Application of infrared spectroscopy for the prediction of nutritional content and quality assessment of faba bean (*Vicia faba* L.)

Joel B. Johnson | Kerry Walsh | Mani Naiker

School of Health, Medical and Applied Sciences, CQUniversity, Bruce Hwy, North Rockhampton, Queensland, Australia

Correspondence
Joel B. Johnson, School of Health, Medical and Applied Sciences, CQUniversity, Bruce Hwy, North Rockhampton, QLD 4701, Australia. Email: joel.johnson@cqumail.com

Funding Information
Australian Government Research Training Program Scholarship

**Abstract**

With growing consumer interest and demand for health-benefiting functional foods such as faba beans, particularly evident in developed countries, commercial production of this crop is increasing. In concert with increased production levels comes an equally great need for the inexpensive rapid measurement of nutritional parameters for quality determining purposes. As an analytical tool, near-infrared spectroscopy has been well explored for the quantification of proximate nutritional composition, such as protein, starch and oil contents in faba bean and faba bean-derived products. Near-infrared spectroscopy has also been demonstrated to have potential for the noninvasive prediction of low-level micronutrients such as the total polyphenol content in faba bean and faba bean-derived products, although further exploration in this area is required to provide a more acceptable model. In some instances, the authors may be inadvertently measuring micronutrient concentrations through a secondary correlation with certain macronutrients. It is particularly difficult to determine if this is the case if exacerbated by the lack of an independent validation test set in the paper in question. The associated technique of mid-infrared spectroscopy shows particular promise for the rapid, noninvasive characterisation of structural components of faba bean, such as carbohydrates and proteins. Complementary applications of these two technologies are likely to yield a wealth of potential applications.

**KEYWORDS**

Fourier transformed infrared (FTIR) spectroscopy, attenuated total reflectance mid-infrared (ATR-MIR) spectroscopy, near-infrared spectroscopy (NIRS)

1 | INTRODUCTION

The modern consumer is more connected and informed than ever before, and one of the results of this is an increased consumer demand for functional foods. Functional foods are food sources that have the potential to provide health-benefiting effects in addition to those expected from traditional staple crops such as wheat. Faba bean (*Vicia faba* L.), also known as broad bean or fava bean, is one of the oldest cultivated crops (Singh & Bharati, 2013) and is an example of such a species that has benefited from this shift in consumer attitudes towards accepting functional foods. In particular, the high levels of antioxidant and phenolic compounds present in faba bean seeds have been linked to their health-benefiting effects, which include protection against radical species, antihypertensive and anticancer activity.
The seed material from faba bean may be consumed as is or utilized as an adjunct additive for the creation of novel value-added foodstuffs (López-Barrios, Gutiérrez-Uribe, & Serna-Saldívar, 2014; Vioque, Alaiz, & Girón-Calle, 2012). Traditionally consumed in countries in the Middle East and South East Asia (AEGIC, 2017), commercial production of this leguminous crop is steadily increasing in developed countries such as Australia, particularly over the last few decades (Siddique et al., 2000). Australia is now regarded as one of the top five producers and the largest exporter of faba bean in the world (Australian Export Grains Innovation Centre [AEGIC], 2017).

With both increased commercial production and greater consumer awareness and acceptance, there is an increasing need to assess and predict the overall nutritional quality of this crop in terms of protein and carbohydrate composition, as well as other bioactive/functional compounds of commercial crops for quality assurance purposes. However, most analytical procedures are either too expensive, require extensive sample preparation, or take long periods to perform. Hence, the rapid assessment and prediction of the overall nutritional quality of faba bean could be extensively applied amongst producers, buyers and manufacturers alike. Development of such nondestructive analytical tools would reduce the overall analytical costs and time, allowing for more representative sampling and subsequent analysis of the crops.

2 | INFRARED SPECTROSCOPY

Infrared spectroscopy is a noninvasive, nondestructive and rapid spectrophotometric analytical tool that is gradually being applied across many disciplines and to a range of matrices worldwide (Cozzolino, 2014; Johnson, 2020; Johnson & Naiker, 2019; Roggo et al., 2007). Infrared spectroscopy works on the principles of absorption due to molecular vibrations, using electromagnetic wavelengths with a lower frequency than light. The infrared spectrometer emits the full spectrum of infrared wavelengths, which penetrate the sample, with certain wavelengths absorbed by specific chemical bonds present within the sample. The amount of light energy absorption is directly proportional to the concentration or quantity of bonds present in the sample. From the reflected or transmitted wavelengths, the identity and quantity of the compounds present in the sample may be deduced.

There are two principal types of infrared spectroscopy used in food analysis, namely, near-infrared spectroscopy (NIRS) and mid-infrared (MIR) spectroscopy (MIRS) (Figure 1). MIRS is generally defined as the range of wavenumbers between 4,000 and 400 cm⁻¹ (or wavelengths of 2,500 to 25,000 nm), while NIRS incorporates the spectrum from wavelengths of 750 to 2,500 nm (Pasquini, 2003). Although the wavelengths below 1,000 nm mainly result from overlapped, relatively weak third overtones of chemical bonds (Dowell, Throne, & Baker, 1998), the longer pathlengths and opportunity to work with intact samples have made this technology invaluable for food researchers. For example, shortwave NIRS may penetrate centimetres into a sample. Transmission spectroscopy is often used with shortwave NIRS for the analysis of whole grains, as allowed for a more representative spectra to be collected from the sample. On the other hand, longwave NIRS penetrates only millimetres into the sample and hence is often used when analysing surface layers or homogenous materials. Reflectance or interactance spectroscopy is used in the latter instances, as the spectral information is collected from only the sample surface.

