also tested for the presence of three common bacterial contaminants (*Escherichia coli*, *Salmonella* and *Listeria*).

B4. Results of the physical seed quality assessment

The graded seed yield (Table B1) ranged from 0.87-1.32 tonne/ha. Higher seed yield was recorded for AVTMB 1 (1.32 ± 0.10 t/ha), AVTMB 3 (1.32 ± 0.14 t/ha) and AVTMB 4 (1.32 ± 0.15 t/ha), followed by Jade-AU (1.16 ± 0.16 t/ha). A significantly lower yield was found for AVTMB 2 (0.87 ± 0.13 t/ha). The grower and exporter of these mungbean samples were asked to qualitatively rank their preference of these varieties at the point of harvest, with the higher yield and seed quality (size, uniformity, and colour) for three of the new varieties (AVTMB 1, AVTMB 3, AVTMB 4) being reported as the most attractive characteristics. For these reasons, the grower also expressed preference in choosing these varieties for planting in the following season.

The long-term average yield for the mungbean crop in Australia is approximately 0.9 t/ha, with annual averages ranging from 0.6 to 1.1 t/ha between 2005 and 2017 (Chauhan & Williams, 2018). However, in 2019 – the year that this trial was conducted – the Australian mungbean industry average yield was exceptionally low due to severe drought stress, ranging from 0.3-0.5 t/ha (Lyon, 2019). Hence the high mean yields found in the present study (>1.3 t/ha) demonstrate the viability of mungbean cropping in the northern Australian region, if grown under irrigation. Several of the new varieties demonstrating greater adaptation to warmer temperatures may provide improved and more stable yield under the subtropical and tropical north Australian environments, consistent with their yield performance in earlier seasons at Biloela, Rockhampton, and Georgetown (Queensland) (unpublished data).

The seed size (TKW) of the five varieties ranged from 61.6-70.6 g / 1000 seeds (Table B1). One of the higher yielding mungbean varieties (AVTMB1) recorded a significantly smaller seed size (61.6 g/1000) compared to all other new varieties and Jade-AU. However, the sprouting rates were quite comparable between varieties (Table B1). Combined with the absence of microbial contamination across all samples, this indicated high suitability of all varieties for sprouting purposes.

Table B1: Physical characteristics of the five mungbean varieties (given as mean ± 1 SD). Where applicable, results are expressed on a dry weight basis. Varieties with the same superscript were not statistically different according to a post hoc Tukey test at α = 0.05.

Parameter	**AVTMB 1**	**AVTMB 2**	**AVTMB 3**	**AVTMB 4**	**Jade-AU**
Yield (t/ha)	1.32 ± 0.10^{a}	0.87 ± 0.13^{b}	1.32 ± 0.14^{a}	1.32 ± 0.15^{a}	1.16 ± 0.16^{a}
TKW (g/1000 kernels)	61.6 ± 1.01^{a}	67.3 ± 1.04^{b}	$69.2 \pm 1.12^{b,c}$	$69.8 \pm 1.10^{b,c}$	70.6 ± 1.63^{c}
Sprouting suitability (%)	80-90%	90-100%	90-100%	90-100%	90-100%
Microbiological contamination	Nil detected	Nil detected	Nil detected	Nil detected	Nil detected

In terms of the seedcoat colour, a significant difference between varieties was found for luminosity (L*), with AVTMB 4 found to be lighter overall and Jade-AU darker (Table B2). No significant differences were observed for a* (red-green colouration). The colour of AVTMB 4 was significantly less yellow compared to AVTMB 2 (lower value of b*), but not to any other variety. In general, lighter coloured samples (higher L* values) were positively correlated with increased yellowness (higher b*) (Pearson linear correlation; $r_{75} = 0.435$, $P<0.001$) and increased greenness (lower a*) ($r_{75} = -0.335$, $P<0.01$). Yellowness (higher b*) and greenness (lower a*) were also correlated ($r_{75} = -0.437$, $P<0.001$). Overall, there was no significant difference (at $\alpha = 0.05$) between the seed colour of Jade-AU and the varieties AVTMB 2 and 3 for any colour parameter, while the varieties AVTMB 1 and 4 were both significantly lighter in colour compared to Jade.

