University of Nottingham
UK | CHINA | MALAYSIA

Winged bean – a new soybean for the tropics? Genomic analysis for improving nutritional value and breeding efficiency in *Psophocarpus tetragonolobus* seeds

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

October, 2024

Niki Tsoutsoura MSc.

14336257

Supervisors: Prof. John Brameld and Dr Sean Mayes

School of Biosciences

Faculty of Science

University of Nottingham

Table of Contents

Abstract

Climate change, population growth, and lack of nutritional diversity in diets present significant challenges to food security and human health. The need for alternative protein sources is urgent, and the underutilised winged bean (*Psophocarpus tetragonolobus* (L.) DC.) grown in tropical regions emerges as a promising candidate due to its high nutritional profile. With high protein content ranging from 30-40% and oil content of 15-20%, winged bean could play a critical role in enhancing dietary diversity and addressing nutritional deficiencies. This thesis assessed the nutritional composition and in vitro digestibility of winged bean seeds; identified for the first time QTLs linked to nutritional traits such as protein and oil; and performed transcriptomic analysis on developing pods and seeds for the first time.

The winged bean seeds analysed were sourced from field trials in Malaysia, revealing significant variation in protein and fat content among different accessions, with protein levels between 35.4% and 42.6% and fat content ranging from 14.2% to 21.8%. Notably, genotype-environment interactions significantly influenced fat content ($p=0.002$), highlighting the complexity of factors affecting crop nutrition. The amino acid profile analysis indicated that methionine is the limiting amino acid, resulting in a digestible indispensable amino acid score (DIAAS) of 0.14 to 0.21, significantly lower than the DIAAS of casein, which stands at 0.77. This suggests that while winged bean seeds are a good source of protein, supplementation with other amino acid sources high in methionine may be necessary. For the use of winged bean seeds in animal feed, phytic acid content and total phenolics were measured. However, a more accurate assessment of the impact of the antinutritional factors on digestibility is needed.

The next step, after evaluating the nutritional composition of winged bean seeds was to identify the quantitative trait loci (QTL). This study is the first to perform QTL analysis on nutritional traits such as protein and oil content, aiming to identify genetic markers associated with key genes contributing to these traits. Sixteen QTLs and several genes were identified, three of which were characterised as significant and linked to fatty acid contents like linoleic and behenic acids. These findings offer valuable insights for breeding programs aiming to improve the nutritional quality of winged bean. More work needs to be done including

research that combines genomics, transcriptomics, and metabolomics data for improved winged bean varieties.

Furthermore, this thesis includes the first transcriptomic analysis of winged bean developing seeds and pods, uncovering differentially expressed genes related to critical pathways, including fatty acid biosynthesis, seed storage proteins, and flavonoid biosynthesis. A total of 7,954 genes were differentially expressed in the pods, and 10,765 genes in the seeds during development. The reported findings provide a baseline for functional and comparative genomic analysis, helping to better understand the developmental process and mechanisms that contribute to and control the nutritional value of winged bean seeds and pods

Collectively, this research highlights the potential of winged bean as an underutilised crop to improve nutrition and diversify agriculture. Emphasis should be given to further genomic studies to optimise its nutritional benefits in response to the challenges posed by climate change and population growth. More research is needed in developing such underutilised crops for enhancing food and nutritional security.

Acknowledgements

This work was part of the Future Protein Platform, Beacon of Excellence programme funded by the University of Nottingham.

Throughout the years, several people have contributed to this work, without whom I would not have been able to complete it. First and foremost, I would like to thank my supervisors, Dr. Sean Mayes and Prof. John Brameld, for their immense support and guidance. Coming from different fields, they each provided valuable expertise and insights, gave me the freedom to explore my ideas, and encouraged me to think critically. Thank you!

I would also like to thank the two postdoctoral researchers who have become dear friends and have supported me from day one, Dr. Wai Kuan Ho and Dr. Alberto Tanzi. I could not have completed this work without your help. Thank you for the countless hours spent teaching, guiding, and supporting me—physically, practically, and mentally. I feel incredibly fortunate to have had you throughout this journey. You truly made this an enjoyable experience!

My sincere thanks to the technical staff at the University of Nottingham in the North Lab, South Lab, and Glasshouses. Over the years, we came across several technical challenges, from frequent equipment failures to flooding, which caused significant delays. Thank you for staying positive, supportive, and committed to finding solutions that helped me complete my work on time. I truly appreciate each one of you!

A big thank you to the friends I made during my PhD years. Together, we faced challenges, supported each other through difficult times, celebrated our successes, and created wonderful memories. I will forever miss our lab laughs, office chats, and shared lunch breaks. Special thanks to Tian, whose help and support were essential in completing this work. Thank you for the fantastic times we shared!

Balancing a full-time role as a new civil servant while writing my PhD thesis has been a challenge. I would like to thank the Futures team at the Government Office for Science for their support, especially my line manager, Poppy Groves, and our team leader, Karen Folkes.

Thank you for being so understanding and supportive over these past months. I would not have been able to submit my thesis on time without your help and support. I feel incredibly fortunate to be part of this team, and I am truly grateful!

I would also like to thank my friends and family back home, especially my mum, dad, and brother, for being a great support system. Being in different countries during the Covid-19 pandemic was not easy, yet you were always there for me. I couldn't have done this without you.

Lastly, I would like to thank the person who has gone above and beyond to support me: my best friend and husband, Dr Theocharis Tsoleridis. A huge thank you to Theo, who has selflessly supported me not only mentally but also practically, serving as my 'work buddy' on weekends when I had to sample winged bean pods and seeds using liquid nitrogen. Thank you, Theo, for spending your weekend afternoons with me, ensuring my safety, and being there through the hardest times. You have been my greatest supporter and source of inspiration.

Covid-19 Impact Statement

The starting date of my PhD was October 2019, with a duration of 36 months, including 12 months for the thesis writing period. A month into my PhD, in November 2019, I was diagnosed with acute back pain and prescribed strong medication that had severe side effects. As a result, I was unable to work for the next four months. In March 2020, when I was feeling better and ready to return to my PhD studies, the UK government announced the first lockdown, which lasted until August 2020. The Covid-19 pandemic emerged during the early stages of my PhD. Since I had no data to analyse, I used this time to improve my background knowledge and write a literature review on the genetics, nutrition, and food science of winged beans. Parts of this literature review were included in the chapter "The Winged Bean Genome: Winged Bean—One Species Supermarket" of the published book *Underutilised Crop Genomes*.

From August 2020, we were allowed access to the laboratories with strict measures for social distancing and limits on the number of people allowed to work in the labs. Access to the laboratory was restricted to four hours a day. This made training in new methods and equipment even more difficult, as only a small group of students, and sometimes just one person could be trained by a single member of the technical staff. Until spring 2021, smaller lockdowns were implemented. In March 2021, the laboratory underwent a planned move, which caused further delays as the equipment had to be shut down and transferred to a new facility. It wasn't until summer 2021 that I was able to conduct my first protein analysis, providing data that are included in this thesis.

List of abbreviations

E: Environment

G: Genotype

GxE: Genotypic and Environmental interaction

QTL: Quantitative Trait Locus

GWAS: Genome-Wide Association Studies

SFA: saturated fatty acids

MUFA: mono-unsaturated fatty acids

PUFA: poly-unsaturated fatty acids

Glossary

Biological Replicate: Each individual plant grown in the field was treated as a separate biological replicate. In this study, seeds from each plant were collected and analysed individually.

Homologous genes derive from a common ancestor and become separated in evolution in two different ways: by species separation or by gene duplication within a lineage.

Orthologues: Homologous genes related by speciation.

Paralogues: Homologous genes related by duplication.

Technical replicate: This refers to repeated measurements taken from the same sample to assess variability and ensure accuracy in the results

List of Figures

List of Tables

Chapter 1: The winged bean genome

1.1 Introduction

Food security is under threat not only from the rising population, but also from the adverse effects of climate change. The rise in temperature and the increased frequency of extreme weather conditions directly affect crop yields and local food supplies ("Policy Brief Changing Policy Concepts of Food Security" 2006). About 60% of the global food consumption is based on major crop cereals: *Triticum* spp. (wheat), *Oryza sativa* (rice), *Hordeum vulgare* (barley) and *Zea mays* (maize) (FAO, 2010). These major crops may not perform well in the future given that the annual temperatures are expected to rise by approximately 1°C in the areas where they are grown (Zhao et al. 2017). The predicted effect of climate change on the ten globally most cultivated crops barley (*Hordeum vulgare*), cassava (*Manihot esculenta*), maize (*Zea mays*), oil palm (*Elaeis guineensis*), rapeseed (*Brassica napus*), rice (*Oryza sativa*), sorghum (*Sorghum bicolor*), soybean (*Glycine max*), sugarcane (*Saccharum officinarum*) and wheat (*Triticum aestivum*) would be mostly negative on production and yield in Europe, Southern Africa and Australia (Ray et al. 2019). Based on production changes and using the average harvested area information over a 5-year period from 2003 to 2008, Ray *et al.*, 2019 estimated the impact of climate change to greatly affect the global yields of oil palm with a decrease of 13.4% but surprisingly the global yield of soybean would be expected to increase by 3.5% per year. Overall, the global yield of barley, cassava, rice and wheat could decrease by 7.9%, 0.5%, 0.3% and 0.9%, respectively, whereas the global yield of rapeseed, sorghum and sugarcane could increase by 0.5%, 2.1% and 1.0%, respectively, with global maize yield most probably not changing (Ray et al. 2019). Overall, yields are expected to be more variable and unpredictable across regions (Ray et al. 2019). The impact of climate change on global food production is real and has already caused negative effects on the yields of major crops (Frolov et al. 2014; Hertel 2016; Porter et al. 2015; Ray et al. 2019). The impact is predicted to be more severe in the most food-insecure countries, affecting both the availability and affordability of food for the most vulnerable groups.

The world's reliance on staple crops for food and nutrition has contributed to the loss of genetic diversity of crops in the fields as well as a decline in dietary diversity. Considering the

number of edible plant species is close to 30,000; only a small fraction of these, around 200 species, are used for human nutrition (Massawe, Mayes, and Cheng 2016; Khoury et al. 2014; Voss-Fels, Stahl, and Hickey 2019). In order to adapt to the challenges of food insecurity and climate change, a broad diversity of crop plants, with their wild relatives and underutilised species should be assessed (FAO 2010). Underutilised crops, also known as 'neglected' or 'orphan' crops, are indigenous crops often closely related to the culture and diet of the growers (Mayes et al. 2012). Underutilised crops, compared to the non-native crops, often carry resilience traits to abiotic and biotic stresses, such as, tolerance to drought and extreme temperatures as well as pests and diseases (Mayes et al. 2012; Massawe, Mayes, and Cheng 2016). Two of the underutilised legume crops that have been researched in the recent years include bambara groundnut (*Vigna subterranea*) and winged bean (*Psophocarpus tetragonolobus* (L.) DC.) (Ebert 2014; Tanzi, Eagleton, et al. 2019; C. S. Mohanty, Singh, and Chapman 2020). The incorporation of underutilised crops into agroecosystems and rotations would not only enhance agrobiodiversity but could also increase harvestable yields in response to climate change and global warming, thus contribute to food and nutritional security (Ebert 2014; Tanzi, Eagleton, et al. 2019).

Winged bean is an underutilised legume with high nutritional value (Yanagi, 1983). Most parts of the plant are edible and highly nutritious, for instance, the immature pods and seeds are commonly stir-fried, boiled, baked or fermented into local cuisine (National Research Council 1981). As described in the review of Eagleton (2020), winged bean is commonly grown in backyards, and it is known as a "minor garden vegetable". The tender immature pods are eaten raw, sautéed or curried, often as substitutes for more commonly used yard-long beans or common beans in traditional dishes. The immature, full-sized green seeds are made palatable by steam-frying or boiling, especially in the well-known sour vegetable soup "sayur asem" (Figure 1. 1). In some cases, the fully mature seeds were roasted and consumed after the hulls were removed Eagleton (2020).

Figure 1. 1 Winged bean is grown as an occasional fresh vegetable crop in most of Southern Asia. A. In rural household compounds (e.g., Khon Kaen, Thailand). B. In urban fringes settings as a local market crop (e.g. Selangor, Malaysia). C. In small, commercial-scale plantings (e.g., Sri Lanka). D. In urban backyards (e.g., Java, Indonesia); E. from which fresh pods may be sold in wet markets; F. or in modern supermarkets; G. to be eaten uncooked in salad dishes; H. or cooked in stir-fries and curries. As featured in Eagleton (2020)

Apart from the seeds and the immature pods, the tuberous roots, leaves and flowers are also edible (Amoo, Adebayo, and Oyeleye 2006; Cheng et al. 2019). Therefore, winged bean could be an important crop, particularly in the humid tropics where the crop is predominantly grown by small scale farmers (Lepcha et al. 2017). Like other legume crops, winged bean has the ability to fix atmospheric nitrogen, improving soil fertility and potentially contributing to sustainable production of other crops, such as rice, in tropical legume crop rotation (Rahman et al. 2014). Analysis of nitrogen in nodules, roots, and shoots of winged bean and other legumes showed that accumulation of nitrogen depends on the Rhizobium strains in root nodules (Yoneyama et al. 1986).

In summary, apart from its nutritionally rich edible parts (immature pods, seeds, tuberous roots and leaves), winged bean is a good nitrogen fixer and can be utilised in intercropping and crop rotation systems to improve soil fertility in low input cultivation systems. However, there are limitations to its large-scale cultivation that need to be addressed. For example, because of its architecture, the plants require staking which limits mechanised harvesting (Figure 1. 2).

Figure 1. 2 Winged bean plant and tubers roots. Winged bean plant growing on a 2 m tall net structure in a shade house at the University of Nottingham Malaysia (top) and winged bean tuberous roots from the accession A13-5 (Photo by Yuet Tian Chong, University of Nottingham Malaysia, 2021).

Winged bean seeds are rich in protein but contain antinutritional factors that lower their digestibility and palatability. These constraints could be minimised through genetic improvement of available germplasm using conventional and genome-based breeding approaches. Molecular tools and omics technologies could contribute in understanding genes controlling traits of interests and mechanisms of resistance in responses to biotic and abiotic stress, leading to development of improved cultivars in winged bean. Genetic markers could also be developed using genetic information and utilised in marker-assisted selection (MAS) breeding. The integration of metabolomics, transcriptomics and proteomics studies enable further understanding of the complex interactions between genes, proteins and metabolites within a desired phenotype.

1.2. Botanical description, origin and domestication

1.2.1 Taxonomy, plant morphology and reproductive development

Winged bean (2n = 2x = 18) (Figure 1. 2) is a dicotyledonous species grown mainly for its tuberous roots, and unripe pods. It is classified in the Fabaceae family, Papilionoideae subfamily, and genus *Psophocarpus* Neck. ex DC. The name Psophocarpus comes from the Greek, psophos (noise), and karpos (fruit), due to the cracking sound it produces when its mature pod bursts open. It is worth noting that the first publication of winged bean was in 1825 by De Candolle, while the first major public report appeared in the New York Times in 1975 (Khan and Erskine 1978; S. Sri Kantha and Erdman 1984a; Maxted 1990; Claydon 1975).

A taxonomic revision of the genus *Psophocarpus* in the 1980s recognised nine species in the genus. Winged bean, a domesticated species in Asia, and eight species of African origin (Figure 1. 3) (Verdcourt and Halliday 1978). In 1990, a herbarium based study of 126 specimens used 97 characteristics and revealed a tenth species, endemic in Africa (Maxted 1990). The same study grouped winged bean (*P. tetragonolobus*) with six African species (*P. scandens, P. palustris, P. grandiflorus, P. lancifolius,* and *P. lukafuensis*) in a subgenus *Psophocarpus*, and three African species (*P. obovalis, P. monophyllus* and, and *P. lecomtei*) in a subgenus *Vignopsis* (Figure 1. 4).

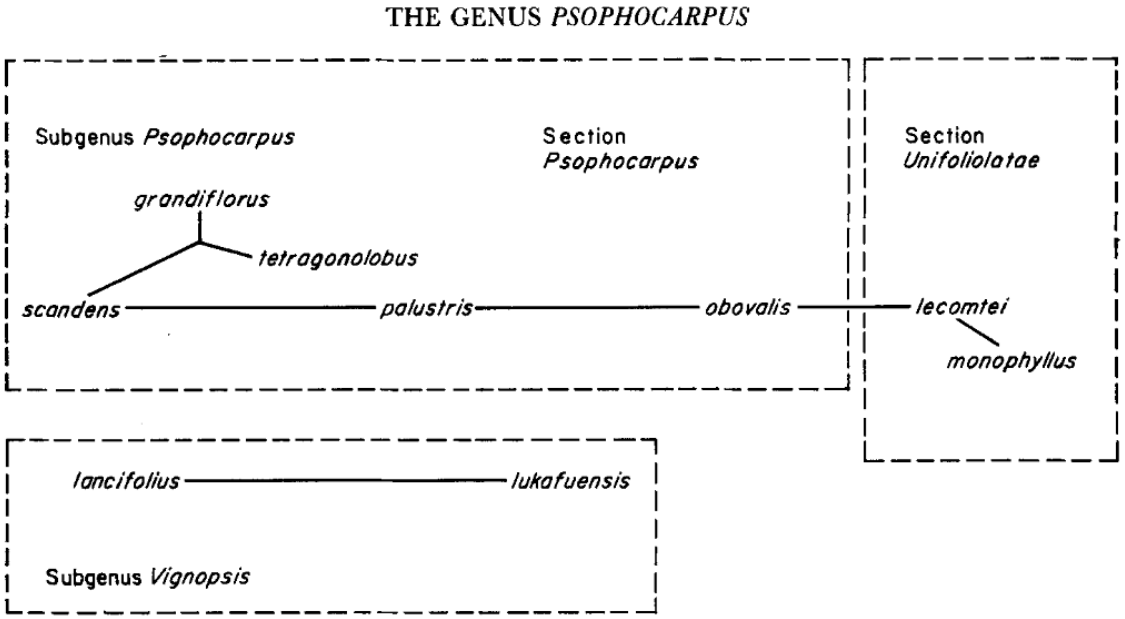


Figure 1. 3 Suggested classification of species of Psophocarpus, as featured in Maxted (1990) modified by Verdcourt & Halliday, (1978).

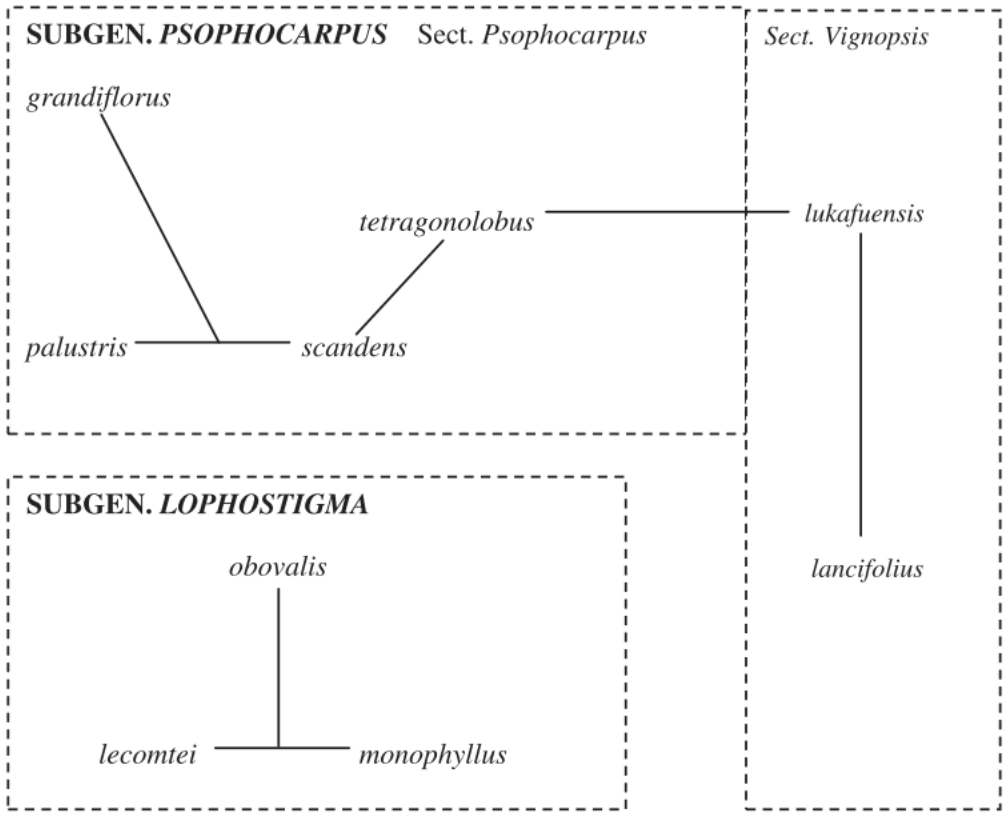


Figure 1. 4 Classification of Psophocarpus species, from Maxted (1990) as featured in Fatihah et al., (2012).

Research on winged bean and its wild relatives has led to a better understanding of the phylogenetic relationships within the genus in recent years. In 2012, a cladistic analysis of morphological traits of herbarium specimens of the nine species and species from three related genera (*Vigna*, *Otoptera* and *Dysolobium*), suggested that the genus *Psophocarpus* is monophyletic with its species been classified into four subclades, with *P. lancifolius,* and *P. lukafuensis* classified under the subgenus *Vignopsis* and *P. obovalis*, *P. monophyllus* and *P. lecomtei* grouped under the subgenus *Lophostigma* (Fatihah, Maxted, and Rico Arce 2012). More specifically, the results proposed: subgen. *Psophocarpus* sect. *Psophocarpus* (*P. palustris*, *P. tetragonolobus*, and *P. scandens*); subgen. *Psophocarpus* sect. *Vignopsis* (*P. lancifolius* and *P. lukafuensis*); subgen. *Lophostigma* (*P. obovalis*, *P. monophyllus and,* and *P. lecomtei*); as well as a new subgen. *Longipedunculares* (*P. grandiflorus*) separating *P. grandiflorus* from the subgenus *Psophocarpus* (Figure 1. 5) (Fatihah, Maxted, and Rico Arce 2012).

Monophyletic subgroup (Subgen. *Longipedunculares*)	
*P. grandiflorus**	
Monophyletic subgroup (Subgen. *Lophostigma*)	
P. obovalis	
P. monophyllus	
P. lecomtei	
Monophyletic subgroup (Section *Vignopsis*)	(Subgen. *Psophocarpus*)
P. lancifolius	
P. lukafuensis	
Monophyletic subgroup (Section *Psophocarpus*)	
P. palustris	
P. tetragonolobus	
P. scandens	

* *P. grandiflorus* was basal to other *Psophocarpus* species.

Figure 1. 5 Psophocarpus species monophyletic subgroups, as featured in Fatihah et al., (2012).

Later on, a phylogenetic analysis based on based on chloroplast genome (cpDNA) regions and internal transcribed spacer (ITS) partially supported these findings, although it separated the four analysed winged bean accessions from all the other species (with *P. palustris*, and *P. scandens* being relatively closely related) (Shuyi Yang, Grall, and Chapman 2018). Interestingly, the authors also reported the successful hybridization of winged bean with *P. scandens*. Further investigations into the genus *Psophocarpus* will likely benefit from this and from the increasing availability of molecular tools (S. C. Mohanty et al. 2014; Vatanparast et al. 2016a; Abdullah, Ho, and Wagstaff 2017; Cheng et al. 2017; Quin Nee Wong, Tanzi, Ho, et al. 2017).

Winged bean can be grown as an annual or a perennial crop (Erskine 1979). It is grown unsupported for tuber production, or traditionally upon vertical structures for pods and seeds, where it can reach up to 4 meters (Figure 1. 2) (S. C. Mohanty et al. 2014). Seeds germination varies greatly within the same accession and between accessions, and scarification of the seed coat is recommended to achieve a more uniform seedling emergence (National Research Council 1981). Wide differences in the shape and colour of leaves, stems and flowers have been observed. Leaves are typically trifoliate in form, with the leaflet shape being ovate, deltoid, or lanceolate. Depending on the variety, the stem colours can vary between green, purple, and greenish purple, with green being the most common colour (Khan 1976). As a legume, winged bean flowers are papilionaceous, with a cleistogamous floral system. While predominantly considered self-pollinating, flowers are often visited by bees and other insects that could lead to a certain degree of cross-pollination (Karikari 1972; Erskine 1980). The flower colour varies from near white to blue and deep purple (Figure 1. 6) (National Research Council 1981; Eagleton 2019). Depending on the variety, and planting conditions, location and photoperiod, plants usually start flowering from five weeks after planting onwards (Herath and Ormrod 1979; Raai et al. 2020).

Figure 1. 6 Winged bean flowers and seeds. Purple winged bean flower from the accession W103 (top left) and white winged bean flower from the accession T53 (top right). Dark brown winged bean seeds (bottom left) and light brown winged bean seeds (right) (Photo of flowers by Niki Tsoutsoura, University of Nottingham, 2021; Photo of seeds by Yuet Tian Chong, University of Nottingham Malaysia, 2021).

In terms of podding, it takes two to four months after sowing for pods to set (Figure 1. 7) (Erskine and Khan 1980; Eagleton 2019). Pod shape is rectangular, semi-flat, flat on the sides, or flat on the suture ('winged') (International Board for Plant Genetic (IBPGR) 1982). Variation of the colour of pods and wings has been observed, with the central portion of the pod being cream, green, pink or purple and the pod wings being green or purple (Erskine and Khan 1977).

Figure 1. 7 Winged bean pods. Purple flowers and green and purple pods of winged bean. Photo taken two months after planting (Photo by Yuet Tian Chong, University of Nottingham Malaysia, 2021). Purple winged bean pod from the accession W120 (in the middle) and green-winged bean pod from the accession W040 (on the right) (Photo of flowers by Niki Tsoutsoura, University of Nottingham, 2021)

In addition, the seed colours can be cream, different shades of brown, deep purple, black or mottled (Khan 1976; International Board for Plant Genetic (IBPGR) 1982; Eagleton 2019). Pod size can reach 30 to 40 cm and contain between five and 21 seeds (Poole, 1978). After successful fertilisation, the developing seeds reach maturity (Figure 1. 6) around sixty-five to eighty-five days after flowering (S. S. Kadam 1984a; Higuchi, Fukumoto, and Iwai 1988; S. S. Kadam et al. 1982).

1.2.2 Origin, distribution and germplasm collection

Winged bean is cultivated throughout Asia; mainly in India, Southeast Asia, in the highlands of Papua New Guinea and in Africa (Verdcourt and Halliday 1978; Shuyi Yang, Grall, and Chapman 2018). Its centre of origin remains enigmatic with two hypotheses prevailing. The first hypothesis supports Africa as the centre of origin based on the morphological similarities among *P. tetragonolobus* and other African species, such as *P. grandiflorus* (Smartt 1980). Possibly, the winged bean progenitor could have originated on the African side of the Indian

Ocean and then been transferred to the east as a wild plant, where it was domesticated (Figure 1. 8) (Lepcha et al. 2017).

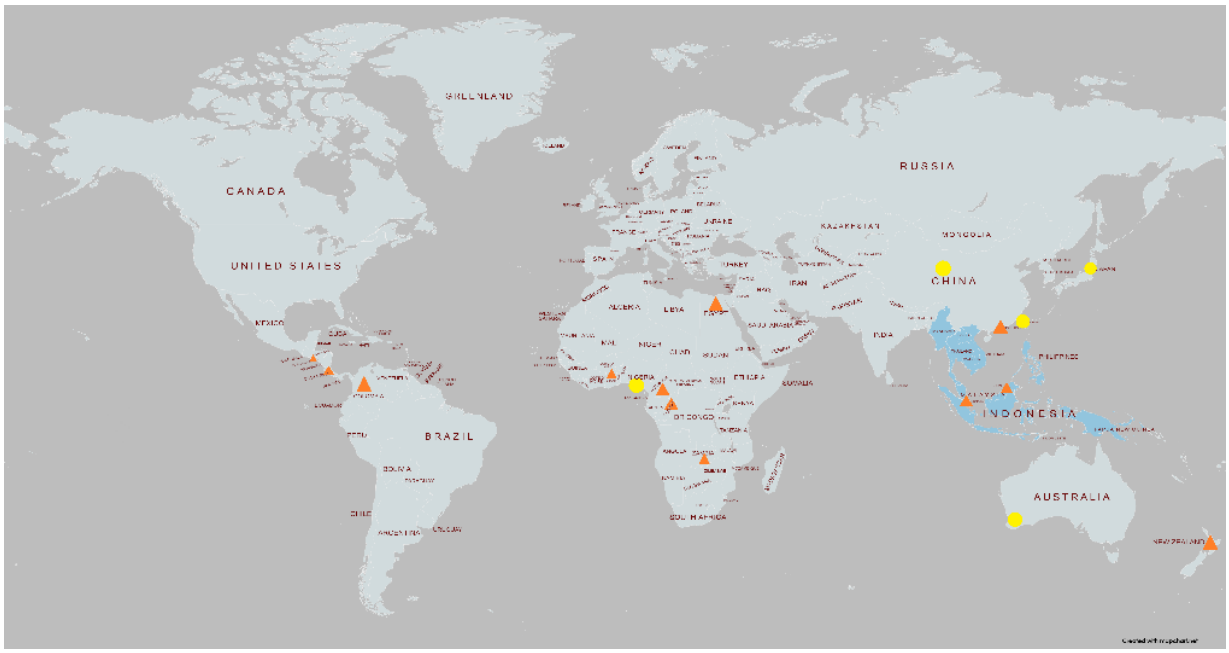


Figure 1. 8 Distribution map of winged bean. The countries where winged bean is cultivated as a crop, mainly in South and South-East Asian and the Pacific regions, are shown with light blue colour. The regions in countries where winged bean research and experiments have been carried out are shown with yellow dots. Whereas winged bean specimens have been preserved in regions and countries shown with the orange triangles (Abe and Nakamura 1987; Amoo, Adebayo, and Oyeleye 2006b; Reddy and Reddy 2015; Tanzi, Eagleton, et al. 2019b; J. Yang and Tan 2011; PFAF 2021; GBIF 2021); (Created by Yuet Tian Chong, University of Nottingham Malaysia, 2021).

The second hypothesis suggests that winged bean arose in Asia from an unknown progenitor that has now become extinct (Verdcourt and Halliday, 1978). The results of the phylogenetic analysis from Yang *et al.* (2018) showed a large genetic distance between winged bean and the closest African relatives of the *P. scandens* and *P. palustris* group, supporting the conclusion made by Verdcourt and Halliday in 1978 that the wild progenitor of winged bean is now extinct (Verdcourt and Halliday, 1978) or at least unsampled. The identification of wild ancestors would shed light not only on the genetic changes related to domestication, but also provide resources for the improvement of winged bean (Prohens et al. 2017). The identification of novel genes from wild relatives and their introgression into winged bean varieties via breeding or gene editing methods could ultimately lead to the development of

more resilient and efficient winged bean accessions with improved yield in an uncertain and rapidly changing environment.

Most winged bean germplasm collections are kept in genebanks maintained either by national and international organisations. Yang *et al.* (2018) named IITA (International Institute of Tropical Agriculture), U.S. Department of Agriculture (USDA), and the National Agriculture and Food Research Organization (NARO, Japan) genebanks as the source of the materials used in their study on the origin and diversification of winged bean (Shuyi Yang, Grall, and Chapman 2018). Approximately 51 accessions are kept by the U.S. National Plant Germplasm System (USDA), 12 by NARO as well as 271 accessions in International Institute of Tropical Agriculture, IITA, Africa.

1.2.3 Genetic diversity

Extensive genetic diversity in a germplasm is of importance as it can act as a great source of variation for improving crops for higher yield and resilience (Prohens *et al.*, 2017). Mohanty et al. (2014) analysed the genetic relationship between 24 winged bean genotypes using Random Amplified Polymorphic DNA (RAPD), and Inter Simple Sequence Repeat (ISSR) molecular markers. The varieties fell into two distinct clusters and seven sub-clusters. Overall, they detected high levels of polymorphism and genetic distances across the varieties, which suggested a wide genetic base for the winged bean germplasm. Chen et al. (2015) used ISSR markers to evaluate the genetic distances in 45 accessions of winged bean which were cultivated in eight countries. However, the ISSR analysis showed little genetic variation and did not detect a significant correlation between the genetic distance and origin of the accessions. Mohanty et al. (2019) investigated the genetic diversity of 95 winged bean accessions from six countries using AFLP and the ITS of nuclear ribosomal DNA markers. The population structure analysis revealed five sub-populations based on estimation from the frequencies of the alleles. Among the accessions, the genetic diversity was at medium to low levels with the maximum similarity detected among accessions from India, indicating the existence of a possible ancestral origin. The cluster analysis grouped the accessions into four groups with accessions from different origins groups in the same sub-cluster. These results

agreed with Chen *et al.,* (2015), Yang *et al.*, (2018) and with later results from Ojuederie et al. (2020).

When trying to assess the amount of genetic diversity, especially in material held in genebanks across the world, relying too heavily on the declared geographic origin, or on the assumption that genetic relationship can be inferred from the phenotype should be avoided (Tanzi, et al. 2019). Following on from this, Tanzi and colleagues suggested that the utilisation of high-throughput technologies, such as genotyping-by-sequencing, could provide more information on the geographic origin and genetic diversity among the germplasm held in seedbanks. This would be crucial to revamp efforts towards preserving winged bean genetic diversity, and to provide access to truly diverse material in breeding programmes.

1.3 Food and Nutritional Value

The seeds of winged bean are rich in protein, vitamins and minerals, as well as secondary metabolites such as phenolics and flavonoids, which act as antioxidants (Gross Rainer 1983; S. Sri Kantha and Erdman 1984a; Lepcha et al. 2017a; Adegboyega et al. 2019b; Bassal et al. 2020). However, winged bean also has a considerable content of antinutritional factors (ANFs) such as proteinase inhibitors, tannins and phytic acid. These ANFs have the ability to inhibit the absorption of various nutrients either by preventing their release during digestion or by binding to proteins or other nutrients, resulting in them passing through the gastrointestinal tract (Adegboyega et al. 2019b; M. Singh et al. 2019).

1.3.1 Protein

The crude protein content of mature seeds of winged bean ranges from 30% to 40% (Černý et al. 1971; Claydon 1975; S. S. Kadam 1984b; Adegboyega et al. 2019). Adegboyega et al. (2019) compared 25 winged bean accessions and found a crude protein content between 34% and 40% in processed seeds, and 28.5% to 31% in unprocessed seeds. The processed seeds were cleaned, slightly roasted under low heat, coarse-milled and winnowed to remove the seed coat, then the decorticated grains were milled into fine powder and sieved, whereas the unprocessed seeds were cleaned and milled into fine powder directly. There were significant

differences observed for the moisture, fat, crude fibre and carbohydrate contents, as well as crude protein, in both the processed and unprocessed seeds. The observed protein levels in unimproved material offers opportunities for selection to target high protein contents for inclusion in genetic improvement programmes. Overall, winged bean has higher protein content than other legumes grown in the tropic regions, such as lentil (*Lens culinaris*) (24.6%), cowpea (*Vigna unguiculata*) (23.8%), chickpea (*Cicer arietinum*) (20.5) and pigeon pea (*Cajanus cajan*) (21.7%) (US Department of Agriculture 2019). Comparing winged bean to soybean, types of soybean meal that include parts of the hulls contain less than 47% protein, whereas high protein types of soybean meal is obtained from dehulled seeds contain 47-49% of protein (Heuzé, Tran, and Kaushik 2020).

1.3.2 Amino Acid composition

Winged bean seeds resemble soybean in amino acid composition. Mnembuka and Eggum, (1995) compared the nutritive value of winged bean with soybean and other legumes grown in Tanzania, including green gram (*Vigna radiata*), bambara groundnut, pigeon pea, field pea (*Pisum sativum*) and cowpea. The results showed methionine to be the most limiting amino acid in winged bean seeds, followed by tryptophan, histidine and cysteine, while winged bean seeds were rich in lysine and threonine. These results agree with other studies (Okezie and Martin 1980; Ekpenyong and Borchers 1982; R. King and Puwastien 1987; Wan Mohtar et al. 2014), suggesting that winged bean is a good alternative to soybean, based on the amino acid profile. The study of (Prakash et al., (1987) showed variation in protein and amino acid composition between 16 strains of winged bean obtained from the Indian Institute of Horticulture Research, Banglore, one NBRI selection and four strains from Sri Lanka which were cultivated under uniform conditions as shown in Table 1. 1.

Table 1. 1 Protein and amino acid composition (g/100 g protein) of winged bean and soybean seeds. Cow milk values are expressed as amino acid composition per total amino acids (g/100 g amino acids).

Amino Acids	**Winged bean [1]**	**Soybean [1]**	**Cow milk[2]**
Glutamic acid	8.2-12.00	9.9	19.66
Aspartic acid	7.5-11.0	10	7.60
Leucine	6.2 – 7.7	7.2	9.44
Lysine	4.2 – 6.5	5.4	8.96
Arginine	4.0-6.2	5.3	4.06
Proline	6.1-8.4	6.8	8.99
Serine	6.1-8.4	6.6	5.24
Valine	4.7-6.4	4.9	5.24
Tyrosine	3.3-4.5	2.2	5.67
Isoleucine	4.7-6.7	5.2	4.54
Phenylalanine	3.8-5.3	4.9	4.73
Alanine	6.2-8.3	6.6	3.41
Threonine	4.6-6.9	5.4	4.11
Glycine	6.5-7.9	7.3	1.75
Histidine	2.8-4.1	3.8	3.30
½ Cysteine	0.6-1.7	1.5	0.82*
Methionine	0.1-1.0	1.2	2.48
Protein (%)	38.1-45.0	43.7	2.82

Values of winged bean[1] and soybean[1] seeds from Prakash *et al*., (1987); values of cow milk[2]; *responds to full cysteine from Ceballos et al. (2009).

1.3.3 Minerals and Vitamins

The rich mineral and vitamin concentration of winged bean seeds, pods and leaves adds to its nutritional value. Winged bean seed flour could be an important source of minerals with the levels of phosphorous, calcium and magnesium being similar to soybean (Mnembuka and Eggum 1995; Amoo, Adebayo, and Oyeleye 2006a). Okezie and Martin (1980) compared 20 winged bean varieties from Puerto Rico and the United States of America by nutrient analysis of seeds, seed hulls and fresh leaves, detecting significant differences in nutrient contents between the genotypes and the different parts of the plant. The results showed high calcium (0.28-0.86%), phosphorus (0.36-0.72%) and iron (58-308 ppm) levels in the dry dehulled and whole seeds, with the seed hull containing the highest levels of iron. Interestingly, the calcium content in leaves (3.21-4.41%) was considerably higher than the seeds, and higher levels of iron (126-298 ppm), potassium (0.62-1.66%) and magnesium (0.28-0.36%) were detected in leaves as well. Leaves of winged bean are used in soups and salads, making them an important source of calcium, phosphorus and iron in tropical countries. Leaves also contain high levels of carotenoids (5,240-20,800 IU/100 g fresh weight), vitamin C (14.5-128 µg/100 g fresh weight) and folic acid (67 µg/100 g dry weight), while the seeds contain adequate levels of folic acid and tocopherols (B. O. de Lumen, Fiad, and Fiad 1982; Kantha and Erdman 1984). Tuberous roots are also edible and rich in protein with the protein content ranging from 12-19%, with sulphur containing amino acids being limited. However, antinutritional factors such as trypsin and chymotrypsin inhibitors as well as haemagglutinins were present in the tubers (S. Sri Kantha and Erdman 1984c; Kortt and Caldwell 1984; Adegboyega et al. 2019).

1.3.4 Lipids

Lipids are major components of winged bean seeds and the content is dependent upon the genotype, environment, location and soil type where the cultivar is grown (Worthington, Hammons, and Allison 1972; Garcia and Palmer 1980b; R. D. King and Puwastein 1987; Lepcha et al. 2017). The oil content of mature winged bean seeds varied from 15% to 20.4%, with the saturated and unsaturated fatty acids ranging from 30-40% and 60-70%, respectively (Khor, Tan, and Wong 1982a; S. Sri Kantha and Erdman 1984a). Mohanty et al. (2014) determined the fatty acid composition in immature, mature and fully mature seeds of winged bean. They found that the immature seeds had the highest percentage of saturated fatty acids (61.3%),

whereas the fully mature seeds had the highest percentage of unsaturated fatty acids (75.5%), with nearly equal ratio of mono-unsaturated fatty acids (38.6 %) and poly-unsaturated fatty acids (36.9 %) (Mohanty *et al.*, 2014).

1.3.5 Antinutritional Factors (ANFs)

Legumes are an important part of human diets due to their high nutritional value. They are rich in protein, fibre and vitamins, but they also contain ANFs, such as phytic acid, tannins and proteinase inhibitors. These ANFs are characterised as non-nutritional or toxic compounds that can inhibit the absorption of nutrients and have deleterious effects, particularly when the pulses are consumed raw or under-cooked. In humans and animals, antinutritional factors such as lectins can cause diarrhoea, vomiting, inflammation and blood agglutination (Peumans and Van Damme 1995). In plants, these compounds are used as defence mechanisms against fungi, insects, and herbivores, as well as being an energy storage for the plants to continue their growth under extreme environmental conditions (Bessada, Barreira, and Oliveira 2019).

The ANFs can be classified into two groups based on their structure, with the first group containing proteins such as lectins, agglutinins, protease inhibitors and bioactive compounds; while the second group contains non-protein compounds such as phytic acid, tannins and saponins (Sánchez-Chino et al. 2015). Processing methods, such as germination, dehulling, moist heat and soaking in water or alkali solutions, are commonly used to reduce the levels or activities of ANFs in legumes and thereby eliminate their negative effects on digestion and absorption of nutrients, without compromising the nutritional value of the pulses (Samtiya, Aluko, and Dhewa 2020).