The majority of modern MIRS systems use a Fourier transform in order to simultaneously measure all wavelengths across the MIR spectrum, thus the term FTIR (Fourier transform infrared) spectroscopy will be used throughout this review to refer to mid-infrared spectroscopy. To provide greater signal amplitude and clarity, attenuated total reflectance (ATR) sampling is often used with FTIR spectroscopy. This necessitates direct contact between the sample and the ATR platform. As the signal amplitude depends to a large degree on the pressure applied, it can be quite difficult to obtain quantitative results from ATR–FTIR. The particle/grain size can have an impact on the spectra obtained, particularly for ATR–FTIR spectroscopy (Lee, Liong, &

FIGURE 1 Three commercial infrared spectrophotometers. (a) Nicolet Antaris FT-NIR (b) Bruker Alpha II FTIR with ATR platform (c) FOSS NIRSystems Model 6500
Jemain, 2017; Udvardi et al., 2017). However, grain size also impacts upon the spectra obtained from NIRS (Rinnan, Berg, & Engelsen, 2009; Wiley, Tanner, Chandler, & Anderssen, 2009). Data preprocessing methods performed on the spectra, such as the use of derivatives or multiplicative scatter correction, are often used to overcome this (Cozzolino, 2014; Lee et al., 2017). Nevertheless, there do not appear to be any systematic, controlled studies investigating the specific effects of particle size on the resultant spectra in cereal/pulse matrices.

Historically, NIRS has been the dominant form of infrared spectroscopy used for food analysis, largely due to the low instrumentation cost, high signal to noise ratio of the detector and the greater penetration of the infrared wavelengths into the sample matrix, due to the longer wavelengths used. Contemporary applications are increasing in their use of portable and in-line instrumentation. On the other hand, the MIR spectrum contains a larger array of more specific and characteristic absorption peaks for a range of functional group chemical bonds present (Cozzolino, 2014; Johnson, Collins, Skylas, & Naiker, 2019), hence has the potential to provide a more detailed fingerprint of the sample being analysed (Figure 2). The application of FTIR for the analysis of food crops has been steadily increasing, particularly over the past decade. For instance, over the past year, studies have reported using FTIR in the analysis of grain crops such as wheat (Johnson et al., 2019), mungbean (Johnson, Collins, Power, Chandra, Portman, Blanchard et al., 2020) and faba bean (Johnson, Collins, Skylas, Quail, Blanchard, & Naiker, 2020). This review focuses on faba bean and the historical and emerging application of both forms of infrared spectroscopy for the determination of its overall nutritional quality. This area was reviewed very briefly by Rodríguez Espinosa, Guevara-Oquendo, Sun, Zhang and Yu (2019); however, these authors focused only on selected applications of FTIR and did not include NIRS in their review.

3 | PROXIMATE NUTRITIONAL COMPOSITION

Proximate nutritional composition refers to the broad classes of macronutrients that make up the majority of foodstuffs and is usually determined through relatively basic analytical procedures. Aspects of the proximate composition include moisture, crude protein content, ash, crude fat and crude fibre content. NIRS has been widely used for the determination of numerous aspects of proximate nutritional composition in most crops, including faba bean (Font, del Río-Celestino, & de Haro-Bailón, 2006; Williams, Stevenson, Starkey, & Hawtin, 1978).

3.1 | Protein

As has been the case with other grain crops such as wheat, the quantification of protein content was one of the first applications of NIRS in faba bean. By the late 1970s, Williams et al. (1978) had reported the development of a reflectance NIRS model (utilising the ratio of absorption at 2,180 nm to that at 2,100 nm) that was used to predict the level of protein in ground faba bean samples (Table 1). The calibrated protein prediction model had an $R^2$ of 0.96, indicating that nearly all points fell on a straight line. The root mean square error (RMSE), a measure of the differences between the predicted and true values, was 0.56. The coefficient of variability was around 1.5%, comparing favourably to 1.2% for the Kjeldahl method. Analysis speed

![Figure 2](https://via.placeholder.com/150)

**Figure 2** Locations of selected functional groups in FTIR spectra collected from milled faba bean flour. Note the spectral variation between the three commercial Australian faba bean varieties illustrated here
was reported as >200 samples per instrument per day, although with use of in-line apparatuses this could easily exceed thousands of samples per day. Nowadays, commercial instruments including in-built calibration models are available for the determination of protein, fibre, fat, ash and moisture (BUCHI, 2019).

El-Sherbeeny and Robertson (1992) used NIRS (wavelength range not stated) to measure the protein content of 840 faba bean lines. The calibration sample size was quite small, at only 50 samples, with validation performed using the Kjeldahl method for total protein determination for every tenth sample. In addition, the authors of this work did not report whether whole or powdered faba bean samples were used. Protein concentrations for the samples included in the calibration curve ranged from 18.0% to 31.1%. The results obtained from these researchers showed NIRS to be quite accurate, with the standard deviation between the values obtained from NIR and Kjeldahl methods at 0.28%, while the coefficient of variability was 1.13%. Lepse, Dane, Zeipinca, Dominguez-Perles and Rosa (2017) also determined protein contents via NIRS in their attempt to determine optimise crop combinations of faba bean and a range of vegetables (e.g., onion, cabbage and carrot) for the greatest protein yield per hectare, but did not report any measures of the error of prediction.

Wang, Liu and Ren (2014) reported a NIRS model for the determination of protein in ground faba bean seed powder with an $R^2$ of 0.94 and root mean square error of cross-validation (RMSECV) of 0.34% (leave-one-out cross-validation). They also demonstrated that protein content could also be predicted from the intact seed, but with lower accuracy ($R^2 = 0.76$, RMSECV = 0.60%). This can be attributed to the greater heterogeneity of the intact seed, as only the outer layer of the seed volume is sampled, which in turn reduces the reproducibility of the infrared spectra and decreases the prediction accuracy of the model created.