The flour colour may also play an important role in the consumer acceptability of mungbean flour products (Liu et al., 2018b); hence it was also assessed in this study. The luminosity of the flour was much lighter than that found for the seed colour (Table B2), while the other parameters (a* and b*) were relatively similar to those found for the seed colour. In contrast to the correlations observed for mungbean seedcoat colour, lighter coloured flours (higher L*) tended to have greater blueness (lower b*) ($r_{75} = -0.587$, $P<0.001$) and greenness (lower a*) ($r_{75} = -0.396$, $P<0.001$). Increased levels of yellowness (higher b*) were positively correlated with redness (higher a*) ($r_{75} = 0.579$, $P<0.001$). No significant differences were observed in the flour colour from Jade-AU and any other variety, however, that of the AVTMB 1 variety was significantly darker compared to AVTMB 3 and 4 (Table B2).

Table B2: Colour of the mungbean samples (n=5 field replicates for each variety). Varieties with the same superscript were not statistically different according to a post hoc Tukey test at α = 0.05.

Variety	Seed material			Flour		
	L*	a*	b*	L*	a*	b*
AVTMB 1	35.86 ± 0.78[a,b]	-4.03 ± 0.38	24.53 ± 1.00[a,b]	76.87 ± 1.41[a]	-3.39 ± 0.16	24.98 ± 0.58
AVTMB 2	37.93 ± 1.35[c]	-4.47 ± 0.28	26.69 ± 1.31[b]	78.01 ± 1.22[a,b]	-3.29 ± 0.31	25.26 ± 2.40
AVTMB 3	37.52 ± 1.02[b,c]	-4.29 ± 0.58	24.50 ± 1.87[a,b]	80.44 ± 0.60[b]	-3.49 ± 0.06	23.40 ± 0.86
AVTMB 4	34.12 ± 1.27[a]	-3.91 ± 0.72	22.12 ± 1.64[a]	80.24 ± 1.10[b]	-3.55 ± 0.11	22.93 ± 0.36
Jade	38.56 ± 0.53[c]	-4.68 ± 0.45	24.09 ± 1.27[a,b]	78.84 ± 1.92[a,b]	-3.43 ± 0.09	24.46 ± 1.36
P value	<0.001	0.145	0.002	0.002	0.203	0.050

L = 100 (white), L = 0 (black); +a = red, -a = green; +b = yellow, -b = blue

B5. Summary of field trial results

Overall, the field trial results suggest that new mungbean lines tested for adaptation to warmer northern Australia environment performed equal to or better than a currently established commercial line. Furthermore, the physical seed quality analysis demonstrated desirable characteristics in several of the new lines (e.g., large seed size, bright green colouration) which should make them highly desirable in the domestic and international markets.

Appendix C – Development of 96-well microplate methods

C1. FRAP microplate method development

As the benchtop FRAP assay used a measurement wavelength of 593 nm but the microplate reader was fitted only with a 595 nm filter, the first stage of method development was to ascertain whether there was any significant difference in absorbance values obtained at these two wavelengths. This was conducted by preparing triplicate sets of 7 Trolox standards (between 0-300 mg/L). These were analysed following the standard benchtop protocol for FRAP (as described in Section 3.3.5), with the resultant absorbances measured at 593 and 595 nm using the Genesys 10S UV-Vis spectrophotometer. There was no significant difference in the average absorbance of each standard concentration at 593 compared to 595 nm (paired samples *t*-test, $t_6 = -2.44$, $P>0.05$), indicating that a wavelength of 595 nm could be successfully used in the FRAP assay.