This study focused on the adverse effects of the ANFS. However, it is worth mentioning that when consumed in small amounts, ANFs can have a positive effect on human health. For example, studies have shown that proteinase inhibitors and isoflavones can act as antioxidants, anticancer and anti-diabetic agents (Messadi et al. 1986; DeClerck and Imren 1994; Gurfinkel and Rao 2003; Dong and Qin 2011; Sánchez-Chino et al. 2015; Lopez-Corona et al. 2022; Urbano et al. 2000; Beninger and Hosfield 2003; Rambaran 2020).

Phenolic compounds are broadly classified into flavonoids and non-flavonoids and are known for their antioxidant properties (Lopez-Corona et al. 2022). In legumes, polyphenols such as flavanols, anthocyanins, and tannins (Figure 1. 9) are mainly concentrated in the seed coat, with lower levels found in the cotyledons (Beninger and Hosfield 2003; B. Singh et al. 2017; B. J. Xu, Yuan, and Chang 2007; Oomah et al. 2010; B. Xu and Chang 2008). Polyphenols are bioactive secondary metabolites playing an important role in plants' defence mechanism, contributing to antioxidant activities (Beninger and Hosfield 2003; B. Singh et al. 2016; Rambaran 2020; B. Singh et al. 2017; Adebamowo et al. 2005).

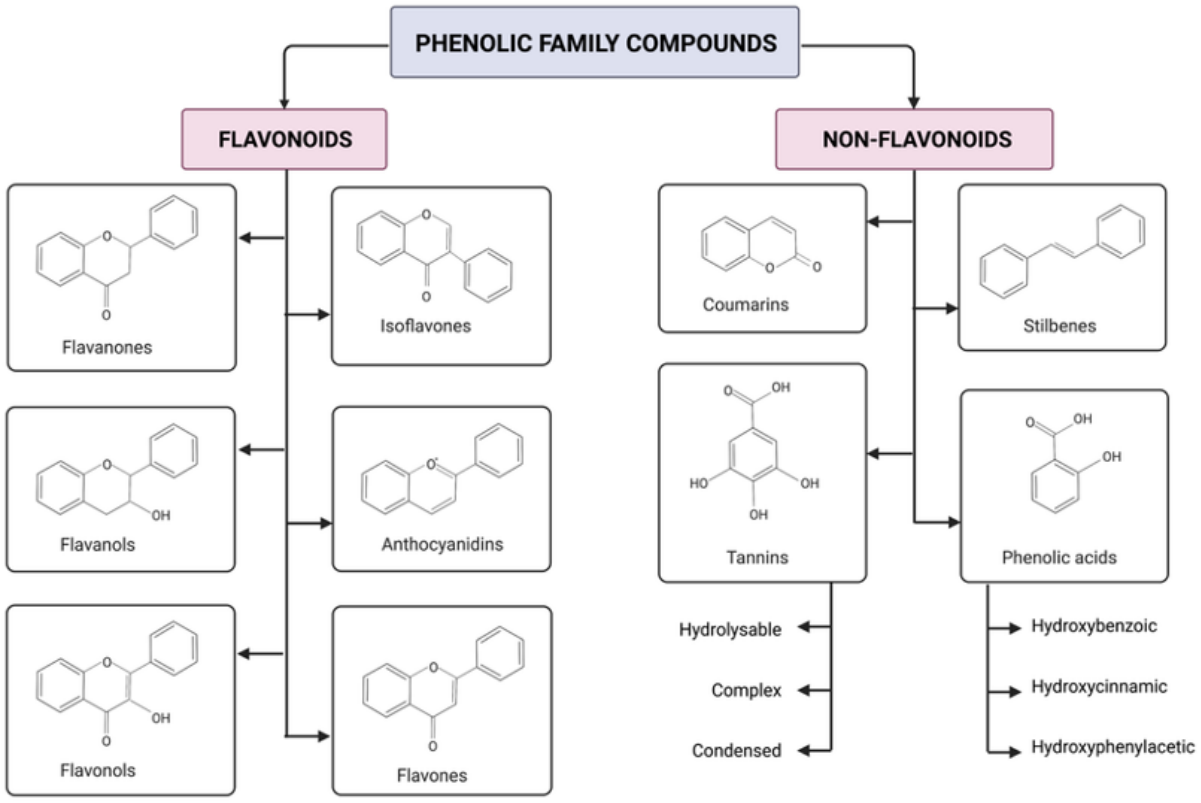


Figure 1. 9 Phenolic family compound basic skeletal structure and their classification in subclasses. There are two major groups: flavonoids and non-flavonoids. Flavonoids have six subclasses: anthocyanidins, flavonols, flavanones, flavanols, flavones and isoflavones. Anthocyanidins become anthocyanins when sugars are linked in their chemical structures. Non-flavonoids are subdivided into four subclasses: stilbenes, tannins, coumarins and phenolic acids. Tannins are categorized into condensed, hydrolysable and complex. For phenolic acids, a more specified classification is to divide them into three groups: hydroxybenzoic, hydroxyphenylacetic and hydroxycinnamic acids. Created using licensed BioRender (2022). As described and featured in Lopez-Corona et al. (2022) .

Adegboyega *et al.* (2019) showed significant variations in tannin content ranging from 1.8% to 2.5% in processed seeds and, 1.3% to 3.4% in unprocessed winged bean seeds. The processing methods used not only increased the protein content in the flour, but also lowered the tannin content, probably due to the removal of the seed coat. The high protein content of winged bean seeds, combined with use of appropriate processing methods, suggests that winged bean flour could be added in food formulation in order to increase protein content.

It is important to understand the genetic mechanisms and biosynthetic pathways related to the production and accumulation of antinutritional factors in winged bean seeds. For example, in soybean, where genetic improvement of seed protein content is desirable, major quantitative trait loci (QTL) for soybean protein have been detected and mapped on chromosomes 20 (LG-I), and 15 (LG-E) (Patil et al. 2017). Shedding light on the genes and mechanisms regulating the synthesis of seed storage proteins and ANFs in winged bean, would be useful in selecting accession or improved accessions of high nutritional value and reduced levels of ANFs. Genotyping with molecular markers is a requirement for quantitative trait loci (QTL) mapping and genome-wide association studies (GWAS). The identification and utilisation of genetic markers would assist breeding programmes (Chapman 2015a); while gene editing methods could also be a valuable tool in silencing genes related to undesirable traits, thereby contributing to the genetic improvement of winged bean.

1.3.5.1 Proteinase (trypsin and chymotrypsin) inhibitors

Legumes contain serine protease inhibitors that inhibit digestive enzymes, trypsin and chymotrypsin, by competitive binding. The protease inhibitors are categorised into two groups, the Kunitz trypsin inhibitor (KTI) of 20-24 kDa, and the Bowman-Birk inhibitor (BBI) with molecular weight of 7-8 kDa (Birk 1985; Wati et al. 2010; Muzquiz et al. 2012). Kortt (1983) examined 27 varieties of winged bean from six regions across South-East Asia for proteinase inhibitor contents and found that levels of trypsin and chymotrypsin inhibitors varied between genotypes, with Malaysian and Indonesian varieties showing the lowest levels. The study also showed that the average chymotrypsin inhibitory activity ranged between 30-48 mg chymotrypsin inhibited per g of defatted seed and was higher than trypsin inhibitory activity (23-36 mg trypsin inhibited per g of defatted seed). It is also important to note that the stoichiometry of inhibition for the winged bean chymotrypsin inhibitor activity

was 1:2 (i.e. 1 molecule inhibits 2 enzymes,) whereas it was 1:1 for trypsin inhibitor activity (Kortt 1979; 1980; 1981).

From winged bean, there have been three BBI trypsin inhibitors isolated and nine KTI (four trypsin, four chymotrypsin and one trypsin/chymotrypsin inhibitor) (Shibata et al. 1986; Giri et al. 2003). Giri *et al.* (2003) purified seven winged bean trypsin inhibitors and showed different binding potentials against gut proteinases of *Helicoverpa armigera. H. armigera* is a major bollworm pest of cotton (*Gossypium*), legumes and other plant species. *H. armigera* has evolved insecticide resistance, significantly lowering yields in countries like India, Australia, Indonesia, and Thailand, resulting in annual losses of $300–500 million (Srinivas et al. 2004). The 28kD winged bean trypsin inhibitor showed at least a three-fold higher inhibitory activity in the gut of *H. armigera* than the bovine version. In addition, the putative Kunitz-type chymotrypsin inhibitor genes, WCI2 and WCI5, isolated from winged bean, were shown to inhibit the gut proteinases of *H. armigera* larvae. These results suggest that the proteinase inhibitors of winged bean could be a fruitful target of further studies for the development of transgenic lines resistant to *H. armigera*, a pest that affects many important crops and develops fast resistance to pesticides (Giri et al. 2003; Telang et al. 2008).

1.3.5.2 Phytohemagglutinins or lectins

Lectins (hemagglutinins or phytohemagglutinins) are a group of proteins found in various organisms that have the ability to bind carbohydrates. When these proteins bind specifically to known sugars and agglutinate red blood cells, they are referred to as lectins. Lectins have at least one non-catalytic domain, which can bind reversibly to specific monosaccharides or oligosaccharides (Lagarda-Diaz et al., 2017). In plants, lectins are found abundantly in the cotyledons and endosperm of legume seed, accounting for 2-10% of the total protein (Lis and Sharon 1986). They contribute to physiological regulation, defence against microorganisms, transport of carbohydrates, mitogenic stimulation and recognition of nitrogen-fixing bacteria from the genus *Rhizobium* (Sharon and Lis 1990; 2004; Chrispeels and Raikhel 1991; Nasi, Picariello, and Ferranti 2009). Lectins apart from their deleterious effects in humans and animals (Peumans and Van Damme, 1995). can also have positive effects. Based on their activities, the production of lectins may have practical applications, as they have numerous

positive effects in human health, such as antitumor (Kwan Lam and Bun Ng 2011), antifungal and antiviral activities (Sánchez-Chino et al. 2015; Lagarda-Diaz, Guzman-Partida, and Vazquez-Moreno 2017).

The presence of lectins in winged bean seeds relate to its toxicity Kortt (1983). When a 30% raw winged bean diet was fed to rats, they showed significant growth depression, morphological changes in the small intestine and 100% mortality within 10 to 20 days. The lethal effect was eliminated by autoclaving winged bean seeds at 120^{o}C for 30 minutes. Rats fed with autoclaved winged bean seeds gained body weight comparable to the rats fed with casein diet, suggesting that the lethal action was eliminated by autoclaving (Higuchi, Suga, and Iwai 1983).

1.3.5.3 Tannins

Tannins, as well as other phenolic compounds, are secondary metabolites widely produced by plants, that play an important role in defence strategies against insects, birds and fungi. Tannins also contribute to colour, flavour and astringency of fruits, but tannins seem to negatively impact the nutritional quality of the food (Chiba 2003). Tannins can be classified into four groups depending on the structure of the monomer: proanthocyanidins (or condensed tannins), hydrolysable tannins, complex tannins and phlorotannins (Serrano et al. 2009). Hydrolysable tannins are hydrolysed by enzymes, acids or alkalis; whereas condensed tannins are resistant to hydrolysis. Condensed tannins are the major polyphenols in commonly consumed food (Salunkhe and Chavan 1990; Gilani, Cockell, and Sepehr 2005). Tannins have the ability to bind and precipitate proteins, therefore reducing protein and amino acid digestibility in monogastrics, such as pigs and poultry (Smulikowska et al. 2001). The complexes formed with glycoproteins and the astringent properties of tannins, reduce their palatability and lower the nutritional value of pulses (Gilani, Cockell, and Sepehr 2005; Bessada, Barreira, and Oliveira 2019).

Unfortunately, heat is not the best method for the reduction of tannin content, as tannins are heat resistant (B. O. De Lumen and Salamat 1980). Instead, soaking winged bean seeds in different salts has been effective in decreasing the tannins. Like many pulses, the seed coat

of winged bean has the highest tannin levels, but its removal is difficult and is not commonly practiced when cooking (B. O. De Lumen and Salamat 1980; Tan, Wong, and de Lumen 1984). The high tannin content and indigestible fibre of the seed coat are thought to be responsible for the lower metabolisable energy and the poor response of broilers fed winged beans (Benitto O. De Lumen, Gerpacio, and Vohra 1982). Therefore, winged bean accessions with lower tannin content in the seed coat and improved techniques for seed coat removal during food processing should be investigated to avoid the ingestion of high tannin contents.

1.4 Barriers to the greater utilisation

In this context, a leguminous species such as winged bean has the potential to contribute in different ways, especially considering the nutritional values mentioned in the previous section, although several constraints limit its wider utilisation.

Winged bean has the potential to be a cash crop with limited input requirements, but when grown as a horticultural crop it requires a vertical structure (trellising or staking) to increase pod productivity, leading to increase in the cultivation cost in the short term and also limit mechanised harvesting (Q. N. Wong, Massawe, and Mayes 2015). On the other hand, winged bean could make use of structures set up for other major crops (i.e. tomatoes) in rotation systems. In addition, improving its plant architecture and yield-component traits could allow an increase in planting density, thus further improving the final harvest per unit of land (Tanzi, Ho, et al. 2019).

As a pulse crop, ideotypes with early maturing, “bushy” architecture, dwarf type with side branches and short or few internodes that produces many pods would be desirable for increased seed production and cultivation in large scale. Kesavan and Khan (1978) isolated winged bean mutants with determinate and dwarf growth habits using gamma rays and ethyl methyl sulphonate (EMS) on a number of pure line genotypes. No determinate mutants had been isolated from that experiment, however, mutants with single cotyledon, albino types and darker colour foliage were reported (Jugran et al. 1986; Quan et al. 2011). The generation or identification of a dwarf variety is still considered an objective in winged bean research, probably to facilitate harvesting through mechanisation and improve its utilisation as pulse

crop, by adapting to arable cropping. However, it will be crucial to assess the final impact on pods and seed productivity that such a change may cause.

Another limiting factor to the greater utilisation of winged bean is the hard-to-cook seeds it produces. The seeds contain antinutritional factors and a thick seed coat with cooking time lasting for several hours, leading to significant loss of protein quality by lowering protein digestibility and bioavailability as well as minerals such as potassium and magnesium (Ekpenyong and Borchers 1980a; Henry, Donachie, and Rivers 1985). However, processing methods such as seed coat removal, heating, soaking, boiling, and pressure cooking the seeds can improve nutritional quality by reducing ANF concentration, thus increasing protein digestibility (Ekpenyong and Borchers 1980a; S. S. Kadam and Smithard 1987; Gilani, Xiao, and Cockell 2012). The quantification of the ANFs and the effect of seed treatments (such as seed coat removal) should be also evaluated in view of the quality and quantity of seed protein. Protein quality is affected by the digestibility and quantity of the essential amino acids. For example, the dehulling process would increase protein content and reduce the antinutritional factors contained in the seed coat such as tannins. This would improve the palatability, taste and digestibility of pulses as well as decrease the cooking time (Bessada, Barreira, and Oliveira 2019). The dehulling of winged bean seeds would have a positive impact on protein quality and quantity but it is practically difficult.

1.5 Genome

1.5.1 Genome sequencing

Winged bean has a diploid (2n = 2x = 18) genome of around 1.22 Gbp (Vatanparast et al. 2016). Ho et al. (2024) utilised the combination of Illumina (for accuracy) and Oxford Nanopore Technologies (ONT) (for long reads) platforms to generate genomic resources for winged bean. Combining with Bionano Genomics optical mapping, the current genome assembly is comprised of 48 hybrid scaffolds covering 536,131,541 bp in total with a N50 size of 23,875,316 bp (N90 = 6,932,124 bp). The scaffolds generated from this approach range from 122,770 to 38,637,442 bp. Two genetic maps in winged bean using a common paternal parent have been used to place 38 hybrid scaffolds onto pseudo chromosomes, resulting in a

draft genome with nine pseudochromosomes encompassing 530,283,461 bp and 26,354 protein coding genes annotated. The pseudochromosome length ranges from 24,607,972 to 85,053,349 bp.

1.5.2 Transcriptome assembly and molecular markers

Utilising transcriptome sequences for genetic studies of crop origins, genetic diversity and identifying molecular markers for genetic mapping can be useful tools to accelerate the genetic improvement of crops through breeding and/or gene editing. Chapman (2015b) sequenced, assembled and annotated the transcriptomes of four underutilised crops, hyacinth bean (*Lablab purpureus* (L.) Sweet), grasspea (*Lathyrus sativus* L.), winged bean (accession 477137 from Nigeria) and bambara groundnut (*Vigna subterranea* (L) Verdc.). The aim was to identify SSR markers and conserved orthologous set of markers across the legume species that could be used to investigate the genetic variation. The Illumina-based transcriptome by Chapman (2015) was used in the study of Vatanparast et al. (2016) to identify contigs from the transcriptome assembly corresponding to the Kunitz trypsin inhibitor (KTI) gene family within the *Psophocarpus* transcriptome.

A *de novo* transcriptome assembly and annotation of two winged bean accessions (CPP34 and CPP37) from Sri Lanka were produced by Vatanparast *et al.* (2016). In this study, using single-end 454 pyrosequencing that produces long reads (300–800 bp). The genotype CPP34 produced 369,820 single-end reads (136,943,216 bp), and the genotype CPP37 produced 334,639 single-end reads (92,126,948 bp). When comparing the independent reads from the transcripts of CPP34 and CPP37, less than 200 high-confidence SNPs were detected, corresponding to approximately one SNP every 150,000 bp and indicating a high similarity between the two accessions. Combining the reads from the independently sequenced accessions, a single assembly CPP34-7 was produced. Unassembled reads, notated as singletons post-assembly, were included to the final assembly of CPP34-7 as they could possibly be full-length mRNA transcripts and used in the Gene Ontology (GO) and SNP analyses. The assembled contigs from the CPP34-7 were compared to protein sequence databases from chickpea (*Cicer arietinum* L.), pigeon pea (*Cajanus cajan* (L.) Huth), soybean (*Glycine max* (L.) Merr.), common bean (*Phaseolus vulgaris* L.), *Medicago truncatula* Gaertn., and *Lotus japonicus* (Regel) K. Larsen using the BLASTX (translated nucleotide sequence

searched against protein sequences). The BLASTX revealed that from CPP34-7, 96.5% of the contigs had significant sequence similarity to one or more protein sequences used, with most of the contigs (57.3%) being most similar to soybean.

From the GO analysis, 274 transcripts were annotated as transcription factors and 176 putative winged bean transcription factor genes were identified and classified in at least ten different families with their overall distribution being similar to other legumes (Vatanparast *et al.*, 2016). The top five categories were: basic leucine zipper (bZIP; 32), Teosinte-Branched1/Cycloidea/ PCF (TCP; 19), MADS (17), MYB (11) and WRKY (9). In terms of identifying SSRs, the analysis showed that of a total of 12,956 SSRs. From those, 10,984 were perfect SSRs, consisting of a single motif repeats; 13 were imperfect SSRs, containing a base pair not belonging to the motif between repeats and 1,959 compound SSRs composed of two or more adjacent individual repeats. From the perfect SSRs, 7,933 SSRs were hexamers with only two repeats and the remaining 2,994 perfect SSRs were 405 di-, 1288 tri-, 482 tetra-, 211 penta- and 608 hexamer SSRs. The repeat motif type (AG/GA/TC/CT)n accounted for the 77.7% of all the dinucleotide repeats, whereas the motif types (AT/TA)n and (AC/CA/GT/TG)n accounted for the 13.7% and 8.6%, respectively.

A high confidence set of 5,190 SNPs were identified with 96% being one-to-one point mutations between the Sri Lankan samples and a geographically separated Nigerian winged bean genotype. The latter was sequenced by Chapman (2015). From the 5,190 SNPs, around 4% (211 SNPs) were length variants with one or more point mutation. The one-to-one polymorphisms were found to be 4,979 SNPs, with 3,433 (68.9%) being transitions and 1,546 (31.1%) being transversions, making a transition: transversion ratio of 2.22 (Vatanparast *et al.*, 2016). SNPs can be used to access variation among genotypes and utilised in quantitative trait loci (QTL) mapping, linkage maps as well as breeding studies as reported in the study of Vatanparast *et al.* (2016). However, further research needs to be done for validation of SNPs from this study.

Vatanparast *et al.* (2016) also analysed the soybean trypsin inhibitor (STI) gene family and found similarity to winged bean KTI. Understanding the evolution and diversity of the Kunitz-type trypsin inhibitors gene family in winged bean, the STI sequences were used to generate

a gene tree and showed that 28 out of the 32 putative *Psophocarpus* STI regions were clustered with soybean (*Glycine max)* and at least eight Kunitz trypsin inhibitor loci were linked within 68 kbp on chromosome 8 (positions 44,850,000-44,918,000). Lineage-specific amplification of *Psophocarpus* Kunitz-type trypsin inhibitors sequences was suggested, considering the conserved synteny between soybean and *Psophocarpus*. Identification of molecular markers linked to the trypsin inhibitor genes family in winged bean would assist genotyping and breeding. In addition, the investigation of structure and regulation of winged bean Kunitz trypsin inhibitor genes could be utilised not only in breeding selection, but also in gene editing. Genes, transcription factors and markers identified in these accessions could be utilised in breeding to reduce the amount of antinutritional factors, such as trypsin and chymotrypsin inhibitors as well as tannins, and improve the nutritional value of the seeds.

The first set of validated SSR markers in winged bean was reported by Wong *et al.* (2017), with 18 genic-SSRs, 9 of which were further used by Yang et al., (2018) in their population genetic analysis. Wong et al. (2017) developed a *de novo* transcriptome assembly from leaf, root, pod and reproductive tissues of six Malaysian winged bean accessions. From the 198,554 contigs (with a N50 of 1462 bp), 138,958 contigs (70.0%) were annotated. The majority of the SSR motifs identified were AAG/AGA/GAA/CTT/TCT/TTC trinucleotide-repeats (4855), followed by dinucleotide repeats (4500) with dimer motifs AG/GA/CT/TC type and AT/TA. These results are similar to the study of Vatanparast *et al.*, (2016) and Jayashree *et al.*, (2006) which reported SSRs distributed in Expressed Sequence Tags (ESTs) from soybean (*Glycine max*), barrel medic (*Medicago truncatula*) and lotus (*Lotus japonicus*). The study of Wong *et al.* 2017 , identified 18 SSR markers with 8 of them consisted of dinucleotide and 10 consisted of trinucleotide repeated motifs. The 18 SSR markers were validated as polymorphic across the nine winged bean accessions originated from 5 countries. For the 18 SSR markers the individual Polymorphism Information Content (PIC) ranged from 0.16 to 0.67. More specifically, for the 8 dimer SSR markers an average number of 2.5 alleles per locus was observed with an average PIC value 0.37. In addition, the 10 trimer SSR markers amplified on an average of 2.4 alleles per locus, with an average PIC value 0.39. The limited number of accessions used could be one of the reasons why the validation rate of the polymorphic markets was low. Wong *et al.* suggested that increasing the number of screened accessions from different geographical origins could possibly lead to a higher validation rate of polymorphic markers. Interestingly, the study of Yang *et al*., (2018) used five primer pairs of

these markers in their study aiming to identify genetic clusters of winged bean accessions and relations to their geographic origin. In addition, the study of Singh *et al.* (2017) on two winged bean accessions with high and low condensed tannin levels, reported that in the total number of sequences examined 2237 and 1618 SSRs were revealed in the high and low condensed tannin accessions, respectively. In the high condensed tannin accession, 881 SSRs were trinucleotide repeats whereas there were 663 trinucleotide SSRs in low condensed tannins. The development and validation of genic SSR markers provides greater information on the winged bean genome. The construction of linkage map with the molecular markers would help to identify QTLs and assist plant breeding.

The study by Singh et al. (2017) examining the leaf transcriptome of two winged bean accessions, containing different condensed tannin content, revealed more than 1200 contigs that were differentially expressed. They were selected based on low and high condensed tannin content in leaf tissues and variable metabolite concentration in the seeds of the contrasting accessions. The transcriptome and pathway analysis revealed that the anthocyanidin synthase and chalcone synthase genes were less expressed in the low condensed tannin accession, whereas they were highly expressed in the high condensed tannin winged bean accession. Singh *et al.* (2017) proposed that the condensed tannin biosynthesis could take place in the leaves and then be transported to the seeds.

The study of Singh *et al.* (2017) also identified genes and contigs responsible for the biosynthesis of the condensed tannins. The RNA-seq data, using Illumina Nextseq 500 sequencer, generated 102,586 contigs for high condensed tannin accession and 88,433 contigs for the low condensed tannin winged bean accession. Contig generation using the same hash length resulted in 87925 and 69464 contigs for the high and low condensed tannin winged bean accessions, respectively. From both samples, the total number of contigs after clustering at 95% identity and query coverage was 44,972. The total of 44,972 contigs were annotated *with Arabidopsis thaliana, Glycine max* and *Lycopersicum esculentum*. These contigs were assembled and mapped to the reference canonical pathways in KEGG. The similarity searches on the databases gene ontology (GO) and Kyoto encyclopaedia of genes and genomes (KEGG) showed that 5210 contigs were involved in 229 different pathways. The results also revealed differential expression of 1235 contigs detected between the two

accessions. KEGG analysis showed that 10 condensed tannin biosynthesis genes, including anthocyanidin synthase (ANS), 4-coumarate-CoA ligase (4-CCL), chalcone synthase (CHS), chalcone—flavonone isomerase (CHFI), chalcone isomerase (CHI), cinnamyl alcohol dehydrogenase (CAD), dihydroflavonol 4-reductase (DFR), cinnamoyl CoA reductase (CCR), phenylalanine ammonia-lyase (PAL) and anthocyanidin 3-O-glucosyltransferase (A3GT) had lower expression in the leaves of the low condensed tannin winged bean accession compared to the high condensed tannin winged bean accession.

In the same study of Singh *et al.* (2017), the *de novo* assembly of contigs revealed 15 different types of transcription factors families with 33 and 5 contigs encoding for high and low condensed tannin accessions of winged bean, respectively. The transcription factors were more frequently present in high condensed tannin accession than in low condensed tannin accession, with the bHLH group of transcription factors only present in the high condensed tannin accession. However, it would also be interesting to perform transcriptome sequence analysis in the growing pods of winged bean accessions with high and low tannin content to reveal more information about the genes involved in the pathways of condensed tannin biosynthesis.

Tannin content in the seeds has been correlated with the colour of the seeds and flowers as tannins are part of the flavonoid biosynthesis pathway (Klu, Jacobsen, and Van Harten 1997; Smulikowska et al. 2001). Indirect selection on a distinguishable trait like the colour of flowers and seeds could improve selection of accessions with lower levels on antinutritional factors, like tannins, in the winged bean seeds. The study of Klu et al. (1997) generated four mutants with altered tannin content using gamma radiation in the parental accessions and F_1 and F_2 seeds with altered seed coat colour were selected. Interestingly, all the plants from the F_3 generation produced only seeds with altered seed coat colour, providing a strong indication that a recessive mutation could be involved. Identifying genes, transcription factors and markers in high and low condensed tannin accessions of winged bean could be a useful tool in creating new accessions with silenced genes, or in breeding programmes with the aim of reducing the levels of the antinutritional effect of the tannins and improve the nutritional value of the winged beans.

1.6 Future Prospects

It is important to understand the genetics underlying the desirable quantitative and qualitative traits, as it could assist breeding selection and contribute to the development of winged bean varieties with desirable plant architecture and high-quality nutrition products. In the 1980s, Erskine and Khan studied qualitative traits and found that a single gene controlled the shape of the pod and the colour of stem, calyx, pod and pod wings (Erskine and Khan, 1977). They also studied the overall variability within and between landraces of winged bean collected in the Highlands of Papua New Guinea for three quantitative characters (flowering, pod length and seed weight). Significant differences between the landraces were detected for the three traits, suggesting differences in the selection pressures on the landraces probably caused by the local farmers. These differences, especially in adjacent landraces, could be maintained by a low or absence, of gene flow among the landraces (Erskine and Khan 1980). This is of high probability as winged bean is a cleistogamous, self-fertilising crop with limited gene flow even in close proximity (Tanzi, Eagleton, et al. 2019). In the small sample size of 14 landraces examined for variability in stem colour, pod specking, pod wing colour and pod shape, all the loci examined showed allelic polymorphism (Erskine and Khan 1980). Genetic diversity among the accessions could be utilised in breeding programmes to obtain higher variability and introduce desirable traits.

Next generation sequencing technologies and high throughput phenotyping techniques, when combined in genome wide association studies, have the potential to reveal genetic loci that are associated with key traits (D'Agostino and Tripodi 2017). The integration of genomic studies from other legumes, such as soybean and common bean, with the transcriptome sequences of winged bean could be useful in identifying molecular markers. Comparative genomics, the development of molecular markers, linkage maps, and QTL analysis would contribute. By identifying the genetic bases of desirable traits, for the improvement of winged bean plant architecture, yield and nutritional value (Wong *et al.*, 2017).

The best winged bean ideotype, for growing it as a pulse crop, is proposed to be early maturing, with a "bushy" and dwarf architecture, producing seeds with high nutritional value and reduced antinutritional factors without the need for trellising (Klu, Jacobsen, and Van

Harten 1997). Aiming to improve the protein content of winged bean seeds, linkage analysis and genome wide association studies (GWAS) should be combined with quantitative trait loci associated with protein content in winged bean. This will help identifying genes related to protein content and closely linked markers, which would also serve as useful tools for plant breeders.

As an underutilised and largely unimproved crop, it will be important to couple genomics and marker-assisted selection with new approaches to breeding species with flat genetic structures. Moreover, the desirable ideotype for winged bean will depend on the end use and context of cultivation, so a number of ideotypes may be required and a series of selection indices are also needed, ranging from genomic to protein functionality and processing. Such a flat genetic structure together with complex selection indices may provide opportunities for Genomic Selection (GS) models, supplemented by high throughput phenotyping (Montesinos-López et al. 2021). At the least, it is clear that multiple alleles will need to be combined to achieve the breeding objectives, so Multiple Advanced Generational Inter Crosses (MAGIC); (Huynh et al. 2018) may be appropriate, combined with Nested Association Mapping (NAM); (Gangurde et al. 2020) populations to elucidate and identify the desirable alleles in a hybrid-GS approach.

Once desirable ideotypes are in development, they need to be coupled with processing methods to decrease the antinutritional factors; processing methods such as boiling and autoclaving are widely used, with the time and temperature varying based on the amount of the antinutritional factors. In addition, antinutritional factors that are heat resistant are more difficult to decrease, adding extra steps to the processing methods, such as seed coat removal, of the seeds would raise the cost of food products. Therefore, improving the nutritional value of winged bean by identifying varieties with lower levels of antinutritional factors and understanding their genetic control would assist breeding selection, but will take longer to achieve. Breeding should focus on developing not only on greater protein quantity but also quality, taking into account amino acid content as well as reducing levels of antinutritional factors.

In addition, processing methods and production of fermented products such as tempe in winged bean would be more appealing to the consumers. Fermentation and germination of seeds are also methods that can reduce the amount of antinutritional factors (Samtiya, Aluko, and Dhewa 2020). As the winged bean protein is limited in S-amino acids such as methionine and rich in lysine, combining winged bean flour or protein isolates with cereal flour from rice, maize or wheat that have higher methionine and lower lysine would increase the protein quality of the flour. The quality of these two proteins combined will be higher than either of the two components, increasing the biological value due to protein complementation.

1.7 Conclusion

The domestication of crops over thousands of years has led to significant changes in their morphology, plant architecture and yield as compared to their wild ancestors. Nowadays, increased food demand drives the need for higher production of quality food. Donald (1968) suggested the design of a crop ideotypes that would have a predictable performance in a specific environment. Winged bean as a crop has multiple uses and multiple parts of the plant are eaten (e.g. favouring tubers in Thailand and immature pods in Malaysia). This is a major advantage, as tubers, pods and seeds are of high nutritional value.

Winged bean has been cultivated by indigenous communities in Asia and sold in local markets. As an underutilised crop, it has received limited research for the improvement of its vining plant architecture and the utilisation of its high nutritional value. Recent advances in transcriptomics could assist genomic research and accelerate breeding selection using genetic markers. Winged bean could play an important role in food security and more research needs to be done in order to explore the potential of winged bean to become a new soybean for the tropics.

Chapter 2: Materials and Methods

This chapter outlines the research methods used to achieve the aims of this thesis. It provides details on the techniques used, along with the rationale for selecting these methods.

2.1 Samples Preparation

Standard laboratory methods were used to determine the nutritional profile of winged bean seeds, that are in line with the AOAC International Standards. The winged bean seeds were first ground using a coffee grinder and then passed through a 0.5 mm sieve (Ultra-Centrifugal Mill ZM 200, Retsch). This initial step was crucial in ensuring the seeds were ground into a fine powder, minimising the risk of high variability of technical errors in the subsequent analysis. The samples were stored at -80°C overnight and then freeze-dried to remove any moisture using a freeze dryer (Martin Christ Freeze Dryer, Germany). The freeze-drying method was selected to preserve the nutrients and prevent any chemical changes without the use of heat. The nutrient analyses were performed on dry matter.

2.2 Crude protein

The protein content was determined by multiplying the nitrogen content detected using a Protein Analyzer (FlashEA® 1112 N/Protein, Thermo Scientific) with the standard conversion factor of 6.25, meeting the performance requirements of AOAC. 990.03 (Thermo Fisher Scientific, 2007). Across the literature, the conversion factor of 6.25 is commonly used to ensure comparable results. There are a few recommendations for using the factor of 5.7 for legumes (Mossé 1990; Mosse and Baudet 1983). However, this has not been widely adopted, as it could lead to confusion, and results may not be comparable with published studies.

The protein analyser uses a modified Dumas combustion method (Dumas 1831), where the samples are combusted at a high temperature in an oxygen-rich environment, converting all nitrogen to nitrogen gas (N_2). The nitrogen gas was then measured using a thermal conductivity detector and the amount of nitrogen detected was converted to protein content using the conversion factor 6.25 and the weight of the sample was used to calculate the

nitrogen content of the sample as shown in Equation 1; while the protein content as a percentage of the dry sample was calculated as shown in Equation 2.

Equation 1

$$Nitrogen\ detected\ in\ the\ sample\ x\ 6.25 = Nitrogen\ content\ of\ sample$$

Equation 2

$$\frac{Nitrogen\ content\ of\ sample}{Weight\ of\ sample} x\ 100$$

For the quality controls and standards, pure aspartic acid (10.52 N% (w/w)) (D5055, Elemental Microanalysis) was used. They were prepared first to avoid any cross-contamination with the samples. For the standards, quality controls and samples, 50 mg were weighed. Bypasses (empty tin capsules) were also used as a blank. To ensure the quality of the analysis, two quality control samples of aspartic acid were run after every 18 test samples.

2.3 Crude fat

Total fat was determined using 1.5 g of dried winged bean seed powder in the Gerhart SOXTHERM® extraction system, which works on the principle of solvent extraction. Each sample was added to a thimble that was then covered with cotton wool and placed into a metal holder in a glass flask. In the flasks, boiling stones were placed at the bottom to assist with the boiling of the solvent, and weighed (this is then W1). For the solvent extraction, 140 ml of petroleum ether was used. The solvent was boiled and as it evaporated, it was condensed and dripped over the sample dissolving the fat. The system continuously rinsed the sample with condensed solvent, ensuring maximum extraction of fat. The extracted fat was concentrated at the bottom of the flask. The flasks with the extracted fat were placed in desiccators to cool before they were weighed (W2). The weight of fat extracted was calculated as shown in Equation 3. The fat content as a percentage was calculated as shown in Equation 4.

Equation 3

$$W2 - W1 = Weight\ of\ extracted\ fat$$

Equation 4

$$\frac{Weight\ of\ extracted\ fat}{Weight\ of\ sample} x\ 100$$

The fat extracted was then stored in 3 ml hexane at -20°C for fatty acid analysis.

2.4 Fatty acid analysis

The relative amount of fatty acids was determined using GC-MS. The extracted fat (stored in hexane) was used for the esterification of fatty acids to methyl esters (FAMEs). More specifically, 200 µl of the extracted fat in hexane were added in test tubes. Then 700 µl of 10M KOH and 5.3 ml of methanol were added. The test tubes were vortexed for 30 seconds before being placed in a water bath at 55°C for 90 minutes. Every 20 minutes, they were vortexed for 20 seconds. After the 90 minutes, the samples were placed in an ice bath for 10 minutes to lower their temperature. Then 580 µl of 12M H_2SO_4 was added to the test tubes, and the samples were vortexed for 30 seconds before returning to the same water bath at 55°C for 90 minutes, where samples were vortexed every 20 minutes for 5 seconds. The samples were then cooled in an ice bath for 10 minutes, and 3 ml of 98-100% analytical-grade hexane was added. The samples were vortexed and centrifuged for 10 minutes at 1763 x g, room temperature (ThermoFisher Scientific). The sample's solution formed two phases, with the top layer being the organic phase and the bottom containing water and methanol. The top layer containing the esterified fatty acids (FAMEs) was removed and stored in tubes at -20°C.

The FAMEs were detectable by gas chromatography-mass spectrometry (GCMS) as described by O'Fallon et al., 2007. From the FAMEs solution, 400 µl was added to vials and external calibration standard curve was prepared were 200 µl was added to each vial for the GCMS ISQ7000 (GC: Thermo Scientific Trace 1300, MS: Thermo Scientific ISQ 7000). For the GC main

column, the CP-Sil 88, Agilent, 100 m, 0.25 mm, 0.20 µm, was used with inlet temperature at 260°C. The split ratio was 1:10 for standards calibration and 1:50 for samples, and the oven was held at 140°C for 5 minutes and then increased by 4°C every minute until 240°C where it remained for 10 minutes. For the MS the MAS transfer line temperature was set to 250°C and the ion source temperature was set to 200°C. The fatty acids were identified using GC retention time and MS scan data quantified by the Chromeleon™ software (Thermo Scientific). The software provided the relative amount of each fatty acid as a percentage of the total fatty acids detected in the samples.

2.5 Amino acid analysis

Amino acid composition, except tryptophan which is degraded during acid digestion, was determined by HPLC-MS/MS using Vanquish™ Column Compartments and TSQ Altis Quadrupole Mass Spectrometer, Thermo Scientific™ as described by Muleya et al. 2023. Methionine and cysteine were measured as methionine sulfone and cysteic acid respectively due to their oxidation during the processing of the samples. For the amino acid analysis method, 30 mg of each sample was weighed and placed in a tube followed by the addition of 2.5 ml of chilled oxidation solution (10 ml of 30% hydrogen peroxide and 90 ml of formic acid 98-100%). Samples were placed again in the fridge at 4^{o}C for 16 hours. After oxidation, 0.42 g of sodium metabisulphite was added to each sample to degrade any excess oxidation reagent. Then, 2.5 ml 12M HCl and 0.5 ml of hydrolysis reagent (6M HCl with 1% phenol) were added to each sample. The samples were then placed in an oven at 110 ^{o}C for 24 hours.

After hydrolysis, the samples were cooled on ice for 10 minutes and quantitatively transferred to 50 ml falcon tubes and ammonium formate (pH 2.8, 20 mM) was used to rinse the content into the tubes. Then 16 ml of 4M ammonium formate was added to the falcon tubes. The pH was adjusted to 2.8 using 2 concentrations of ammonium formate at 4 M or 100 mM, as needed. The volume was up to 50 ml using 20 mM ammonium formate with pH 2.8. The samples were then centrifuged at 4996 x g for 10 minutes at room temperature, and the supernatant was passed through a 2.22 µm filter into an HPLC vial. The dilutions were made to ensure the nitrogen content level was equivalent to 1-10 µg/ml before being placed on the HPLC-MS.

In HPLC-MS/MS (High-Performance Liquid Chromatography coupled with Tandem Mass Spectrometry), positive ion mode was used for the amino acids due to their basic functional groups, which can be easily protonated in solution resulting in enhanced detection and reliable mass spectrometric analysis. From the sample, 1 µl was injected into the 30 ^{o}C Trinity P1 mixed mode column (150 mm x 2.1 mm, 3 µM, Thermo Scientific™ Acclaim™). The mobile phase A consisted of 20 mM ammonium formate in water at pH 2.75, and mobile phase B consisted of 100 mM ammonium formate in water and acetonitrile (80:20 v/v). Chromatographic separation was achieved through gradient elution as shown in Table 2. 1.

Table 2. 1 Conditions of chromatographic separation via gradient elution

Time (min)	**Flow Rate (ml/min)**	**Mobile phase A (%)**	**Mobile phase B (%)**
0	0.3	100	0
5	0.3	100	0
7	0.3	0	100
14	0.3	0	100
14.5	0.35	100	0
16.5	0.35	100	0
17	0.3	100	0
18 0	0.3	100	0

For the MS, the triple quadrupole mass spectrometer has an ion source with a heated electrospray ionisation probe. Sheath gas was set at 45 arbitrary units, auxiliary gas at 15 arbitrary units. The spray voltage for positive ionisation was at 3500 V and for negative at 2700 V. Vaporiser temperature was set to 370^{o}C and transfer tube temperature was set at 270^{o}C. Source fragmentation was applied at 15 V. For the data collection, Selected Reaction Monitoring (SRM) mode was used in a resolution of 0.7 full width at half maximum for both quadrupoles.

The TSQ Altis™ triple quadrupole mass spectrometer was controlled by TSQ Altis Tune Application software while the operation and sample run were controlled by XCalibration software. For the data integration and extraction, Chromeleon™ (Thermo Scientific) software was used. The data was normalised using the recovery of the Standard Reference Material® 3234 Soy Flour and gross amino acid compositions of the substrates were expressed as mg/g material. The percentage of each amino acid was calculated as well as their mg per 1 g of protein, for each sample. It is worth noting that methionine and cysteine were measured as methionine sulfone and cysteic acid, respectively due to their oxidation during the processing of the samples. Additionally, it was not possible to measure tryptophan as it was degraded during the acid digestion

2.5.1 Autoclaving of samples

The winged bean seed samples that were selected for the in vitro digestion were first autoclaved to reduce the effect of the antinutritional factors, mainly protein inhibitors, before the in vitro digestion. Winged bean seed powder (0.5 g) was autoclaved at 120^{o}C for 15 minutes. The heating cycle was 35 minutes from 20^{o}C to 120 ^{o}C and another 15 minutes at 120 ^{o}C. The cooling phase lasted for approximately 30 minutes until the temperature dropped to 75 ^{o}C. The door was opened, and the samples were left to cool at room temperature.