Another aspect that deserves consideration in terms of protein content is the structural and conformational composition, which may influence the availability and digestibility of the protein when consumed, and hence the bioavailability of potential nutritional health benefits of the faba bean crop. Rodriguez Espinosa (2018) applied FTIR spectroscopy to quantify the amounts of various protein structures in ground faba bean seeds (Table 2). Significant differences in most measures of protein molecular structure were found between low tannin and normal tannin containing faba bean genotypes, particularly in the ratios of Amide I: Amide II bonds and the amount of $\beta$-structures, highlighting the potential interaction between tannins

| TABLE 1 | Applications of NIRS in the faba bean crop |
|---|---|---|---|---|---|---|
| Analyte | Matrix | Geometry | Wavelengths | Accuracy | References |
| Protein | Milled seed | Reflectance | 2,180 & 2,100 nm | RMSE 0.56% | (Williams et al., 1978) |
| Protein | Not stated | Reflectance | 875–2,357 nm | RMSECV 0.34% | (Wang et al., 2014) |
| Protein | Milled seed | Reflectance | 875–2,357 nm | RMSECV 0.34% | (Wang et al., 2014) |
| Protein | Whole seed | Reflectance | 875–2,357 nm | RMSECV 0.60% | (Wang et al., 2014) |
| Protein | Not stated | Reflectance | Not stated | CV 1.13% | (El-Sherbeeny & Robertson, 1992) |
| Protein | Milled seed | Reflectance | 1,319–2,013 nm | RMSECV 0.72% | (Wang et al., 2014) |
| Protein | Whole seed | Reflectance | 1,319–2,013 nm | RMSECV 0.72% | (Wang et al., 2014) |
| Protein | Whole seed | Reflectance | 1,319–2,013 nm | RMSECV 0.17% | (Wang et al., 2014) |
| Starch | Milled seed | Reflectance | 1,639–1,836 nm | RMSECV 0.17% | (Wang et al., 2014) |
| Starch | Whole seed | Reflectance | 1,639–1,836 nm | RMSECV 0.18% | (Wang et al., 2014) |
| Tannins | Whole seed | Reflectance | 1,100–2,500 nm | SEP 0.54% | De Haro et al. (1988) |
| Total polyphenols | Milled seed | Reflectance | 801–1,641 nm | RMSECV 0.40 mg/g | (Wang et al., 2014) |
| Total polyphenols | Whole seed | Reflectance | 801–1,641 nm | RMSECV 0.42 mg/g | (Wang et al., 2014) |
| Glycine betaine | Leaflets | Reflectance | Not reported | RPD 1.81 | Ali et al. (2016) |
### TABLE 2  Applications of mid-infrared spectroscopy (MIRS) in the faba bean crop

| Application                                                                 | Specific analytes | Matrix            | Key wavenumbers                      | References                                      |
|----------------------------------------------------------------------------|-------------------|-------------------|--------------------------------------|------------------------------------------------|
| Determining effects of steam pressure heat treatment on protein secondary structure | Amide I & II bonds α-helices β-structures | Milled seed       | 1,718–1,480 cm\(^{-1}\) 1,647 cm\(^{-1}\) 1,627 cm\(^{-1}\) | (Rodriguez Espinosa, 2018)                       |
| Detecting protein molecular conformation changes following in vitro rumen digestion | Amide I & II bonds α-helices β-sheets | Milled seed       | 1,641, 1,535 cm\(^{-1}\) 1,644 cm\(^{-1}\) 1,637 cm\(^{-1}\) | (Deng et al., 2019)                              |
| Determining effects of ultrasound treatment on protein secondary structure  | Amide I, II & III bonds α-helices β-sheets β-turns Anti-parallel β-sheets Inter- and intra-molecular aggregates | Protein isolate | 1,636, 1,524 and 1,250 cm\(^{-1}\) 1,654 cm\(^{-1}\) 1,634 cm\(^{-1}\) 1,666 cm\(^{-1}\) 1,680 cm\(^{-1}\) 1,618 and 1,694 cm\(^{-1}\) | (Martínez-Velasco et al., 2018)                   |
| Determining effects of high pressure homogenization on protein secondary structure | β-sheets α-helices β-turns | Protein isolate   | 1,618, 1,632, 1,639, 1,683, 1,693 cm\(^{-1}\) 1,648, 1,658 cm\(^{-1}\) 1,670, 1,693 cm\(^{-1}\) | (Yang et al., 2018)                              |
| Characterisation of polymers produced from faba bean protein               | None              | Polymers of isolated protein and glycerol | 4,000–500 cm\(^{-1}\) | (Montalvo-Paquini et al., 2014)                  |
| Investigation of changes in protein structure in two new varieties following in vitro rumen digestion | Amide I and II bonds α-helices β-sheets | 6 μm cross-sections of seed | 1,662–1,641, 1,552–1,529 cm\(^{-1}\) 1,660–1,641 cm\(^{-1}\) 1,635–1,621 cm\(^{-1}\) | (Rahman et al., 2019b)                           |
| Comparing starch crystallinity of faba bean to that of black bean and pinto bean | Ordered double helix domains Amorphous domains | Isolated starch | 1.047 cm\(^{-1}\) 1.022 cm\(^{-1}\) | (Ambigaipalan et al., 2011)                      |
| Determining effect of gamma irradiation on short-range crystalline order of starch | Amorphous domains | Isolated starch | 1.018 cm\(^{-1}\) | (Sofi et al., 2013)                             |
| Investigating responses of starch to combined heat and moisture treatment  | Double helix domains Amorphous domains | Isolated starch | 1.048 cm\(^{-1}\) 1.016 cm\(^{-1}\) | (Ambigaipalan et al., 2014)                      |
| Investigating changes in starch granular architecture following various annealing and heat moisture treatment | Inter- and intra-molecular protein interactions Crystalline/amorphous starch | Flour | 1,625, 1,616 cm\(^{-1}\) 1,047, 1,022 cm\(^{-1}\) | (Chávez-Murillo et al., 2018)                     |
| Comparing carbohydrate structure between regular and low tannin varieties | Structural carbohydrates Cellulosic carbohydrates Total carbohydrates | Milled seed | 1,518, 1,445, 1,390 cm\(^{-1}\) 1,235 cm\(^{-1}\) 1,145, 1,076, 1,015 cm\(^{-1}\) | (Rodriguez Espinosa, 2018)                        |
| Investigating differences in carbohydrate structure between two varieties  | Structural carbohydrates Cellulosic carbohydrates Total carbohydrates | Flour | 1,451, 1,394, 1,238 cm\(^{-1}\) 1,237 cm\(^{-1}\) 1,147, 1,075, 1,012 cm\(^{-1}\) | (Rahman et al., 2019a)                           |
| Profiling fatty acid composition of oil extracted from Sudanese faba bean   | cis double bonds cis olefinic groups | Oil extracted from seed | 3,008 cm\(^{-1}\) 1,376, 792, 765, 723 cm\(^{-1}\) | (Khalil et al., 2017)                            |
| Discrimination of white and green faba bean via protein profiles           | α-helices β-sheets Unordered structures |                | 1,700–1,600 cm\(^{-1}\) | (Xu et al., 2015)                               |
| Discrimination of three cultivars                                          | None              | Pods              | 2,892–2,643, 1,711–1,032 cm\(^{-1}\) | Terouzi et al. (2017)                            |
| Discrimination between regular and low tannin varieties                    | None              | Whole seed        | Not stated                          | (Rodriguez Espinosa, 2018)                       |
| Discrimination between two growing years                                    | None              | Milled seed       | Greatest influence at 1688 cm\(^{-1}\) | (Johnson et al., 2020)                           |