The second step in the method development process was to determine the impact of incubation temperature on the resultant absorbances. For this, two sets of Trolox standards (0-300 mg/L) were prepared following the benchtop protocol, with one set incubated in a waterbath at 37°C, and the other incubated at room temperature. Again, there was no significant difference between standards incubated at 37°C and those incubated at room temperature (paired samples *t*-test, $t_6 = -1.73$, $P>0.05$), demonstrating no significant effect of incubation temperature.

The third stage of method development was assessment of the repeatability of the microplate reader and microplate FRAP method. Duplicate absorbance measurements from the microplate reader – conducted on five Trolox standards (50-400 mg/L) each analysed in triplicate (n=15 standards in total) – demonstrated an average coefficient of variation (CV) of 0.30%, indicating a high level of instrument repeatability.

The repeatability of the FRAP microplate method was determined by calculating the average CV for the chickpea samples (each of which was analysed in duplicate). The average CV across these samples was 9.21%, indicating acceptable repeatability for the method.

Finally, the Trolox standards showed a high level of linearity ($R^2 = 0.998$; Figure C1) and their absorbances were strongly correlated between the benchtop and microplate methods ($R^2 = 0.987$; Figure C2). This provided a high level of certainty in the accuracy of the values obtained using this method.

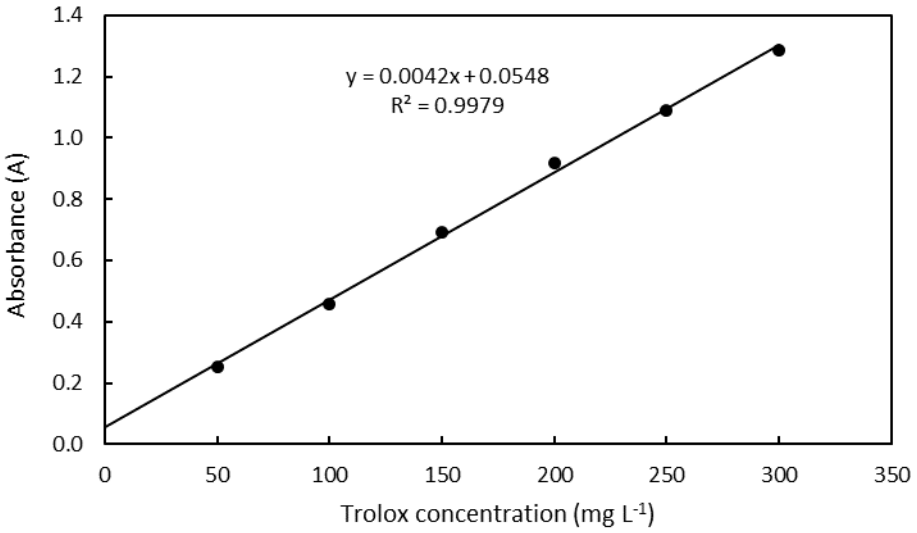

Figure C1: Linearity of the Trolox standards analysed using the FRAP microplate method.

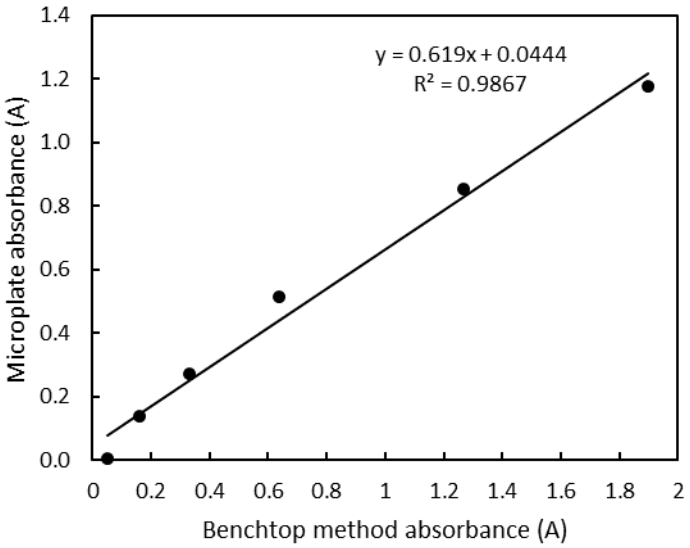

Figure C2: Correlation of the absorbance readings for Trolox standards using the benchtop and microplate methods for FRAP.