2.6 Digestible Indispensable Amino Acid Score (DIAAS)

2.6.1 In vitro digestion

For the in vitro digestion, around 80 mg of protein per sample was required. Therefore, the amount needed for each sample was calculated using the protein content, previously detected using the protein analyser. After the calculations, the autoclaved winged bean samples were accurately weighed, with an average of approximately 0.2 g of powdered winged bean seed used from each sample for digestion. A pH adjustment test was run to determine the volume of 2 M HCl and 2 M NaOH required for each sample to adjust the pH to 3 (for gastric phase) and 7 (for intestinal phase), respectively.

Blank and positive control samples were used in triplicate. For the blank, a pre-made protein-free cookie (PFC) was used and each weighed around 1 g. A highly digestible protein, casein (skim milk powder protein 42.34%, lactose 49.8% and fat 0.89% w/w Fonterra) was used as a positive control. Around 0.092 g of casein was mixed with 0.25 g of the PFC. The PFC consisted of 40.8 g purified corn starch, 15.7 g sucrose, 4.9 g cellulose, 0.7 g baking powder, 0.5 g ground ginger and 36.9 g margarine. The margarine was melted, and all the ingredients were mixed. Then 35 g portions were baked at 175^{o}C for 30 minutes. In order to mimic the different phases of the digestive tract in the static digestion procedure, enzymes in a buffer of a specific pH were added to the sample at each phase. The enzymes were prepared, and the procedure was followed as described by Brodkorb et al. (2019).

For the oral phase, the simulated digestion fluids were preheated to 37°C. Each sample was mixed with 800 µL of warm simulated salivary fluid (SSF) electrolyte stock solution, which was adjusted to a pH of exactly 7. Next, 10 µL of 0.3M CaCl2 and 90 µL of MilliQ water were added to the mixture. For winged bean seeds, as a starch-containing food, 100 µL of α-amylase stock enzyme solution was added. The mixture was then incubated at 37°C for 2 minutes on an end-over-end rotator placed in an incubator.

For the gastric phase of digestion, a pre-determined volume of 2M HCl (recorded from the pH adjustment test), was added to each sample. Next, a master mix was prepared where 1.6 mL of pre-warmed simulated gastric fluid (SGF) electrolyte stock solution (pH 3) was added, followed by 1 µL of 0.3M CaCl2 and 79 µL of MilliQ water. Subsequently, 320 µL of pepsin stock solution, prepared in water, was introduced. The master mix was prepared based on the number of samples and 2 mL of this master mix was added to each sample. The samples were incubated for 2 hours at 37°C in an end-over-end rotator within an incubator. After digestion, the pepsin activity was inactivated by adjusting the pH to 7.0 ± 0.1 with 2M NaOH, using the volume recorded from the previous pH adjustment test.

Approximately 30 minutes before the completion of the gastric phase, bile was dissolved in simulated intestinal fluid (SIF) and vortexed until fully dissolved, then brought to 37°C. Pancreatin was also dissolved in SIF and vortexed for at least 10 seconds, followed by placement in an ultrasound bath for 5 minutes at room temperature (45 Hz, 130 W). Then,

the pancreatin solution was centrifuged for 5–10 minutes at 3,000 x g at room temperature. The resulting supernatant was transferred into a fresh Falcon tube and kept on ice until needed for the next step of the digestion process.

For the intestinal phase, the gastric digestion mixture with pH already adjusted to 7 ± 0.1 was used. To this, a master mix was prepared with 784 µL of MilliQ water and 16 µL of $CaCl_2$ solution (0.3M). Then, 1.6 mL of bile dissolved in SIF and 1.6 mL of pancreatin dissolved in SIF were introduced to the master mix. For efficiency, the master mix was again prepared for the number of samples to be analysed, and 4 mL of this master mix was added to each sample. The samples were incubated for 2 hours at 37°C in an end-over-end rotator placed in an incubator at 37°C.

For the precipitation process, first empty 50 mL falcon tubes with their caps were weighed and the weight was recorded as W1. An absolute methanol bottle was placed on ice a few minutes before intestinal digestion was completed. Immediately after digestion, the samples were placed on ice. Quantitatively the digesta were transferred to the pre-weighed 50 mL falcon tubes and filled to 40 mL with ice-cold absolute methanol. The tubes were weighed again and the new weight was recorded as W2. Then the tubes were incubated at -20°C overnight. The following day, they were centrifuged at 4°C at 4,000 x g for 15 minutes. The supernatant was carefully transferred to new falcon tubes using a pipette, leaving about 5 mL of supernatant with the pellet to avoid collecting the interface. The supernatant was stored at -20°C until further use. To each pellet, 5 mL of ice-cold 80% methanol was added to the tubes and vortexed to resuspend the pellets. Then, it was centrifuged again for 10 minutes at 4,000 x g and 4°C, and the supernatant was removed by pipetting. This step was repeated by adding another 5 mL of 80% methanol, followed by centrifugation at 4,000 x g and 4°C for 10 minutes. The supernatant was discarded, and the tubes were left open under a fume hood overnight to evaporate the methanol. The tubes were placed at -80°C for storage and in liquid nitrogen if possible, before freeze-drying. After freeze-drying, the tubes were stored in a desiccator and the final weight of the dried pellets was recorded as W3. The pellets were stored at -20°C until further analysis. The weight of the supernatant is calculated as (W2 - W3), and the weight of the dried pellet is (W3 - W1).

To calculate AA digestibility the following equation was used:

$$AA\ digestibility\ \% = \frac{Fs - Cs}{Fs - Cs + \max(0; Fp - Cp)} x\ 100$$

Where:

Fs is the amount (mg) of amino acids in the digesta supernatant of the test sample

Cs is the amount (mg) of amino acids in the digesta supernatant of the cookie blank/control

Fp is the amount (mg) of amino acids remaining in the test sample pellet

Cp is the amount (mg) of amino acids remaining in the pellet of the cookie blank/control

If Fp – Cp gives negative values, the value was truncated to 0

2.6.2 Hydrolysis procedure of digesta samples for amino acid analysis

2.6.2.1 MeOH digesta extracts (Supernatants)

To prepare the samples for analysis, 500 µL of the supernatant solution was transferred into a 2 mL Eppendorf tube. The samples were dried in a vacuum concentrator. After drying, 2 mL of chilled oxidation solution (10 ml of 30% hydrogen peroxide and 90 ml of formic acid 98-100%) was added and the sample was left to dissolve for a few minutes. The solution was gently mixed by swirling the tubes until the sample was completely dissolved and then covered lightly with cling film. The samples were incubated at 4°C for 16–18 hours with the caps opened to allow for proper oxidation and the resulting solution was carefully transferred to hydrolysis tubes. Then the same steps as described above in the amino acid analysis were followed.

2.6.2.2 Digesta pellets

To prepare the hydrolysis samples for amino acid analysis, the entire pellet was weighed into hydrolysis vessels. Next, 2.5 mL of the oxidation solution was added, and the mixture was vortexed gently until the sample was completely dissolved. The solution was then incubated at 4°C for 16–18 hours to ensure thorough oxidation. If necessary, the solution was carefully

transferred to hydrolysis tubes. Following this, 0.42 g of sodium metabisulfite was added and the procedure was followed as described above in the amino acid analysis.

Both the supernatant and pellets were used to estimate the amount of amino acids present. To determine the amino acid digestibility, the amino acid data from undigested winged bean seed samples, which were previously analysed for total amino acid content, were used. The digestible indispensable amino acid score (DIAAS) was calculated as described by Muleya et al., 2023.

2.7 Determination of Phytic acid content

For the determination of phytic acid, the protocol from the phytic acid (phytate) kit from Megazyme was used (K-PHYT, Megazyme 2019). For the method, 1 g of winged bean seed powder and 20 ml of 0.66 M HCl were stirred vigorously overnight.

The method involved acid extraction of inositol phosphates, and treatment with phytase that specifically hydrolysed phytic acid and the lower myo-inositol phosphate forms, as shown in Equation 5.

Equation 5:

$$Phytic\ acid\ +\ \ H_2O\ \ \xrightarrow{phytase}>\ myo-inositol\ (phosphate)n\ +\ Pi$$

Following the first reaction, alkaline phosphatase (ALP) further hydrolysed myo-inositol (phosphate)n producing myo-inositol and Pi, as shown in Equation 6. The treatment with ALP is useful to release the final phosphate from myo-inositol phosphate which can be relatively resistant to phytase.

Equation 6:

$$myo - Inositol\ (phosphate)n \xrightarrow{ALP} myo - inositol + Pi$$

Pi and ammonium molybdate reacted to form 12-molybdophosphoric acid, which was subsequently reduced under acidic conditions to molybdenum blue, as shown in Equations 7 and 8.

Equation 7

$$Pi + ammonium\ molybdate\ 12 - molybdophosphoric\ acid$$

Equation 8

$$12 - molybdophosphoric\ acid + H_2SO_4/ascorbic\ acid \rightarrow molybdenum\ blue$$

The amount of molybdenum blue was measured by the increase in absorbance at 655 nm and was proportional to the amount of Pi present in the samples. Pi was quantified as phosphorus from a calibration curve generated using standards of known phosphorus concentrations provided by the kit.

2.8 Determination of Total Phenolic Content

For the determination of total phenolics the method from Makkar (2003) was used. The total phenolics were expressed as tannic acid equivalents (TAE). For that purpose, 1 g of tannic acid (TA) was diluted in 10 ml of distilled water (1:10 w/v) and vortexed. A standard curve was made as shown in Figure 2. 1.

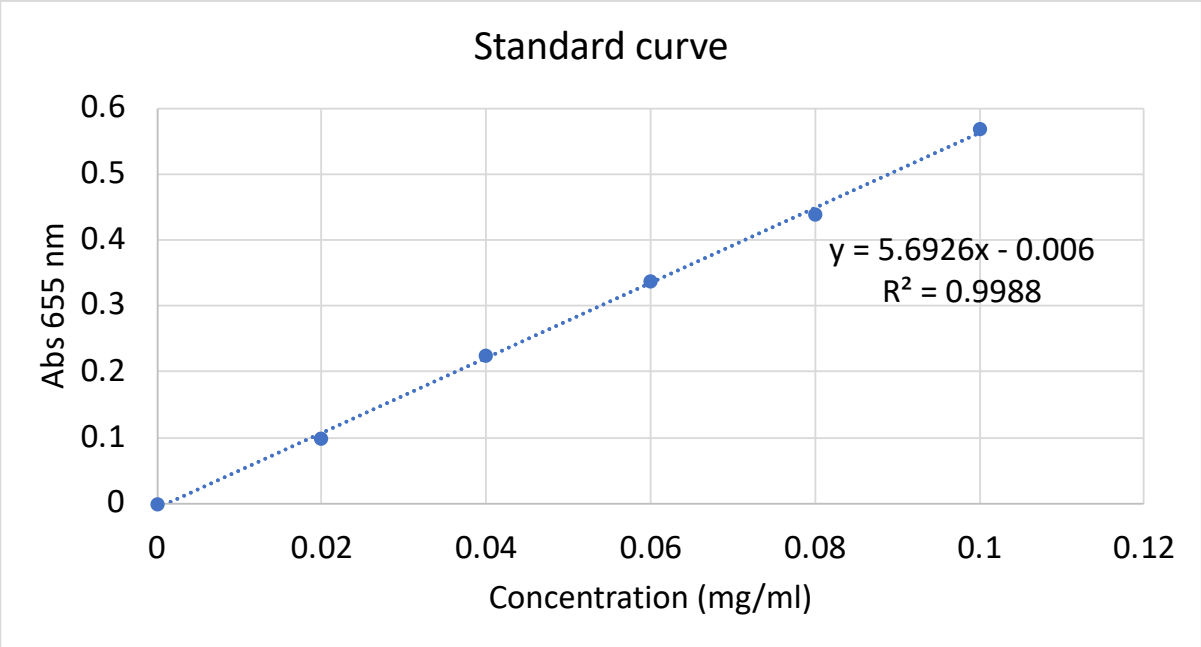


Figure 2. 1 Standard curve of TA using 0-0.1 mg/ml.

Winged bean seed powder of around 0.1 g was diluted 1:50 (w/v) in 5 ml of 0.1M NaOH. The samples were vortexed for 60 seconds and then were stored overnight in the fridge at 4°C. The next day, the samples were vortexed again for 60 seconds. A colour change was noted in the solutions from white to brown. From each sample, 50 µl was placed into 1.5 ml tubes and further diluted 1:6 (v/v) using 0.1 M NaOH. They were then centrifuged at 9,520 x g for 10 minutes at room temperature. The supernatants were transferred into fresh 1.5 ml tubes and stored at 4°C until used. From each sample, 25 µl were dispensed on a 96 well plate in triplicate. Then, in each well the following were added: 200 µl of distilled water, 25 µl of Folin-Ciocalteu reagent (1: 3 (v/v) in distilled water) and 25 µl sodium carbonate (20% w/v). A colour change from clear to blue was observed. As it is sensitive to light exposure, the plate was covered in foil and left at room temperature for 1 hour, before measuring the absorbance at 655 nm at room temperature. The amount of total phenolics was calculated using the standard curve by the plate reader software Microplate Manager 5.0 from Bio-Rad Laboratories Inc.

2.9 Determination of phenolic compounds - condensed tannins

The winged bean seed and pod samples were ground using a freezer mill and then freeze-dried using a vacuum freeze drier. Depending on sample availability, the samples were weighed out accurately as shown in Table 2. 2. As winged bean samples were collected during pod and seed development, the sample material available varied. The volume of the extraction solvent was adjusted as shown in the table below (Table 2. 2).

Table 2. 2 Volume of extraction solution used according to the samples weight.

Sample weight (mg)	Volume of extraction solution (ml)
10	0.25
20	0.5
40	1
50	1.25

The samples were weighed into a 1.5 ml microcentrifuge tube, and the extraction solvent of 2% formic acid in absolute methanol including internal standard 1 (IS1) was added as shown in Table 2. 3, and vortexed for 30 seconds. The samples were sonicated for 10 minutes and centrifuged at 4 ^{o}C for 20 minutes. Then, the supernatant was filtered using a 0.22 nm filter into a collection tube and stored at -20^{o}C. The extraction was repeated four times, and the extracts were all collected into the same collection tube. From there, 400 µl of the extract was added into a fresh tube with 400 µl of distilled water and vortexed to mix. From this mixture, 360 µl was placed into a vial and 40 µl of internal standard (IS2) was added in a ratio of 1:10. For the standard curve, both the IS1 and IS2 were included at a final concentration of 0.0015 mg/ml in a 1% formic acid 50% methanol solution. The following standards were used: catechin, epicatechin, procyanidins B1, B2 and C1. Condensed tannins were expressed as the total of these compounds. To correlate the colour changes of the pods to the concentration of the anthocyanins, anthocyanin standards, cyanidin 3-O glucoside and pelargonidin 3-O glucoside, were also used. However, the results of this analysis are not included as the method requires further optimisation.

Table 2. 3 Internal standards and their concentration. As the IS2 was diluted, the end concentration of both IS1 and IS2 was 0.0015 mg/ml.

Internal standard (IS)	Compounds	Concentration (mg/ml)
IS1	Taxifolin	0.003
	3-methulcatechol	0.003
IS2	Vanillin	0.0015
	Trans-cinnamic acid	0.0015

The LCMS analysis was performed using an Agilent 1260 Infinity II UHPLC, coupled to an Agilent 6546 tandem Quadrupole – Time of Flight mass spectrometer (Agilent Technologies, Cheadle, UK). For the chromatographic separation, the following were used: a Kinetex Biphenyl column (1.7 u, 100 x 2.1 mm; Phenomenex, Macclesfield, UK) held at 40 ^{o}C. Solvents A (type 1 water) and B (methanol), each including 0.1% (v/v) formic acid were used. The gradient flowing at 0.3 ml/min, was held at 5% B for the first 2 minutes, then increased to 19% at 8 minutes, and at 82.5% at 17 minutes. Solvent B was raised to 95% by 18 minutes, held for 2 minutes, and then returned to initial conditions over 2 minutes with a further 3 minutes of reconditioning.

The electrospray ionisation (ESI) source used drying gas at 320 ^{o}C flowing at 8 L/min, and sheath gas at 350 ^{o}C flowing at 11 L/min. The nebuliser was set to 35 psi, and VCAP and nozzle voltages were at 3500 V and 1000 V respectively. The voltages for the Fragmenter, Skimmer and Octopole RF were 110, 65 and 750 V respectively. The MS1 data was acquired in negative form (-)ESI, between 50 m/z-1700 m/z.

2.10 QTL analysis

QTL analyses for protein, oil and fatty acid contents were performed using MapQTL v6 on F_3 seeds from F_2 individuals of the XB2 cross between Ma3 and FP15. For the QTL analyses both non- parametric and parametric tests were used. The data was tested for normality using Shapiro-Wil test, and a Permutation test was used to calculate the genome-wide (GW)

significant logarithm of the odds (LOD) threshold ($\alpha = 0.05$) for protein content, oil content and fatty acids. Then, each trait was analysed through a Kruskal-Wallis (KW) test to establish single marker-trait associations (at $p< 0.01$), followed by Interval Mapping (IM) analysis using the GW LOD as a threshold. QTLs consistent between the two tests were reported. A QTL was considered significant when equal to or above the GW-LOD threshold and explained ≥ 10% of phenotypic variance (PVE%). Multiple-QTL model (MQM) mapping was utilised for QTL analysis of seed protein, oil and fatty acid contents. The QTLs on the genetic map were placed using MapChart v2.32, including markers with high LOD score and the flanking markers of a 2-LOD drop. The amino acid sequences of the genes from soybean and Arabidopsis were blasted on winged bean amino acid sequences, using the CLC software by Wai Kuan Ho. The winged bean genes were considered as homologues to genes in soybean and Arabidopsis when the E-value $\leq 10^{-30}$and $\geq$ 70% in sequence similarity (soybean). Further information on whole genome sequencing, genetic mapping and bidirectional BLASTP has been reported by Ho et al. 2024.

2.11 RNA sequencing

Three winged bean plants (FP15) were grown in the growth chamber at 26°C/20°C day/night temperature with a 12-hour day length at the University of Nottingham Sutton Bonington Campus. Each plant was defined as a biological replicate. The pods and seeds of each plant were sampled on days 7, 15, 22, 30, 37 and 45 after flowering (DAF). As the plants were grown in a 12-hour day length, sampling was carried out at mid-day, around the 6th hour of their day.

As the winged pods are quite long, particularly at Days 30 and 45, each pod was divided into three equal parts (top, middle, and end). The samples were rapidly frozen in liquid nitrogen and stored at –80 °C. To separate the seeds from the pods before RNA extraction, a pestle, mortar, and a set of forceps were pre-cooled using liquid nitrogen. Each part of the pod was placed in a mortar and cracked open with a pestle. The seeds were removed using the forceps and placed into a pre-cooled Falcon tube submerged in liquid nitrogen.

The winged bean pod and seed samples were ground using a pestle and mortar, followed by a freezer mill. Total RNA was extracted separately from the three pod sections and from the seeds using the RNeasy® Plant Mini Kit, following the manufacturer's instructions. Before

being sent for sequencing, the RNA extracts from each pod section were pooled according to concentration to ensure that RNA from each part of the pod (top, middle, and end) was represented in the final sample.

Pod and seed samples from Days 15, 30, and 45 after flowering, and pod samples from Day 37 were sent for sequencing, due to cost restrictions. Days were selected at 15-day intervals starting from the day of flowering, considering that winged bean seeds typically reach maturity and are harvested around 60 days after flowering. Library preparation and transcriptomic sequencing were conducted by Novogene (Cambridge) using an Illumina NovaSeq6000 sequencer (Novogene, Cambridge). The 150-bp paired-end reads were mapped using HISAT2 v2.0.5 followed by differential quantification analysis at fragments per kilobase of transcript per million fragments mapped (FPKM) level using feature Counts 1.5.0-p3 and DESeq2 v1.20.0. The screening criteria for differential genes were |log2(FoldChange)| >= 1 & padj<= 0.05 as commonly used empirical values in actual projects. Where:

- log2FoldChange: the ratio of gene expression level between the treatment group and the control group was processed by the shrinkage model of the differential analysis software, and finally the logarithm was taken with 2 as the base
- padj: the corrected p-value of multiple hypothesis test.

Chapter 3 – Winged bean: nutritional profile and in vitro digestibility

3.1 Introduction

Legumes are an important source of protein and oil for many people around the world (Maphosa & Jideani, 2017). *Psophocarpus tetragonolobus* (L.) DC., is a self-pollinated legume grown in tropical regions. It is commonly known as winged bean, goa bean and asparagus pea (Lepcha et al., 2017). The mature seeds of the winged bean have a relatively high protein content, similar to soybean. The protein content ranges from 29 to 37%, the oil from 15 to 18% and the carbohydrate from 25-38% (Kadam, 1984). Even though a large proportion of the seed is carbohydrates, the carbohydrates in legumes are digested slowly and can be part of a low glycemic index diet (Jenkins et al. 2012).

As in most legume seeds, the sulphur-containing amino acid methionine is the limiting amino acid with cysteine following; whereas lysine is in abundance (Wan Mohtar et al., 2014). Maize (as an example of cereals) is low in lysine and has a higher amount of methionine (Das et al., 2021), compared to winged bean. Therefore, this indicates that winged bean (low in methionine and high in lysine) could potentially complement well with cereals. Winged bean seeds contain saturated fatty acids at 30–40% and unsaturated at 60–70% (Khor et al., 1982). The oil extracted from winged bean seeds contains a favourable amount of unsaturated fatty acids, mainly linoleic acid (18:2) at around 35% (Mohanty et al., 2015). However, little is known about the environmental impact on the nutritional value of winged bean seeds as well as their digestibility and the Digestible Indispensable Amino Acid Score (DIAAS). This information would be highly beneficial for incorporating winged bean into an animal feed plan, ensuring that it meets the necessary nutritional requirements.

As a leguminous crop, winged bean seeds contain antinutritional factors (ANFs). The ANFs are classified into two categories based on their structure: protein and non-protein-containing groups. The protein-containing ANFs are more heat-liable and can be

deactivated by several processing methods (such as protein inhibitors) compared to the non-protein-containing ANFs such as condensed tannins and phytic acid. The antinutritional factors interact with the proteins, inhibiting their absorption and therefore lowering the protein digestibility and bioavailability of winged bean seeds. In the mature winged bean seeds, it has been reported that condensed tannins comprise between 1.36% and 3.43% of dry seed weight, and phytate between 4.09% and 9.96% of dry seed weight in unprocessed seeds (De Lumen & Salamat, 1980; N. H. Tan et al., 1983).

Proanthocyanidins or condensed tannins are poorly absorbed through the gut barrier and metabolised by the intestinal microflora (Gonthier et al., 2003; Ruiz-Aquino et al., 2023). In the small intestine, procyanidins can form complexes with proteins, starch and digestive enzymes. Proanthocyanidin-protein complexes are less soluble and digestible to enzymes, where one mole of proanthocyanidins is reported to bind 12 moles of protein (Gonthier et al., 2003; Serrano et al., 2009).

A major form of phosphorus storage in legumes and cereals is phytate. Phytates or phytic acid, has the ability to chelate micronutrients such as magnesium, iron, zinc, calcium and potassium (Kies et al., 2006). Monogastric animals like pigs and poultry, as well as humans, lack the digestive enzyme phytase that catalyses the hydrolysis of phytic acid, preserving the bioavailability of the micronutrients (Kishor Gupta et al., 2013).

Digestion is a complex process where several enzymes are involved to break down carbohydrates, fat and proteins. This study focuses on the enzymes used in the in vitro digestion of proteins INFOGEST. Starting from the oral phase, α-amylases initiate the digestion of starch by acting on the α-1,4 glycosidic linkages in amylose, resulting in maltose and glucose, not acting on maltose, a disaccharide composed of two glucose subunits linked by an α-1,4 linkage. They are then deactivated by the low-pH gastric acid in the stomach. In the gastric phase, pepsin cleaves peptide bonds formed by aromatic amino acids (phenylalanine, tyrosine, and tryptophan) as well as before proline, according to recent findings (Rodriguez et al., 2008; Smith & Morton, 2010; Suwareh et al., 2021). In the intestinal phase of digestion, pancreatin was used. Pancreatin contains several enzymes, in this case mainly focusing on trypsin (100 U/mL trypsin activity in the

final mixture) where trypsin cleaves peptide linkages on the carboxylic acid group of lysine and arginine. This specificity results in smaller peptides, which is essential for their absorption. Other enzymes are contained in the pancreatic juice such as chymotrypsin which cleaves after the aromatic amino acids, phenylalanine, tyrosine, and tryptophan; and elastase which breaks down elastin (Smith & Morton, 2010).

This study aimed to investigate the nutritional composition of various winged bean accessions grown in two distinct environments, focusing on three key areas: (1) assessing the nutritional profile, including the influence of environmental factors, genotype, and their interaction; (2) examining the presence of antinutritional factors such as phytic acid and total phenolics; and (3) evaluating protein digestibility using the in vitro digestion model, INFOGEST. This aims to provide a comprehensive understanding of the impact of environment and genotype on the nutritional value of winged bean seeds and its potential as a nutritious food source.

3.2 Materials and Methods

3.2.1 Seed collection

The winged bean seeds were harvested from two different field trials in Malaysia, supervised by the University of Nottingham in Malaysia. The winged bean seeds were harvested, and the field trials were performed by the PhD student Yuet Tian Chong. Information about the environmental conditions in the field trials can be found in the supplementary materials (Supplementary Figure 3. 1, Supplementary Figure 3. 2) as mentioned in Tian's (Chong 2024). The field trial at the Future Research Centre (FRC) (2°55'51 N, 101°52'34 E) was conducted between May 2019 and March 2020 with triplicates in randomised complete block design, and the second field trial in a commercial organic farm named Firefly Farm (FF) (2º56'21''N 101º54'08''E) started in January 2021 and was completed in October 2021 with five replicates in randomised complete block design. In terms of fertiliser input, at the FRC location foliar spray of N-P-K (nitrogen-phosphorus-potassium) was used at 6-days after transplanting and 10 g of N-P-K 15-15-15 mixed with lime were added to each plant every month. During the flowering and podding stages, the fertiliser input was increased to 20 g per plant, with N-

P-K 12-12-17. In contrast, for the FF trial, poultry manure was used whenever required following organic farming practices. The seeds sown in the two trials came from different seed batches, and as there is always a possibility of cross-pollination in winged bean seeds. The winged bean seeds were collected from mature pods, which had turned completely brown and dry at the field were subjected to another round of oven drying at 30°C until constant weight was obtained for measurements and stored at in tightly sealed barrels at room temperature with silica beads, before they were sent to the University of Nottingham, UK, for nutritional analysis. From the selected 20 winged bean accessions (representing a mini core collection) grown in each location, not all produced sufficient seed material from at least three biological replicates (one biological replicate is one individual plant). In the FF location, only 19 winged bean accessions provided sufficient seed material, whereas in FRC only 11 winged bean accessions (Figure 3. 3). These accessions were selected from 91 accessions (Figure 3. 2, Figure 3. 3), originating from World Vegetable Cenre (AVRDC), International Institute of Tropical Agriculture (IITA), Malaysian Agricultural Research and Development Institute (MARDI), East-West Seed company (EWS) and a personal donor, Dr Graham Eagleton. The mini core collection represents the genetic diversity as well as the viable seeds collected after a round of single seed descent (SSD) purification and multiplication in the field, except those from EWS (dried leaf provided for genotyping), as mentioned in the thesis of Alberto Stefano Tanzi (Tanzi 2018).

Continent	Country (Code)	N. of Accessions	Sources
Africa	Liberia (LBR)	2	IITA
Africa	Nigeria (NGA)	4	EWS, IITA
America	Colombia (COL)	5	EWS
Asia (East)	China (CHN)	1	EWS
Asia (South East)	Indonesia (IDN)	10	WVC, EWS, IITA
Asia (South East)	Malaysia (MYS)	8	WVC, Donors, EWS, MARDI
Asia (South East)	Myanmar (MMR)	2	Donor
Asia (South East)	Philippines (PHL)	23	WVC, EWS
Asia (South East)	Singapore (SGP)	1	Donor
Asia (South East)	Thailand (THA)	22	WVC, EWS
Asia (South)	Bangladesh (BGD)	2	EWS, IITA
Asia (South)	Sri Lanka (LKA)	6	EWS
Oceania (South Pacific)	Fiji (FJI)	1	WVC
Oceania (South Pacific)	Papua New Guinea (PNG)	4	EWS, IITA

Figure 3. 1 Summary of countries divided by continent and the number of accessions from each of them as featured in the thesis of (Tanzi 2018).

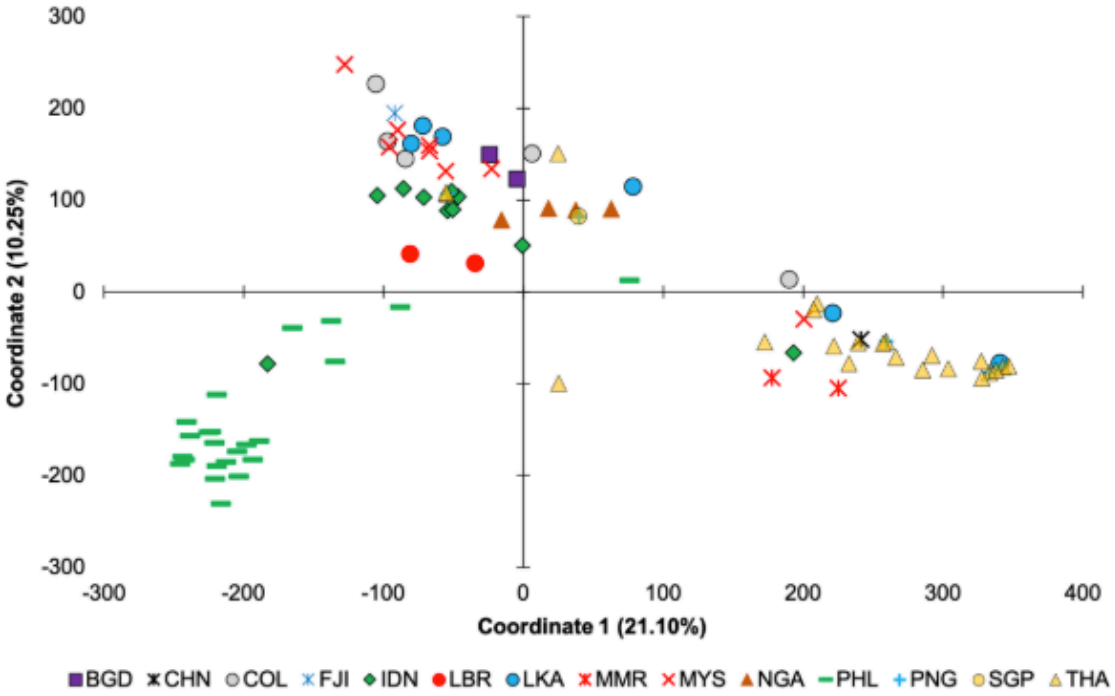


Figure 3. 2 Principal Coordinate Analysis (PCoA) from pairwise individual-by-individual genetic distance (GD) analysis of accessions from 15 countries. 31.35% of cumulative variation was explained by the 2 axes. Legend reports the symbols for each country of origin (see country codes in the table above), as featured in the thesis of (Tanzi 2018).

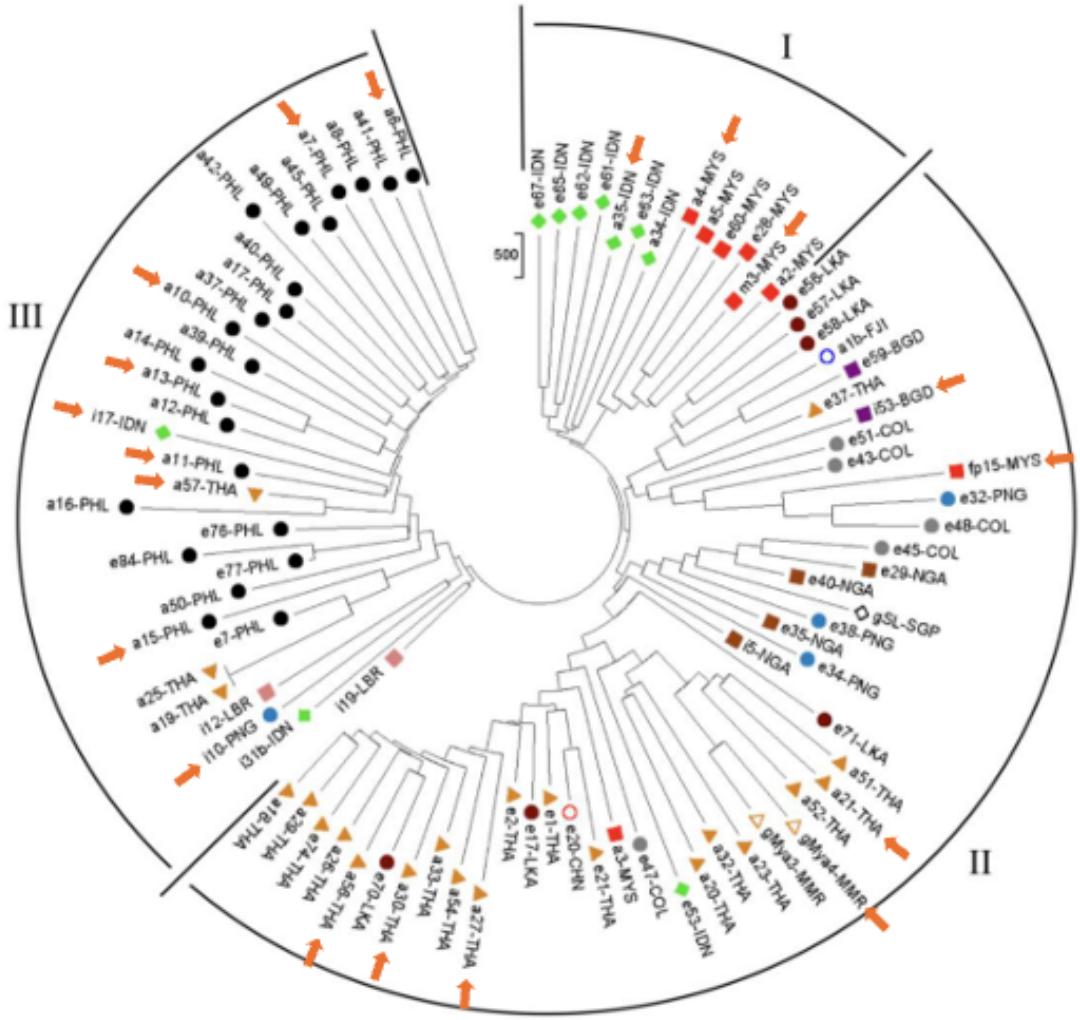


Figure 3. 3 Neighbour-Joining (NJ) method tree generated using the GD matrix in MEGA 7. Branch lengths are to the same scale as genetic distances, as featured in the thesis of Tanzi (2018). The winged bean accessions used in this study are shown with orange arrows.

3.2.2 Analytical Methods

Standard laboratory methods were used to analyse the winged bean seeds. The methods used were in line with the AOAC International Standards. The winged bean seeds were first ground using a coffee grinder and then passed through a 0.5 mm sieve (Ultra-Centrifugal Mill ZM 200, Retsch). The samples were stored at -80°C overnight and then freeze-dried. This study focused on the protein and fat content of mature winged bean seeds. Carbohydrate and fibre contents were not measured due to time constraints and technical limitations related to equipment availability and sample quantity.

The **protein content** was determined using a Protein Analyzer (FlashEA® 1112 N/Protein, Thermo Scientific) and the standard conversion factor of 6.25, meeting the performance requirements of AOAC. 990.03 (Thermo Fisher Scientific, 2007). The **fat content** was quantified using the Gerhart SOXTHERM® extraction system which works on the principle of solvent extraction. For each sample, 1.5 g of ground seeds were added to a thimble that was then covered with cotton. For the solvent extraction, petroleum ether was used. The oil extract was calculated by subtracting the weight of the pre-weighed flasks from the flasks containing the oil extract. The oil extract was stored in hexane at -20°C for fatty acid analysis.

3.2.3 Determination of fatty acid profiles

The relative amount of fatty acids was determined using GC-MS. After the oil was extracted from the winged bean seeds, it was resuspended in hexane and stored at -20°C. The oil extract was used for the esterification of fatty acids to esters (FAMEs) using methanol, 10M KOH, 12M H_2SO_4 and hexane. The FAMEs were detectable by the GC-MS as described by O'Fallon et al., 2007.

3.2.4 Amino Acid Analysis

Amino acid composition was determined by HPLC-MS using Vanquish™ Column Compartments and TSQ Altis Quadrupole Mass Spectrometer, Thermo Scientific™. Methionine and cysteine were measured as methionine sulfone and cysteic acid,

respectively due to their oxidation during the processing of the samples. Note that it was not possible to measure tryptophan as it was decomposed during the acid digestion.

3.2.5 Autoclaved

Winged bean seed powder 0.5 g was autoclaved at 120°C for 15 minutes to reduce the effect of the antinutritional factors, mainly protein inhibitors, before the in vitro digestion. The heating cycle was 35 minutes from 20°C to 120 °C and another 15 minutes at 120 °C. The cooling phase lasted for approximately 30 minutes until the temperature dropped to 75 °C. The door was opened, and the samples were left to cool at room temperature.

3.2.6 In Vitro Digestibility (INFOGEST)

From the 20 winged bean accessions grown in the FireFlies Farm, 8 accessions were selected to estimate their protein digestibility in vitro, based on the protein, oil, and antinutritional factor contents. For each winged bean accession, 3 biological replicates were used. A highly digestible protein, casein, was used as a positive control and a protein-free cookie was used as a blank. In the static digestion procedure, in order to mimic the different phases of the digestive tract, enzymes in a buffer of a specific pH were added to the sample. Starting with the **oral phase**, salivary fluid (pH 7), of $CaCl_2$ (1.5 mM in the final mixture) and α-amylase were added to each sample. Then the **gastric phase** was initiated by the addition of gastric fluid (pH 3) containing $CaCl_2$ and pepsin. To stop the incubation, the pH was increased to 7, by adding 2M NaOH into the mixture. For the **intestinal phase**, the simulated intestinal fluid containing pancreatin, bovine bile and $CaCl_2$ was added to the mixture. After the incubation, the samples were placed on ice and ice-cold absolute methanol was added, to precipitate any remaining proteins or long peptides. The samples were placed in the -20°C freezer overnight and the next day, 80% (v/v) methanol was added to help separate the absorbable peptides and free amino acids (still in the supernatant) from the non-digested proteins and unabsorbable peptides (present in the pellet). Both the supernatant and pellets were used to estimate the amount of amino acids present. The digestible indispensable amino acid score (DIAAS) was calculated as described in detail in Chapter 2: Materials and Methods, and as described by Muleya et al., 2023.

3.2.7 Antinutritional factors

3.2.7.1 Condensed tannins

In this study, the amount of condensed tannins was estimated as the total of the phenolic compounds measured by HPLC using the equivalent standards for catechin, epicatechin, procyanidin B1, B2 and C1, as described in the Methods Chapter 2.

3.2.7.2 Phytic acid

For the estimation of phytic acid content in the winged bean seeds, the phytic acid assay kit from Megazyme was used. A calibration curve was performed at the same time for the colourimetric determination of phosphorus in the winged bean samples (Megazyme, 2019).

3.2.7.3 Total Phenolics

The total phenolics were measured in the 8 selected winged bean accessions. The method was adapted from Makkar, 2003. The total phenolics were calculated as tannic acid equivalent, using a calibration curve.

3.2.8 Statistical analysis

Statistical analysis was performed on the winged bean accessions using two software Genstat23 and GraphPad Prism10. One- and two-way ANOVA were applied accordingly, followed by Tukey's post hoc test. For the correlations, Pearson's correlation was used.

3.3 Results

3.3.1 Fat and fatty acid content

Since not all accessions were grown at both locations, the two locations were analysed independently. For the winged bean seeds grown in the FF, the fat content varied significantly ($p<0.0001$) among the 19 winged bean accessions, with Ma3 having the highest fat content followed by FP15 with 21.8% and 20.6%, respectively. The lowest amount of fat was in A27 at 14.2% (Figure 3. 4). In the FRC location, the difference in fat content was again significant ($p=0.0024$), between the 11 winged bean accessions, the

highest amount of fat was 18.9% in A30 followed by 18.5% in Ma3; while A4 had the lowest fat content at 12.5% (Figure 3. 5). Following that, the interaction between the genotype and the environment was investigated for the 10 winged bean accessions grown at both locations. A significant genotypic and environmental GxE interaction (p=0.002) on the fat content was detected (Figure 3. 6). As shown in Figure 3. 6, only A4 and Ma3 had a higher fat content in the FF location, while the rest of the winged bean accessions had a higher fat content in the FRC location. Further research is needed to determine the environmental impact on the fat and fatty acid content of winged bean accessions. Based on the fat content results in both locations, the three winged bean accessions that showed the highest and the three winged bean accessions that showed the lowest amount of fat content as well as one accession with an average fat content were selected for fatty acid analysis, as shown in the blue boxes in Figure 3. 6.