(Continues)
and nutritional quality. The Amide I bond results mainly from C=O
stretch and its level of absorption is modulated by the protein sec-
dondary structure (Byler & Susi, 1986). Amide II is produced by C–N
stretch and N–H bend, and its absorption varies with the level of
hydrogen interactions. Steam pressure heat treatment had a greater
effect on the protein structure of low tannin containing varieties com-
pared with normal tannin varieties. In general, steam pressure
increased the levels of amide bonds, while microwave irradiation
resulted in a decrease. Using traditional techniques such as X-ray crys-
tallography would be prohibitive for studies such as this. Fortunately,
FTIR provides a cheap, rapid alternative of gaining insight into the
behaviour of such functional molecules.

Subsequently, Deng et al. (2019) investigated the relative
amounts of α-helices, β-pleated sheets, Amide I and Amide II bonds
present in faba bean following a simulated in vitro rumen digestion.
The ratio of Amide I: Amide II peak area was positively correlated with
the soluble fraction, but negatively correlated with the degradable
fraction, rumen bypass feed crude protein and rumen undegradable
protein. The study highlighted the strength of FTIR for detecting
molecular conformational changes.

Martínez-Velasco et al. (2018) used FTIR to determine changes in
the secondary structure of faba bean proteins following ultrasound
treatment. This included the quantification of Amide I, II and III bonds,
and the relative proportions of intermolecular aggregates, α-helices,
β-pleated sheets, β-turns and anti-parallel β-sheets. Quantification of
the latter protein secondary structures (α-helices and β-structures) is
particularly important, given that these structures affect the gelling
strength and emulsion properties of the subsequent flour (De la Rosa-
Millan et al., 2015; Doiron, Yu, McKinnon, & Christensen, 2009). It
was found that high-intensity ultrasound treatment increased the
levels of β-sheets and marginally increased the levels of α-helices, but
markedly decreased the amount of inter-molecular aggregates.

In a similar study, Yang, Liu and Zeng (2018) used FTIR to identify
protein conformational changes following the high-pressure homoge-
nization of faba bean protein. Combined with sizing distribution stud-
ies that used dynamic laser scattering, the authors demonstrated that
high-pressure homogenization modulated interaggregate hydrophobic
interactions lead to dissociation of large protein aggregates into more
soluble supramolecular aggregates.

FTIR has also been used in the characterisation of polymeric films
produced from faba bean protein concentrate and glycerol (Montalvo-
Paquini, Rangel-Marrón, Palou, & López-Malo, 2014). Analysis of peak
characteristic of protein and glycerol showed no change, indicating
that the protein was not covalently reacting with the glycerol in any
major fashion.

Rahman, Refat, Zhang, Zhang and Yu (2019) used synchrotron
radiation-based Fourier transform infrared microspectroscopy to
model the changes in protein structure and composition associated
with digestion in a model ruminant system. Due to the intensity of the
synchrotron light source, this technique provides a greater spectral
resolution and higher accuracy over regular FTIR spectroscopy. The
apparatus used 6-μm cross-sections of sample for transmission spec-
troscopy, hence requiring somewhat more sample preparation over
techniques such as ATR–FTIR. Based on the protein structural regions
of the spectrum, two faba bean varieties could be discriminated using
principal component analysis and hierarchical cluster analysis. Both
crude protein and true digestibility of the samples could be predicted
from the ratio of α-helix and β-sheets.

| Application | Specific analytes | Matrix | Key wavenumbers | References |
|-------------|------------------|--------|-----------------|------------|
| Characterisation of chemical composition of polyurethane foam manufactured from faba bean stalks | None | Polyurethane foam/faba bean stalks | 3,421–863 cm⁻¹ | (Zhang et al., 2014) |
| Exploring physiological response of faba bean saplings to arsenic toxicity | None | Freeze-dried and lyophilized sapling roots | 3,000, 2,800, 2,400, 2000, 1,540, 1,160, 830 cm⁻¹ | (Boccia et al., 2013; Sturchio et al., 2012) |
| Exploring physiological response of faba bean plants to aluminium toxicity | None | Roots, stems and leaves | 3,375, 2,928, 1,645, 1,428, 1,245, 1,060 cm⁻¹ | (Wang et al., 2011) |
| Exploring physiological response of faba bean saplings to essential oil treatment | α-helices, β-turns, β-sheets | Freeze-dried and lyophilized sapling roots | 1,645–1,662 cm⁻¹, 1,613–1,637 cm⁻¹, 1,662–1,685 cm⁻¹, 1,637–1,645 cm⁻¹ | (Mecozzi & Sturchio, 2015) |
| Discrimination between faba bean roots and wheat roots collected from soil cores | None | Dried, ground roots | 3,400–2,400, 1,750–1,200, 1,070–950, 860–400 cm⁻¹ | (Streit et al., 2019) |
1,940 to 1,800 nm) for the prediction of moisture content in ground faba bean samples, giving a mean coefficient of variation of 2.9% weight/fresh weight. Considering the limited precision of the NIRS instrumentation available at the time, this is a remarkably high accuracy. Although no recent studies appear to have reported the use of NIRS for the prediction of moisture content in faba bean, it is expected that the accuracy would have improved greatly over the past 40 years.