C2. CUPRAC microplate method development

The development of the CUPRAC microplate method involved similar steps to those described for the FRAP microplate method. However, as the microplate reader had a 450 nm filter, no change in the measurement wavelength was required.

To determine the impact of incubation temperature on the resultant absorbance, two sets of six Trolox standards (0-500 mg/L) were analysed using the benchtop method, with one set

incubated at 50°C for 30 minutes and one set incubated at room temperature. Although those incubated at 50°C showed significantly higher absorbance values (paired samples *t*-test, t_5 = 16.44, P<0.001), the magnitude of this difference was very small, at approximately 1.2% of the total absorbance. Furthermore, the absorbance values were highly correlated between the two incubation temperatures (R^2 = 0.994; Figure C3). Consequently, incubation at room temperature was considered appropriate for the microplate method.

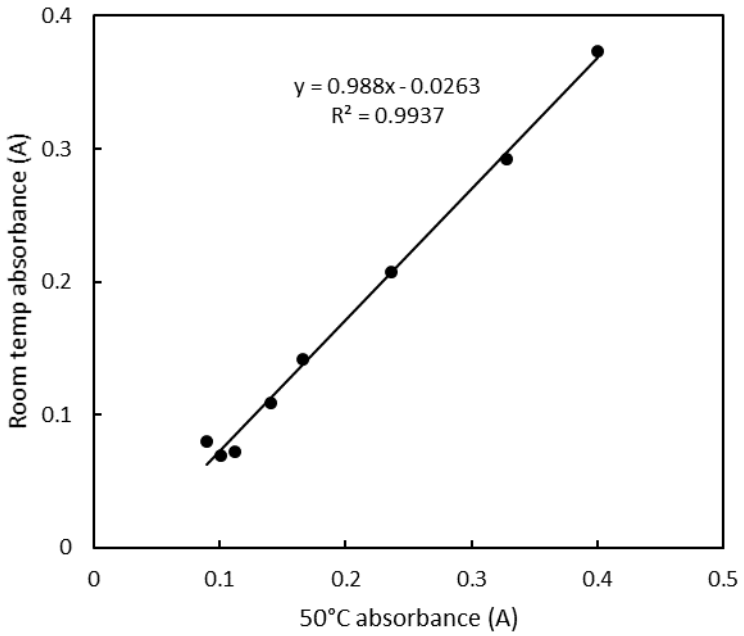

Figure C3: Correlation of the absorbance readings for Trolox standards incubated at 50°C and room temperature using the CUPRAC method.

The second stage of method development was determining the effect of incubation time, for incubation at room temperature. For this, nine Trolox standards (0-1000 mg/L) were analysed in quintuplicate using the microplate method (n=45 samples in total), with their absorbances measured every 5 minutes for one hour. The results are shown in Table C1.

While the absorbance values of the higher Trolox standards (300-1000 mg/L) showed very little variation throughout the 60 minute period, the absorbances of the blank (0 mg/L) and 100 mg/L standard increased considerably over this time. Most of this change took place in the first 45 minutes for 0 mg/L and the first 25 minutes for 100 mg/L. Consequently, 30 minutes incubation was chosen as a trade-off between time efficiency and the maximum absorbance development.

Table C1: Time dependence of the absorbance of Trolox standards using the CUPRAC microplate method. Results are given as the ratio of absorbances compared to the initial absorbance measured after 5 minutes incubation.