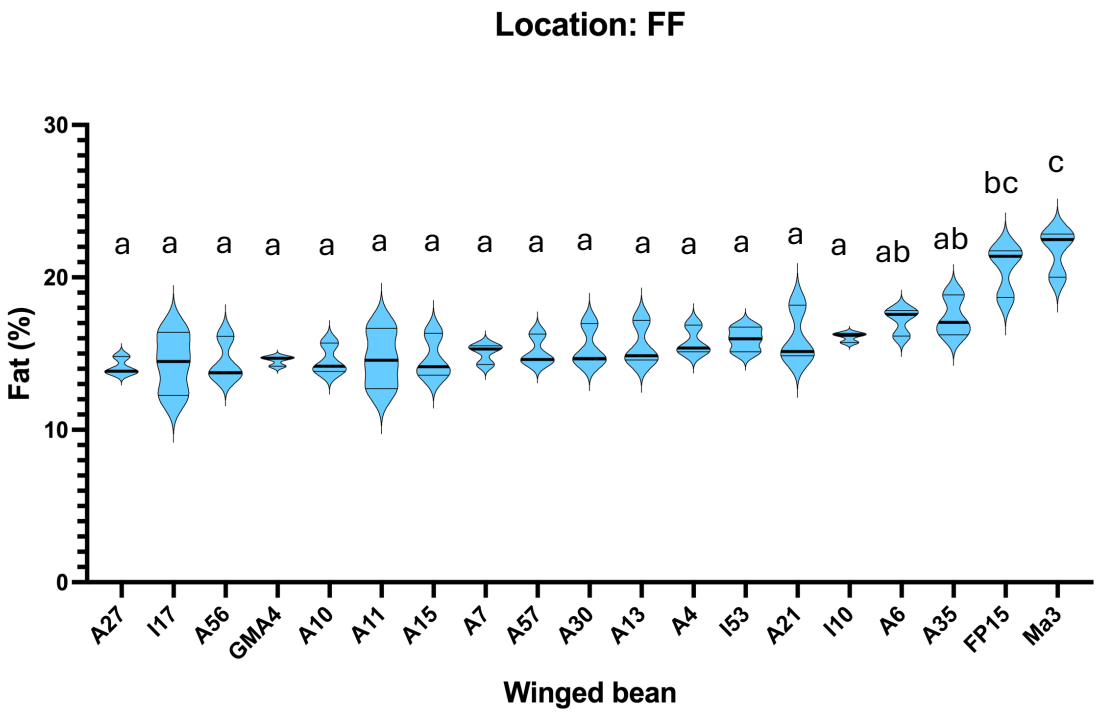


Figure 3. 4 Fat (%) of winged bean accessions (n=3) in FF. The fat content varied significantly p<0.0001. For the statistical analysis, one-way ANOVA and Tukey's posthoc test were used. The different letters represent significant differences among the winged bean accessions.

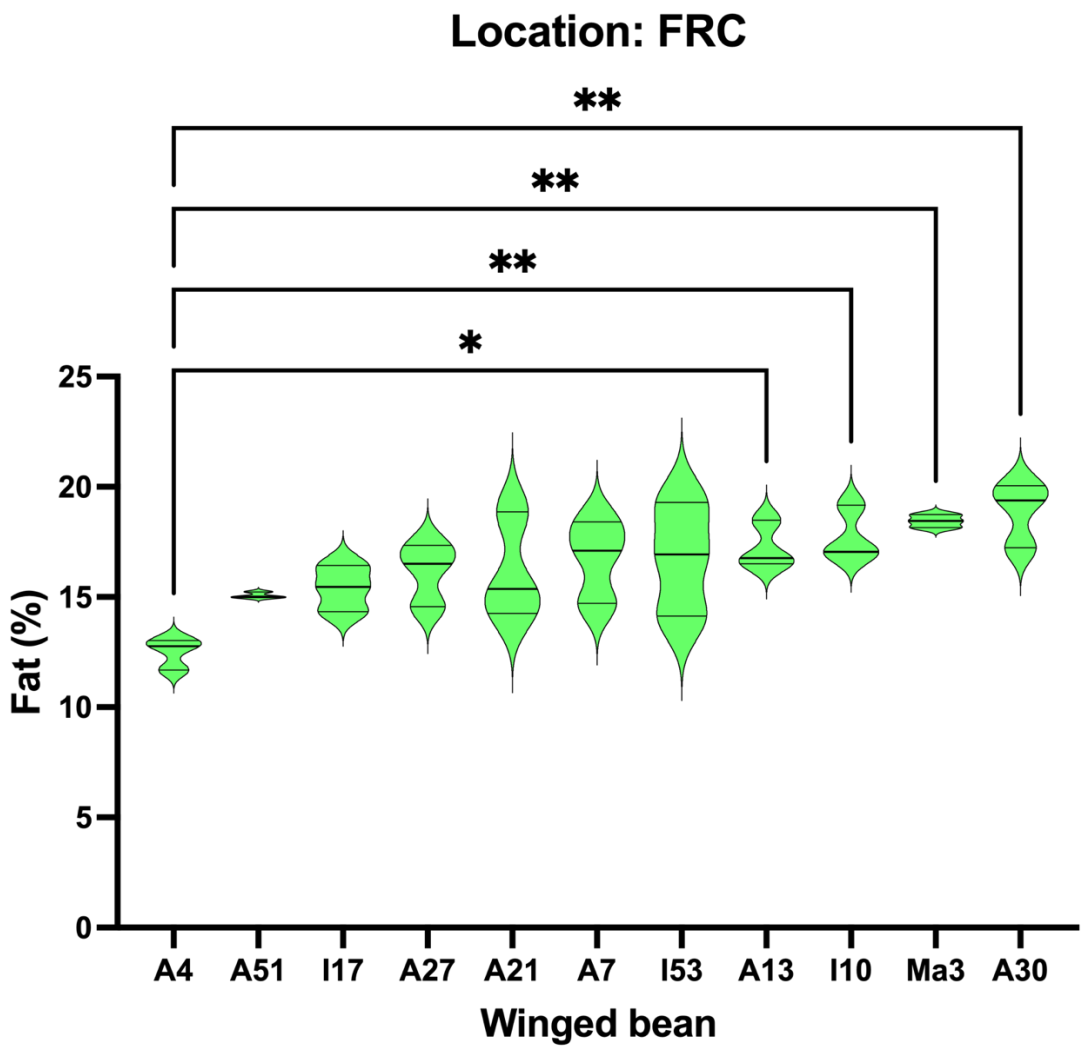


Figure 3. 5 Fat (%) of winged bean accessions (n=3) in FRC. The fat content varied significantly p=0.0024. For the statistical analysis, one-way ANOVA and Tukey's posthoc test were used.

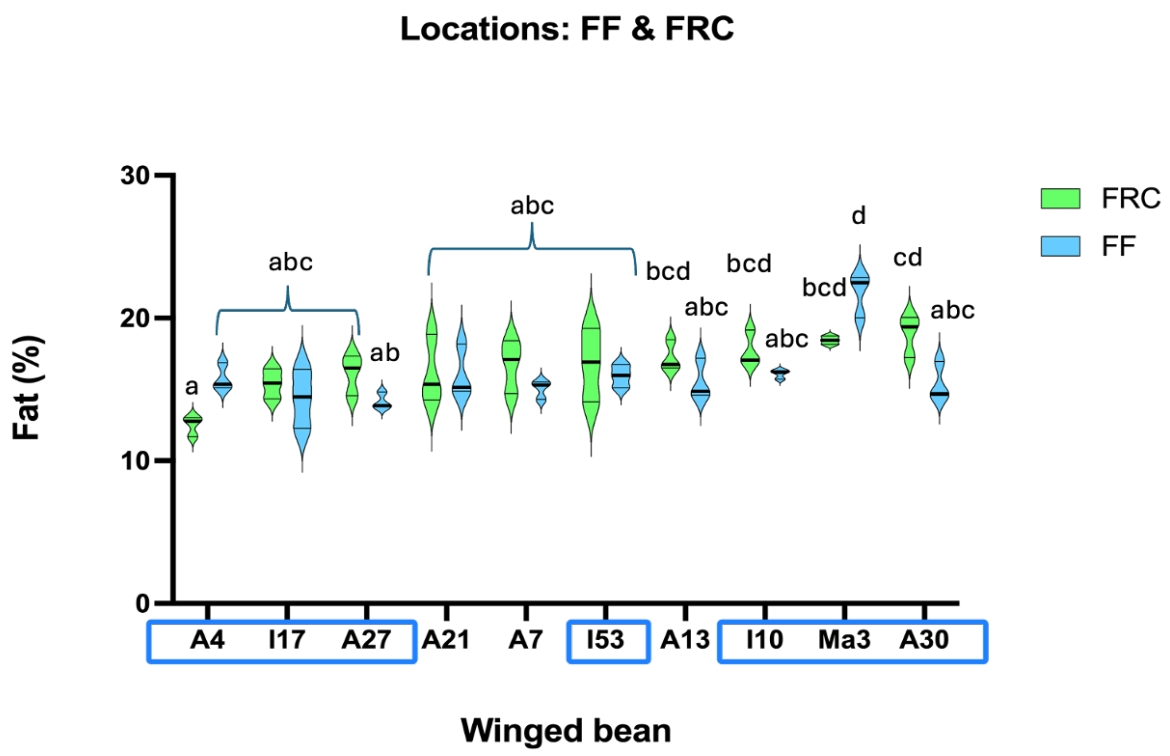


Figure 3. 6 Fat (%) of winged bean seeds (n=3) in FF and FRC. There was a significant interaction (p=0.002) between the Genotype (winged bean accessions) and the Environment (locations FF and FRC). The Genotype had a significant effect (p<0.0001), however, the Environment did not (p=0.1191). For the statistical analysis, two-way ANOVA and Tukey's posthoc test were used. The different letters represent significant differences among the winged bean accessions. The winged bean accessions in the blue boxes were selected for fatty acid analysis.

Fatty acid analysis was performed in the seven winged bean accessions that were grown in FF and FRC locations to identify if the genotype, environment and their GxE interaction had an effect on their content (Figure 3. 6). The relative amount of the fatty acids was quite similar between the two locations where oleic was in abundance at around 35%, followed by linoleic at 23% and behenic at 20%. Overall, the saturated fatty acids (SFA) were 37% of the total fatty acids, with mono-unsaturated fatty acids (MUFA) comprising 38% and poly-unsaturated fatty acids (PUFA) 24% of the total fatty acids (Table 3. 1). For the fatty acids with a relative amount above 5% (palmitic, stearic, oleic, linoleic, behenic), a two-way ANOVA was performed to detect the effect of genotype, location and their interaction (Table 3. 2).

Table 3. 1 Relative amount of fatty acids (%). The relative amounts of fatty acids in the two locations (FF and FRC) are presented as the mean (n=3) ± SEM. The numbers next to the fatty acids correspond to: 1 = Saturated fatty acids, 2 = Monounsaturated fatty acids, and 3 = Polyunsaturated fatty acids.

Fatty Acids		Locations	
		FF	FRC
C12:0	**Lauric**[1]	0.01±0	0.01±0
C14:0	**Myristic**[1]	0.06±0	0.04±0
C16:0	**Palmitic**[1]	8.63±0.17	8.03±0.2
C16:1	**Palmitoleic**[2]	0.21±0.01	0.16±0
C17:0	**Heptadecanoic**[1]	0.06±0	0.05±0
C18:0	**Stearic**[1]	5.35±0.12	5.47±0.14
C18:1n9t	**Elaidic**[1]	0.02±0	0.02±0
C18:1n9c	**Oleic**[2]	34.08±0.55	35.89±0.58
C18:2n6c	**Linoleic**[3]	23.76±0.49	22.01±0.59
C20:0	**Arachidic**[1]	1.8±0.05	1.75±0.04
C20:1	**cis-11-Eicosenoic**[2]	2.32±0.06	2.5±0.06
C18:3n3	**3n3 Linolenic**[3]	1.18±0.04	1.06±0.04
C20:2	**cis-11,14-Eicosadienoic**[3]	0.1±0.01	0.06±0
C22:0	**Behenic**[1]	19.53±0.34	20.07±0.31
C20:3n3	**ETE**[3]	0.36±0.02	0.39±0.02
C22:3n9	**Euricic**[3]	0.03±0	0.04±0
C24:0	**Lignoceric**[1]	2.2±0.07	2.23±0.07
	SFA	37.65	37.64
	MUFA	37.03	38.99
	PUFA	25.04	23.13
	Total	99.71	99.77

As mentioned, the fat content was significantly affected by the interaction of the genotype and the environment ($p<0.001$). However, that was not the case for any of the fatty acids (Table 3. 2). The environment had a significant effect on the relative amounts of palmitic ($p<0.001$), oleic ($p=0.005$), and linoleic acids ($p=0.006$). Specifically, in the winged bean accessions A27, I17, A4 and Ma3, the relative amounts of palmitic and linoleic acids were significantly higher at the FF location, whereas for these accessions, the relative amount of oleic acid was significantly higher at the FRC location. Genotype had a significant effect ($p<0.001$) on the relative amounts of all fatty acids, except behenic acid, which appeared to not be affected by either environment, genotype, or their interaction. However, both genotype and the interaction between genotype and environment significantly influenced the actual amounts of fatty acids ($p<0.05$) (Table 3. 2).

Table 3. 2 Relative amount of fatty acids (%). The relative amounts of fatty acids in the two locations (FF and FRC) are presented as the mean (n=3) ± SEM, analysed using a two-way ANOVA and Tukey's post-hoc test. The different letters represent significant differences among the winged bean accessions. The numbers next to the fatty acids correspond to: 1 = Saturated fatty acids, 2 = Monounsaturated fatty acids, and 3 = Polyunsaturated fatty acids.

Winged bean accessions	Location	Fat (%)	Palmitic[1]	Stearic[1]	Oleic[2]	Linoleic[3]	Behenic[1]
A27	**FF**	14.15±0.33ab	8.83±0.34bcdef	5.29±0.27ab	35.08±1.47abc	23.76±1.07b	18.81±0.25
	FRC	16.14±0.82abc	8.13±0.26bcd	5.94±0.14b	38.26±0.71bc	21.07±0.68ab	18.84±0.72
I17	**FF**	14.38±1.19ab	9.07±0.5def	4.77±0.1ab	33±0.4ab	25.44±1.71b	19.3±0.67
	FRC	15.41±0.61abc	8.1±0.14bcd	5.18±0.16ab	34.35±0.24abc	24.55±0.5b	19.72±0.71
A30	**FF**	15.41±0.78abc	9.65±0.33ef	5.14±0.2ab	32.42±1.07a	24.33±1.4b	19.76±1.01
	FRC	18.89±0.85cd	9.84±0.13f	5.24±0.28ab	34.42±0.54abc	21.79±0.67ab	19.65±0.9
A4	**FF**	15.79±0.55abc	8.86±0.18cdef	4.99±0.28ab	32.59±0.61ab	24.53±0.84b	20.15±0.96
	FRC	12.49±0.41a	7.59±0.04ab	4.35±0.22a	36.12±0.76abc	23.96±1.22b	19.27±0.68
I53	**FF**	15.94±0.47abc	7.89±0.23abcd	5.75±0.48b	36.99±2.5abc	22.81±1.78ab	18.21±1.38
	FRC	16.78±1.49bc	7.71±0.04abc	5.98±0.19b	35.77±0.97abc	22.23±0.42ab	19.9±0.43
I10	**FF**	16.09±0.18abc	8.45±0.13bcde	5.48±0.13ab	32.74±0.59ab	24.56±0.71b	20.49±0.47
	FRC	17.75±0.71bc	8.04±0.18bcd	5.83±0.37b	32.63±0.61ab	23.33±0.24b	21.29±0.75
Ma3	**FF**	21.78±0.89d	7.66±0.06abc	6.05±0.2b	35.7±0.98abc	20.91±0.14ab	19.97±1.34
	FRC	18.45±0.17cd	6.77±0.32a	5.78±0.19b	39.66±1.79c	17.16±1.87a	21.82±0.31
p-value	**G**	<0.001	<0.001	<0.001	<0.001	<0.001	0.106
	E	0.414	<0.001	0.380	0.005	0.006	0.228
	GxE	<0.001	0.085	0.194	0.209	0.695	0.643

3.3.2 Protein content

The amount of protein varied significantly among the 19 winged bean accessions (p=0.0299), in the FF location; with the winged bean accession GMYA4 having the highest protein content at 42.67%, while Ma3 had the lowest at 35.44% (Figure 3. 7). Significant differences in the protein content were also detected at the FRC location, where different winged bean accessions had the highest and lowest protein content. In the FRC location, the highest protein content was in A51 at 38.42% compared to the lowest in A13 at 35.32% (Figure 3. 8).

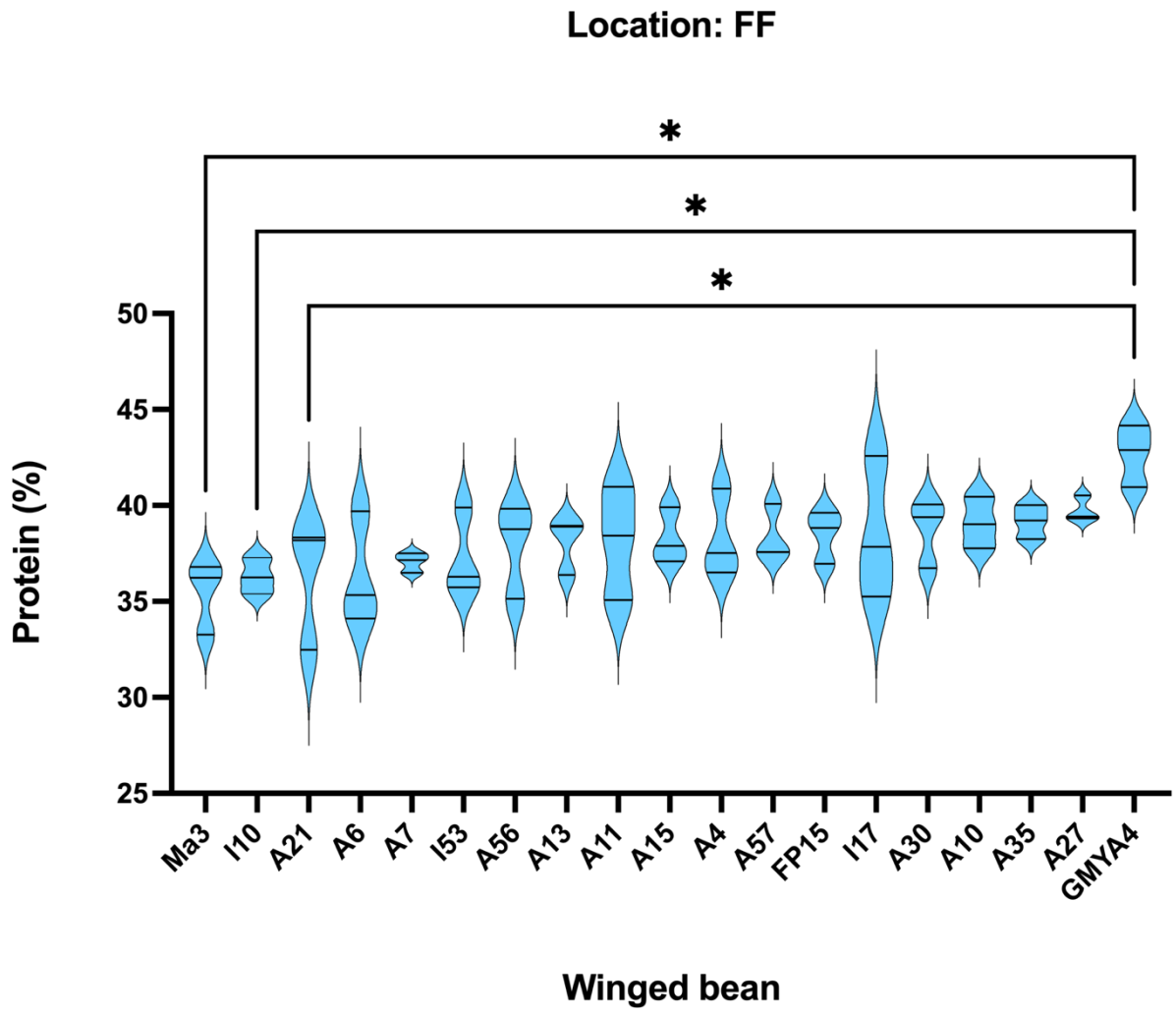


*Figure 3. 7 Protein (%) in FF. The difference in the amount of protein between the winged bean accessions was significant p=0.0299. The data (n=3) was analysed using one-way ANOVA and Tukey's test. ($F_{(7,16)}$=3.069, p=0.0299). Protein (%) was expressed as the percentage of protein in 100 mg of dry seed weight. The significance (p<0.05) is mentioned with the star sign *.*

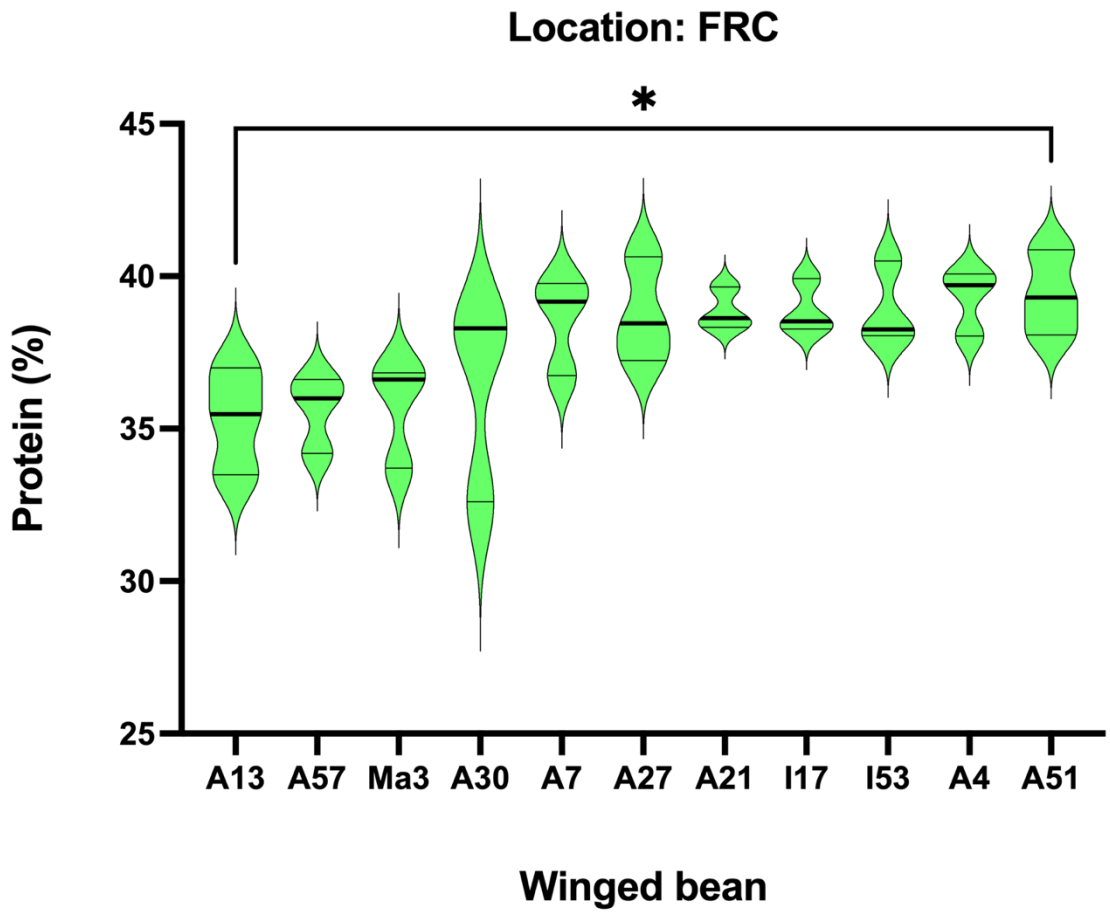


*Figure 3. 8 Protein (%) in FRC. The protein (%) in the winged bean seeds grown in the FRC location varied significantly (p=0.0161). The data (n=3) was analysed using one-way ANOVA, and Tukey's test. Protein (%) was expressed as the percentage of protein in 100 mg of dry seed weight, The significance is mentioned with the star sign * p<0.05.*

The 11 winged bean accessions that were grown in both locations, FF and FRC, and had three biological replicates (n=3) were analysed using a two-way ANOVA, to investigate a GxE interaction. For the winged bean accessions examined, it was shown that the GxE interaction and the environment had no significant effect on the protein content (Figure 3. 9a). Therefore, the data from both locations was re-analysed using considering only the genotypic effect (n=6) (Figure 3. 9b). Interestingly, the winged bean accession Ma3 still showed the lowest protein content at 35.65%. The genotypic effect appeared to play a more decisive role in determining protein content, rather than the environmental conditions.

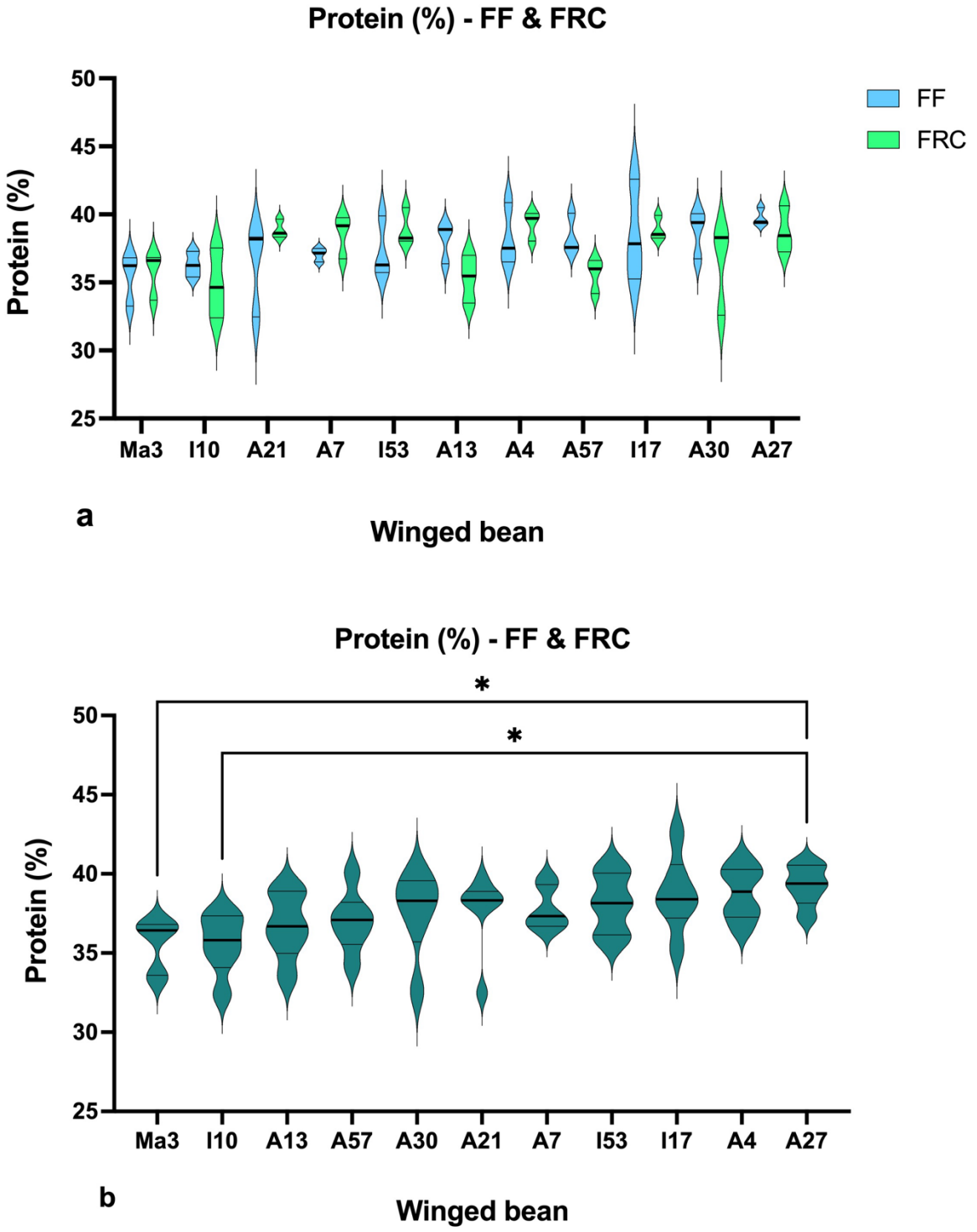


*Figure 3. 9 Protein (%) of winged bean seeds in FF and FRC. a, The interaction between the environment and the genotypes was not significant (p=0.2032). The genotype had a significant effect on protein content (p=0.0226), while the environment had no significant effect (p=0.5642). The data (n=3) was analysed using two-way ANOVA and Tukey's test. b, Protein (%) in both locations FF and FRC. The protein (%) varied significantly among the winged bean accessions (p=0.0497). The data (n=6) was analysed using one-way ANOVA and Tukey's test. The significance is mentioned with the star sign * p<0.05.*

3.3.2.1 Fat – Protein Correlation

In the winged bean seeds, protein and fat were strongly and negatively correlated in both locations (Figure 3. 10). The negative correlation between protein and fat content could be due to the metabolic trade-offs and competition for resources during seed development, as well as genetic factors that contribute to the nutritional composition of the seeds. Understanding the relationship between fat and protein content could be very useful for breeding programmes aiming to improve the nutritional quality of winged bean seeds.

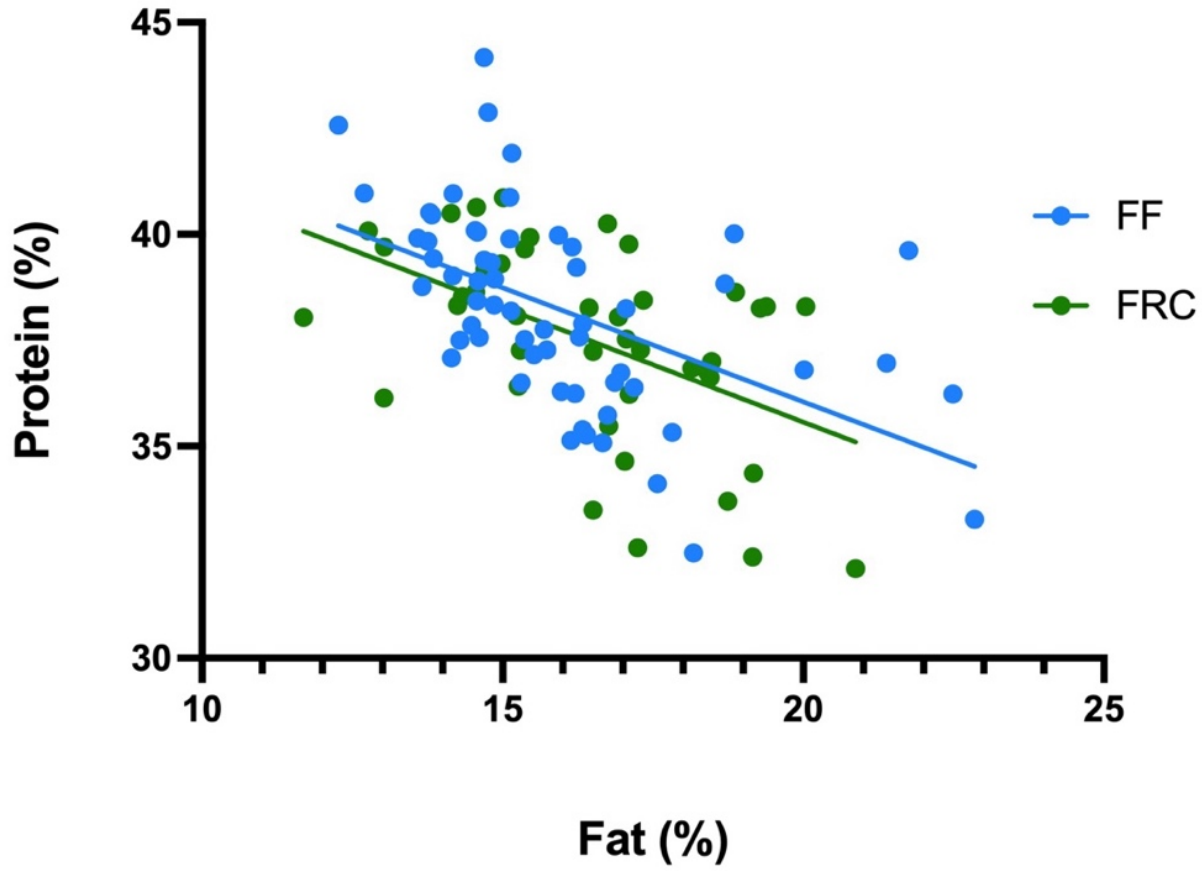


Figure 3. 10 Protein (%) and fat (%) correlation in the two locations FF and FRC. Pearson's correlation was used. There was a significant and negative correlation between protein (%) and fat (%) in both locations. For the FF, the negative correlation was significant at $p<0.0001$ and $R^2 = 0.2704$; and for the FRC the negative correlation was significant at $p=0.0007$ and $R^2 = 0.2545$.

3.3.2.2 Amino acid composition

The quality of a protein source is influenced by its amino acid composition. For the winged bean seeds examined, the amount of amino acids was expressed as mg of amino acid per 100 mg of dry seed weight; glutamate was significantly higher in GMYA4 and I17 (Table 3. 3). When looking at the protein content, the winged bean line GMYA4 had significantly higher protein content compared to the other winged bean lines (Table 3. 3), resulting in an overall higher amount of glutamate.

The amount of amino acids was expressed as mg of amino acid per gram of protein without significant differences observed (Figure 3. 11, Supplementary Table 3. 1). All amino acids, except tryptophan, were measured and expressed as milligrams per gram of protein, which was estimated using the protein analyser. Since tryptophan was not measured, the total amino acid content was expected to be less than 1000 mg per gram of protein. The amino acid content showed no significant variation across the different winged bean accessions (Figure 3. 11, Supplementary Table 3. 2).

Table 3. 3 Amino acids mg/100 mg of dry seed weight. The amount of the amino acids shown in the table ± SEM (n=3), were analysed using a one-way ANOVA and Tukey's post-hoc test, p-values are shown. The different letters represent significant differences among the winged bean lines, and the p-value is highlighted with bold letters.

WB Lines	Ma3	I10	A6	I53	FP15	I17	A30	GMYA4	ANOVA p-value
Protein (%)	**35.44±1.1a**	**36.3±0.54a**	**36.38±1.7a**	**37.3±1.3ab**	**38.47±0.79ab**	**38.56±2.14ab**	**38.73±1.01ab**	**42.67±0.93b**	**0.0299**
Total AAs	34.46±1.14	33.93±1.48	33.29±1.3	34.1±2.12	35.34±1.9	39.78±3.72	38.12±0.99	39.64±0.75	0.1198
Histidine	1.06±0.03	1±0.05	0.98±0.05	0.99±0.06	1.07±0.07	1.2±0.1	1.15±0.03	1.2±0.06	0.0745
Isoleucine	1.6±0.05	1.63±0	1.61±0.08	1.57±0.11	1.67±0.07	1.88±0.16	1.79±0.11	1.85±0.03	0.1414
Leucine	2.92±0.17	2.85±0.09	2.91±0.05	3.01±0.22	2.96±0.19	3.43±0.32	3.29±0.21	3.45±0.06	0.1721
Lysine	2.55±0.06	2.53±0.11	2.52±0.07	2.48±0.13	2.63±0.18	3.01±0.29	2.9±0.1	2.94±0.14	0.1120
Phenylalanine	1.67±0.02	1.66±0.07	1.63±0.08	1.69±0.11	1.69±0.11	1.92±0.18	1.82±0.04	1.89±0.04	0.2493
Threonine	1.38±0.04	1.37±0.08	1.32±0.05	1.36±0.1	1.38±0.08	1.49±0.17	1.5±0.03	1.44±0.04	0.7971
Tyrosine	1.37±0.06	1.28±0.06	1.28±0.05	1.26±0.08	1.35±0.09	1.51±0.16	1.46±0.06	1.59±0.03	0.0970
Valine	1.88±0.1	1.96±0.09	1.82±0.09	1.75±0.07	1.77±0.06	2.21±0.22	2±0.11	2.11±0.14	0.1496
Methionine*	0.45±0.02	0.43±0.01	0.42±0.02	0.42±0.02	0.44±0.02	0.47±0.03	0.44±0.02	0.46±0.01	0.6839
Cysteine*	0.35±0.03	0.35±0.05	0.35±0.03	0.33±0.03	0.28±0.03	0.41±0.05	0.4±0.03	0.44±0.01	0.1099
Alanine	1.41±0.18	1.34±0.05	1.27±0.1	1.36±0.07	1.38±0.03	1.53±0.2	1.62±0.07	1.54±0.09	0.4052
Arginine	2.72±0.08	2.65±0.09	2.58±0.07	2.72±0.13	2.9±0.15	3.08±0.24	2.86±0.11	3.13±0.06	0.0736
Aspartate	3.73±0.06	3.71±0.29	3.61±0.12	3.87±0.46	3.95±0.21	4.34±0.35	4.02±0.13	4.38±0.18	0.3450
Glutamate	**6.08±0.18a**	**6±0.3a**	**5.84±0.24a**	**6.11±0.42a**	**6.57±0.45ab**	**7.39±0.73b**	**7.02±0.13ab**	**7.36±0.1b**	**0.0395**
Glycine	1.69±0.07	1.64±0.1	1.63±0.07	1.63±0.07	1.67±0.06	1.74±0.12	1.83±0.15	1.68±0.04	0.8053
Proline	2.06±0.06	2.05±0.09	1.97±0.09	2.06±0.13	2.1±0.14	2.49±0.31	2.33±0.08	2.48±0.05	0.0947
Serine	1.55±0.06	1.47±0.07	1.55±0.11	1.48±0.11	1.55±0.08	1.69±0.18	1.67±0.05	1.71±0.01	0.4855

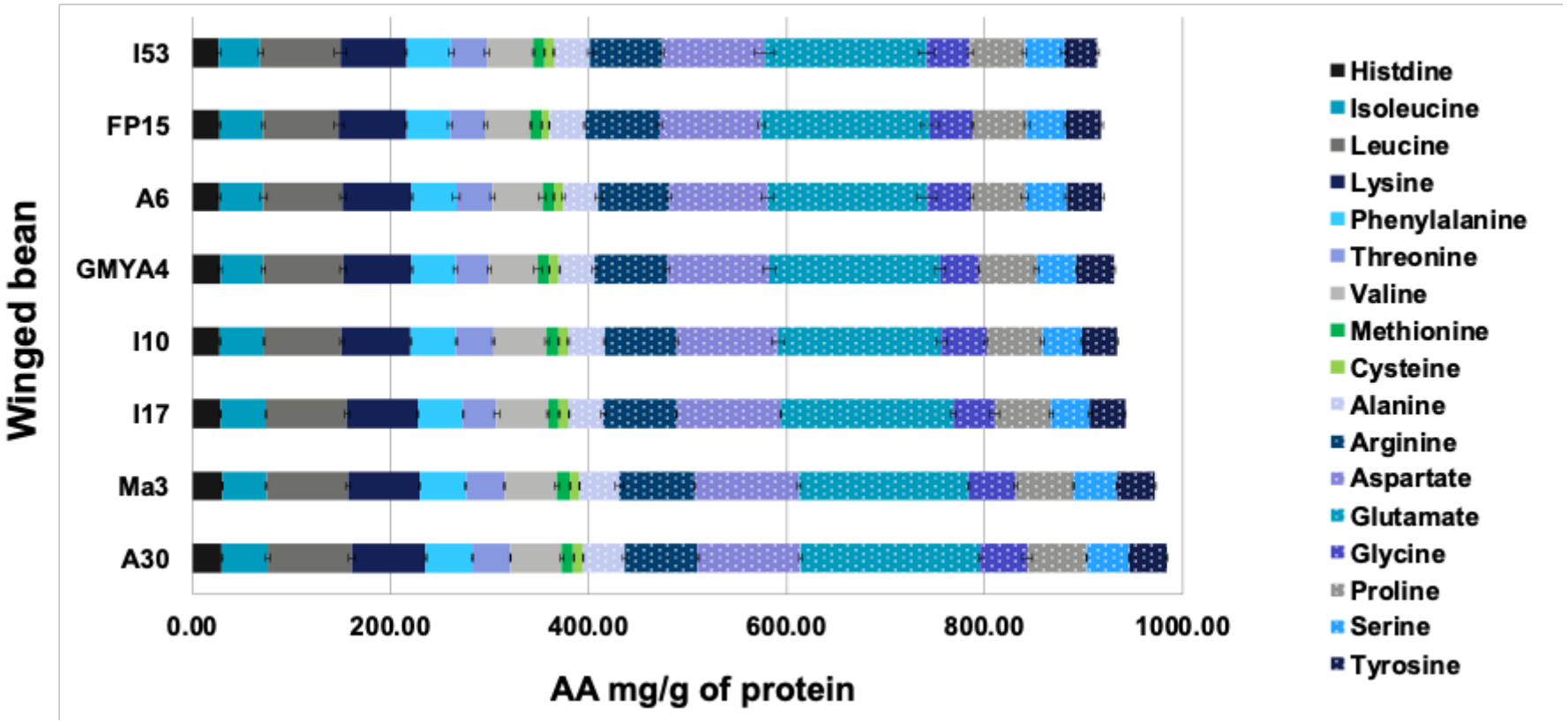


Figure 3. 11 Amino acid mg/g of protein. The amount of amino acids (excluding tryptophan) was expressed as mg per 1 gram of protein, as estimated by the protein analyser. The data were analysed using one-way ANOVA (n=3, except I17 where n = 2).

3.3.3 Antinutritional factors

Even though winged bean seeds were grown in two locations, only from the FF location were enough biological replicates (n=3) and seed material for the analysis of antinutritional factors. Taking that into consideration, the analysis for the antinutritional factors and in vitro digestion was performed on seeds from the location: FF.