### 3.3 Starch and other carbohydrates

As one of the most important functional biopolymers, the structural properties of starch plays an integral role in determining the overall nutritional quality of all grain and legume-based products, including that for faba bean. Wang et al. (2014) used NIRS to predict the starch concentration of ground faba bean with reasonable accuracy \((R^2 0.86, \text{RMSEC} 0.72, \text{RPD} [\text{ratio of performance to deviation}] 2.64)\). However, although the sample sizes were reasonable (calibration set of 203 samples and a validation set of 41 samples), only leave-one-out cross-validation was used in this study.

Ambigaipalan et al. (2011) used ATR–FTIR to quantify the relative amounts of ordered crystalline double helix domains to that of amorphous domains in starches isolated from faba bean, through using the ratio of absorbances at 1,047 and 1,022 cm\(^{-1}\). ATR–FTIR is particularly suitable to this work, not only due to the rapid analysis times but also as it has been demonstrated to respond to changes in molecular structure in short range order (i.e., the double-helical order) (Sevenou, Hill, Farhat, & Mitchell, 2002). However, at such long wavelengths the mid-infrared wave only penetrates around 2 \(\mu m\) into the sample, thus the results are only representative of the starch granule surfaces (Ambigaipalan, Hoover, Donner, & Liu, 2014). The level of crystallinity in faba bean starch (mean ratio of crystalline to amorphous domains of 0.886 ± 0.016) was shown to be less than that of black bean or pinto bean.

Sofi, Wani, Masoodi, Saba and Muzaffar (2013) found that gamma irradiation produced a breakdown of glycosidic bonds and reduced short-range crystalline order (i.e., double helices) of starch composition in faba bean. Both properties were simultaneously measured from the starch powder using FTIR spectroscopy.

Subsequent work, also using FTIR on starches isolated from faba bean seeds, investigated the combined effects of heat and moisture treatment on starch structural composition (Ambigaipalan et al., 2014) using a slightly different wavelength ratio to quantify crystalline and amorphous starch (1,048 to 1,016 cm\(^{-1}\)). Variation between faba bean genotypes was found in both the initial orderliness of the starch present (with ratios ranging from 0.772 to 0.889) and the level of decreased orderliness in response to heat and moisture treatment (11.3% to 13.2% decrease).

Other researchers have also used ATR–FTIR to assess changes in starch granular architecture following various hydrothermal treatments, in addition to their molecular interactions with the proteins present (Chávez-Murillo, Veyna-Torres, Cavazos-Tamez, de la Rosa-Millán, & Serna-Saldívar, 2018). Notably, it was found that the heat treatment produced new physical interactions between starch and proteins at a molecular scale, with the level of interaction correlated with the amount of resistant starch present. Once again, it should be noted that such real-time, submolecular studies would remain cost-prohibitive for most research facilities without the use of ATR–FTIR.

As with proteins, the molecular structure of carbohydrates can influence their digestibility. Hence, the determination of changes in carbohydrate structure between different varieties or following various treatments. Rodríguez Espinosa (2018) used FTIR as a rapid means of assessing carbohydrate structure in ground faba bean seeds, including the levels of structural carbohydrates, cellulosic carbohydrates and total carbohydrates. For most measures, minimal difference in carbohydrate structure was found between regular and low tannin faba bean varieties, although the regular tannin varieties did typically have slightly higher levels of total carbohydrates.

Rahman et al. (2019) used ATR–FTIR to investigate and quantify structural carbohydrates, cellulosic carbohydrates and total carbohydrate structures in whole faba bean flour. Although only two faba bean varieties were included in the study, it was observed that the higher tannin-containing variety had higher levels of structural and cellulosic carbohydrates compared to the low tannin variety.

### 3.4 Fatty acids and oils

Khalil, Salih and Mustafa (2017) used FTIR to assist in their investigation of the fatty acid composition of oil extracted from faba bean seeds that included the identification of both the cis-double bonds and cis-olefinic groups. A number of spectral variations were noted between different faba bean varieties, attributed to differences in the composition and chain length of fatty acids, including the degree and position of any double bonds present.

Wang et al. (2014) reported the development of an acceptable NIRS model for the prediction of oil content in ground faba bean samples \((R^2 0.66, \text{RMSEC} 0.17, \text{RPD} 1.72 \text{for cross-validation})\). The typical range and coefficient of variation of oil concentrations found in faba bean are 0.48%–1.99% and 0.27%, respectively. As previously noted in Section 3.3, this study used a relatively large sample size \((n = 203 \text{ for calibration})\) but only leave-one-out cross-validation for their models. Quite similar results were found for whole faba bean samples \((R^2 0.66, \text{RMSEC} 0.18, \text{RPD} 1.72)\).

### 4 QUANTIFICATION OF TANNINS AND POLYPHENOLS

One of the major groups of phytochemicals present in faba bean is the polyphenols. While most polyphenols are considered beneficial due to their positive cardiovascular effects, certain polyphenols such as condensed tannins are considered antinutritive as they can decrease the efficiency of nutrient uptake and metabolism (Chung, Wong, Wei, Huang, & Lin, 1998).
4.1 | Total polyphenols

Polyphenols are phytochemical micronutrients, most of which are well known for their health-benefiting effects (Shahidi & Ambigaipalan, 2015). As a traditional measure of polyphenol content, the total polyphenols in a sample are usually quantified using the Folin–Ciocalteu assay, which can be quite time consuming (Johnson et al., 2019). Wang et al. (2014) used NIRS to predict the total polyphenol content in ground faba bean, with an $R^2$ of 0.79, RMSECV of 0.40 and RPD of 2.20 for leave-one-out cross-validation. In contrast to proximate components such as moisture, protein and carbohydrates, no commercial instruments appear to be marketed for the purpose of polyphenol determination.