Time (mins)	Trolox concentration (mg/L)								
	0	100	200	300	400	500	600	800	1000
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
10	1.006	1.055	1.004	0.993	0.985	0.990	0.986	0.987	0.986
15	1.036	1.117	1.022	1.001	0.988	0.990	0.976	0.993	0.985
20	1.046	1.143	1.030	1.007	0.998	1.000	0.977	0.995	0.989
25	1.078	1.156	1.045	1.023	0.995	1.004	0.981	0.990	0.997
30	1.092	1.140	1.043	1.019	0.996	1.003	0.981	0.992	0.997
35	1.102	1.111	1.042	1.010	0.996	1.023	0.986	0.989	0.987
40	1.127	1.118	1.036	1.012	0.999	1.019	0.986	0.991	0.987
45	1.165	1.118	1.044	1.011	1.002	1.013	0.992	0.993	0.985
50	1.160	1.114	1.050	1.016	1.006	1.011	1.001	0.992	0.985
55	1.170	1.160	1.063	1.012	1.007	1.017	0.994	0.992	0.985
60	1.185	1.165	1.073	1.013	1.010	1.024	0.999	0.991	0.984
% change	18.5	16.5	7.3	1.3	1.0	2.4	-0.1	-0.9	-1.6

Again, the microplate reader possessed a high level of repeatability in the absorbance readings, with duplicate readings of the 45 Trolox standards showing an average CV of 0.59%. The repeatability of the CUPRAC microplate method was determined from duplicate analysis of the 97 chickpea samples, with the CV calculated to be 8.5%.

Finally, the microplate method showed good linearity for the Trolox standards (R^2 = 0.999; Figure C4) and the absorbance values obtained through this method were strongly correlated with those obtained using the benchtop method (R^2 = 0.971; Figure C5). Consequently, the optimised microplate CUPRAC method was deemed suitable for the analysis of antioxidant activity in the chickpea extracts.

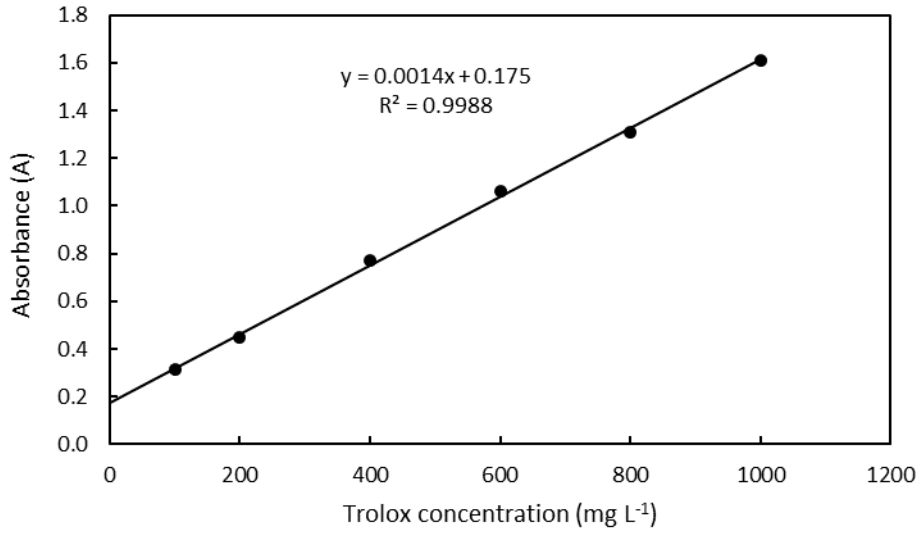

Figure C4: Linearity of the Trolox standards analysed using the CUPRAC microplate method.