3.3.3.1 Phytic acid content

The phytic acid content varied, but not significantly, among the winged bean accessions. GMYA4 had a higher amount of phytic acid 1.77 mg, compared to I53 which had the lowest value at 0.83 mg (Figure 3. 12). Even though GMYA4 had almost double the amount of phytic acid compared to I53, the difference was not significant due to the high variation detected among the biological replicates. Two reasons that could explain this variation are that the analysis relies on the colourimetric determination of phosphorus, which is sensitive to time and temperature. Additionally, the three biological replicates used for each of the winged bean accessions analysed were derived from different plants. Therefore, this likely contributed to further variation in the results.

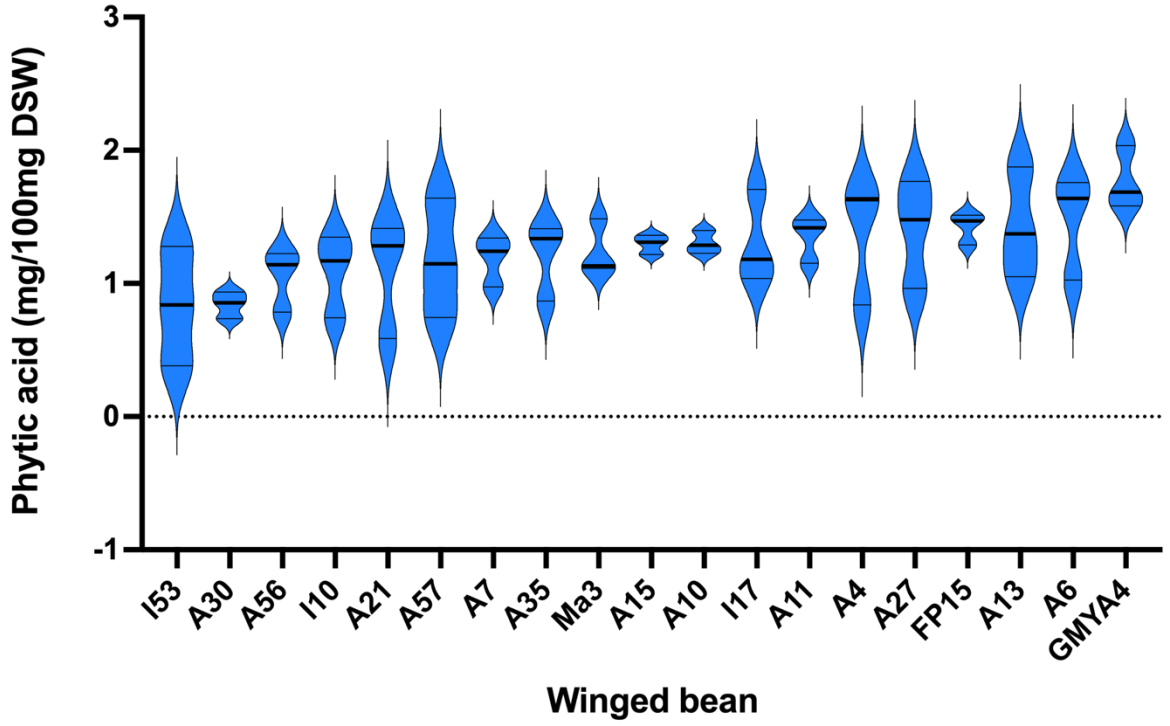


Figure 3. 12 Phytic acid mg/100 mg of dry seed weight. The amount of phytic acid among 19-winged bean accessions (n=3). For the statistical analysis, one-way ANOVA was used. No significant difference was detected (p=0.1493).

3.3.3.2 Total Phenolics

The total phenolics were expressed as mg of tannic acid equivalent (TAE) per 100 mg of dry seed weight. The winged bean accession A30 had the lowest value of TAE at 1.35 mg, while A6 and FP15 had significantly ($p<0.05$) higher TAE at 1.64 mg and 1.66 mg, respectively (Figure 3. 13). The winged bean accessions A6 and FP15 high in TAE had a purple seed coat colour while the low TAE accessions such as A30 and GMYA4 had a cream seed coat colour. The rest of the winged bean accessions had a brown seed coat colour and an intermediate TAE value (Figure 3. 14f).

The correlations between total phenolics and protein content, fat content and phytic acid were not significant (Figure 3. 14c,d,e). While the correlation between phytic acid and protein content was significant ($p<0.036$) (Figure 3. 14a), fat content was not correlated to phytic acid ($p=0.908$) (Figure 3. 14b).

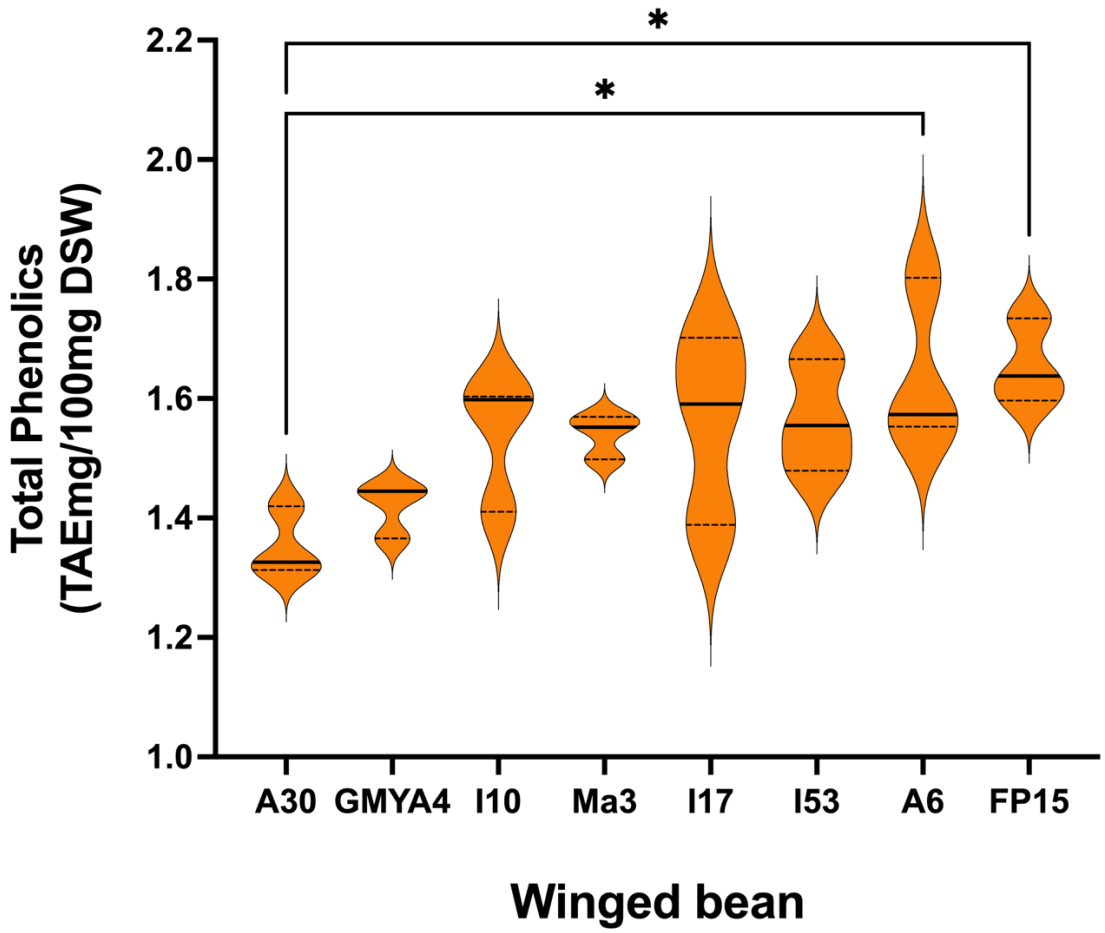


*Figure 3. 13 Total phenolics (mg TAE/100 mg dry seed weight). Between the winged bean accessions, there was a significant difference in the amount of total phenolics (p=0.0222). The total phenolics were estimated as equivalents of tannic acid in mg per g of dried seed. The data (n=3) was analysed using one-way ANOVA and Tukey's test. The significance is mentioned with the star sign * p<0.05.*

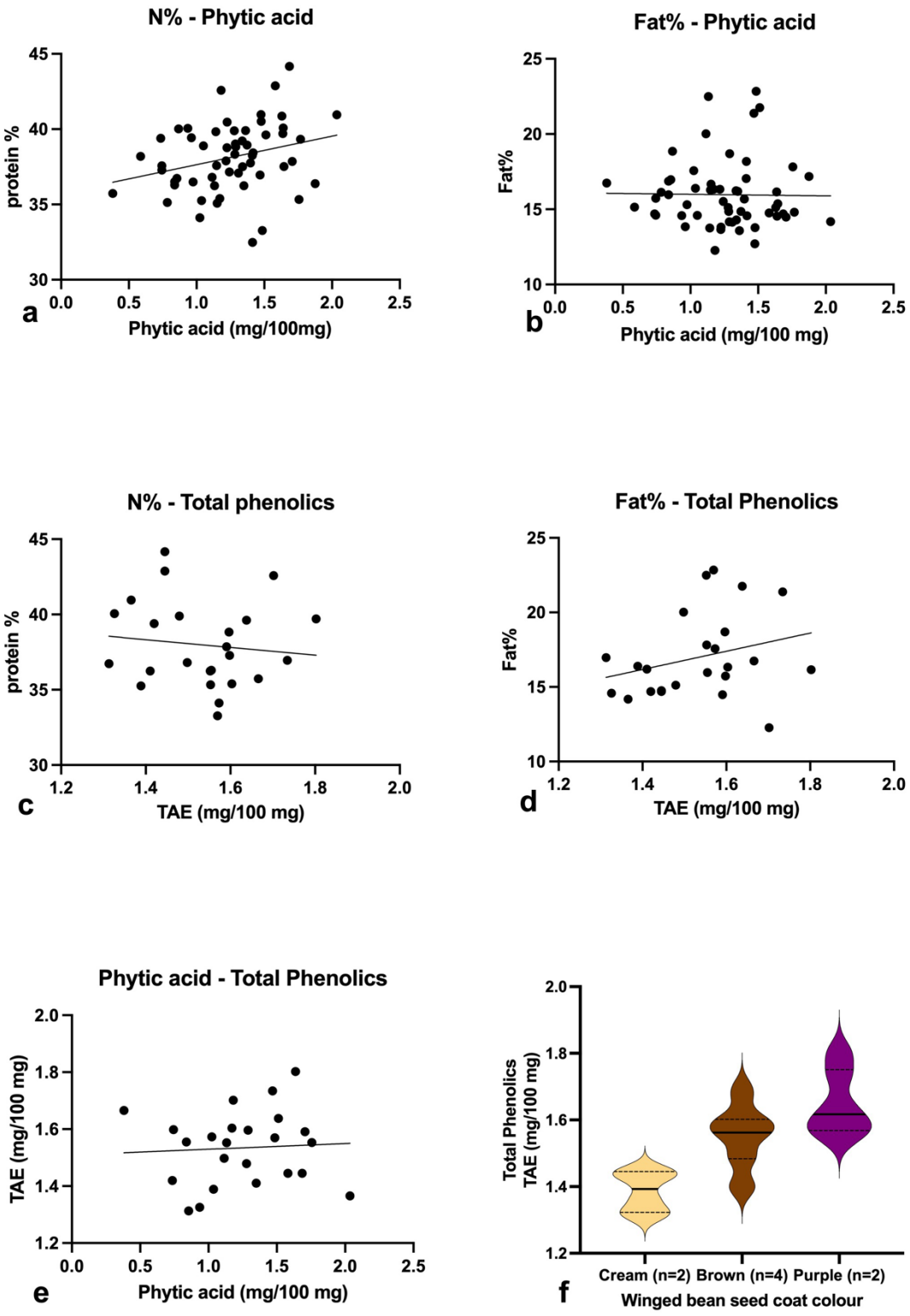


Figure 3. 14 Correlations. Pearson's correlation was used. a, the correlation between protein content and phytic acid p=0.0360 and R^2=0.0775. b, the correlation between fat content and phytic acid p=0.9080 and R^2=0.0002. c, the correlation between protein content and total phenolics p=0.59 and R^2=0.0134. d, the correlation between fat content and total phenolics p=0.1923 and R^2=0.0760. e, the correlation between total phenolics and phytic acid content is p=0.7782 and R^2=0.0037.

3.3.3.3 Condensed Tannins – Phenolic Compounds

Total phenolics, more specifically tannins, are in the same biosynthetic pathways as anthocyanins. Therefore, a correlation between the colour of the seeds and the amount of condensed tannins could be possible (Figure 3. 15f). For the determination of the condensed tannins in winged bean seeds, the LCMS was used. While further optimisation is needed, these results are still a good indication of the condensed tannin content in the winged bean seeds.

The polyphenols measured were significantly different among the winged bean accessions. Overall, the winged bean accessions A30 and GMYA4 have a cream seed coat and a lower amount of catechin, epicatechin, and procyanidins B1, B2 and C1. The winged bean accession A6 with the dark purple seed coat had a high level of polyphenols 41.144 µg/100 mg DW. While the low protein Ma3 winged bean accession had the highest amount of polyphenols at 68.533 µg/100 mg DW (Table 3. 4). The values of the condensed tannins

Table 3. 4 Condensed tannins µg/100mg of dry seed weight. The amount of polyphenols in mature winged bean seeds shown in the table ± SEM (n=3), was analysed using a one-way ANOVA and Tukey's post-hoc test, p-values are shown. The different letters represent significant differences among the winged bean accessions.

Winged bean	Catechin	Epicatechin	Procyanidin B1	Procyanidin B2	Procyanidin C1	Total
GMYA4	0.004±0.003a	0.043±0.021a	0.004±0.002a	0.111±0.006a	0.101±0.01a	0.263±0.023a
A30	0.004±0.002a	0.123±0.082a	0.006±0.002a	0.111±0.005a	0.098±0.021a	0.341±0.089a
I17	0.412±0.117ab	9.906±3.868ab	0.144±0.007ab	0.994±0.127ab	0.854±0.059a	12.31±4.157a
I53	0.478±0.119ab	15.37±4.737ab	0.196±0.073ab	1.286±0.376ab	1.068±0.224a	18.398±5.517ab
FP15	0.68±0.091ab	16.65±1.603ab	0.112±0.025ab	1.149±0.12ab	0.957±0.091a	19.548±1.565ab
I10	1.298±0.374ab	33.087±8.777ab	0.422±0.105b	3.102±0.946ab	2.1±0.509ab	40.009±10.69ab
A6	1.29±0.077ab	35.592±5.372ab	0.318±0.032ab	2.393±0.186ab	1.552±0.065ab	41.144±5.511ab
Ma3	1.655±0.684b	58.202±25.163b	0.502±0.185b	4.444±1.758b	3.73±1.353b	68.533±29.136b
P-value	0.005	0.009	0.003	0.007	0.003	0.008

3.3.4 In Vitro Digestibility

To select winged bean accessions for in vitro digestion analysis, their nutritional profiles were carefully considered. The accessions were ranked based on protein content, from highest to lowest, and eight accessions were chosen to represent a broad nutritional range. This selection included accessions with varying levels of protein, fat, phytic acid, condensed tannins, and total phenolics, ensuring a diverse set for comparative analysis. Seed coat colour was also taken into account, with both purple and cream-coloured accessions included alongside the more common brown types. This approach was intended to capture the nutritional and biochemical diversity within the winged bean accessions, allowing for a more comprehensive assessment of how these factors may influence in vitro digestibility. In addition, accessions that were included in other experiments as well as ongoing experiments that are not mentioned in this thesis were prioritised (Table 3. 5)

Table 3. 5 Nutritional profile of winged bean accessions. The average values are shown. The winged bean accessions have been ranked from high to low protein (%). The accessions highlighted have been selected for the in vitro digestion.

Winged Bean accessions	Protein (%)	Fat (%)	Phytic acid (mg/100 mg)	Total Phenolics (TAE mg/100 mg)	Condensed Tannins (µg/100mg DW)	Seed Coat Colour
GMYA4	42.67	14.54	1.77	1.42	0.263±0.023	Cream
A27	39.76	14.15	1.40	-		Cream
A35	39.16	17.38	1.21	-		Brown
A10	39.08	14.56	1.30	-		Brown
A30	38.73	15.41	0.84	1.35	0.341±0.089	Cream
I17	38.56	14.38	1.31	1.56	12.31±4.157	Brown
FP15	38.47	20.61	1.42	1.66	19.548±1.565	Purple
A57	38.41	15.14	1.18	-		Brown
A4	38.3	15.79	1.37	-		Brown
A15	38.29	14.69	1.30	-		Brown
A11	38.16	14.65	1.35	-		Light Brown
A13	38.08	15.55	1.43	-		Dark Brown
A56	37.91	14.52	1.05	-		Dark Brown
I53	37.3	15.94	0.83	1.57	18.398±5.517	Brown
A7	37.05	15.04	1.19	-		Brown
A6	36.38	17.18	1.47	1.64	41.144±5.511	Purple Black
A21	36.33	16.06	1.09	-		Brown
I10	36.3	16.09	1.09	1.54	40.009±10.69	Brown
Ma3	35.44	21.78	1.24	1.54	68.533±29.136	Brown

The 8 winged bean accessions selected for the in vitro digestion had a significant difference in their protein content (Figure 3. 16c). The amount of essential amino acids (EAA) mg per gram of total amino acid (TAA) was not significantly different among the accessions (Figure 3. 16d), neither was the total amino acid (TAA) digestibility (%). The winged bean accession FP15 had the lowest TAA digestibility (%) at 40.75%, while I17 had the highest value at 58.15% (Figure 3. 16a). There is a great variation between the TAA digestibility (%) between the two accessions, that should be further investigated. The limiting amino acid for the winged bean is the sulphur-containing amino acid methionine, followed by cysteine, as in most pulses. The digestible indispensable amino acid score (DIAAS) ranged among the winged bean accessions from 0.14 to 0.21 in the accessions FP15 and I53, respectively (Figure 3. 16b) with the sulphur containing amino acid methionine being the limiting one (Supplementary Table 3. 6). Interestingly, the purple seed winged bean accession FP15 has a relatively high protein content between 38-39%, with low digestibility.

No significant correlation was found between phytic acid content and total amino acid digestibility, nor between total phenolics and total amino acid digestibility. This could be explained due to the increased variability among the biological replicates for both total amino acid digestibility and phytic acid content (Figure 3. 17).

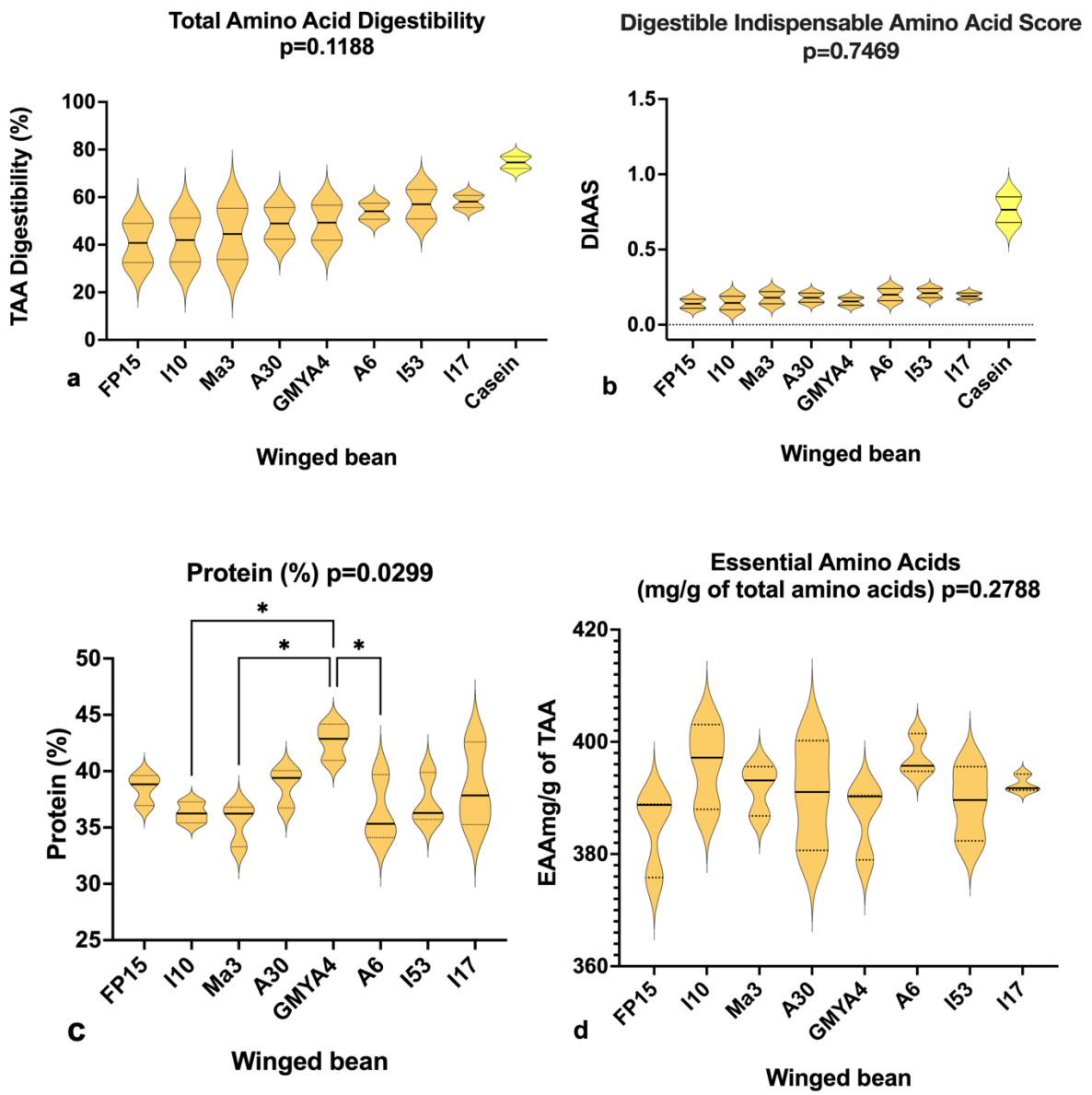


*Figure 3. 16 In Vitro Digestibility. a, total amino acid digestibility of the 8 winged bean accessions did not show a significant variation p=0.1188. For the statistical analysis one-way ANOVA was used and only 2 out of the 3 biological replicates were included (n=2). b, DIAAS, digestible indispensable amino acid score of the 8 winged bean accessions did not show a significant variation p=0.7469. For the statistical analysis one-way ANOVA was used and only 2 out of the 3 biological replicates were included (n=2). c, the protein content of the 8 winged bean accessions used for the In vitro digestion system. The protein (%) varied significantly among the winged bean accessions used p=0.0299. The data (n=3) were analysed using one-way ANOVA and Tukey's test. The significance is mentioned with the star sign * p<0.05, d, the essential amino acids expressed in mg per 1 gram of total amino acid measured in the winged bean seeds did not significantly vary p=0.2788. The data (n=3) were analysed using one-way ANOVA.*

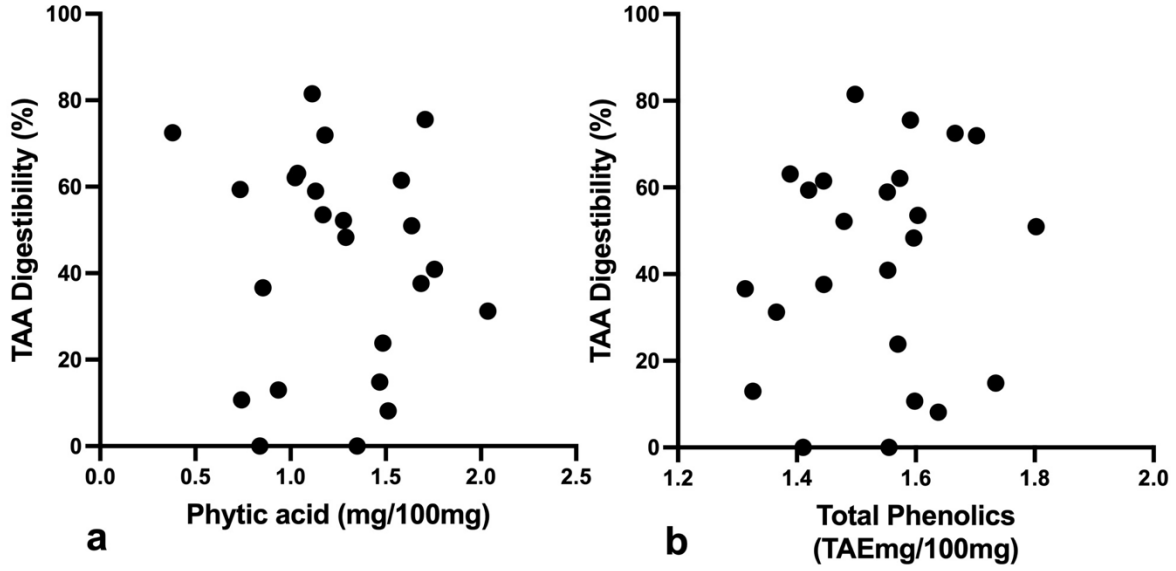


Figure 3. 17 Total amino acid digestibility. Pearson's correlation was used. a, the correlation between TAA digestibility and phytic acid p=0.701 and R^2=0.007. b, the correlation between TAA digestibility and total phenolics p=0.571 and R^2=0.015.

3.4 Discussion

This study is the first to investigate the nutritional composition of these winged bean accessions, along with the impact of the environment, genotype, and their interaction. Additionally, it is the first time that antinutritional factors such as phytic acid and total phenolics have been measured in these accessions, as well as their protein digestibility using the INFOGEST model.

The results suggest that the genotype significantly influences the nutritional value of winged bean seeds. The genotype had a significant effect on the protein, fat and fatty acid content, while the GxE interaction had a significant effect on the fat content only. The antinutritional factors phytate and total phenolics varied between the winged bean accessions, with the total phenolics having a significant variation. In these accessions, the digestibility varied between 40-58% however that was not statistically significant; and the DIAAS ranged from 0.14 to 0.21 with methionine being the limiting amino acid.

3.4.1 Nutritional composition and the effect of G, E and GxE interaction

3.4.1.1 Fat content and Fatty acids

The winged bean accessions had a significantly different amount of fat in their seeds, which ranged from 14.15% to 21.78% in FF (Figure 3. 6) and 12.49% to 18.45% in FRC (Figure 3. 5). A similar amount of fat content in winged bean seeds has been reported (Adegboyega et al., 2019; Amoo et al., 2006; Garcia & Palmer, 1980). Even though the environment did not appear to significantly affect the fat content, there was a significant interaction between the genotype and the environment (GxE) (Figure 3. 6). The GxE has also been reported to affect the fat content in other crops such as soybean and winter rapeseed (Hou et al., 2006; Shafii et al., 1992). It would be useful to determine the nutritional composition of several winged bean accessions in different environments. Environmental factors such as soil quality, climate, and farming practices can influence the levels of protein, oil, and fatty acids in the seeds. Understanding these variations could help optimise winged bean seed production for specific nutritional outcomes, ensure consistency in quality, and support breeding programmes. This could potentially support the nutritional demands of different regions, for animal feed or human consumption.

Fat content is an important trait nevertheless, it is the fatty acid composition that determines the quality of the oil extract. In this study, the fatty acids measured accounted for more than 99% of the total fatty acids in winged bean oil extract. The winged bean seeds contained saturated fatty acids at around 37% and unsaturated at around 62%, with roughly 38% being monounsaturated and 24% polyunsaturated fatty acids (Table 3. 1). More specifically, the fatty acid in abundance was oleic (~35%) followed by linoleic (~23%) and behenic (~20%); with palmitic (~8%) and stearic (~5%) having lower concentrations (Table 3. 2). These five fatty acids comprised more than 90% of the total fatty acids. These results are consistent with what was reported in the literature (Ekpenyong & Borchers, 1980; Garcia & Palmer, 1980; Khor et al., 1982). Recently, studies on winged bean oil extracted using hexane, showed the composition as well as the physiochemical properties of its fatty acids meet the edible oil characteristics (Lepcha et al., 2017; Mohanty et al., 2015, 2021). Interestingly, the fatty acid composition

in winged bean seeds seems to change as they mature, with the percentage of palmitic acid decreasing from 35% to 7% and the percentage of unsaturated fatty acids rising (Mohanty et al., 2015).

Winged bean seeds are rich in unsaturated fatty acids, particularly oleic and linoleic acids (Table 3. 1). Oleic acid, a mono-unsaturated fatty acid, is linked to reduced coronary heart disease risk and provides improved thermal and oxidative stability during oil refining and storage (Jesú Carrero et al., 2007). Linoleic acid, a polyunsaturated essential fatty acid, plays an important role in plasma membrane structure and the production of metabolic regulatory compounds (Omode et al., 1995). Replacing saturated fatty acids with polyunsaturated fats like linoleic acid has been shown to reduce cardiovascular disease risk by nearly 30%, comparable to cholesterol-lowering drugs such as statins (Mensink & Katan, 1989; Sacks et al., 2017; Whelan, 2008).

In this study, high levels of behenic acid (20%) were detected (Table 3. 2). Behenic acid is a long-chain saturated fatty acid which is poorly absorbed and has been related to increased cholesterol in humans (Cater & Denke, 2001). In winged bean seeds, the amount of linoleic and behenic acid seemed to vary in the literature, with some reporting linoleic at ~25% and behenic at ~20% (Ekpenyong & Borchers, 1980) while others found that linoleic was ~35% and behenic acid <10% (Mohanty et al., 2015, 2021). The differences in the relative amounts of linoleic and behenic fatty acids may be attributed to both genetic variations and environmental factors. In soybean seeds, it has been shown that higher temperatures could affect fat content, with the average air temperature during the reproductive and filling stages playing an important role in the concentration of fatty acids (Carrera et al., 2011). Therefore, more research is required to determine the effect of the genotype, environment and their interaction on the fatty acid content of winged bean seeds.

3.4.1.2 Negative correlation of protein and fat

In this study, winged bean seeds were grown in two different locations at two different times. In both experiments, a significant negative correlation between protein and fat

content was reported $p<0.0001$ for FF and $p=0.0007$ for FRC location (Figure 3. 10). A negative correlation between protein and oil content has been repeatedly shown in soybean seeds and summer rape (Bolon et al., 2010; GRAMI et al., 1977; Hwang et al., 2014). Apart from the genetic and environmental factors that have a significant impact on the oil and protein content in winged bean seeds, the developmental stages play an equally important role. Another parameter affecting protein and oil content is related to energy requirements. The average calorific values for oil and protein are around 9.4 and 4.6 Kcal/g, with oil providing almost double the amount of energy compared to protein. Both the synthesis of oil and protein require energy and carbons, in the form of carbohydrates. Therefore, the hypothesis proposing competition between oil and protein synthesis for the same sources of energy and carbohydrates was formulated (Hanson et al., 1961).

The accumulation of nutrient compounds in seeds during their filling and development is influenced by the supply of photoassimilates, the strength of the sink, as well as the partitioning and remobilisation of carbon (Aguirre et al. 2018). Ultimately, embryo development relies on the continuous supply of photoassimilates from the maternal tissues (Aguirre et al. 2018). The study of Allen and Young (2013) proposed that the flux through pyruvate to acetyl-CoA may serve as a key regulatory point that could potentially contribute to the negative correlation of oil and protein content observed in soybean. The proportion of pyruvate directed toward making acetyl-CoA for fatty acid biosynthesis (for oil), versus being used for amino acid biosynthesis (for protein) could influence the protein levels.

Plants first synthesise sugars for their metabolism and growth. The quality and quantity in which sugars are stored in the plants depends on the plant's genotype and environment. Therefore, other yield traits such as seed size, weight, seed number per pod, and number of seeds per plant should be taken into consideration when investigating the negative correlation between protein and oil content (Assefa et al., 2018; Hanson et al., 1961).

3.4.1.3 Protein content and Amino acids

Protein content varied significantly among the winged bean genotypes from 35% to 38% in the FRC and 35% to 42% in FF (Figure 3. 9). Similar results have been reported with protein content varying in winged bean seeds between 34% and 40% (Adegboyega et al., 2019; Kantha & Erdman, 1984). It has been reported that nitrogen supplementation increases the protein content in soybean seeds (Rotundo & Westgate, 2009). At the FRC location there was an application of chemical fertilisers and at the FF location organic farming practices were followed, and poultry manure was applied. The protein content was not significantly affected by the environment nor the GxE interaction (Figure 3. 9), suggesting that the genotype could be having a greater influence on the protein accumulation in the seeds. However, other parameters such as soil fertility and cultivation practices should be taken into consideration.

According to a meta-analysis of environmental effects on soybean seed composition, higher temperatures had a negative effect on oil accumulation in soybean seeds, while protein was not significantly affected. Similar responses have been reported in other crops such as rapeseed and sunflower (Ne Triboi & Triboi-Blondel, 2002). A possible explanation proposed was that the higher temperatures would reduce the seed-filling period and, therefore, increase the rate at which the components accumulate in the seed. In soybean, higher temperatures may increase the rate of protein accumulation in the seeds. Even if the time of seed filling is reduced and the total nitrogen uptake or fixation is shorter, nitrogen remobilisation from the leaves may occur and contribute to the accumulation of protein in the seeds. In contrast, oil accumulation in the seeds is negatively affected by higher temperatures. During the shorter seed-filling periods, oil accumulation would rely on the current photoassimilation production and consequently be reduced. There is no compensation for the oil accumulation as there is for the protein, where nitrogen is remobilised from the leaves. Therefore, temperature has been shown to have a negative impact on oil concentration (Ne Triboi & Triboi-Blondel, 2002; Rotundo & Westgate, 2009). In this work, the environment and or the GxE interaction. had a significant effect on fat content and or the fatty acids (Table 3. 2).

Both protein quantity and quality are important traits. Protein quality is related to the capacity of a protein source to meet the requirements of essential amino acids of an organism, for its physiological needs. To estimate the protein quality of a source, the amount of amino acids would be compared to a reference protein. A widely used reference protein is casein as its amino acid composition meets the minimal requirements for humans. Therefore, identifying essential amino acids in the protein source that are in a quantity lower than desired (limiting amino acids) is key (FAO/WHO/UNU expert consultation, 2007; Kurpad, 2013)

In legumes, the limiting amino acid seems to be the sulphur-containing methionine, while they tend to be rich in lysine. Similar is the case for winged bean seeds, with methionine and cysteine being the limiting amino acids (Supplementary Table 3. 6). Based on winged bean seeds' amino acid profile, it compares well with soybean and previous research has suggested that winged bean seeds could be a good alternative to soybean seeds (Ekpenyong & Borchers, 1982; King & Puwastien, 1987; Okezie & Martin, 1980). Looking at the proportion of amino acids as mg per gram of protein, overall the proportion was similar in the different winged bean accessions (Figure 3. 11). This could be explained by the similar genetic background and the biological functions, especially when considering seed germination. The amount of amino acids is mainly determined by the seed storage protein. For example, the low amount of methionine in legumes and lysine in cereals could be explained by the different amounts of these amino acids accumulated in the seed storage proteins (Shewry et al., 1995). In winged bean seeds the low amount of methionine and high content of lysine complement well the higher methionine content and low levels of lysine in cereals for example maize (Muleya et al. 2023). Combining winged bean seeds with cereals would improve the protein quality in diets and animal feeds. As a next step, it would be valuable to determine the amino acid composition and protein digestibility of animal feed where winged bean seeds have replaced soybean seeds. Additionally, other important factors, such as the levels of antinutritional compounds in winged bean seeds and the effects of processing methods on digestibility and nutrient bioavailability, should be assessed to gain a better understanding of winged bean's potential to substitute soybean.

3.4.2 Antinutritional Factors

3.4.2.1 Total Phenolics - Condensed Tannins

In this study, total phenolics were expressed as mg of tannic acid equivalent (TAE). The purple winged bean seeds of the A6 and FP15 (1.64 mg and 1.66 mg TAE/100 mg) accessions had significantly higher amounts of total phenolics (p=0.022) compared to the cream seed coat accession A30 (1.35 mg TAE/100 mg) (Figure 3. 13). Similar results have been reported by Adegboyega et al. (2019) who analysed 25 winged bean accessions with tannin content ranging from 1.36% – 3.43% of unprocessed winged bean seeds. The amount of condensed tannins was estimated as the total phenolic compounds by HPLC using the equivalent standards. The winged bean accession Ma3 had a significantly higher amount of phenolic compounds, while the accessions A30 and GMYA4 had the lowest (Table 3. 4).

Even though there were significant differences in the amount of protein and total phenolics among the winged bean accessions, these differences were not reflected in the total amino acid digestibility or the DIAAS (Figure 3. 16). Ma3 accession with brown seed coat colour had a higher content of condensed tannins compared to FP15 and A6 which have a purple and dark purple seed coat colour, respectively. Similar results have been reported by Kotaru et al. (1987), who analysed 12 winged bean varieties with tannin content ranging 1.35 - 6.75 mg D-catechin equivalents/g of bean. In their study, a black seed coat variety showed lower tannin content compared to a variety with the pale brown seed coat, observing no correlation between seed coat colour and tannin content.

In seeds with pigmented seed coats, the differential allocation of carbon through the phenylpropanoid pathway may result in brown seeds accumulating higher levels of condensed tannins (proanthocyanidins), while black seeds may accumulate anthocyanins, which could mask or replace tannin-derived pigmentation. This could suggest a trade-off in the biosynthesis of phenolic compounds, where the pathway diverges after shared flavonoid intermediates, such as leucocyanidin, favouring either tannin or anthocyanin production depending on genotype and the expression of pigmentation-related regulatory genes. Further research is needed to explore potential correlations between seed coat colour and tannin

content in winged beans. For the winged bean accessions studied in this thesis, further investigation is needed to determine the amount of condensed tannins and their effect on protein digestibility. Additionally, the impact of processing methods on total phenolic content, condensed tannins, and overall protein digestibility should be further investigated.

3.4.2.2 Phytic acid or Phytate

In terms of phytic acid, there were no significant differences between the winged bean accessions (Figure 3. 12). The winged bean accessions A6 (1.47 mg/100 mg) and FP15 (1.42 mg/100 mg) had a higher amount of phytic acid compared to A30 (0.84 mg/100 mg). Similar results have been reported by Kantha & Erdman, 1986. Kotaru et al., (1987) showed that the phytic acid content of 12 winged bean seeds, from Papua New Guinea, Indonesia and Japan, ranged from 7-12.03 mg/g of bean (0.7-1.2 mg/100 mg).

As phytate is heat stable, none of the heat treatments of seeds such as autoclaving, microwave, infrared, hot air oven, and cooking in boiling water affected its content (Kadam et al., 1987). In this study, the winged bean seeds were autoclaved before the total amino acid digestibility was determined. However, no significant correlation between the phytic acid content and the in vitro digestibility (Figure 3. 17). One explanation could be the great statistical error on phytic acid content due to the sensitivity of the method and the small number of replicates (n=3).

Even though in this study the phytic acid content did not seem to have a significant difference among the winged bean accessions nor an impact on the digestibility, it is worth mentioning methods that could reduce its amount and effect on nutrient bioavailability, as phytic acid can low the bioavailability of zinc in soybean seeds (Erdman et al., 1980; Zhou et al., 1992). Pre-treatment methods of seeds have been used to reduce the phytate content, such as fermentation, soaking, germination and enzymatic treatment of grains with phytase enzyme (Gupta et al., 2015). Gene editing methods that produce cereal seeds with altered chemistry of seed phosphorous and low levels of phytic acid have been explored (Raboy et al., 2001). This could provide information on genetic mechanisms for improved varieties with lower levels of phytic acid. Another way

to improve phosphorous bioavailability is the addition of microbial phytase to the diets of monogastric livestock. The addition of microbial phytase increased the availability of phosphorus in the low-phosphorous diet of broilers and decreased the amount of phosphorus in their faeces (Simons et al., 1990).

3.4.3 In Vitro Digestion

In this study, the protein quality was measured using the in vitro digestible indispensable amino acid score (DIAAS). Among the winged bean accessions, the DIAAS ranged from 0.14 for the FP15 to 0.21 for the I53 accession (Figure 3. 16). Methionine was found to be the limiting amino acid, as expected for legumes and in line with the literature (Černý et al., 1971; Okezie & Martin, 1980; Wyckoff S. & Vohra P., 1982). In contrast, the DIAAS for casein was 0.77. The limiting amino acid was the aromatic amino acid phenylalanine (Figure 3. 16). Even though the total amino acid digestibility ranged from 40-60%, it was not statistically significant due to the small number of replicates (n=2) and the quite substantial statistical error for the winged bean accessions with a digestibility below 50%. In addition, it seemed that the error was smaller when the digestibility was above 50%, suggesting a possible limitation of the INOGEST system and/or of the LCMS to detect the small amount of amino acids in the supernatants that could be at the bottom end of the standard curve.

The winged bean seeds were autoclaved before the in vitro digestion. Overall, the DIAAS and TAA digestibility (%) were lower when compared to the untreated-raw samples (Supplementary Figure 3. 3). It has been reported that the antinutritional factors such as trypsin inhibitors are heat liable and their activity is decreased with heat treatments (De Lumen & Salamat, 1980; Esaka et al., 1987). There have been reports where the protein quality was improved by heat treatment (Kadam & Smithard, 1987; Saadi et al., 2022). In several processing methods such as boiling and autoclaving, there is also the chance of protein aggregation and denaturation which can lead to lower protein digestibility. The protein quality of faba bean was significantly lower after boiling (Martineau-Côté et al., 2023) and the in vitro protein digestibility of winged bean seeds after 10 minutes of autoclaving was lower than autoclaving for 5 minutes (Tan et al., 1984). Therefore, this

could be one explanation for the decreased total amino acid digestibility and lower DIAAS of the autoclaved winged bean seeds. Another reason could be a technical error, as the raw and autoclaved seeds were not handled by the same person.