4.2 | Tannins

Tannins are a group of complex phenolic polymers derived from flavonol. As an important antinutritive component of faba bean seeds, the condensed tannin concentration, which can range from negligible to up to 7% w/w, is thus one factor that must be taken into consideration when developing new faba bean varieties (Helsper, Hoogendijk, van Norel, & Burger-Meyer, 1993; Marquardt, Ward, & Evans, 1978). As such, development of a reliable, rapid method for quantification of the tannin contents in faba bean would be greatly beneficial to the industry, particularly for the plant breeders. In order to assist in this area, De Haro, López-Medina, Cabrera and Martin (1988) demonstrated that NIRS can noninvasively be used to measure the tannin content in whole faba bean seeds. Tannins are largely located in the seed coat, ensuring that NIRS is measuring nearly all of the tannins present. Across the range of 0.01–7% w/fw tannin in the faba bean seeds measured, the best cross-validation results obtained had an $R^2$ value of 0.93 and standard error of prediction of 0.54%. The authors used a calibration set of 60 samples, with a similar-size validation set, but did not report the cross-validation method used. It should be noted that in this study, no calibration samples were included with tannin concentrations in the range of 1%–3.5%.

4.3 | Vicine and convicine

Two other important antinutritional factors present in faba bean are vicine and convicine, both of which are alkaloid glycosides (Burbano, Cuadrado, Muzquiz, & Cubero, 1995). If consumed in high levels by individuals suffering from a genetic mutation in the red blood cell enzyme glucose-6-phosphate (estimated to occur in around 5% of the world population), this can lead to a form of haemolytic anaemia known as favism (Burbano et al., 1995). Vicine and convicine are currently quantified using a lengthy HPLC procedure (Skylas et al., 2019), so a rapid analytical method would save considerable time. The major challenge results from the typically low concentrations of these compounds (around 0.6%–0.9% by weight). Nevertheless, Puspitasari (2017) did report the prediction of these compounds in faba bean flour using NIRS. The best performing prediction model gave an $R^2$ of 0.968 and standard error of cross-validation of 0.094%. Subsets of the calibration samples were used for validation purposes. The RPD, which should ideally be at least 3, ranged from 2.67–3.14 for the five validation subsets, leading the authors to suggest that NIRS would be suitable for screening purposes.

5 | ANALYSIS OF OTHER SPECIFIC COMPOUNDS

The majority of infrared-utilising studies incorporating the analysis of specific types of compounds have focussed on qualitative identification of the isolated compounds, rather than the characterisation and quantification of them within the faba bean matrix. However, with the invention of more precise spectrometers and powerful chemometric tools, it is envisaged that the measurement of specific compounds present within the matrix could be possible without any prior extraction and purification. For example, El-Haramein, Moneim, Nakkoul and Williams (1998) reported the successful quantification of the neurotoxin beta-N-oxalyl-amino-L-alanine in chickling vetch (Lathyrus spp.) at concentrations between 0.1%–0.8% using NIRS (sample size $n = 88$). Although the standard error of prediction was just 0.05% for whole grains, it does not appear that the authors used an independent test set for validation.

Infrared spectroscopy, particularly MIRS, is frequently used for the confirmation and quantification of individual compounds isolated from faba bean. A number of confirmatory identification utilising MIRS applications reported to date include that for (−)-jasmonic acid from faba bean pericarp (Dathe et al., 1981), saponins (Amarowicz, Yoshiki, & Okubo, 1998), (3Z)-nonenal (Gardner & Hamberg, 1993), a range of unsaturated fatty acids (Hamberg & Hamberg, 1990) and oligomeric proanthocyanidins (Deng, Chu, Liu, & Zhang, 2013).

6 | DISCRIMINATION BETWEEN VARIETIES AND GROWING LOCATIONS

Determining the variety or growing location of a faba bean sample can be an important aspect of quality assurance or authentication. In contrast to the previously mentioned studies, which report either quantification or identification of a compound using infrared spectroscopy, determination of variety/growing location is a qualitative method based on broad chemical differences. The principle behind the discrimination of different varieties and growing location/season is dependent on the different chemical compositions produced by the genotype × environment interaction. Both factors play a role in influencing the resultant chemical composition of the harvested faba bean seed, sometimes to varying degrees depending on the compounds in question (Skylas et al., 2019).

Wang et al. (2014) used Fourier transform near-infrared (FT-NIR) spectroscopy to identify faba bean cultivars grown in different locations across China, based on spectral features corresponding to
protein, starches, oil and polyphenols. The study was a once-off observation with no validation, with the clustering accuracy at around 79.5%, hence could be refined considerably. However, spectral responses were observed in crops grown at differing altitudes, longitudes and latitudes.

Also working on Chinese faba bean, Xu et al. (2015) reported the discrimination of white and green varieties using a combination of their polysaccharide and protein absorption bands, measured using FTIR, and their mineral contents, measured using ICP-MS (inductively coupled plasma mass spectrometry).

Terouzi, Gorfti and Oussama (2017) performed a similar study discriminating between three faba bean cultivars, but with discrimination solely based on ATR–FTIR-derived spectral data obtained from the faba bean pods. Distinctive clusters were recognised through PCA, while PCA–LDA (principal component analysis coupled with linear discriminant analysis) classified the three cultivars with 100% accuracy.

While the aforementioned study by Terouzi et al. (2017) incorporated spectral information from across the entire MIR spectrum, Rodríguez Espinosa (2018) applied principal component analysis (PCA) to solely the protein or carbohydrate regions of FTIR spectra obtained from faba bean seeds. This allowed discrimination between varieties to some degree, with cluster analysis showing partial separation between low tannin and regular tannin-containing varieties.

Finally, Johnson et al. (2020) used ATR–FTIR as a rapid means of profiling phytochemical differences between 10 varieties of Australian faba bean. Partial least squares discriminant analysis (PLS-DA) was able to classify samples by growing year with acceptable accuracy (~87%). However, attempts to classify the growing location with PLS–DA were met with less success (~60% accuracy). Further work is required in this area, particularly using a greater number of growing locations.

7 | APPLICATION OF INFRARED SPECTROSCOPY ON OTHER PORTIONS OF THE FABA BEAN PLANT

Several studies have investigated the use of infrared spectroscopy on other, nonedible portions of the faba bean plant. These are included here to highlight the versatility of the technology and illustrate the range of potential applications to which it can be applied.