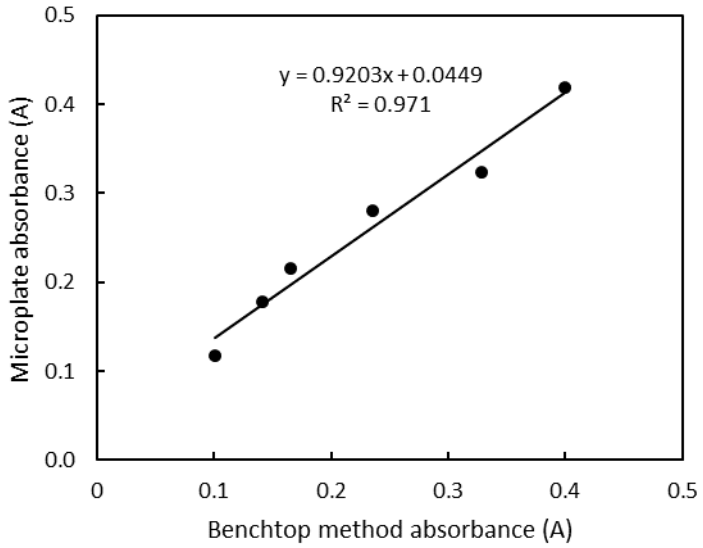

Figure C5: Correlation of the absorbance readings for Trolox standards analysed using the benchtop and microplate methods for CUPRAC.

Appendix D – Chickpea sample details

The sample details and physical parameters (hundred kernel weight and flour colour) for the 97 desi chickpea samples analysed are provided in Table D1.

Table D1: Sample details and physical parameters for the 97 desi chickpea samples.

ID	Location	Year	Variety	Trial type^	Treatment	HKW (g/100)	Flour colour		
							L*	a*	b*
CPJ-01	Rupanyup	2017	Howzat	D	Chem1	21.20	79.86	2.14	25.19
CPJ-02	Rupanyup	2017	Howzat	D	Chem2	19.42	79.83	2.21	23.51
CPJ-03	Rupanyup	2017	PBA Maiden	D	Chem2	21.85	81.21	1.61	23.93
CPJ-04	Rupanyup	2017	PBA Maiden	D	Chem2	20.77	82.34	1.33	24.53
CPJ-05	Rupanyup	2017	PBA Slasher	D	Chem1	20.81	81.70	1.45	23.56
CPJ-06	Banyena	2017	Kyabra	P	PathA	21.56	81.40	1.34	25.68
CPJ-07	Banyena	2017	Kyabra	P	PathA	21.44	78.62	2.11	24.89
CPJ-08	Banyena	2017	Kyabra	P	PathA	23.02	80.75	1.56	25.42
CPJ-09	Banyena	2017	Kyabra	P	PathA	22.84	80.87	1.19	24.72
CPJ-10	Banyena	2017	Kyabra	P	PathA	22.25	80.66	1.91	26.02
CPJ-11	Banyena	2017	Kyabra	P	PathB	21.87	79.71	1.57	24.51
CPJ-12	Banyena	2017	Kyabra	P	PathB	25.05	81.02	1.33	25.68
CPJ-13	Banyena	2017	Kyabra	P	PathB	18.75	80.49	1.19	24.93
CPJ-14	Banyena	2017	Kyabra	P	PathB	26.64	79.95	1.69	26.03
CPJ-15	Banyena	2017	PBA Slasher	P	PathA	19.17	80.43	1.62	23.48
CPJ-16	Banyena	2017	PBA Slasher	P	PathA	18.54	80.16	1.74	24.28
CPJ-17	Banyena	2017	PBA Slasher	P	PathA	17.24	79.98	1.85	25.34
CPJ-18	Banyena	2017	PBA Slasher	P	PathA	17.85	79.31	1.84	24.27
CPJ-19	Banyena	2017	PBA Slasher	P	PathA	18.33	79.33	1.93	24.59
CPJ-20	Banyena	2017	PBA Slasher	P	PathB	18.45	79.79	2.07	25.31
CPJ-21	Banyena	2017	PBA Slasher	P	PathB	19.40	80.12	1.70	25.19
CPJ-22	Banyena	2017	PBA Slasher	P	PathB	18.85	78.93	1.68	25.20
CPJ-23	Banyena	2017	PBA Slasher	P	PathB	17.25	80.23	1.60	25.14
CPJ-24	Banyena	2017	PBA Striker	P	PathA	22.01	80.77	1.48	25.82
CPJ-25	Banyena	2017	PBA Striker	P	PathA	19.28	78.88	1.66	24.73
CPJ-26	Banyena	2017	PBA Striker	P	PathA	21.67	80.88	1.44	24.54
CPJ-27	Banyena	2017	PBA Striker	P	PathB	18.44	79.25	1.52	25.57
CPJ-28	Banyena	2017	PBA Striker	P	PathB	15.93	76.97	1.84	24.20
CPJ-29	Banyena	2017	PBA Striker	P	PathB	17.73	78.45	1.79	24.36
CPJ-30	Banyena	2017	PBA Striker	P	PathB	18.76	78.06	1.98	23.83
CPJ-31	Horsham	2019	Kyabra	C	Nil	23.83	80.99	2.04	28.42
CPJ-32	Horsham	2019	Kyabra	C	Nil	24.30	81.66	2.11	27.89
CPJ-33	Horsham	2019	Kyabra	C	Nil	23.95	81.30	1.89	26.03
CPJ-34	Horsham	2019	PBA Boundary	C	Nil	21.03	80.99	1.15	27.80
CPJ-35	Horsham	2019	PBA Boundary	C	Nil	18.77	79.29	1.86	27.27
CPJ-36	Horsham	2019	PBA Boundary	C	Nil	18.69	80.00	1.72	28.45
CPJ-37	Horsham	2019	PBA Hattrick	C	Nil	17.90	80.00	1.48	27.94
CPJ-38	Horsham	2019	PBA Hattrick	C	Nil	19.73	79.87	1.30	28.49
CPJ-39	Horsham	2019	PBA Hattrick	C	Nil	19.00	79.49	2.21	28.94