Even though there were significant differences in the amount of protein and the amount of total phenolics among the winged bean accessions, these differences were not reflected in the total amino acid digestibility or the DIAAS (Figure 3. 16). Interestingly, the TAA digestibility (%) did not relate to the amount of phytic acid or phenolic content in the winged bean accessions analysed (Figure 3. 17). This could be explained by the increased variation among the limited number of biological replicates, the food matrix that could affect the digestibility measurements, the autoclaving that could possibly have deactivated protein inhibitors but also aggregated seed storage proteins, lowering their digestibility.

Processing methods are used to improve digestibility and palatability of pulses. For example, the dehulling process could improve the palatability, taste and digestibility as well as reduce the tannin content and cooking time as the water uptake improves (Bessada et al., 2019). Additional work and optimization are needed to increase the accuracy of results in determining TAA digestibility (%) and the amount of antinutritional factors, and to identify genes related to seed storage proteins and antinutritional factors. Aiming to determine and improve the protein quality of winged bean seeds, further research is required to understand the impact of antinutritional factors on the digestibility of winged bean seeds across different genotypes, as well as the role of processing methods in mitigating these effects.

3.5 Conclusion

In developing countries, the main sources of protein in people's diets are pulses and beans. Winged bean seeds have a high protein content and are rich in lysine. Combining winged bean seeds with cereals in animal feeds and human diets could provide a good protein source. However, the winged bean seeds as many pulses, contain antinutritional factors. Increased condensed tannin intake and its effect on protein digestibility could be

considerable and further aggravate protein malnutrition, especially in the tropical regions where winged bean is grown.

The nutritional profile of winged bean seeds varied significantly among the accessions studied, highlighting their genetic diversity, which is highly valuable for plant breeding and enhancing resilience to climate change. Improving the nutritional profile of winged bean seeds requires a diverse approach, from plant genetics to food processing methods. Processing methods can have a beneficial impact on the nutritional quality of winged bean seeds. Apart from soaking, boiling and autoclaving other methods such as popping, fermentation and germination could contribute positively too. Heat-stable antinutritional factors might require enzyme supplementation in the feed such as phytase, and processing methods such as seed coat removal. However, these methods might be pushing higher the cost of food production and animal feed. Still, there is a lot of research that needs to be done on optimising the parameters of the different processing methods to maximise their effect to improve the nutritional quality of the seeds.

Further research is essential to uncover winged bean's potential to become an alternative protein source to soybean in the tropic regions. In terms of plant breeding, the genetic differences among the winged bean genotypes have a significant effect on protein variability. Therefore, a deeper understanding of the environmental effect and its significance as well as the genetic variation of winged bean accessions and their interaction, is essential for breeding programmes. Growing winged bean genotypes in different environmental conditions to record their performance and the nutritional profile of the seeds is essential. This information would be the baseline to design desirable ideotypes while taking into consideration the GxE interaction. Aiming to increase the protein quality of the winged bean seeds, breeding for high protein content with a higher amount of the limiting amino acids and a lower level of heat-stable antinutritional factors is crucial. Recent progress in genomics and transcriptomics holds promise for advancing research and breeding efforts through the utilisation of genetic markers. Methods such as Quantitative Trait Locus (QTL) and Genome-Wide Association Studies (GWAS) analysis could help identify molecular markers and genes that could improve the nutritional quality of

winged bean seeds would be a useful tool in the hands of plant breeders for marker-assisted selection.

Chapter 4: QTL analysis for protein, fat content and fatty acids

4.1 Introduction

In a fast-growing population under climate change, there is an increased pressure on food resources. Animal-based protein production has been associated with negative environmental impacts, by increasing greenhouse gas emissions, water consumption and land use (Henchion et al. 2017). The cropland is less than half of the global pasture area, and at the same time, approximately a third of the crops harvested are consumed as feed by livestock animals (Alexander et al. 2017). Legumes are considered to be an important source of protein across the world (Bessada et al., 2019). Winged bean is an underutilised tropical legume that has a high protein content, and the nutritional composition of winged bean seeds has been repeatedly compared to soybean, as an alternative protein and oil source (Prakash et al., 1987; Makeri et al. 2017). Nevertheless, winged bean seeds have high levels of antinutritional factors such as protein inhibitors, tannins and phytic acid (Adegboyega et al. 2019a; S. Sri Kantha and Erdman 1984b; M. Singh et al. 2019). Proximate composition analysis of 25 winged bean accessions revealed significant variation in crude protein, oil content, carbohydrates and crude fibre highlighting the underlying genetic variation in the winged bean germplasm (Adegboyega et al. 2019).

Winged bean has a diploid (2n = 2x = 18) genome of around 1.22Gbp (Vatanparast et al. 2016). The development of SSR-markers and genetic linkage groups provide the necessary tools for molecular breeding and genetic improvement of winged bean germplasm, for both qualitative and quantitative traits (Wong et al. 2016; Wong et al. 2017; Ho et al. 2024). Only recently, the first genetic linkage map and QTLs on pod, flower and seed-related traits was reported on winged bean (Chankaew et al. 2022; Ho et al. 2024).

In this study, the F3 population from Tanzi *et al.*, 2019 was used. The F3 population was generated from the parents Ma3 and FP15, which vary in morphology and seed content. The objectives of this study were: (1) to map QTLs for the protein, oil and five fatty acids (palmitic,

stearic, oleic, linoleic and behenic acids); (2) to identify putative genes that could assist in breeding selection. It is worth keeping in mind that the composition of the seeds is heavily influenced by the environment, such as the soil nutrients and agronomic practices, as mentioned in Chapter 3. While this study has served as a foundation, it is necessary that winged bean populations from different crosses are grown in several environments to get a better understanding of the genetic basis of differences in seed protein, oil and fatty acid composition.

4.2 Materials and Methods

4.2.1 Proximate analysis

The winged bean seeds used for the nutritional analysis were sent from the University of Nottingham, Malaysia following physiological traits evaluated in the F2 progenies as detailed Tanzi et al. (2019). The F2 population (XB2 cross) was generated by crossing Ma3 (paternal) with FP15 (maternal) winged bean accessions (individual FP15-10-3-2xMa3-8b) (Tanzi et al. 2019). Biological replicates of Ma3 and FP15 accessions were grown alongside the F2 accessions, from June to November 2017, Malaysia. The F3 winged bean seeds produced were stored and sent to the UK for analysis.

Figure 4. 1 Winged bean pod and seeds from the parental accessions. On the left, the pod and seeds from the FP15 and on the right from Ma3 (photo edited by Niki Tsoutsoura University of Nottingham, UK and Yuet Tian Chong, University of Nottingham, Malaysia).

Standard laboratory methods were used to analyse the winged bean seeds. The applied methods were in line with the AOAC International Standards. The winged bean seeds were ground using a centrifugal mill that passed them through a 0.5 mm sieve (Ultra-Centrifugal Mill ZM 200, Retsch). The samples were stored at -80°C overnight and then freeze-dried. The protein content was determined based on the modified Dumas method using a Protein Analyzer (FlashEA® 1112 N/Protein, Thermo Scientific) with the conversion factor of 6.25. The seed protein content result has been published by Ho et al. (2024). The oil content was

quantified by the Soxhlet method in a Gerhart SOXTHERM® extraction system using petroleum ether for the extraction. The oil extract was then used for the esterification of fatty acids to methyl esters (FAMEs) using methanol, 10M KOH, 12M H2SO4 and hexane, as described in the Methods Chapter 2. The relative amount of fatty acids was determined using GC-MS as described by O'Fallon et al., 2007.

4.2.2 QTL analysis

F3 seeds from F2 individuals of the XB2 cross between Ma3 and FP15 were analysed for protein content (n = 161), oil content (n = 93), and fatty acids (n = 93), the number of accessions was limited by the seed weight. The QTL analysis for protein, oil and fatty acids content was performed using the software MapQTL v6 with both non-parametric and parametric tests. Firstly, the data was tested for normality using Shapiro-Wil test, and a Permutation test was used to calculate the genome-wide (GW) significant LOD threshold (α = 0.05) for protein content (GW LOD=3.7), oil content and fatty acids (GW LOD = 3.8). The Permutation test calculates the significance threshold based on the actual data rather than on assumed normally distributed data (Ooijen 2009). Then, each trait was analysed through a Kruskal-Wallis (KW) test to establish single marker-trait associations (at $p < 0.01$), followed by Interval Mapping (IM) analysis using the GW LOD as a threshold. QTLs consistent between the two tests were reported. A QTL was considered significant when equal to or above the GW-LOD threshold and explained ≥ 10% of phenotypic variance (PVE%). Multiple-QTL model (MQM) mapping was utilised for seed protein, oil and fatty acids content QTL analysis. The QTLs on the genetic map were placed using MapChart v2.32, including markers with high LOD score and the flanking markers of a 2-LOD drop. Further information on whole genome sequencing, genetic mapping and bidirectional BLASTP has been reported by Ho et al. 2024.

The amino acid sequences from soybean and Arabidopsis were blasted on winged bean amino acid sequences, using the CLC software by Wai Kuan Ho. The winged bean genes were considered homologues to genes in soybean and Arabidopsis when the E-value $\leq 10^{-30}$ and $\geq$ 70% in sequence similarity (soybean).

4.2.3 Statistical analysis

The biological replicates of the parental accessions Ma3 and FP15 were analysed for the protein content using a Welch t-test due to the unequal sample size (Ma3 n=16, FP15 n=9). For the oil and fatty acid contents, as the Ma3 and FP15 had the same sample size (Ma3 n=3, FP15 n=3), one-way ANOVA was performed. Pearson's correlation was used for the fatty acid correlations based on the approximately normal distribution of the data.

4.3 Results

4.3.1 Agronomic traits

The XB2 population has been used to study traits such as protein (Ho et al. 2024) and yield-related traits (Tanzi et al. 2019). The results for the protein content in winged bean seeds have been published by Ho et al. (2024). In addition, to determine the genetic variation of oil and fatty acids in winged bean seeds, the total oil content (%) and the relative amount of palmitic, stearic, oleic, linoleic and behenic acids were measured. The normality test showed that the data was normally distributed for the protein content and the fatty acids palmitic, oleic and linoleic across the F_3 seeds (Table 4. 1).

Table 4. 1 Test for normality. Shapiro-Wilk test was used for the protein (%), oil (%) and relative amount of the fatty acids Palmitic, Stearic, Oleic, Linoleic and Behenic. The numbers correspond to 1=Saturated fatty acids, 2=Monounsaturated fatty acids, and 3=Polyunsaturated fatty acids.

			Fatty acids (%)				
	Protein (%)	**Oil (%)**	**Palmitic[1]**	**Stearic[1]**	**Oleic[2]**	**Linoleic[3]**	**Behenic[1]**
Test statistic W:	0.9725	0.9789	0.9485	0.9779	0.9607	0.6603	0.9834
Probability:	0.002	0.1369	0.001	0.116	0.007	<0.001	0.289

Although the deviation of oil content and most of the fatty acids (except behenic and stearic acid) was not large among the parental accessions Ma3 and FP15, transgressive segregation was observed among the F_3 seeds, as outlined in (Figure 4. 2). Transgressive segregation was observed in the XB2 population; where protein content ranged from 27.18% to 44.56% in the XB2 population, while the parents, Ma3 and FP15, had an average of 34% and 39.3%, respectively (Table 4. 2). This was also the case for palmitic acid. The oil content in Ma3 was 20.91 and 19.89% in FP15, whereas the oil content in the XB2 population ranged from 16.84% to 23.31% (Table 4. 2, Figure 4. 2). The transgressive segregation observed on the traits measured suggested that these quantitative traits are complex and controlled by several genes.

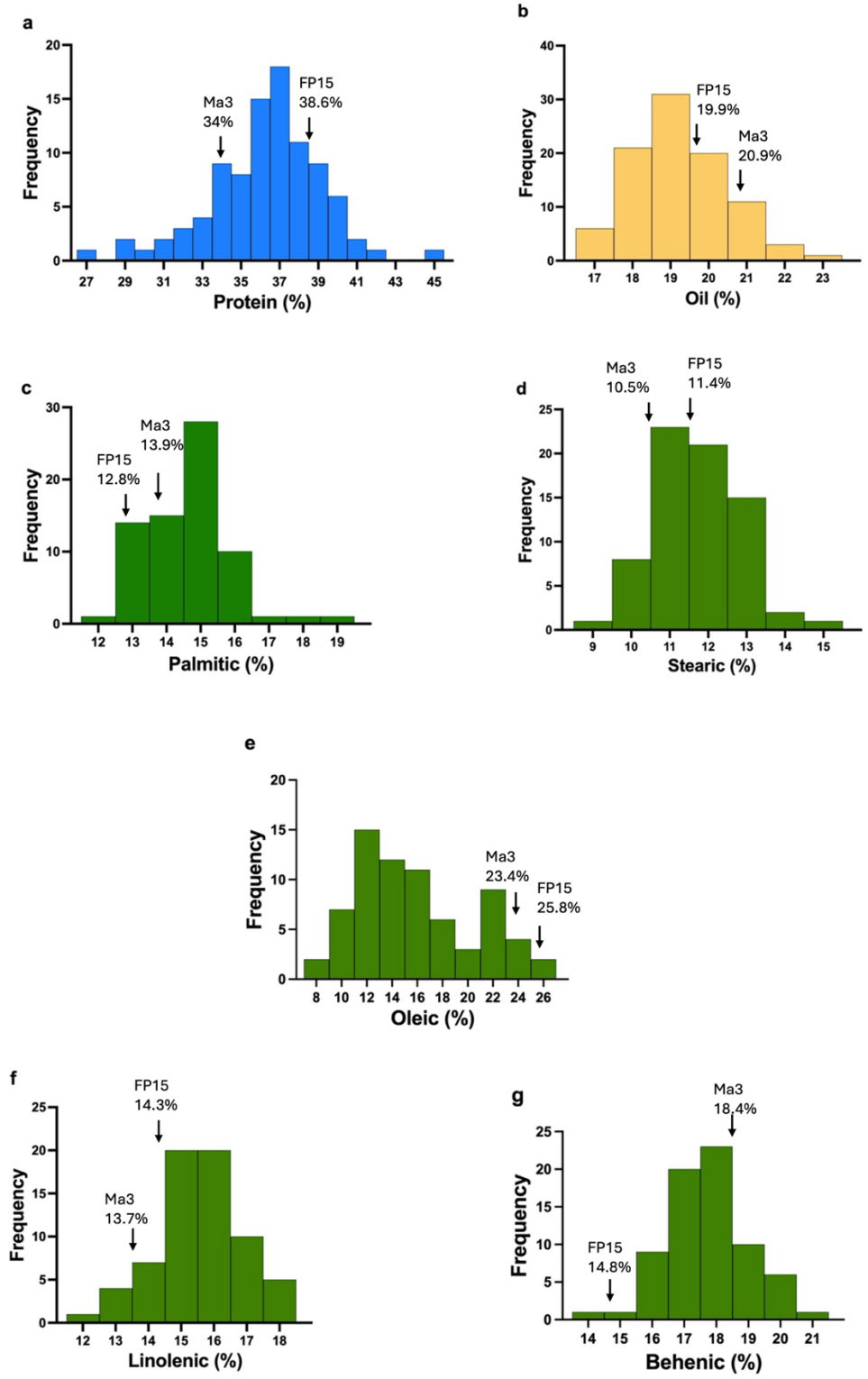


Figure 4. 2 Frequency distribution. a. Protein g/100 g of dry seed weight (n=161); b. Oil g/100 g of dry seed weight (n=93) c. Relative amount of palmitic acid (%) of winged bean oil (n=93); d. Relative amount of stearic acid (%) of winged bean oil (n=93); e. Relative amount of oleic acid (%) of winged bean oil (n=93); f. Relative amount of linoleic acid (%) of winged bean oil (n=93), g. Relative amount of behenic acid (%) of winged bean oil (n=93). The parental accessions FP15 and Ma3 are shown.

*Table 4. 2 Descriptive statistics of protein (%), oil (%) and relative amount of the fatty acids Palmitic, Stearic, Oleic, Linoleic and Behenic. For each trait, the average mean is shown ± the standard error. For the protein (%) content, for Ma3 n=16 and FP15 n=9. To compare the protein content between Ma3 and FP15 as these are two groups with unequal sample sizes, independent samples t-test was used. For the XB2 individuals analysed the size sample was n=161. For oil and fatty acids, for Ma3 and FP15 n=3, and one-way ANOVA was used. For XB2 individuals, the sample size was n=93. sd = Standard deviation, CV (%)= coefficient of variation. 1=Saturated fatty acids, 2=Monounsaturated fatty acids, 3=Polyunsaturated fatty acids. * when p<0.05. , ** when p<0.01*

	Trait	Min	Max	Range	Mean	SD	CV (%)	Ma3	FP15
	Protein (%)	27.18	44.56	17.38	36.31	2.93	3.12	34.0±0.38**	39.03±1.03
	Oil (%)	16.84	23.31	6.47	19.24	1.26	0.01	20.91±0.19	19.89±0.27
C16:0	**Palmitic[1]**	12.18	19.11	6.92	14.63	1.20	1.69	13.93±0.77	12.76±0.17
C18:0	**Stearic[1]**	8.57	14.88	6.31	11.71	1.16	1.63	10.45±0.63*	11.35±0.08
C18:1n9c	**Oleic[2]**	7.17	26.84	19.66	15.72	4.63	6.52	23.45±0.65	25.80±0.43
C18:2n6c	**Linoleic[3]**	12.30	17.78	5.48	15.54	1.24	1.85	13.65±0.29	14.34±0.9
C22:0	**Behenic[1]**	14.20	20.76	6.56	17.75	1.29	1.82	18.41±0.19*	14.78±0.1

The pairwise correlation analysis on the fatty acids showed significant correlations. A significant positive correlation was observed between palmitic with stearic and behenic acids. Stearic and behenic acids were also positively correlated. In contrast, oleic acid was negatively correlated with linoleic acid. Both oleic and linoleic acids were negatively correlated with palmitic acid, stearic acid and behenic acid (Table 4. 3).

*Table 4. 3 . Pearson's correlations of fatty acids and fat content. The R-value from the correlations is shown in the table, with the relative amount of fatty acids. The significance is mentioned with the star sign *** p<0.001, **** p<0.0001. The numbers correspond to: 1=Saturated fatty acids, 2=Monounsaturated fatty acids, 3=Polyunsaturated fatty acids.*

Fatty Acids	**Palmitic[1]**	**Stearic[1]**	**Oleic[2]**	**Linoleic[3]**
Stearic[1]	0.759****			
Oleic[2]	-0.491****	-0.368***		
Linoleic[3]	-0.547****	-0.542****	-0.321**	
Behenic[1]	0.473****	0.404****	-0.437****	-0.428****

4.3.2 QTL Analysis

The QTL analysis resulted in the identification of several QTLs for protein, oil and fatty acids. The QTLs with LOD score equal to or higher than the genome-wide LOD threshold with a PVE>10% were considered significant. Of the 16 QTLs reported, 3 were significant QTLs (qLin-5-1, qLin-9-1 and qBeh-3-1). The rest of the QTLs reported had a high LOD score or a PVE higher or close to 10% (Table 4. 4). The two protein QTLs explained 17.3% of the variation observed in the XB2 population, while the three oil QTLs explained 35.6%. Interestingly, the four linoleic QTLs explained 53.2% of the variation with the two significant QTLs having a PVE of 31.6%. The third significant QTL was reported in behenic, qBeh-3-1, with a PVE of 16.9%. For behenic, the two QTLs detected in chromosome 3 (qBeh-3-1 and qBeh-3-2) had a total PEV of 27.2% (Table 4. 4). For palmitic, 2 QTLs were detected in chromosomes 3 and 5, with PEV equal to 20.1%. For stearic, it was interesting to note that the QTL (qStear-3-1) detected was the same as qPal-3-1 in chromosome 3, with a similar PEV of 10%. For oleic, no significant QTLs were detected, however, the two QTLs in chromosomes 1 and 5 had a total PEV of 23.6% (Table 4. 4).

In chromosomes 4, 6 and 8, no QTLs related to protein, oil and fatty acids were found. Interestingly, two significant QTLs were closely located, qBeh-3-1 and qProt-3-1. The qProt-3-1 was close to the qBeh-3-2, too. As shown in Figure 4. 3, there was an overlap of the QTL regions of these two traits. Similar was the case for the QTLs in chromosome 5, qOil-5-1 and qOl-5-1.

*Table 4. 4 Significant QTLs for fatty acids. Phenotypic Variation Explained (PVE). 1=Saturated fatty acids, 2=Monounsaturated fatty acids, 3=Polyunsaturated fatty acids. The * at the end of the QTLs name means that there is a winged bean gene in close proximity to the market, with high similarity to soybean and Arabidopsis genes involved in the fatty acid biosynthesis pathway. With bold are the significant QTLs and highlighted is the same marker.*

Trait	QTL name	Group	Position	Marker	LOD	PVE (%)	Additive effect	Dominance effect
Protein	qProt-3-1	Chr3	59.104	13560_30:G>T	4.19	9.2	-1.153	0.801
	qProt-7-1	Chr7	90.798	18371_20:A>G	2.96	8.1	-1.004	0.791
Oil	qOil-2-1	Chr2	0	24732_23:G>A	3.3	12.3	-0.552	0.441
	qOil-5-1	Chr5	259.438	8171_14:C>G	3.23	12	0.691	0.055
	qOil-9-1	Chr9	16.211	11154_46:G>T	3.04	11.3	-0.589	0.223
Palmitic[1]	qPal-3-1*	Chr3	7.37	25763_26:G>A	3.38	9.9	-0.307	-0.907
	qPal-5-1*	Chr5	0	14758_43:A>C	3.5	10.2	0.59	-0.386
Stearic[1]	qStear-3-1*	Chr3	7.37	25763_26:G>A	2.27	10	-0.307	-0.907
Oleic[2]	qOl-1-1*	Chr1	83.493	13470_64:C>T	3.5	14.3	3.949	1.551
	qOl-5-1	Chr5	273.897	11489_5:T>C	2.34	9.3	-1.34	2.708
Linoleic[3]	qLin-2-1*	Chr2	25.921	8506_16:A>G	2.7	9.4	0.749	-0.191
	qLin-5-1*	Chr5	37.137	23128_19:A>G	**4.06**	**14.7**	-0.917	-2.083
	qLin-5-2*	Chr5	86.797	19704_36:C>T	3.43	12.2	1.309	1.566
	qLin-9-1*	Chr9	48.053	27258_49:C>T	**4.58**	**16.9**	-0.912	-0.375
Behenic[1]	**qBeh-3-1**	Chr3	45.835	17225_8:C>G	**4.19**	**16.9**	1.234	0.988
	qBeh-3-2	Chr3	65.703	26777_47:A>C	2.66	10.3	-0.78	-0.989

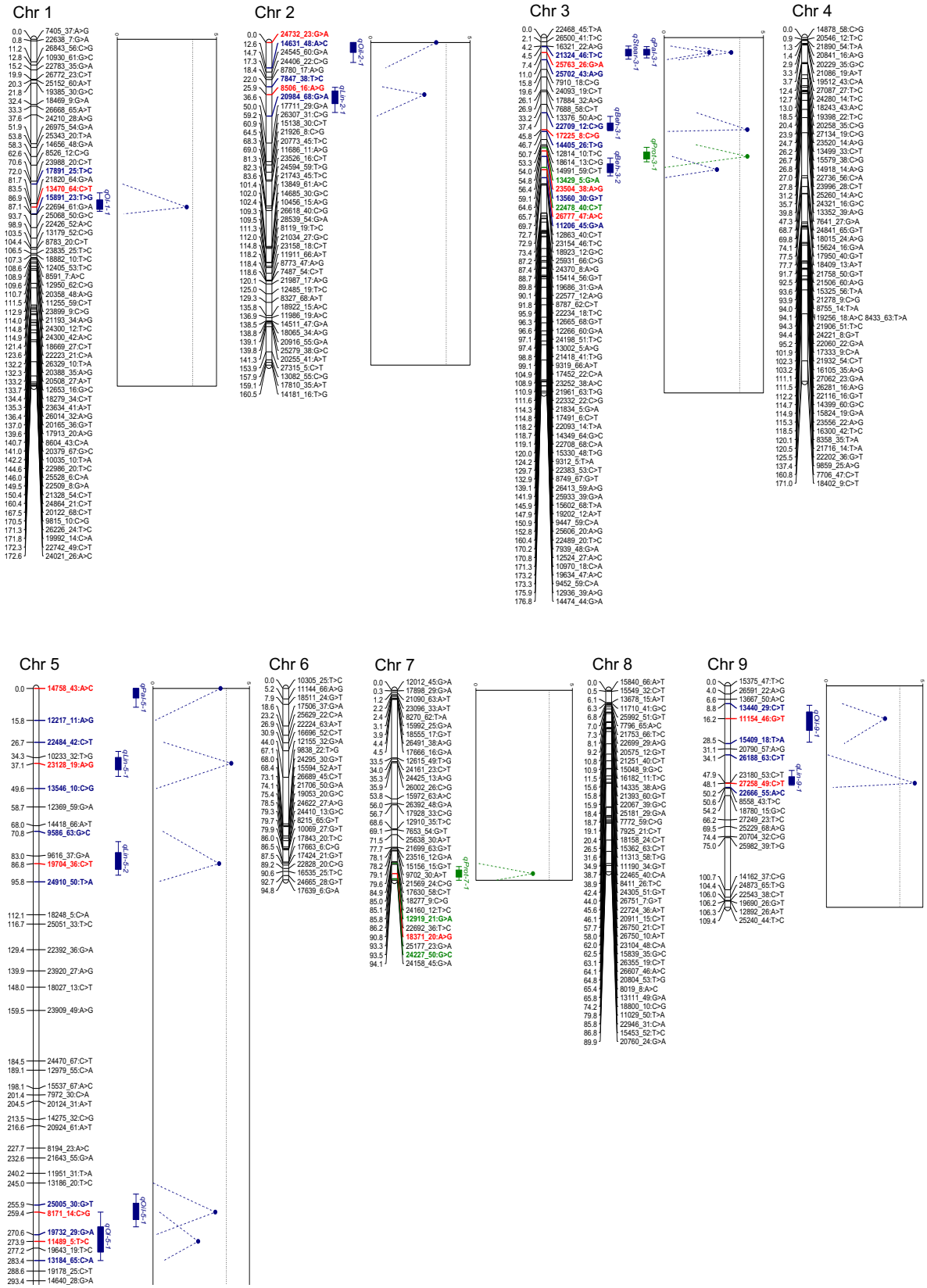
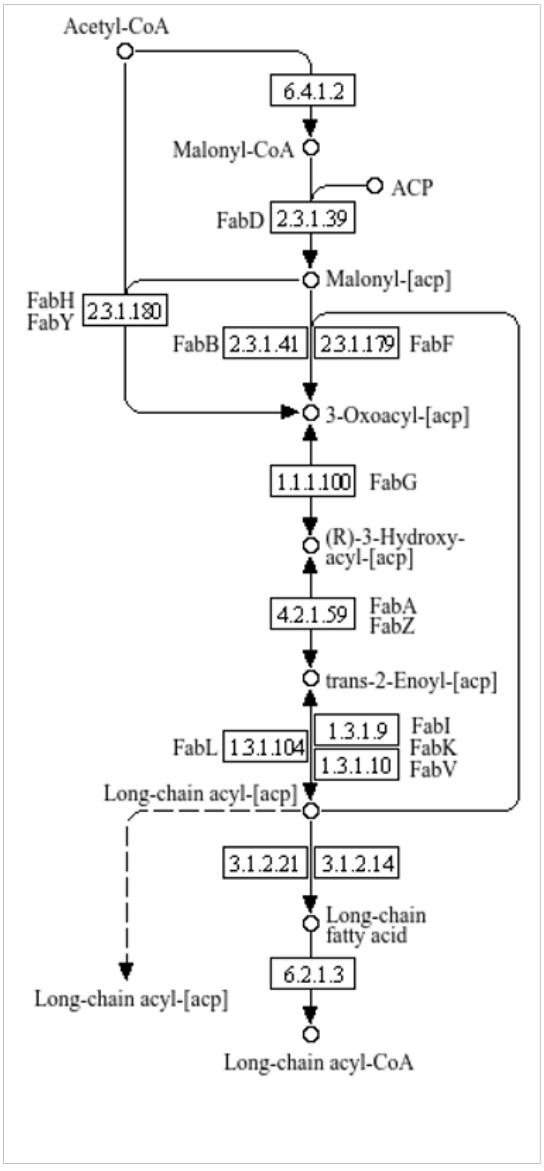


Figure 4. 3 Genetic map of the winged bean genome of the XB2-population of the QTLs for protein and fatty acids. This figure shows the position of the QTLs detected by MapQTLs 6.0 on winged bean genome for protein content (green colour) and fatty acids (blue colour). The marker with the high LOD score is shown with red colour, whereas the flanking markers are shown with green for protein content and blue for fatty acids.

4.3.3 Putative winged bean genes

Genes involved in fatty acid biosynthesis and homologues to soybean and Arabidopsis have been found in the winged bean genome (Table 4. 5). Interestingly, 23 winged bean genes have been identified with similarity higher than 80% to soybean genes and E-value close to zero. Related to protein QTLs, within the qProt-3-1, the Psote03G0119900 homologous to xylem bark cysteine peptidase 3 Glyma.08G116300 was found; and in qProt-7-1, the gene Psote07G0210200.1 homologous to a DELLA protein Glyma.11g216500 in soybean was detected. DELLA proteins, named from five conserved amino acids (aspartic acid, glutamic acid, leucine, leucine, and alanine) in their N-terminal domain, are negative regulators of gibberellin (GA) signalling of the GA receptor (Eckardt 2007).

The winged bean gene Psote05G0083100.1 has homologues to Glyma.09G277400 (*FabH/KASIII* gene) (Figure 4. 4) and was located within the significant QTL qLin-5-1, in chromosome 5. The Psote01G0031000.1 within qOl-1-1 is homologous to *FabI* of soybean (Glyma.11G101400). In chromosome 2, Psote02G0269200.1 was within the qLin-2-1 QTL and homologues to *FabZ* Glyma.15G052500. While in chromosome 3, the gene Psote03G0022300.1 homologous to *ACCC-2* Glyma.05G221100), was within the QTL of qPal-3-1 and qStear-3-1. In chromosome 5, within the qLin-5-2 the winged bean gene Psote05G0096100.1 homologues to *ACCA-2* (Glyma.18G195700) was identified. In terms of protein content, the winged bean genes Psote03G0078700.1 and Psote03G0078800.1 were close to the protein QTL, qProt-3-1, and homologues to the 11S globulin soybean genes Glyma.05G169200 and Glyma.05G169100, respectively.



KEGG Pathway	Winged bean gene
6.4.1.2	**Psote05G0053200.1**
	Psote03G0022300.1
	Psote05G0096100.1
	Psote09G0032300.1
	Psote05G0032900.1
FabD	Psote02G0175100.1
FabH	**Psote05G0083100.1**
FabF	Psote06G0352200.1
	Psote04G0300400.1
	Psote03G0026000.1
	Psote03G0088900.1
FabG	Psote05G0203200.1
	Psote07G0237900.1
	Psote03G0105900.1
FabZ	**Psote02G0269200.1**
FabI	Psote05G0117000.1
	Psote01G0031000.1
MECR	**Psote03G0381000.1**
3.1.2.21	Psote06G0283600.1
3.1.2.14	Psote02G0195400.1
	Psote02G0163100.1
	Psote02G0162800.1
6.2.1.3	Psote04G0059800.1
	Psote02G0093600.1
	Psote07G0243000.1
	Psote04G0191600.1
	Psote01G0429300.1
	Psote04G0339800.1
1.14.19.2	Psote08G0127900.1
	Psote03G0166900.1
	Psote08G0128000.1

Figure 4. 4 Fatty acid biosynthesis pathway. Winged bean genes identified in the oil and fatty acids QTLs were highlighted with bold letters. 6.4.1.2. ACCA-3 = acetyl-CoA carboxylase carboxyl transferase subunit alpha; FabD (2.3.1.39)= malonyl CoA-acyl carrier protein transacylase-like; FabH (2.3.1.180) = 3-oxoacyl-[acyl-carrier-protein] synthase III; FabF (2.3.1.179)= KASII-B, plastid 3-keto-acyl-ACP synthase II-B precursor; FabG (1.1.1.100) = 3-oxoacyl-[acyl-carrier-protein] reductase 4; FabZ (4.2.1.59) = 3-hydroxyacyl-[acyl-carrier-protein] dehydratase; FabI (1.3.1.9-1.3.1.10) = enoyl-[acyl-carrier protein] reductase I; MECR = mitochondrial enoyl-[acyl-carrier protein] reductase / trans-2-enoyl-CoA reductase; 3.1.2.21 = FAT1, fatty acyl-ACP thioesterase B; 3.1.2.14 = FAT1 fatty acyl-ACP thioesterase B; 6.2.1.3 = long-chain acyl-CoA synthetase; 1.14.19.2 = stearyl acyl carrier protein desaturase.

Table 4. 5 Winged bean putative genes. The homologues from Glycine max and Arabidopsis thaliana are shown with their respective E-Values and identity percentages.

WB Genes	QTL	KEGG name	Accession (E-value)	Lowest E-value	Greatest Identity (%)
Psote03G0078700.1	qProt-3-1		Glyma.05G169200	0	86.31
Psote03G0078800.1	qProt-3-1		Glyma.05G169100	0	75.63
Psote03G0119900.1	qProt-3-1		Glyma.08G116300	2.357E-153	77.2
Psote07G0210200.1	qProt-7-1		Glyma.11G216500	0	82.04
Psote01G0031000.1	qOl-1-1	*FabI/ENR1, MOD1*	Glyma.11G101400	0	89.14
Psote02G0085400.1		*ACC1*	Glyma.04G104900	0	93.85
Psote02G0162800.1		*FATB*	Glyma.06G168100	0	88.66
Psote02G0175100.1		*FabD*	Glyma.11G164500	0	93.99
Psote02G0195400.1			Glyma.06G211300	0	85.89
Psote02G0269200.1	qLin-2-1	*FabZ*	Glyma.15G052500	2.94E-138	94.04
Psote03G0022300.1	qPal-3-1 qStear-3-1	*ACCC-2*	Glyma.05G221100	0	95.9
Psote03G0166900.1		*SAD2/ACPD*	Glyma.07G207200	0	96.68
Psote03G0381000.1	qBeh-3-2	*MECR*	Glyma.13G330100	1.41E-162	81.94
Psote05G0053200.1	qPal-5-1	*ACCB, BCCP*	Glyma.09G248900	4.97E-87	90.14
Psote05G0083100.1	qLin-5-1	*FabH - KASIII*	Glyma.09G277400	0	93.95
Psote05G0096100.1	qLin-5-2	*ACCA-3 ACCA-2*	Glyma.18G195700	7.13E-168	86.67
Psote05G0226400.1		*FAD2*	Glyma.19G147400	0	94.33
Psote06G0023300.1			Glyma.07G019100	0	87.32
Psote06G0283600.1		*FAT1 / FATB*	Glyma.05G012300	0	94.23
Psote06G0352200.1		*FabF- KASII-B*	Glyma.13G112700	0	94.21
Psote07G0237900.1		*FabG*	Glyma.18G009200	0	92.17
Psote08G0128000.1		*SACPD-C / SACPD*	Glyma.13G038600	0	83.91
Psote09G0032300.1	qLin-9-1	*ACC2/BCCP*	Glyma.13G057400	1.26E-109	80.38

4.4 Discussion

The nutritional value of winged bean seeds has been reported to resemble soybean (Kantha and Erdman 1986; Amoo, Adebayo, and Oyeleye 2006; Yanagi 1983). Apart from the high protein content ranging between 30-40%, the oil content is also relatively high between 15-20% in winged bean seeds (Worthington, Hammons, and Allison 1972; Garcia and Palmer 1980b; R. King and Puwastien 1987; Lepcha et al. 2017). The nutritional quality of the seeds depends on the protein content and the amino acids as well as the amount of saturated and unsaturated fatty acids. Therefore, it is important to understand the genetics behind the desirable quantitative traits. This information would be very useful in breeding selection and contribute to improving the nutritional quality of winged bean varieties. In recent years, the genomic data provided by sequencing platforms accelerated gene discovery, marker development and marker-trait association aiming to improve marker-assisted breeding for the winged bean. Previous studies have focused on morphological and yield-related traits such as pod and seed colour, days to flowering, pod length, seeds size and number of pods (Tanzi 2018; Tanzi et al. 2019; Wong et al. 2017; Sriwichai et al. 2022).

This is the first time that QTL analysis and putative genes have been reported for both protein content and fatty acids in winged bean seeds. In this study, the plants were not cultivated under standardised conditions, and as a result, environmental factors may have influenced the nutritional composition observed. This should be taken into account when interpreting the data and considering its application in genetic improvement efforts. Out of the 16 QTLs reported, two were related to protein content. Of the 14 fatty acids QTLs, there were three significant QTLs, two for linoleic acid (qLin-5-1 and qLin-9-1) and one for behenic acid (qBeh-3-1). Winged bean genes homologous to soybean genes involved in the biosynthesis of fatty acids have been identified as well as their closest map markers. Protein and fatty acid content are complex quantitative traits, controlled by multiple loci in many crop species. As such, the identification of genetic markers and quantitative trait loci (QTLs) is valuable for breeding programmes aimed at improving nutritional quality (Wang ML et al. 2015; Li et al. 2018a; X. Wang et al. 2012).

4.4.1 Protein QTLs

The parental accessions had significantly different protein contents with Ma3 at 34% and FP15 at 39.3%. Interestingly, transgressive segregation was observed in the XB2 population (Table 4. 2, Figure 4. 2) with similar results reported in soybean (Warrington et al. 2015; Y. H. Zhang et al. 2015). For protein content, 2 QTLs were found - qProt-3-1 and qProt-7-1 in chromosomes 3 and 7, respectively (Table 4. 4). Two tandem genes (Psote03G0078700 and Psote03G0078800) homologous to the 11S globulin in soybean (Glyma.05G169200) were found 6.6 Mbp upstream of the qProt3-1 (Table 4. 5). Most leguminous seeds contain 7S and 11S globulins and albumins as major seed storage proteins (Adachi et al. 2002). Within qProt-3-1, the Psote03G0119900 homologous to xylem bark cysteine peptidase 3 (Glyma.08G116300) was found. In Arabidopsis, the homologous gene (AT1G20850.1) is a cysteine protease XBCP2 (Molina et al. 2021). In qProt-7-1, the winged bean gene Psote07G0210200.1 was found which is homologous to Glyma.11G216500 in soybean and AT3G03450.1 in Arabidopsis. The AT3G03450.1 (RGL2) gene encodes for a DELLA protein; a member of the GRAS superfamily of putative transcription factors and it is a probable transcription regulator that represses the gibberellin (GA) signalling pathway (Rombolá-Caldentey et al. 2014). DELLA proteins don't directly code for seed storage proteins but can affect their expression and accumulation by influencing plant growth, seed development, and environmental response (Gomez et al. 2023; Phokas and Coates 2021; Rombolá-Caldentey et al. 2014).

4.4.2 Fatty acid QTLs

In this study, the two unsaturated fatty acids, oleic and linoleic acids, comprise around 60-70% of the oil in winged bean seeds, while the saturated fatty acids, behenic, palmitic and stearic acids, are estimated to be at approximately 10-20%, 15% and 10%, respectively. These results are in line with the literature (Sekhar Mohanty et al. 2021; Garcia, Palmer, and Young 1979; Higuchi, Terao, and Iwai 1982).

For palmitic acid, 2 QTLs have been reported (qPal-3-1 and qPal-5-1) in chromosomes 3 and 5, respectively (Table 4. 4). Interestingly, the QTL on chromosome 3, qPal-3-1, was also detected during the QTL analysis for stearic acid, qStear-3-1. It is important to note that for

the qPal-3-1, the GW LOD score was 3.38, just at the threshold, and the PVE% was 9.9, just below 10% which was the threshold for identifying a QTL as significant (Table 4. 4). This might be a consistent QTL, however, more research needs to be done. Within the QTL region, the winged bean gene Psote03G0022300.1 homologous (95.9%) to *ACCC-2* Glyma.05G221100 from soybean was detected (Table 4. 5). The ACCC-2 in soybean is an isoform of the enzyme acetyl-CoA carboxylase (ACCase). ACCase plays an important role in fatty acid biosynthesis, as it catalyses the carboxylation of acetyl-CoA to malonyl-CoA, which is the first step in the fatty acid synthesis pathway (Figure 4. 4). ACCase is a multi-subunit enzyme consisting of biotin carboxylase (BC), biotin carboxyl carrier protein (BCCP), and the carboxyltransferase subunits (alpha CT-α and beta subunits CT-β) (Reverdatto, Beilinson, and Nielsen 1999). The winged bean gene (Psote05G0053200.1) homologous (90.14%) to the soybean *BCCP* gene Glyma.09G248900 was shown to be located within the palmitic acid QTL, qPal-5-1, in chromosome 5 (Table 4. 5).

Out of the 16 QTLs mentioned, two were linked to oleic acid, qOl-1-1 and qOl-5-1 on chromosomes 1 and 5, respectively. Within the range of qOl-1-1, the winged bean gene Psote01G0031000.1, homologous (89.14%) to soybean Glyma.11G101400 was regarded as putative (Table 4. 5). The Glyma.11G101400 in soybean is the *FabI* gene. FabI is an enoyl-[acyl-carrier-protein] reductase [NADH] and catalyses a key regulatory step for the biosynthesis of fatty acids in *E.coli* (Bergler et al. 1996). It accepts NADH and NADPH as cofactors and is inhibited by palmitoyl-CoA (Bergler et al. 1996). In Arabidopsis, mutant plants have a reduced lipid level and pleiotropic morphological defects such as abnormal shape of leaves (L. Luo et al. 2019). Additionally, the qOl-5-1 overlapped with the qOil-5-1, suggesting that the oleic acid content might be related to the variation in oil content among the parental accessions. Further analysis should be done to identify genes and mutations within the two QTLs that could be related to oil content.