7.1 | Leaves and stems

Whole faba bean plants were included amongst other plant species investigated by Bruun et al. (2005) to estimate total carbon, nitrogen and mineral contents in them using NIRS. The best performing of all models gave an R² value of 0.93 for carbon content but only 0.446 for nitrogen content. Although this study was performed with the intent of developing rapid analytical techniques to study carbon and nitrogen flux in decomposing plant samples, it is anticipated that considerably better results could be obtained using NIRS on a single species (i.e., faba bean), as this would greatly reduce the heterogeneity of the matrix and allow for spectral features relating to either carbon or nitrogen to be extracted with ease. It is likely that such a model could be created using either seed material or vegetative matter. In the latter instance, this could be used to assess the health of a crop during its growth and monitor responses to certain stressors or growing conditions.

Continuing with this quantitative aspect, Ali et al. (2016) used NIRS to develop a calibrated model to determine the levels of glycine betaine in faba bean leaflets. Ultra performance liquid chromatography (UPLC) coupled with electron spray ionisation (ESI) mass spectrometry (MS) was used as a reference validation method. The NIRS model used a modest-sized calibration set of 80 leaflet samples (ratio performance deviation, defined as SD/RMSEP, of 2.29 for calibration and 1.81 for cross-validation) and used to estimate the glycine betaine concentration in the remaining 109 samples. As glycine betaine level is a stress indicator of the plant, the NIRS-derived concentrations were used to effectively ascertain the most suitable faba bean varieties and strains for the inescapably harsh German climate.

Zhang et al. (2014) used FTIR to characterise the chemical composition of faba bean stalks and used the stalks as a feedstock material for the subsequent manufacture of a polyurethane foam. The foam showed promising compressive strength while retaining a very low density. New peaks identified in the polyurethane foam included those attributed to N–H, hydrogen-bonded carbonyl, δ(N–H), ν(C=N), C–H out-of-plane bending of benzene ring and O–C=O out-of-plane bending.

Isolated studies on faba bean leaves have been conducted through the light measurement tool spectroradiometry, typically using visible and shortwave NIR wavelengths (Malthus & Madeira, 1993). This technique has been used to detect the infestations of Botrytis fabae (Malthus & Madeira, 1993) and determine the radiation usage by the crop (the fraction of absorbed photosynthetically active radiation, commonly abbreviated as APAR) (Ridao, Conde, & Mínguez, M. I., 1998; Ridao, Oliveira, Conde, & Mínguez, 1996). This latter application could be used to remotely assess water stress and crop biomass, as shortwave NIR wavelengths can be obtained as satellite imagery.

7.2 | Roots

FTIR and Fourier transform near-infrared spectroscopy (FT-NIR) have been used to investigate the effective responses of faba bean saplings following exposure to high levels of inorganic arsenic (Bocca, Meconi, Mecozzi, & Sturchio, 2013; Sturchio, Napolitano, Beni, & Mecozzi, 2012). Spectral data were obtained from freeze-dried, lyophilized root meristems of the seedlings. Changes in the content of polysaccharides, lipids, aliphatic compounds, nucleic acids and proteins were determined, as well as altered levels of hydrogen bonding in these samples. Given the level of changes observed, both quantitative and structural changes in these compounds were likely to be occurring. In particular, it was believed that inorganic arsenic would replace...
the phosphate groups within DNA molecules, resulting in the proliferation of DNA damage being observed through genotoxicity assays.

Similar work on faba bean roots, stems and leaves used FTIR to investigate the potential mechanisms of aluminium toxicity on faba bean and determine varietal differences in aluminium tolerance (Wang et al., 2011). Through FTIR spectral features, a faba bean variety with predicted high susceptibility to aluminium toxicity was identified. This result was confirmed through the use of traditional heavy metal toxicity assays (root elongation experiment and chrome azurol dyeing).

With growing concerns surrounding pesticide resistance, essential oils have been proposed as a potential bioherbicidal agent. However, their components may have negative physiological effects on the crops as well as the target pest species. Following on from the previously reported work on arsenic and aluminium toxicity, Mecozzi and Sturchio (2015) explored the effects of essential oil treatments upon the protein structure in freeze-dried, lyophilized faba bean roots using diffuse reflectance FTIR. In addition to transitions in the secondary structure of the proteins (α-helices, β-turns and β-sheets), development of random coil structures and oxidation of proteins was observed. Subsequent work incorporating FT-NIR demonstrated further alterations in bonds relating to DNA and other nucleotides, carbohydrate backbones, proteins and lipids (Mecozzi et al., 2017).

One way to increase crop harvest while limiting resource consumption is through mixed cropping of multiple species. However, the underground interactions between the roots of different species are still poorly understood, in part due to the complexity of identifying the species of a root isolated from a soil core. With this in mind, Streit, Meinen, Nelson, Siebrecht-Schöll and Rauber (2019) applied ATR-FTIR to identify ground, dried roots collected from various soil cores. The faba bean roots were readily discriminated from wheat roots, with no apparent effect of genotype or growing year. The technique was subsequently used to investigate the overyielding effect of different faba bean genotypes grown in combination with winter wheat and thus identify the optimum cropping combination.

### 8 DEVELOPMENTS IN INFRARED SPECTROPHOTOMETERS

One major development that has occurred and continues to occur in this field is the reduction in the cost of spectrophotometers. These almost exclusively use NIR wavelengths, as the apparatus is much cheaper than that for MIRS. The lower prices have made this technology more accessible to small organisations or research groups, increasing the range of potential applications (Kosmowski & Worku, 2018). In concert with the decreasing price have come increasing the range of potential applications (Kosmowski & Worku, 2018). In the latter instance, MIRS can discriminate between secondary structures such as unordered aggregates, α-helices, β-pleated sheets, β-turns and anti-parallel β-sheets. Additionally, the greater sensitivity of MIRS is likely to make it a better candidate for the determination of micro-nutrients. However, this aspect remains largely unexplored. In terms of the faba bean crop, there remains much scope for future studies using these technologies, particularly in the area of functional foods.