CPJ-40	Horsham	2019	PBA Striker	C	Nil	22.35	81.62	1.06	27.39
CPJ-41	Horsham	2019	PBA Striker	C	Nil	23.55	81.13	1.13	27.20
CPJ-42	Horsham	2019	PBA Striker	C	Nil	21.54	80.34	1.68	27.16
CPJ-43	Horsham	2019	Rupali	C	Nil	19.60	79.42	2.19	26.91
CPJ-44	Horsham	2019	Rupali	C	Nil	18.59	81.59	1.77	28.22
CPJ-45	Horsham	2019	Rupali	C	Nil	19.35	81.76	1.87	26.73
CPJ-46	Curyo	2019	Howzat	D	Chem1	17.05	78.53	1.93	28.20
CPJ-47	Curyo	2019	Howzat	D	Chem1	17.14	78.92	1.46	27.93
CPJ-48	Curyo	2019	Howzat	D	Chem1	18.86	75.96	2.90	28.04
CPJ-49	Curyo	2019	Howzat	D	Nil	17.24	75.98	1.23	26.86
CPJ-50	Curyo	2019	Howzat	D	Nil	16.83	76.75	1.61	26.87
CPJ-51	Curyo	2019	Howzat	D	Nil	17.67	78.15	0.15	27.73
CPJ-52	Curyo	2019	PBA Striker	D	Chem1	17.67	77.83	1.85	26.00
CPJ-53	Curyo	2019	PBA Striker	D	Chem1	20.14	79.86	1.39	27.19
CPJ-54	Curyo	2019	PBA Striker	D	Nil	18.43	79.26	0.19	24.87
CPJ-55	Curyo	2019	PBA Striker	D	Nil	17.77	76.86	0.15	28.95
CPJ-56	Curyo	2019	PBA Striker	D	Nil	17.59	77.85	0.25	25.31
CPJ-57	Curyo	2019	Sonali	D	Chem1	15.36	78.50	1.20	26.26
CPJ-58	Curyo	2019	Sonali	D	Chem1	16.31	77.40	1.65	26.49
CPJ-59	Curyo	2019	Sonali	D	Chem1	17.21	77.21	1.69	26.24
CPJ-60	Curyo	2019	Sonali	D	Nil	16.77	78.25	0.83	26.40
CPJ-61	Curyo	2019	Sonali	D	Nil	16.63	77.53	1.16	26.09
CPJ-62	Curyo	2019	Sonali	D	Nil	15.53	76.34	1.18	27.00
CPJ-63	Horsham	2019	PBA Striker	D	Nil	22.77	79.20	1.35	27.88
CPJ-64	Horsham	2019	PBA Striker	D	Chem3	26.08	78.93	1.62	28.03
CPJ-65	Horsham	2019	Sonali	D	Chem4	16.95	76.65	1.56	26.86
CPJ-66	Horsham	2019	Sonali	D	Chem1	17.01	77.57	2.14	26.17
CPJ-67	Horsham	2019	Amethyst	V	Nil	15.17	78.83	1.95	28.58
CPJ-68	Horsham	2019	Genesis 509	V	Nil	16.91	77.30	2.59	28.53
CPJ-69	Horsham	2019	Genesis 836	V	Nil	16.95	80.87	0.99	28.98