Oleic acid is further desaturated by the FAD2 enzyme to linoleic acid; with oleic and linoleic acids showing a significant and negative correlation ($p<0.01$) (Table 4. 3). Similar results have been reported in peanut and soybean even when grown in different environments (Silva et al. 2021; Wang ML et al. 2015; Ting et al. 2016). The ratio between oleic and linoleic acids is an important parameter of oil stability during the frying process, where a high oleic and lower

linoleic acid content is preferable (R. Romano et al. 2021; 2013; K. S. Yu, Cho, and Hwang 2018). In this study, 4 QTLs (qLin-2-1, qLin-5-1, qLin-5-2 and qLin-9-1) were identified for linoleic, with 2 of them qLin-5-1 and qLin-9-1 being significant (PVE% > 10% and LOD score > 3.8) (Table 4. 4, Table 4. 5).

Within the QTL region of qLin-5-1 (PVE% = 14.7% and LOD score at 4.06), the winged bean gene Psote05G0083100.1 homologous (93.95%) to *FAbH* or *KASIII* gene, Glyma.09G277400 in soybean, was detected (Table 4. 5). The KASIII enzyme is responsible for catalysing the condensation of acetyl-coA with malonyl-ACP in dissociated fatty acid synthesis, one of the initial steps for fatty acid elongation (Figure 4. 4) (Tsay et al. 1992). In close proximity to the qLin-5-2 QTL, the winged bean gene (Psote05G0096100.1) homologous (86.67%) to *ACCA-3/ ACCA-2* gene in soybean (Glyma.18G195700) was identified (Table 4. 5). The ACCA refers to the alpha subunit of the ACCase enzymatic complex that converts acetyl-coA to malonyl-coA, as mentioned above (Bilder et al. 2006). Interestingly, within the qLin-9-1 the winged bean gene Psote09G0032300.1 homologous (80.38%) to Glyma.13G057400 in soybean with *ACC2/BCCP* function was detected. The ACCase contains the homomeric ACC2 and the heteromeric complex consisting of CT-α, CT-β, BCCP, and BC subunits. It has been hypothesised that the heteromeric complex of ACCase in castor bean and rapeseed could be related to the higher levels of oil in the seeds compared to maize (Liu et al. 2022). This is along with the increased number of lipid biosynthesis genes and reduced number of lipid metabolism genes regulated during the lipid accumulation in both castor bean and rapeseed embryo, compared to maize embryo; suggesting that these differences might be resulting in the seed oil differences among the species (Liu et al. 2022). In the qLin-2-1, the Psote02G0269200.1 gene homologous (94.04%) to *FabZ* gene, Glyma.15G052500 in soybean as well as At5g10160 and At2g22230 in Arabidopsis (Table 4. 5). In *E.coli*, FabZ is involved in the fatty acid biosynthesis pathway catalysing the dehydration of (3*R*)-hydroxyacyl-ACP to *trans*-2-acyl-ACP (Heath and Rock 1996). In plants, the dehydration of β-hydroxyacyl-ACP to trans-2-enoyl-ACP, is performed by β-hydroxyacyl-[ACP]-dehydratase (HAD) enzyme (González-Thuillier et al. 2016). In Arabidopsis, two genes are encoding for *HAD* genes, At2g22230 and At5g10160. Both genes are upregulated during lipid biosynthesis in seed development (Schmid et al. 2005). Another gene that plays an important role in fatty acid biosynthesis is *FAD2*, converting oleic acid to linoleic acid, as it has been reported for both

Arabidopsis and soybean (Schlueter et al. 2007; Okuley et al. 1994). In winged bean, the gene Psote05G0226400.1 homologous (94.33% and 75.26%) to the *FAD2* genes in soybean (Glyma.19G147400) and Arabidopsis (AT3G12120), was located at chromosome 5. Psote05G0226400.1 was not close to any of the QTLs detected in this study (Table 4. 5).

Behenic acid (C22:0), also known as docosanoic acid or docosanoate, is a long-chain saturated fatty acid which is poorly absorbed and has been related to increased cholesterol in humans (Cater and Denke 2001). In winged bean seeds, the amount of linoleic and behenic acid seemed to vary in the literature, with some reporting linoleic at ~25% and behenic at ~20% (Ekpenyong and Borchers 1980) while others found that linoleic was ~35% and behenic acid less than 10% (C. S. Mohanty et al. 2015a; 2021). In this study, behenic was negatively but not significantly correlated with oleic and linoleic acid (Table 4. 3). It is known that from stearic acid, there are two different directions of fatty acid formation. Stearic acid can be elongated to arachidonic acid and then behenic acid with the addition of acetyl residue reactions, or it can be desaturated to oleic acid and then with a further desaturation in the chloroplast to linoleic acid. *FAD2* encodes for a fatty acid desaturase that forms linoleic from oleic acid, with the mutant allele for *FAD2* blocking this pathway (M. L. Wang et al. 2015). It would be interesting to explore this further.

For behenic acid, two QTLs were detected, qBeh-3-1 and qBeh-3-2, with the former being a significant QTL (LOD score of 4.56 and PVE% at 16.9%) (Table 4. 4, Table 4. 5). Interestingly, the behenic acid QTLs were in succession and as shown in (Figure 4. 3), there was the protein QTL qProt-3-1 between them. The winged bean gene Psote03G0381000.1, within the qBeh-3-2, is homologous (81.4%) to *MECR* gene in soybean (Glyma.13G330100) (Table 4. 5). MECR is a reductase highly expressed in mitochondria for the elongation of fatty acids. The overlap of qProt-3-1 and qBeh-3-2 is demonstrated (Figure 4. 3). Similar results have been reported in soybean, with oil and protein QTLs being significantly and consistently correlated (Zhang et al. 2019).

4.5 Conclusion

Considering that the population size for fatty acid QTL analysis (93 individuals, less than 100) is small, the QTLs identified as significant might be overestimated. According to the "Beavis effect", when the population size is only 100 then the effect of QTLs is often overestimated, when it is 500 (a medium size population), it is slightly overestimated and when it is 1000 the estimation is close to the actual magnitude (S. Xu 2003). Therefore, it is essential to consider that the QTLs mentioned in this study have only been detected in a population size of less than 100 and only tested in one environment. Growing the population in several environments increases the phenotypic variation, especially on traits highly affected by the environment such as oil and protein during seed development. If the plants are grown in different locations over the years, it helps account for environmental effects and thus identifying genes with high heritability value (Zhang et al. 2019). Another important parameter that could improve the accuracy of the QTL analysis is high genotypic variation in a cross combined with a good coverage of the genome by markers. Genetic markers that are spaced across the winged bean genome as well as using several crosses of medium-sized or large populations in different environments and years would result in a more precise QTL identification. Nonetheless, an additional method to identify genes associated with traits is GWAS (Genome-Wide Association Studies) which would require hundreds of winged bean accessions. As protein and oil are negatively correlated and their content changes as the seeds mature (Mohanty et al. 2015), it would be interesting to monitor the gene expression during pod and seed development aiming to identify genes involved in protein and oil accumulation for future genome editing and breeding programmes.

Winged bean is a potential alternative protein of source that could complement or partially replace soybean as a dietary protein source for animal or human consumption. Identifying QTLs related to protein and oil composition could assist breeding selection. However, further research needs to be conducted on improving the protein and oil quality of winged bean seeds while minimising the effects of antinutritional factors. A combination of research focusing on breeding for high protein and low antinutritional factors alongside studies on the food properties and processing methods of winged bean seeds would uncover its potential to become a new soybean for the tropics. Understanding the genetics behind desirable

quantitative and qualitative traits would assist breeding selection and contribute to improving winged bean varieties.

Chapter 5: Transcriptomic profiling of winged bean pods and seeds at different developmental stages

5.1 Introduction

Legume genomes have been shaped by a whole-genome duplication event approximately 58 Mya (Vlasova et al. 2016). Winged bean is most closely related to *Glycine* rather than other legume species and is deduced to have diverged approximately 14–16 Mya, before *Glycine* doubled its chromosome number (Figure 5. 1a) (Ho et al. 2024). The estimated synonymous substitution rate (Ks) distribution of the winged bean paranome (set of paralogous genes) revealed that it has undergone the legume-common tetraploidy duplication event (Figure 5. 1b) in Fabaceae as in other legumes (Z. Wang et al. 2015; Ho et al. 2024). Consistent with this, it was also reported that winged bean shares 58.3% and 55.3% collinearity with *G. max* and *G. soja*, respectively, with the largest collinearity blocks observed between Pt04 and Gm10 (Figure 5. 1c). Nevertheless, extensive chromosomal rearrangement or gene translocation was observed between them (Ho et al. 2024). Interestingly, the winged bean genome appeared enriched for genes involved in isoflavonoid biosynthesis and secondary metabolites (such as phytosterol, castasterone and ergocalciferol), potentially functioning as phytoalexins through jasmonate- and ethylene-mediated pathogen defence mechanisms (Ho et al. 2024).

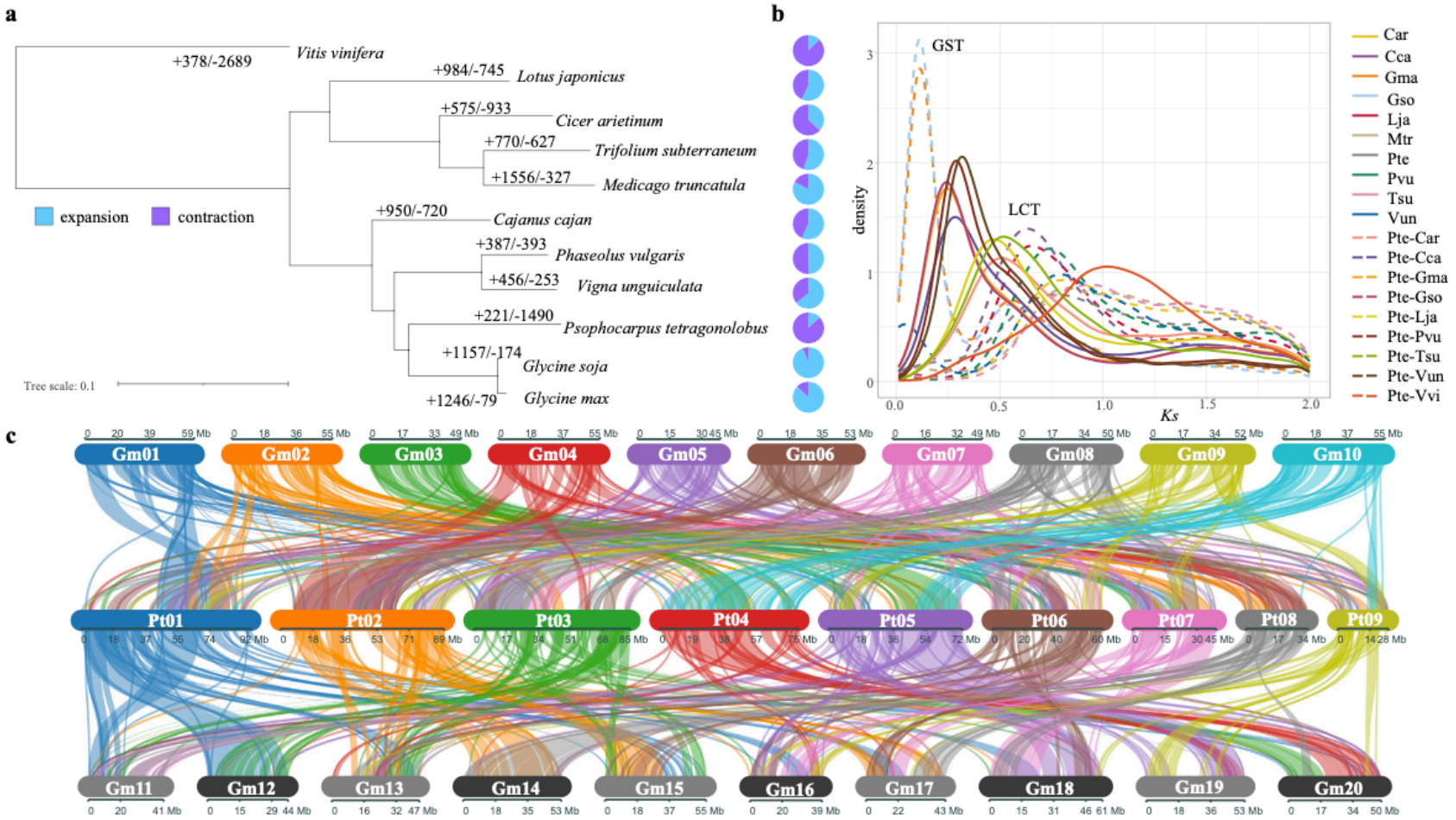


Figure 5. 1 Evolutionary analysis of winged bean with nine legume species. (a) Proportion of expanded (blue) and contracted (purple) gene families in ten legume species. (b) Ks plot of the paralogues and orthologous genes of ten legumes showing the glycine-specic tetraploidy (GST) event in G. max and G. soja and the legume-common tetraploidy (LCT) event experienced in all legumes. Car: Cicer arietinum, Cca: Cajanus cajan, Gma: Glycine max, Gso: Glycine soja, Lja: Lotus japonicus, Mtr: Medicago truncatula, Pte: Psophocarpus tetragonolobus, Pvu: Phaseolus vulgaris, Tsu: Trifolium subterraneum, Vun: Vigna unguiculata. (c) Conserved syntenic blocks between winged bean and G. max. A block of 1000 Ns was added between super-scaffolds within same chromosome for chromosomal level visualisation. As featured in Ho et al., (2024).

Mature winged bean seeds contain between 15% to 21% lipid and 30-40% crude protein with a favourable amino acid balance and micronutrient composition (Adegboyega et al. 2019; National Academy of Science 1981; S. S. Kadam 1984). Seeds accumulate proteins in the cotyledons that provide the free amino acids and carbohydrates needed for germination (Figure 5. 2) (Duranti 2006).

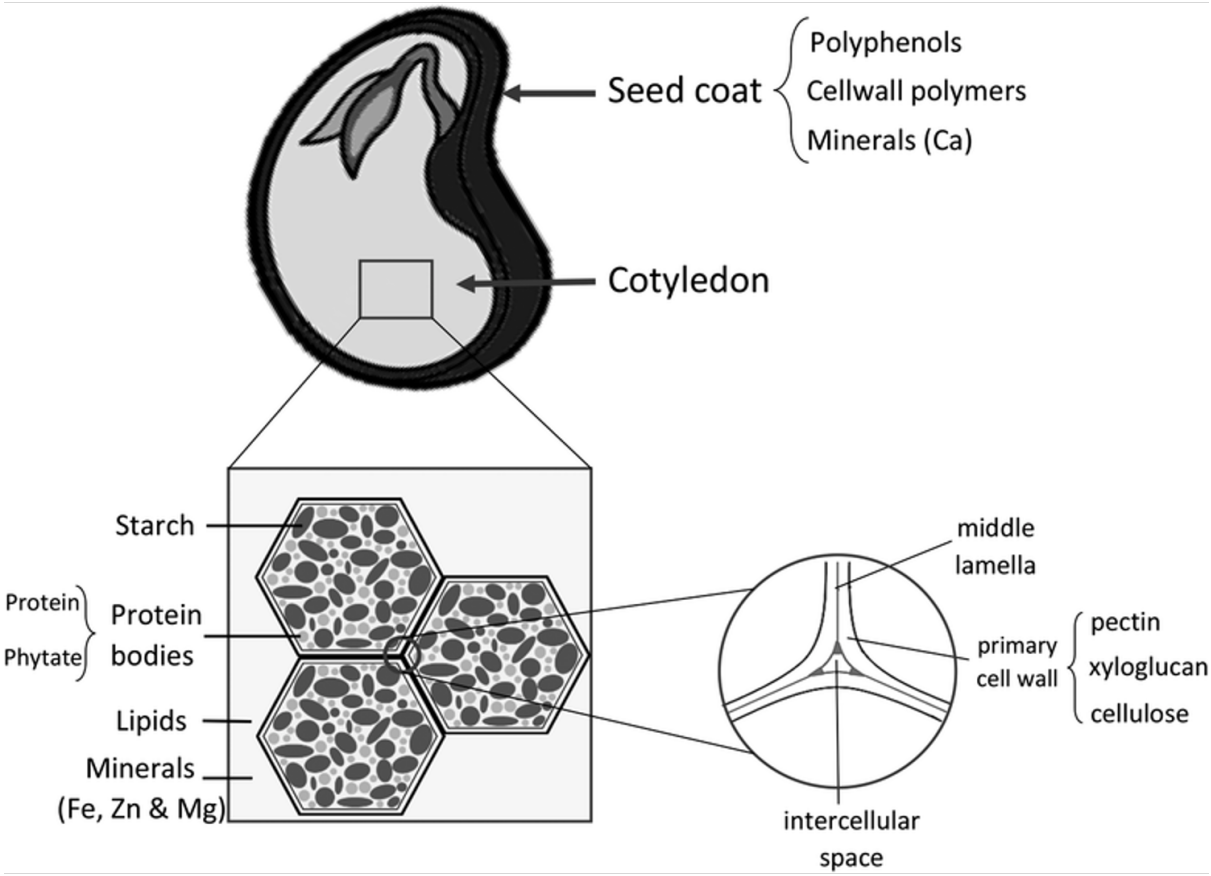


Figure 5. 2 Schematic illustration of the structural composition of bean. As featured in Wainaina et al., (2021), adapted from (Pallares et al. (2021).

The seed storage proteins (SSPs) are classified, based on their solubility, into globulins, albumins, prolamins, and glutelins. In winged bean, the SSPs are mainly composed of 7S and 2S globulins and 2S albumins, which are resolved into three fractions, psophocarpins A, B and C (Figure 5. 3). Psophocarpins A and C are seed storage globulins. Psophocarpin A is a single protein relatively high in sulphur-containing amino acids but it represents a small fraction of the storage proteins. In contrast, psophocarpin C is composed of 7S and 2S globulins and accounts for almost 60% of the seed protein (Blagrove and Gillespie 1978; Gillespie and

Blagrove 1978). Psophocarpin B is an albumin (WBA 1) and makes up about 30% of the seed protein, but antinutritional factors in the form of protein inhibitors make up 35% of psophocarpin B (Kortt 1979; 1980; 1986; 1984). It has also been reported that the fatty acid composition of winged bean seeds changes during maturation, with saturated fatty acids decreasing and mono- and polyunsaturated fatty acids increasing as the seeds mature (Mohanty et al. 2015).

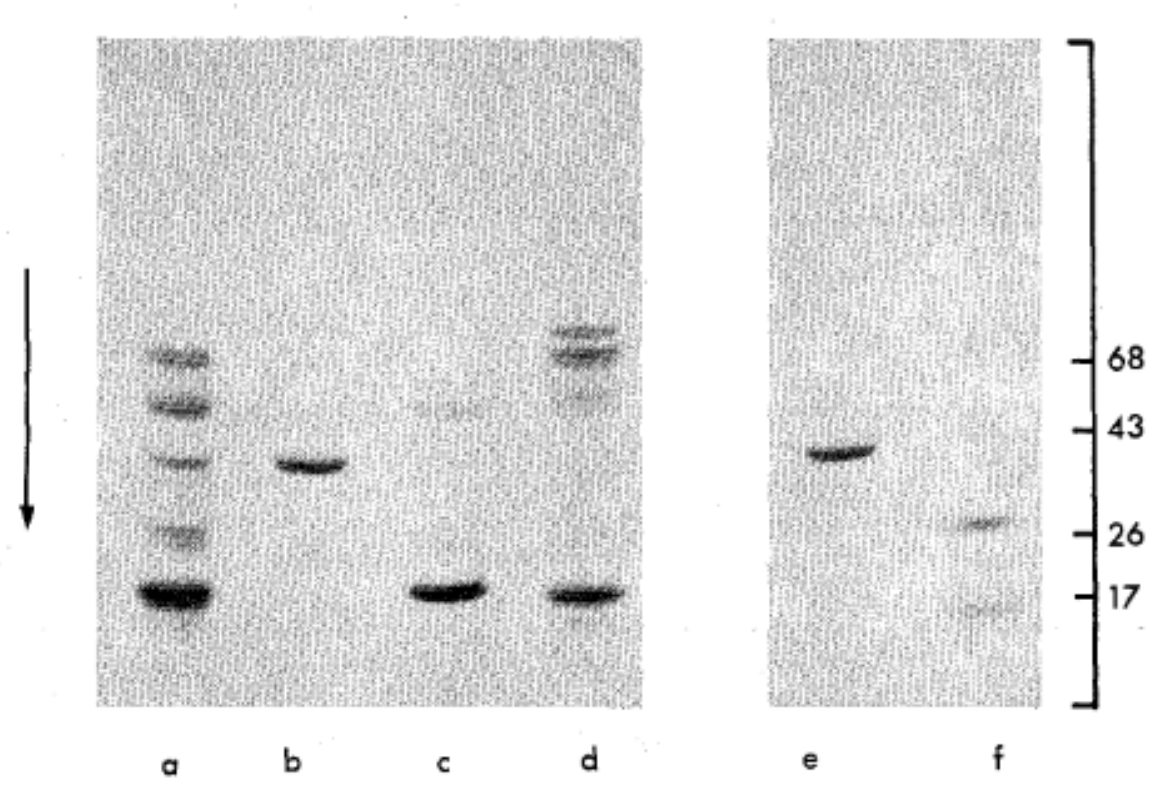


Figure 5. 3 Electrophoresis on SDS-polyacrylamide gradient gels of winged bean proteins. (a) Acetate-chloride extract. (b) Psophocarpin A. (c) Psophocarpin B. (d) Psophocarpin C. (e) and (f) Psophocarpin A before and after reduction with β-mercaptoethanol. Values at right of figure are 10^{-3} x molecular weight. As featured in Gillespie et al., (1978).

Depending on the winged bean genotype, the flower colour varies from near white to deep purple; and the seed colour can be cream, different shades of brown, deep purple and black (Figure 5. 4) (Tanzi, Eagleton, et al. 2019). Apart from its seeds, the immature pods are highly nutritious and more commonly consumed in salads, soups or directly (Sriwichai et al. 2021). The colour of the seeds and flowers has been correlated to the tannin content in the seeds with the darker colours having higher tannin content, as tannins are part of the flavonoid biosynthesis pathway (Klu, Jacobsen, and Van Harten 1997; Smulikowska et al. 2001).

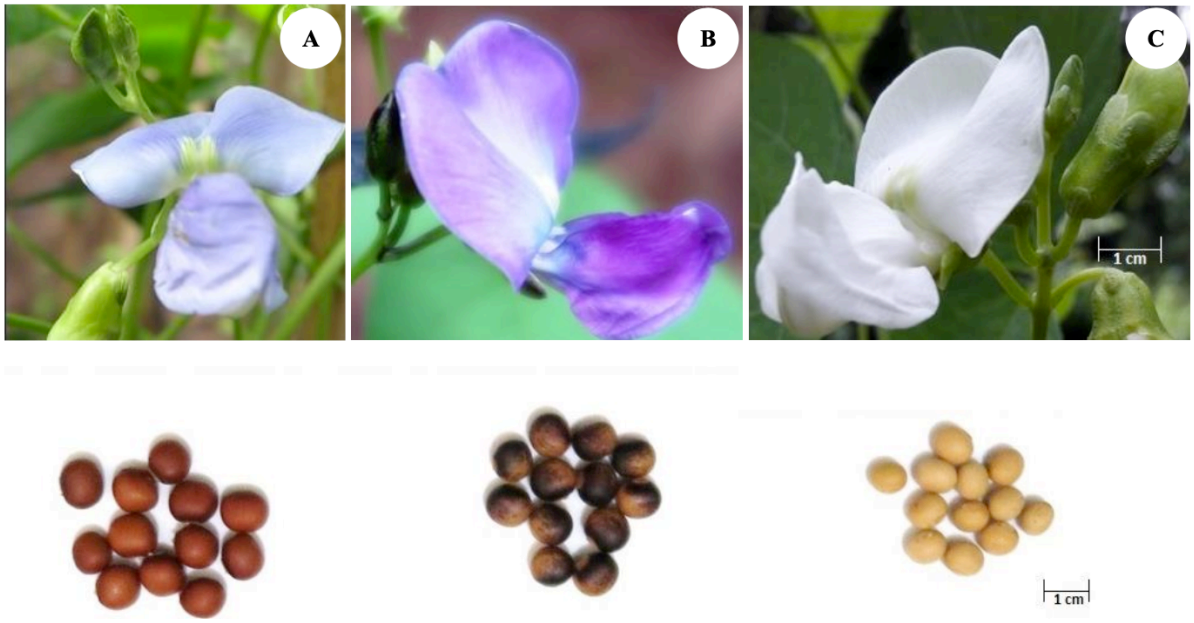


Figure 5. 4 Simple marker characteristics that distinguished the F1 winged bean hybrid plants from its parents: calyx colour, flower colour, colour of seed coat and seed size of (A) Bogor parent, (B) F1 hybrid, and (C) Singha parent. As featured in (Eagleton 2019).

As in many pulses, the highest tannin levels are found in the seed coat of the winged bean seeds. However, its removal is difficult and not commonly practised when cooking (De Lumen and Salamat 1980; Tan, Wong, and De Lumen 1984). Tannins are considered an anti-nutritional factor because they can bind and precipitate proteins, thereby reducing protein and amino acid digestibility in monogastrics (Smulikowska et al. 2001). Tannins are heat-resistant, making heat treatment an ineffective method for their removal. In developing countries, the main sources of protein in people's diets are pulses and beans, which have high tannin content. In these cases, tannin intake and its effect on protein digestibility could be considerable and further aggravate protein malnutrition. When tannin-free varieties of faba bean and lentil seed were compared to tannin-containing varieties, they showed higher in vitro digestibility. Tannin-free varieties underwent better proteolytic digestion, suggesting that an interaction between tannin and protein could lead to lower digestibility of seed proteins (Carbonaro, Virgili, and Carnovale 1996). Therefore, the selection of winged bean accessions with a lower amount of antinutritional factors like tannins in the seeds would be beneficial. The lower level of tannins in winged bean seeds might be linked to a distinguishable trait like the cream colour of the seeds (V. Singh et al. 2017). This would make the screening and selection process easier between cream colour winged bean seeds with

low condensed tannins and dark brown winged bean seeds with high tannin content (V. Singh et al. 2017). However, it is worth mentioning that the antinutritional factors, when consumed in small amounts can have a positive effect on human health. Studies have shown that protease inhibitors and isoflavones can act as anticancer agents (Messadi et al. 1986; DeClerck and Imren 1994; Gurfinkel and Rao 2003; Dong and Qin 2011; Sánchez-Chino et al. 2015).

In recent years, transcriptomic studies on non-model plants have contributed to identifying new functional genes, plant developmental pathways and secondary metabolite pathways, as well as responses to different biotic and abiotic stresses, providing useful information for plant breeders (Tyagi et al. 2022). In soybean seeds, sequencing-based transcriptome profiling (RNA-seq) was effective in identifying genes that play a key role in oil and protein accumulation (Songnan Yang et al. 2019; Qi et al. 2018). Studies on the transcriptomes of progressively developing soybean seeds at different developmental stages defined as days after flowering (DAF) have identified multiple genes and molecular processes. These have provided insights into the mechanisms of seed formation and nutrient accumulation (Ha et al. 2018; Chen et al. 2012; Jones and Vodkin 2013; Qi et al. 2018).

In jack beans, canavalin is the major seed storage protein (Gibbs, et al., 1989). Gibbs et al. (1989) found that an ancient duplication event is responsible for approximately 80% of its amino acid residues. Interestingly, the winged bean genome appears enriched for genes involved in isoflavonoid biosynthesis and secondary metabolites (such as phytosterol, castasterone and ergocalciferol), potentially functioning as phytoalexins through jasmonate- and ethylene-mediated pathogen defence mechanisms (Ho et al. 2024).

In winged bean, recent transcriptome and pathway analyses have revealed different levels of gene expression in leaves with high and low condensed tannin contents (V. Singh et al. 2017). This is the first study to identify transcriptomic variation at different pod and seed developmental stages in winged bean. The winged bean accession FP15 was chosen for this study due to its high protein (39%) and oil content (20%) (determined in Chapter 3) as well as purple flowers, pods and seeds. Focusing mainly on fatty acid biosynthesis, seed storage proteins and condensed tannins, this study aims to identify genes that play a key role in the

nutritional value of the seeds and immature pods of winged bean. These findings could be of great interest to plant breeders or in breeding programmes aiming to improve the nutritional quality of the winged bean pods and seeds.

5.2 Materials and Methods

5.2.1 RNA extraction and sequencing

Three winged bean plants (FP15) were grown in the growth chamber at 26°C/20°C day/night temperature with a 12-hour day length at the University of Nottingham Sutton Bonington Campus. Each plant was defined as a biological replicate. The pods and seeds of each plant were sampled on days 7, 15, 22, 30, 37 and 45 after flowering (DAF). As the plants were grown in a 12-hour day length, sampling was carried out at mid-day, around the 6th hour of their day. The samples were rapidly frozen in liquid nitrogen, and the seeds were separated from the pods before the RNA extraction. The photos below (Figure 5. 5) show size of pods and seeds at an estimated developmental stage on days after flowering.

Total RNA was extracted from the pods and seeds separately, using the RNeasy® Plant Mini Kit, following the manufacturer's instructions. Pod and seed samples from Days 15, 30, and 45 after flowering, and pod samples from Day 37 were sent for sequencing, due to cost restrictions. Days were selected at 15-day intervals starting from the day of flowering, considering that winged bean seeds typically reach maturity and are harvested around 60 days after flowering. Library preparation and transcriptomic sequencing were conducted by Novogene (Cambridge) using an Illumina NovaSeq6000 sequencer (Novogene, Cambridge). The 150-bp paired-end reads were mapped using HISAT2 v2.0.5 followed by differential quantification analysis at fragments per kilobase of transcript per million fragments mapped (FPKM) level using feature Counts 1.5.0-p3 and DESeq2 v1.20.0.

Figure 5. 5 Winged bean pods, seeds and buds of the FP15 accession. A. Early stage of pod and seed development, and flower buds on millimetre paper; B. Immature pod at a later developmental stage, approximately 30 (or even more) days after flowering. The photos were taken by Wai Kuan Ho, University of Nottingham Malaysia.

5.2.2 Protein profiling

Ground-dried winged bean seeds were passed through a 0.5 mm sieve (Ultra-Centrifugal Mill ZM 200, Retsch). The samples were stored at −80°C overnight and were then freeze-dried. The protein content was determined based on the modified Dumas method using an Elemental N Analyzer (FlashEA® 1112 N/Protein, Thermo Scientific) and the conversion factor 6.25.

5.2.3 LCMS analysis of condensed tannins

Condensed tannins were quantified in pods and seeds on days 7, 15, 22, 30, 37 and 45 after flowering, using a standard HPLC method (Calvert et al. 2024); and the following standards: catechin, epicatechin, procyanidins B1, B2 and C1. Condensed tannins were expressed as the total of these compounds. To correlate the colour changes of the pods to the concentration of the anthocyanins, anthocyanin standards cyanidin 3-O glucoside and pelargonidin 3-O glucoside were also used. However, the results for this analysis are not included in this study as the method requires further optimisation. Running HPLC-MS analyses in both positive and negative ion modes separately is an effective approach for accurately detecting anthocyanins

and condensed tannins. Anthocyanins ionise best in positive mode due to their flavylium cation structure, while condensed tannins, such as catechin and procyanidins, are more effectively detected in negative mode because their phenolic hydroxyl groups readily deprotonate. Using separate runs would allow for improved sensitivity and clearer fragmentation patterns for each compound class, although it requires additional injections and careful sample handling. Acidified mobile phases would support ionisation and compound stability, particularly for anthocyanins, which are pH-sensitive.

5.3 Results

5.3.1 Pod and seed development

When the flower fully opened, this marked Day 0 of flowering. As the pod was developing, its colour changed from green with purple specks to purple (Figure 5. 6a). In the winged bean seeds, the total protein content gradually increased during seed development from Day 15 to Day 45. In winged bean seeds, the protein content was 31% on Day 15, 34.5% on Day 30, and 37.5% on Day 45 (Figure 5. 6b).

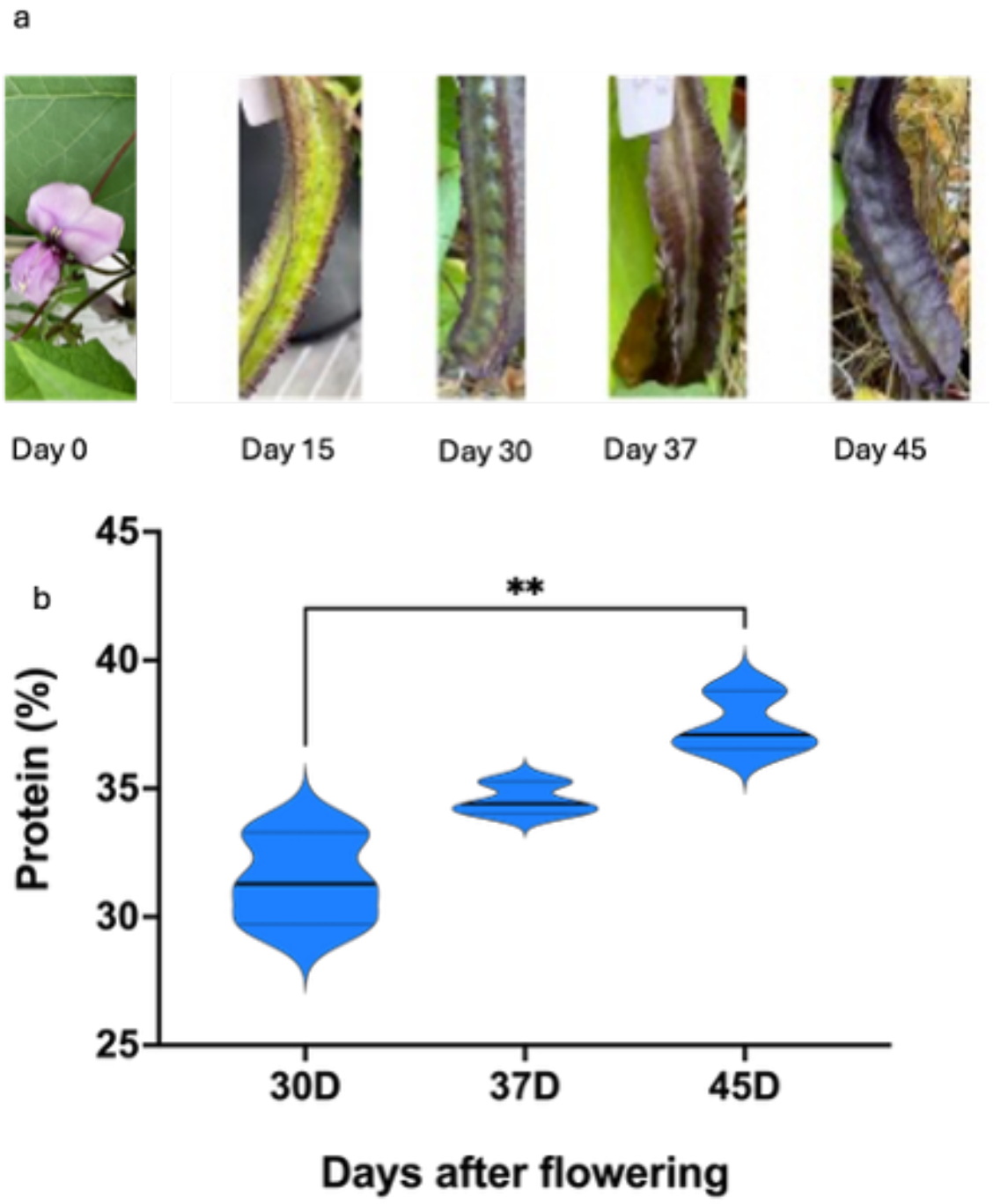


Figure 5. 6 a, Images of winged bean flower and pods at the different developmental stages; b, protein content (% dry weight) in winged bean seeds at different developmental stages. One-way ANOVA (p=0.0037) and Tukey's post-hoc test were used. Different letters above the bars represent significant differences detected. The standard error is shown by the error bars with n=3. (photos edited by Niki Tsoutsoura, University of Nottingham).

5.3.2 RNA Sequencing

The results from Illumina sequencing showed that the clean reads of the samples were above 98%, while the mapped reads ranged from 91% to 94.5% (Supplementary Table 5. 2) and the GC content ranged between 43.88-45.3% (Supplementary Table 5. 1). In the PCA plot (Figure 5. 7), pod and seed samples were separated into two distinct groups along the negative and positive X-axis of PC1, which explained 47.83% of the variation. In PC2, the samples were distributed progressively from negative to positive on the Y-axis following their developmental stage Day15, Day30, Day37 (for the pod only) and Day 45. Overall, the samples within each group were clustered together, indicating that the three biological replicates were at similar developmental stages when collected. However, some samples were wider spread out, such as the seed sample on Day 30 (SFP15_2_30) and the pod samples on Day 30 and Day 37 in replicates 1 and 2, which were closely grouped (Figure 5. 7).

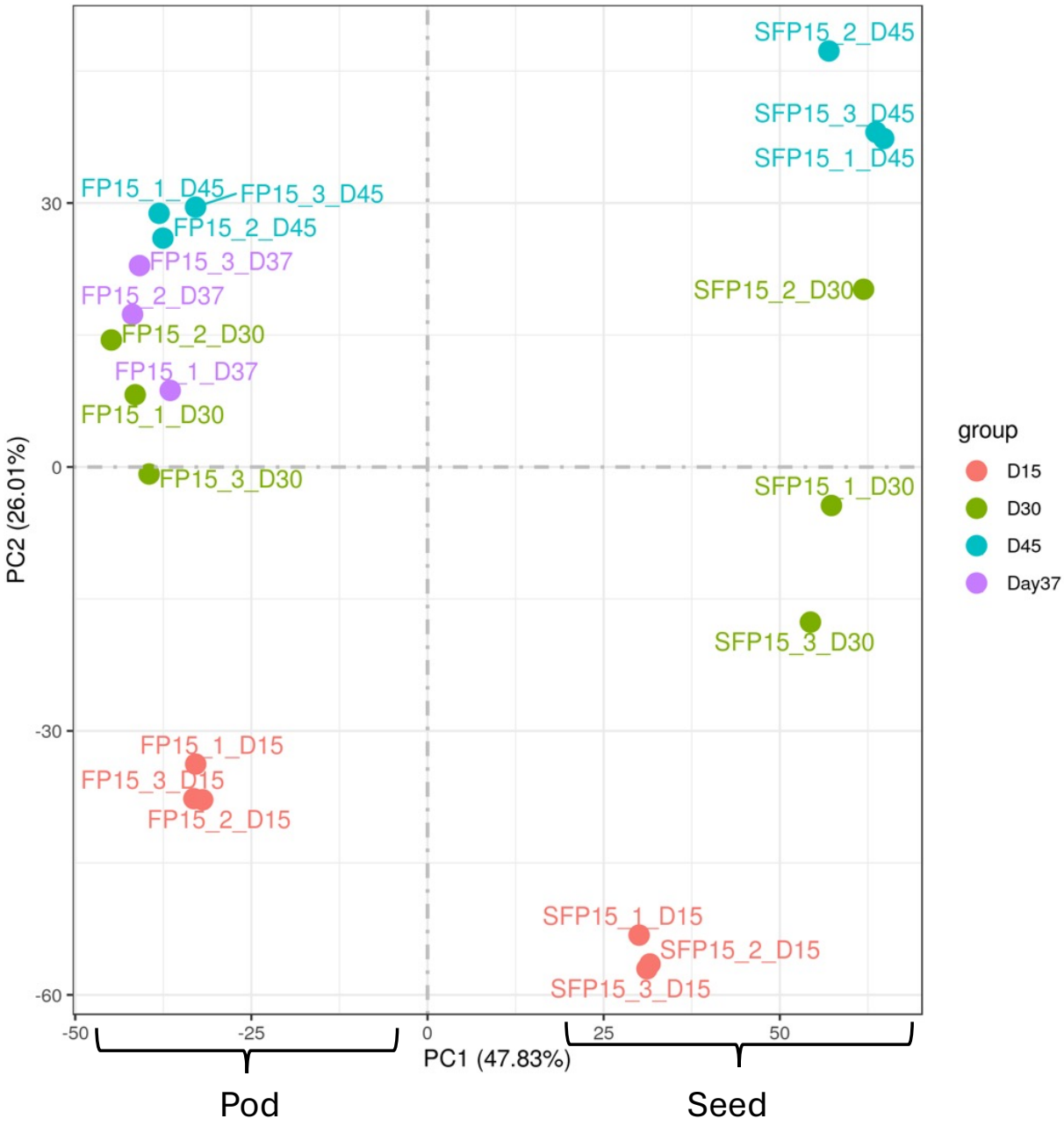


Figure 5. 7 Principal Component Analysis (PCA) shows the correlation among the samples based on gene expression data. The samples have formed clusters based on the stage of development as shown with the different colours as well as the plant material either pod (FP15) or seed (SFP15) on the negative and positive side of the x-axis PC1.

5.3.3 Differentially expressed genes (DEGs)

5.3.3.1 Differences in the gene expression profile

The transcriptome analysis aimed to reveal gene expression changes involved in developing winged bean pods and seeds. Comparative transcriptome analysis was conducted on pod and seed samples collected on Days 15, 30, (37 for pods), and 45. The analysis identified 16,847 commonly expressed genes in the seeds; and 18,047 in the pods. Interestingly, in seeds, over 1,000 genes were expressed at the early developmental stage on Day 15, while only 326 genes were uniquely expressed on Day 30. By Day 45, nearly 500 genes were uniquely expressed (Figure 5. 8a). In pods, 1,107 genes were uniquely expressed on Day 15 (Figure 5. 8b).