### 9 CONCLUSION

Both MIRS and NIRS have been used in a range of applications for the prediction of quality assessment of faba bean and/or faba bean-derived products. While NIRS is a well-established analytical tool for the determination of proximate nutritional parameters such as protein, oil and starch content, MIRS is more proficient for establishing the structural composition of constituents such as starch and protein. In the latter instance, MIRS can discriminate between secondary structures such as unordered aggregates, α-helices, β-pleated sheets, β-turns and anti-parallel β-sheets. Additionally, the greater sensitivity of MIRS is likely to make it a better candidate for the determination of micro-nutrients. However, this aspect remains largely unexplored. In terms of the faba bean crop, there remains much scope for future studies using these technologies, particularly in the area of functional foods.

### ACKNOWLEDGEMENTS

This research was supported by an Australian Government Research Training Program Scholarship awarded to one of the authors (J. J.).

### CONFLICT OF INTEREST

All authors declare that no conflict of interest exists.

### AUTHOR CONTRIBUTIONS

J. J. conceptualised the study, conducted the literature review and prepared the original manuscript draft. All authors contributed to the critical review and editing of the manuscript.

### ETHICAL STATEMENT

This article does not contain any studies with human or animal subjects.

### DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analysed in this study.

### ORCID

Joel B. Johnson https://orcid.org/0000-0002-9172-8587
Kerry Walsh https://orcid.org/0000-0002-3033-8622
Mani Naiker https://orcid.org/0000-0002-6844-8325

### REFERENCES

AEGIC. (2017). Australian pulses: Quality, versatility, nutrition. Retrieved from https://www.aegic.org.au

Agelet, L. E. (2011). Single seed discriminative applications using near infrared technologies. (PhD thesis). Iowa State University, Ames, Iowa. Retrieved from http://lib.dr.iastate.edu/etd/12023 (12023).
All, M., Welna, G. C., Sallam, A., Martsch, R., Baliko, C., Gebser, B., ... Link, W. (2016). Association analyses to genetically improve drought and freezing tolerance of faba bean (Vicia faba L). Crop Science, 56(3), 1,036–1,048. https://doi.org/10.2135/cropsci2015.08.0503

Amarowicz, R., Yoshiki, Y., & Okubo, K. (1998). Two new saponins from faba bean (Vicia faba L.). Zeitschrift für Naturforschung C, 53(9–10), 918–920. https://doi.org/10.1515/znc1998-9-1022

Ambigaipalan, P., Hoover, R., Donner, E., & Liu, Q. (2014). Starch chain interactions within the amorphous and crystalline domains of pulse starches during heat-meal treatment at different temperatures and their impact on physicochemical properties. Food Chemistry, 143, 175–184. https://doi.org/10.1016/j.foodchem.2013.07.112

Ambigaipalan, P., Hoover, R., Donner, E., Liu, Q., Jaiswal, S., Chibbar, R., ... Preiss, A., Schade, W., Sembdner, G., & De Haro, A., López-Medina, J., Cabrera, A., & Martín, A. (1988). A new enol compound in Lycium chinense. Phytochemistry, 27(3), 535–537. https://doi.org/10.1016/S0031-9422(00)89274-1

Boccia, P., Meconi, C., Meccozzi, M., & Sturchio, E. (2013). Molecular modifications induced by inorganic arsenic in Vicia faba investigated by FTIR, FTRNIR spectroscopy and genotoxicity testing. Journal of Toxicological and Environmental Health, Part A, 76(4–5), 281–290. https://doi.org/10.1080/15287394.2013.757211

Bbruun, S., Stenberg, B., Breland, T. A., Gudmundsson, J., Henriksen, T. M., Jensen, L. S., ... Salo, T. (2005). Empirical predictions of plant material C and N mineralization patterns from near infrared spectroscopy, stepwise chemical digestion and C/N ratios. Soil Biology and Biochemistry, 37(12), 2,283–2,296. https://doi.org/10.1016/j.soilbio.2005.04.006

Buchi, L. (2019). ProxiMate™. NIR instrument for the harshest environments. Retrieved from https://www.buchi.com/gb-en/products/nir-solutions/proximate

Burbano, C., Cuadrado, C., Muzquiz, M., & Cubero, J. I. (1995). Variation of nutritional factors in Legume Seeds. Paper presented at the recent advances of research in antinutritional factors and human health: A review.

Chávez-Murillo, C. E., Veyna-Torres, J. I., Cavazos-Tamez, L. M., de la Cruz, J. M., & Gardiner, H. (2016). Examination of the secondary structure of proteins by deconvolved FTIR spectra. Biopolymers: Original Research and Reviews, 104(10), 1,036–1,048. https://doi.org/10.1002/bip.24336

Chavez-Murillo, C. E., Veyna-Torres, J. I., Cavazos-Tamez, L. M., de la Cruz, J. M., & Gardiner, H. (2016). Examination of the secondary structure of proteins by deconvolved FTIR spectra. Biopolymers: Original Research and Reviews, 105(5), 371–383. https://doi.org/10.1002/bip.24391

Chung, K.-T., Wong, T. Y., Wei, C.-I., Huang, Y.-W., & Lin, Y. (1998). Tannins and human health: A review. Critical Reviews in Food Science and Nutrition, 38(6), 421–464. https://doi.org/10.1080/1040869981274273

Cozzolino, D. (2014). 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

Dathe, W., Rönsch, H., Preiss, A., Schade, W., Sembdner, G., & Schreiber, K. (1981). Endogenous plant hormones of the broad bean, Vicia faba L. (–)jasonic acid, a plant growth inhibitor in pericarp. Planta, 153, 530–535. https://doi.org/10.1007/BF00385537

De Haro, A., López-Medina, J., Cabrera, A., & Martin, A. (1988). Determination of tannin in the seeds of Vicia faba by NIR. Wageningen: Paper presented at the recent advances of research in antinutritional factors in legume seeds. Proceedings 1st International Workshop on Anti-nutritional Factors in Legume Seeds.

De la Rosa-Millán, J., Lin, A. H. M., Osorio-Díaz, P., Agama-Acevedo, E., Hamaker, B. R., & Bello-Perez, L. A. (2015). Influence of annealing flows from raw and pre-cooked plantain fruit on cooked starch digestion.