CPJ-70	Horsham	2019	Howzat	V	Nil	21.56	78.24	2.85	29.25
CPJ-71	Horsham	2019	Jimbour	V	Nil	19.51	76.95	2.01	30.26
CPJ-72	Horsham	2019	Kyabra	V	Nil	24.33	81.26	2.42	28.02
CPJ-73	Horsham	2019	Neelam	V	Nil	18.73	77.74	2.37	29.83
CPJ-74	Horsham	2019	PBA Drummond	V	Nil	21.71	76.94	2.71	26.06
CPJ-75	Horsham	2019	PBA Hattrick	V	Nil	19.44	79.71	0.70	28.30
CPJ-76	Horsham	2019	PBA Maiden	V	Nil	25.03	81.37	1.74	28.09
CPJ-77	Horsham	2019	PBA Seamer	V	Nil	21.96	81.81	1.01	28.76
CPJ-78	Horsham	2019	PBA Slasher	V	Nil	20.44	79.97	2.08	28.20
CPJ-79	Horsham	2019	PBA Striker	V	Nil	22.55	80.84	1.64	27.57
CPJ-80	Horsham	2019	Sonali	V	Nil	18.21	77.49	1.99	26.36
CPJ-81	Horsham	2019	Tyson	V	Nil	13.19	79.32	1.39	28.27
CPJ-82	Horsham	2019	Yorker	V	Nil	20.22	80.78	2.57	27.14
CPJ-83	Horsham	2020	Amethyst	B	Nil	14.29	79.85	2.23	28.00
CPJ-84	Horsham	2020	Genesis 509	B	Nil	14.55	76.90	3.14	26.74
CPJ-85	Horsham	2020	Genesis 836	B	Nil	17.17	79.39	2.73	27.59
CPJ-86	Horsham	2020	Howzat	B	Nil	18.92	78.31	2.77	26.84
CPJ-87	Horsham	2020	Jimbour	B	Nil	18.30	79.71	2.39	29.20
CPJ-88	Horsham	2020	Kyabra	B	Nil	21.87	80.55	2.06	27.04
CPJ-89	Horsham	2020	PBA Drummond	B	Nil	20.09	79.78	2.21	26.24
CPJ-90	Horsham	2020	PBA Hattrick	B	Nil	18.60	78.34	2.23	26.40

CPJ-91	Horsham	2020	PBA Maiden	B	Nil	23.54	80.82	2.33	27.59
CPJ-92	Horsham	2020	PBA Seamer	B	Nil	21.30	81.09	2.15	28.52
CPJ-93	Horsham	2020	PBA Striker	B	Nil	21.31	79.07	2.35	27.74
CPJ-94	Horsham	2020	Sonali	B	Nil	15.99	80.25	1.66	26.59
CPJ-95	Horsham	2020	Tyson	B	Nil	13.32	79.77	2.63	27.82
CPJ-96	Horsham	2020	Yorker	B	Nil	18.98	80.17	3.04	28.42
CPJ-97	Curyo	2019	PBA Striker	D	Chem1	18.62	79.48	1.87	27.05

^ B = Variety bulk up trial; C = Cold tolerance trial; D = Disease management trial; P = Pathology trial; V = Variety demo trial