To identify differences in the gene expression profile as the pod and seed mature, Day 15 was compared to the later developmental stages. The number of DEGs (including up- and down-regulated) for each comparison combination is shown in a histogram in Figure 5. 9. In both pods and seeds, the number of up-regulated genes was higher than the down-regulated genes, suggesting that these up-regulated DEGs could be responsible for pod and seed maturation. When comparing Day45vs15, 4583 genes were upregulated in the pods and 6054 were upregulated in the seeds. In both pods and seeds, these were the highest numbers of upregulated genes observed in this study. Equally, the downregulated genes were 3371 in the pod and 4711 in the seed on Day45vs15 (Figure 5. 9). Therefore, the Day45vs15 comparison was selected for further investigation to identify genes involved in the fatty acid biosynthesis pathway, the protein accumulation in seeds as well as the flavonoid pathway.

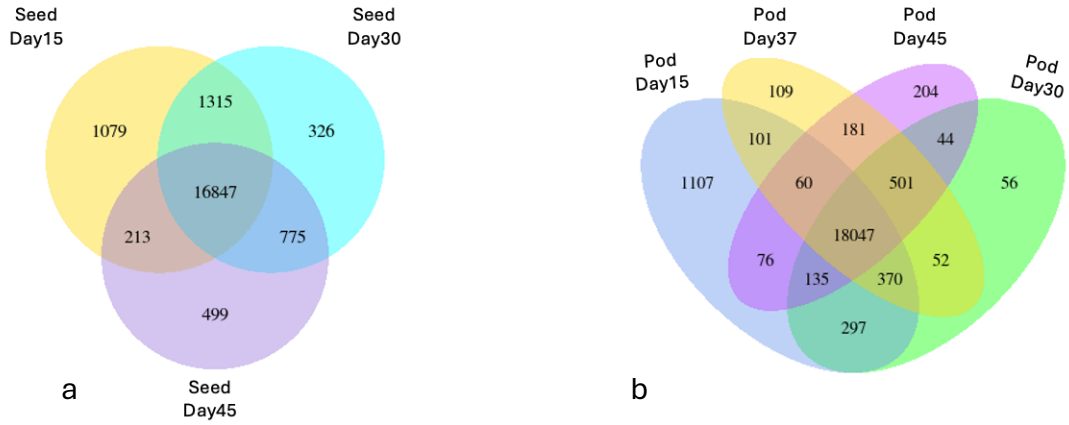


Figure 5. 8 Venn diagrams showing the number of genes that overlap within each group. The overlapping regions show the number of genes co-expressed in two or more groups. a, seed samples forming three groups (Day15, Day30 and Day45); b, pod samples forming four groups (Day15, Day30 and Day45).

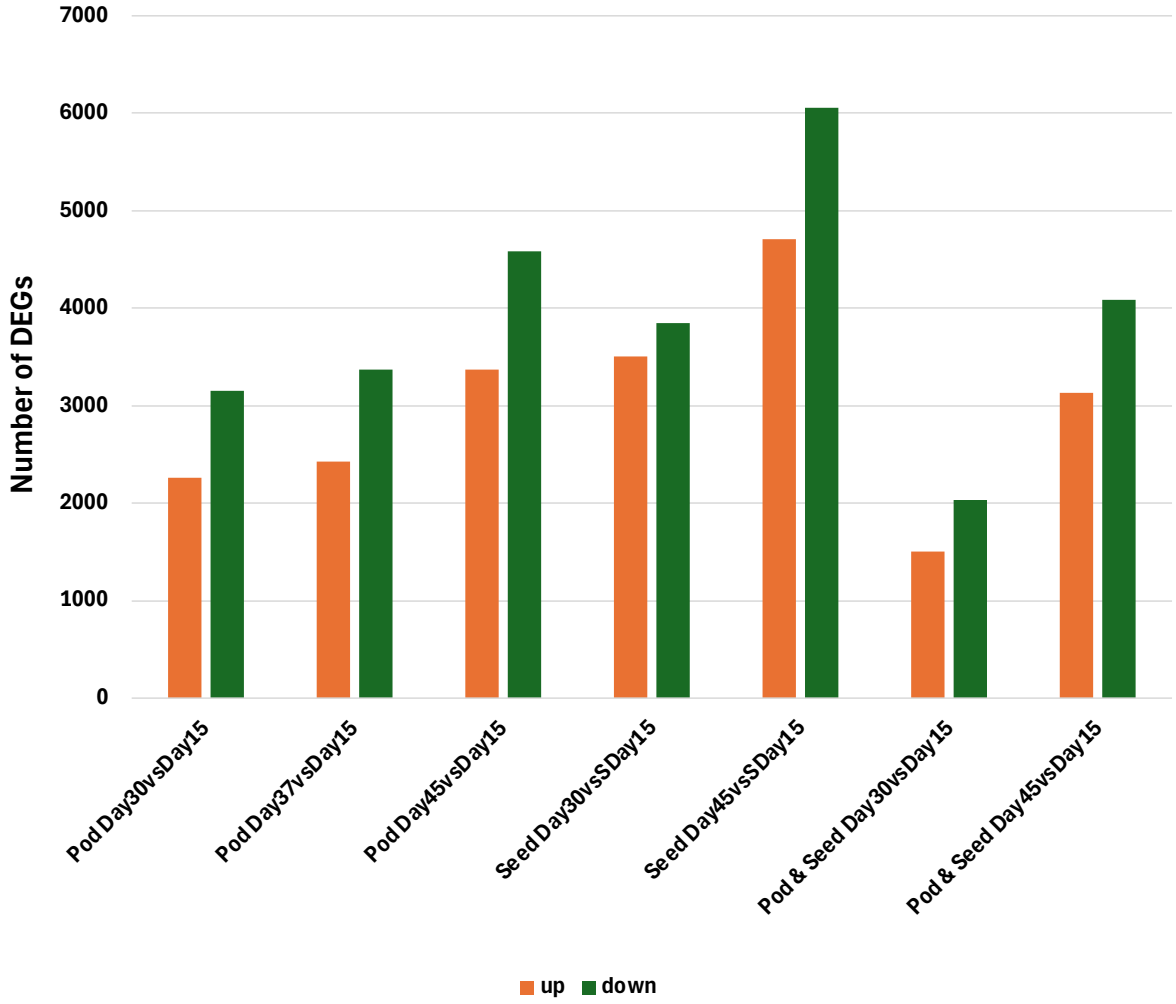


Figure 5. 9 Number of differentially expressed genes (DEG) in winged bean seeds and pods at different developmental stages. For each comparison, the combination is shown on the X-axis. DEGs were determined using the threshold of DESeq2 padj<=0.05 |log2FoldChange|>=1.0.

5.3.3.2 Seed storage protein-related genes in developing pod tissues

When investigating genes related to seed storage proteins, 88 winged bean genes were identified (Figure 5. 10, Supplementary Table 5. 7). The winged bean gene Psote03G0382500.1 coding for the 11S globulin seed storage protein 2 was expressed at similar levels during the maturation period in the pods. However, in the seeds, it was significantly downregulated during seed maturation, with Day 15 exhibiting the highest expression level. The albumin-2 genes, Psote04G0285600.1 and Psote04G0288400.1, were down-regulated in the seeds during the maturation period. In contrast, the gene Psote04G0281900.1 coding for an albumin seed storage protein, was upregulated in pods and seeds, with Day45 having a 12-fold increase. The Psote07G0210200.1 coding for DELLA1 protein (growth regulator protein) was constantly expressed in pod and seed at a similar level across the developmental stages. The winged bean gene Psote04G0062600.1, homologous (87%) to the soybean gene Glyma.20G147600 late embryogenesis abundant (*LEA*) (Figure 5. 10) was upregulated in both the pod and seed during the maturation process, with a more significant increase observed in the seeds (Supplementary Table 5. 3, Figure 5. 10).

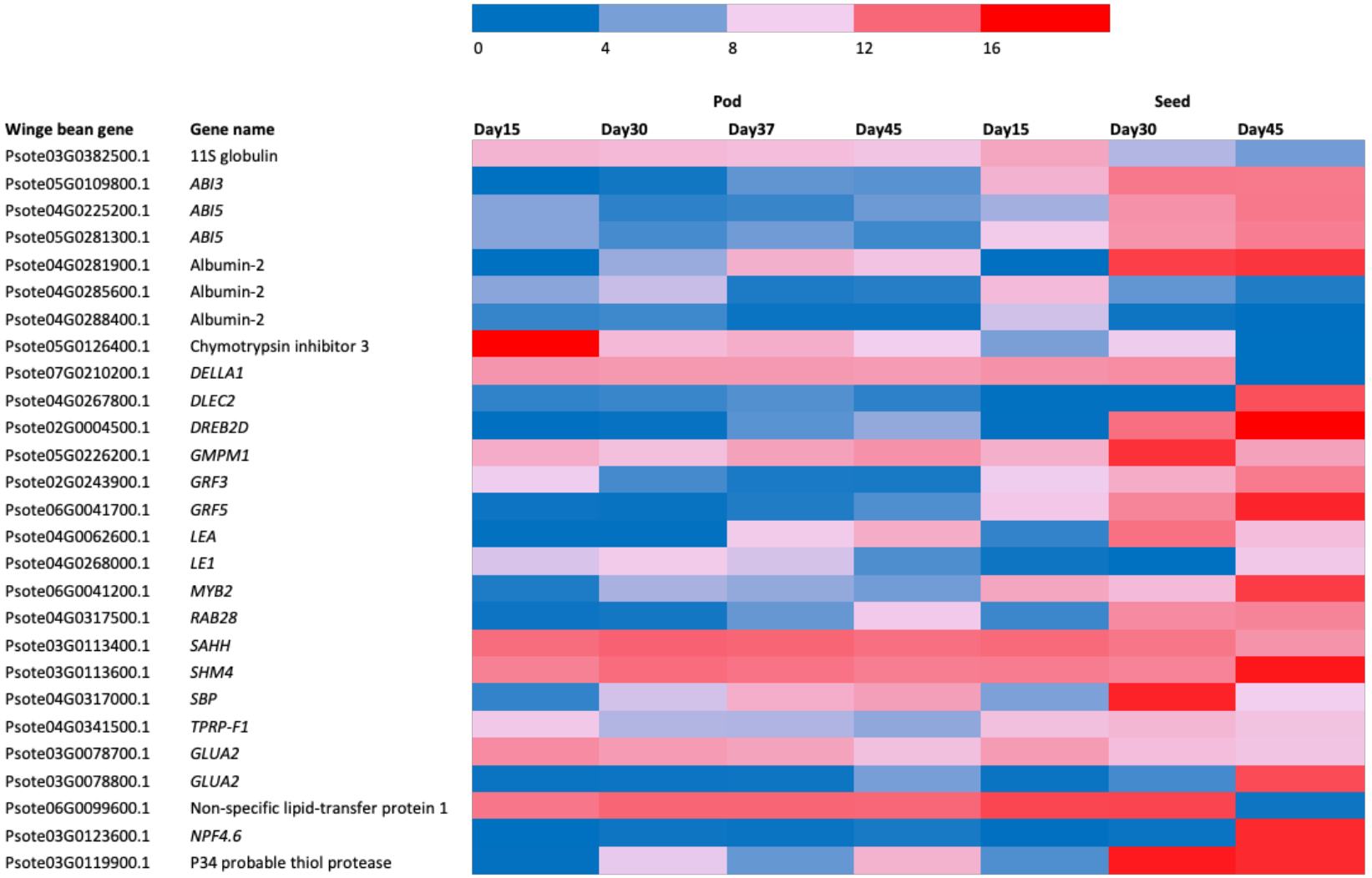


Figure 5. 10 Expression profile of seed storage protein-related genes in developing pod and seed tissues. Expression levels [log2 (FPKM+1)] of albumin, globulin, and related genes in the pod and seed tissues on Day 15, 30, 37 and 45 from left to right. The names shown are based on the gene similarity to other crops during the annotation process.

5.3.3.3 KEGG pathway analysis: Fatty acid biosynthesis and biosynthesis of unsaturated fatty acids

The KEGG enrichment analysis revealed that several genes in the fatty acid biosynthesis (Figure 5. 11) and biosynthesis of the unsaturated fatty acids (Figure 5. 12) pathways were differentially expressed. Furthermore, 42 genes were associated with fatty acid biosynthesis (Supplementary Table 5. 6) and 29 with the biosynthesis of unsaturated fatty acids (UFAs) (Figure 5. 14).

Using the soybean KEGG fatty acid biosynthesis pathway, homologous winged bean genes were identified (Figure 5. 13). The winged bean genes located within the oil or fatty acids QTLs from the previous chapter were highlighted in bold letters. Overall, in the winged bean seeds, the genes involved in the fatty acid biosynthesis pathway were upregulated on Day30; whereas on Day45 their expression levels were lower compared to Day30, and in some cases even compared to Day15, suggesting an increase in the synthesis of saturated fatty acids such as palmitic and stearic at the early stages**.**

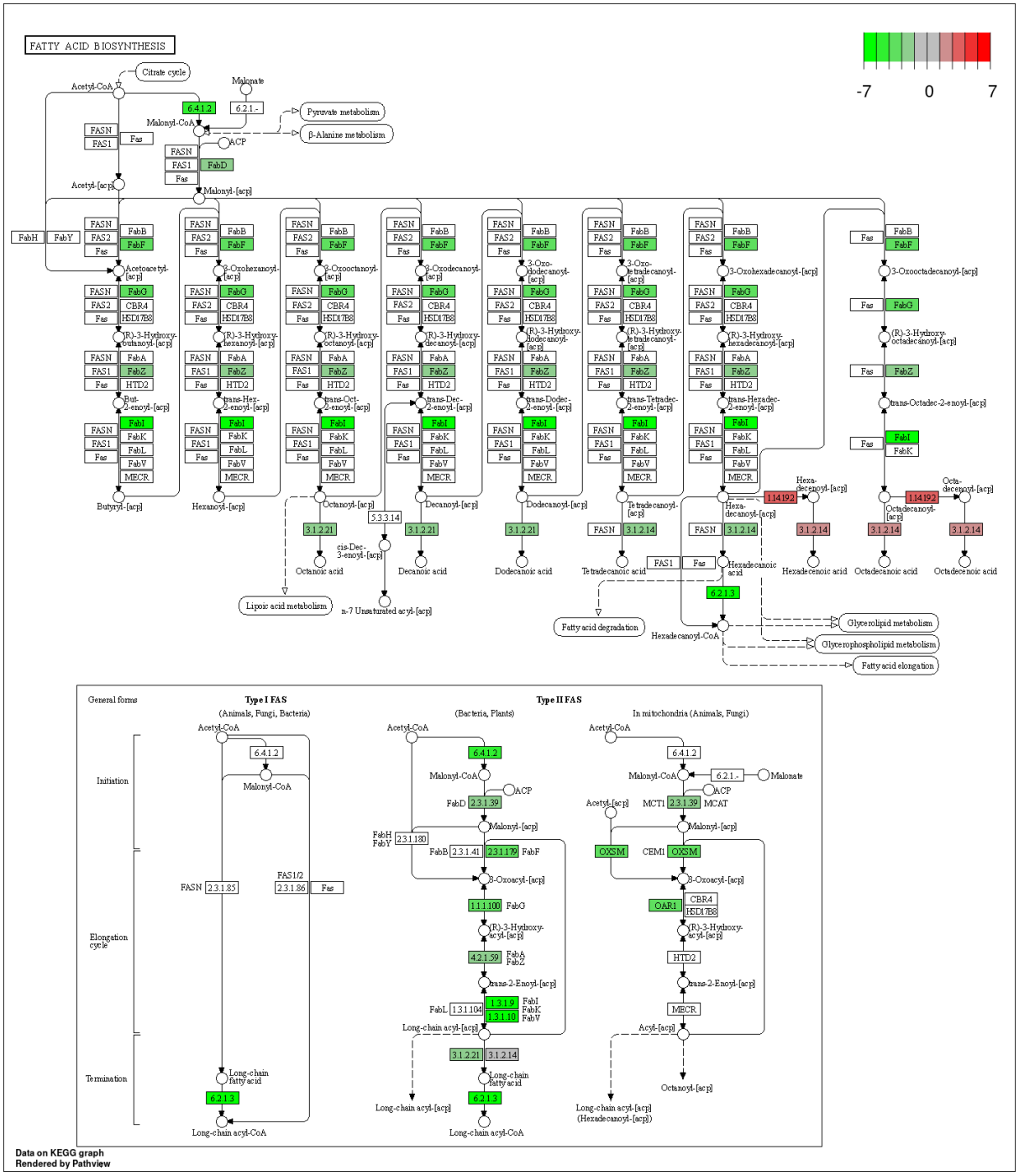


Figure 5. 11 KEGG Fatty acid biosynthesis pathway (var00061). Differential gene expression in winged bean seeds comparing Day45vs15.

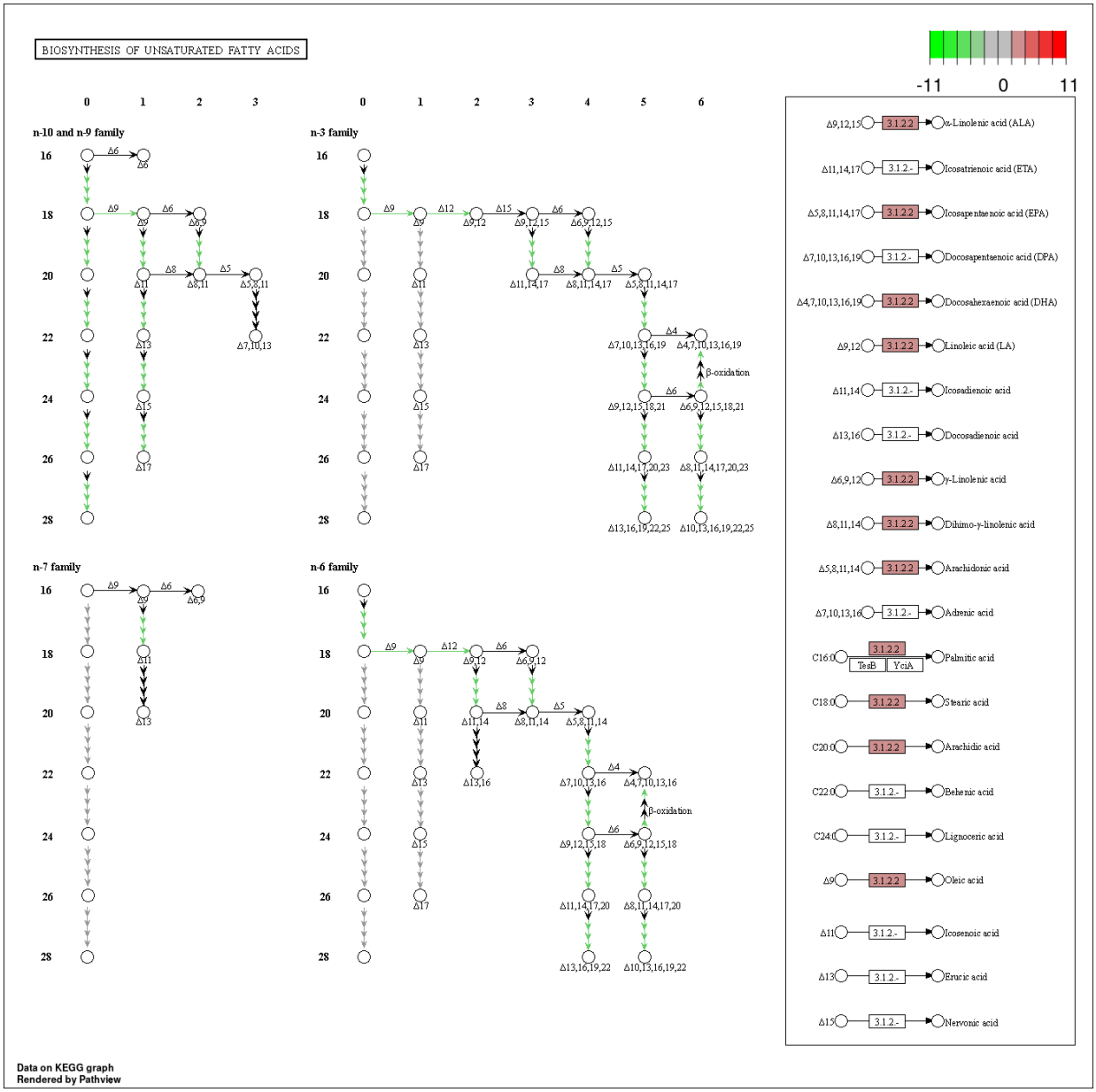


Figure 5. 12 KEGG biosynthesis of unsaturated fatty acids (var01040). Differential gene expression in winged bean seeds comparing Day45vs15.

As the seeds matured to Day45, the winged bean genes involved in the last steps of the fatty acid biosynthesis and the biosynthesis of unsaturated fatty acids pathways were upregulated. The biosynthesis of unsaturated fatty acids involves the elongation of palmitic and stearic acids, and their desaturation to mono-unsaturated (MUFA) and poly-unsaturated fatty acids (PUFA) suggesting the conversion of the saturated fatty acids to MUFA and PUFA as the seeds matured, with similar results been reported in the literature (C. S. Mohanty et al. 2015b) Psote06G0283600.1 (*FAT1*) (palmitoyl-acyl carrier protein thioesterase), Psote04G0339800.1 (*LACS8*) long chain acyl-CoA synthesis (LACS8 Long-chain acyl-CoA synthetase 8), Psote03G0166900.1 (*ACPD*) (ACPD Stearoyl-[acyl-carrier-protein] 9-desaturase) and Psote08G0128000.1 (*S-ACP-DES6*) (S-ACP-DES6 Stearoyl-[acyl-carrier-protein] 9-desaturase).

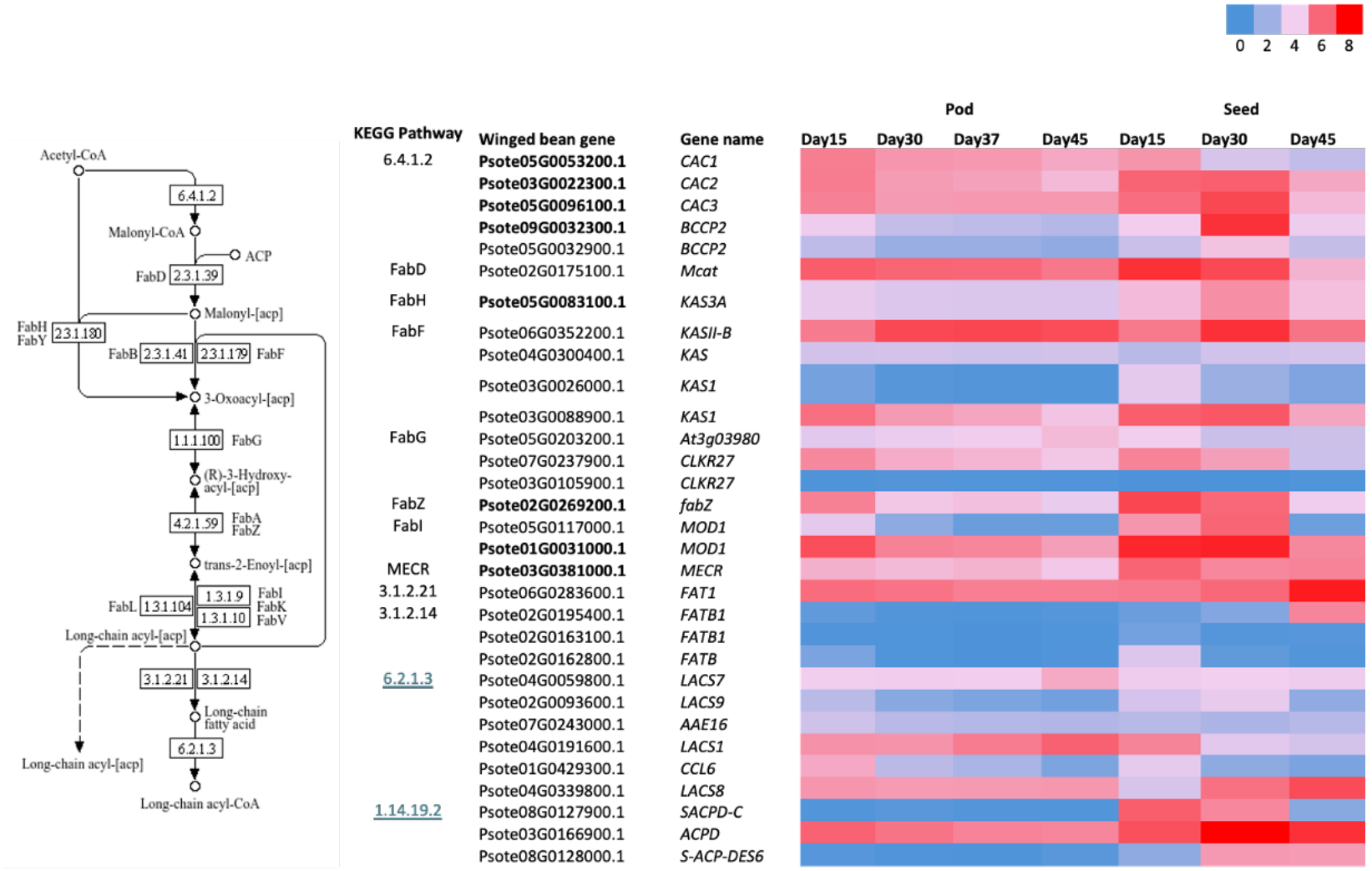


Figure 5. 13 Heatmap of genes in the fatty acid biosynthesis pathway. Transcriptomic profiles of fatty acid biosynthesis genes at Day 15, 30, 37 and 45 (from left to right) in the maturing pods and seeds, with the colour scale reflecting log2(FPKM+1) values. Winged bean genes identified in the oil and fatty acids QTLs in the previous chapter, were highlighted with bold letters.

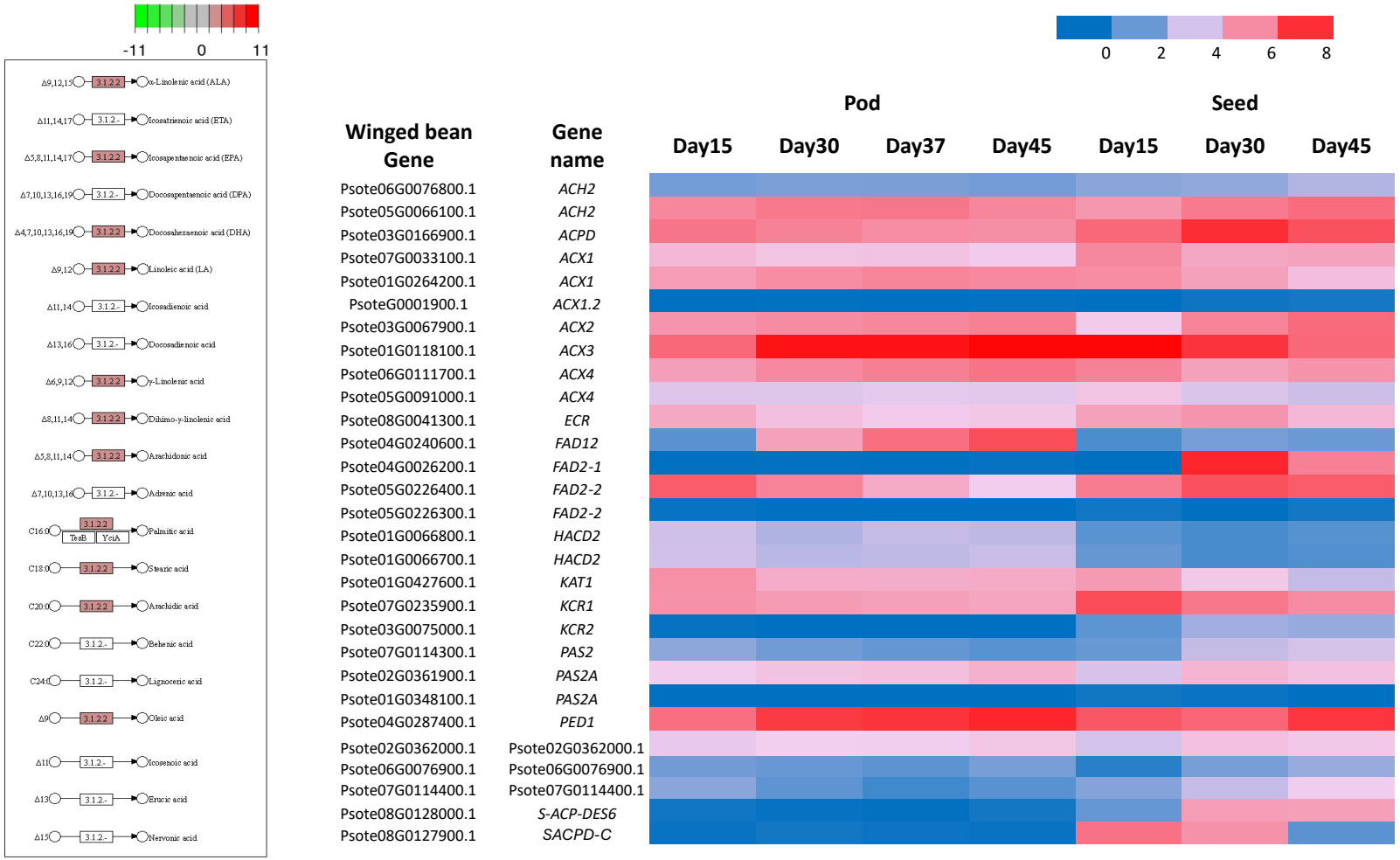


Figure 5. 14 Heatmap of genes involved in the biosynthesis of unsaturated fatty acids. Transcriptomic profiles of biosynthesis of unsaturated fatty acid genes at Day 15, 30, 37 and 45 (from left to right) in the maturing pods and seeds, with the colour scale reflecting log2(FPKM+1) values.

5.3.3.4 KEGG pathway analysis: Flavonoid biosynthesis

Differential gene expression was observed in the flavonoid pathway when comparing the pods from Day 45 to Day 15. (Figure 5. 15). The pod colour changes from green to purple (Figure 5. 6) therefore it is hypothesised that the genes related to the anthocyanins would be upregulated. Analysis of gene expression of enzymes involved in the synthesis of quercetin (FLS- flavonol synthase), cyanidin (ANT17- leucoanthocyanidin dioxygenase), and epicatechin (ANR1- anthocyanidin reductase 1) revealed differential gene expression in pods and seeds at the developmental stages.

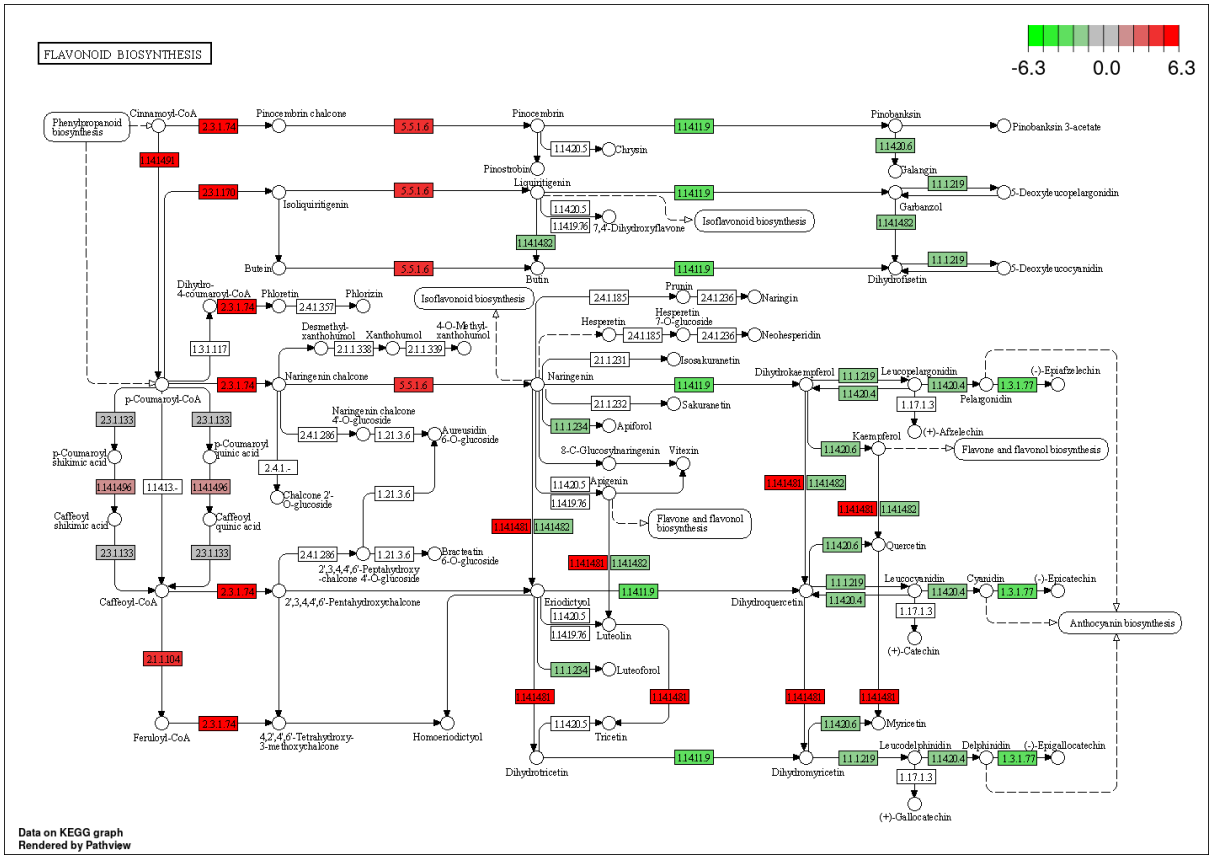


Figure 5. 15 KEGG Flavonoid Biosynthesis pathway. The differential gene expression in pods comparing Day45vs15 is shown with different shades of green and red for down and upregulated genes, respectively. 1.1.1.1219-DFR: bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase-like; 1.1420.4-ANT17: leucoanthocyanidin dioxygenase, 1.3.1.77-ANR1: anthocyanidin reductase 1; 1.17.1.3-LAR1: leucoanthocyanidin reductase 1; 1.14.20.6-FLS: flavonol synthase; 1.14.14.81-CYP75A3: flavonoid 3', 5'-hydroxylase; 1.14.14.82-SF3'H1: flavonoid 3'-monooxygenas

Focusing on the segment of the flavonoid pathway responsible for producing these compounds (Figure 5. 16), the genes involved in anthocyanin synthesis were highly expressed at the early stages, with their expression progressively decreasing in both pods and seeds. Winged bean gene Psote01G0255900.1 (*ANT17*) coding for leucoanthocyanidin dioxygenase, an anthocyanin synthetase for cyanidin, pelargonidin and delphinidin, was continuously expressed; however, the level of expression was declining as the pod and seed developed over time. The genes Psote03G0143100.1 homologous to Glyma.08G062000 coding for anthocyanidin reductase (ANR1) (Lu et al. 2021) and Psote01G0405800.1 homologous to Glyma.10G204800 coding for leucoanthocyanidin reductase (LAR1) (Lu et al. 2021) were expressed in low levels in pods and highly expressed on Day15 in the seeds, with the expression levels of the ANR1 dropping by Day45 in seeds. This would imply a decrease in the anthocyanins synthesis in the seed coat which could be explained by the growth of the seed coat cells as the seeds reach maturity as well as the expression of other transcription factors (TF) that play an important role in the anthocyanins accumulation such as *MYB113*. Although in the seeds, the *LAR1* expression was lower on Days 30 and 45 compared to Day 15, it remained relatively high throughout these stages. *LAR1* expression relates to the formation of flavonols such as catechin, and the *LAR1* gene expression levels were reflected in the catechin content in the seeds. Catechin levels were higher on Day15 and decreased on Days 30 and 45 in the seeds (Table 5. 1). It is worth keeping in mind that the seeds increase in size during the development progresses. Compounds such as condensed tannins (including catechin) and anthocyanins are mainly stored in the seed coat. During seed development, as the cells grow, the seed coat surface increases along with the endosperm but at a different proportion and rate. Lower levels of compounds expressed as a proportion of dry seed weight are associated with changes in seed shape. Similar results of the total flavonoid content (rutin, quercetin and kaempferol) decreasing during seed development have been reported in buckwheat seed (Song et al. 2016).

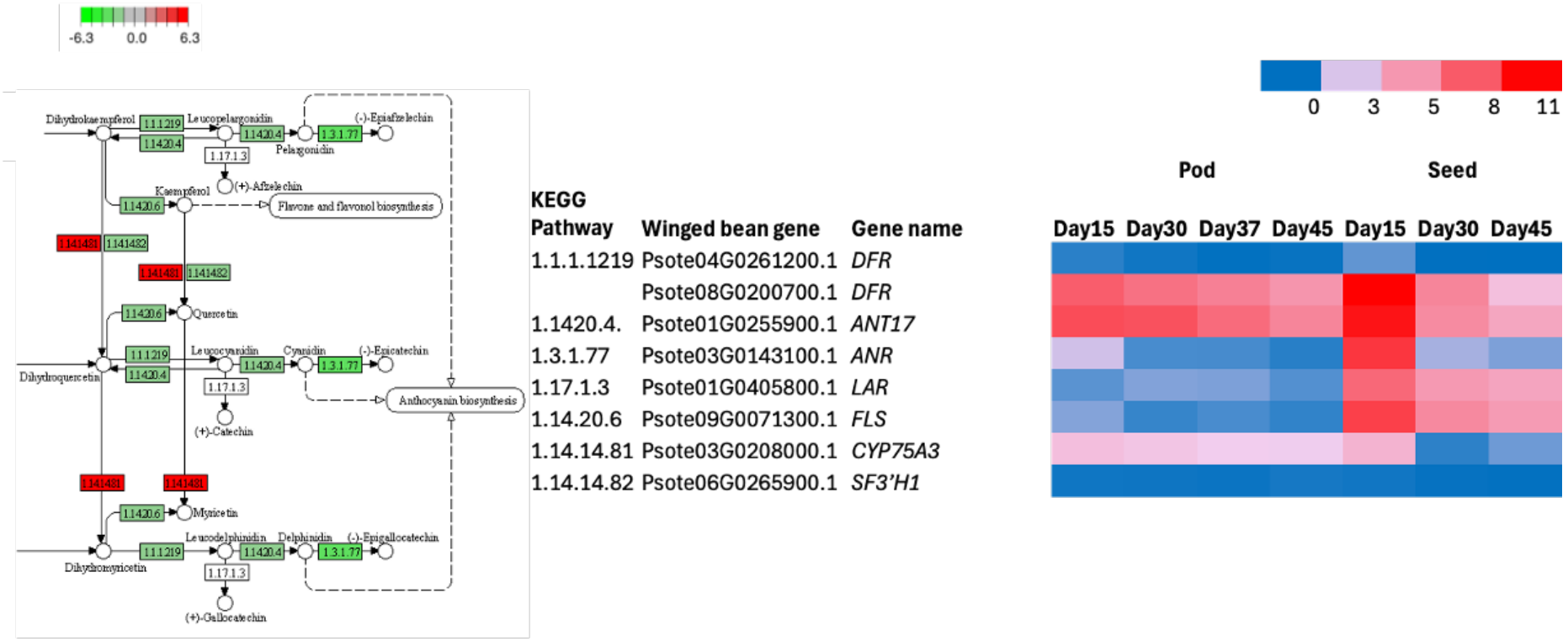
A step before the synthesis of catechins and cyanidin is the expression of flavonol synthase (*FLS*) and dihydroflavonol-4-reductase (*DFR*) genes. It has been reported that the overexpression of *FLS* in tobacco plants promoted flavonol biosynthesis and inhibited the accumulation of anthocyanins, resulting in white flowers. In contrast, the overexpression of *DFR* genes downregulated the endogenous expression of *NtFLS* in in tobacco plants and

promoted anthocyanin biosynthesis (P. Luo et al. 2016). Luo et al. (2016) studied the red and white flowers of different species (*Rosa rugosa* and *Rosa multiflora*, peach (*Prunus persica*), carnation (*Dianthus caryophyllus*), azalea (*Rhododendron simsii*), camellia (*Camellia japonica*), and petunia (*Petunia hybrida*) and noted that in white flowers the *FLS* expression levels were higher, while in red flowers the *DFR* expression levels were significantly higher. As both *FLS* and *DFR* use the same substrates, dihydroflavonols, for the biosynthesis of flavonols and anthocyanin, respectively; it was proposed that the competition between *FLS* and *DFR* genes is crucial for the formation of flower colour. In the winged bean seeds, the expression levels of the gene Psote08G0200700.1 homologous to Glyma.17G252200 coding for *DFR* was highly expressed in both pods and seeds with the higher expression levels on Day15 for both pods and seeds. In contrast, *FLS* was not expressed in pods during development, suggesting the promotion of anthocyanin biosynthesis. However, *FLS* and *DFR* were both expressed in seeds. The *FLS* expression levels were high during the seed development, with Day15 showing the highest level of expression (Figure 5. 16).

Overall, it appears that at the earlier stage of pod development (Day15) the genes related to flavonoid biosynthesis were highly expressed in seeds on Day15 compared to Day45 and this was reflected in the phenolic compounds data where their content mg/g of dried seeds weight decreased from Day15 to Day45 (Table 5. 1). The progressive reduction in gene expression and decrease in phenolic compounds could be attributed to developmental maturation and metabolic shifts within the plant tissues.

Table 5. 1 Phenolic compounds mg/ 100 mg of plant material. This is an estimate as the data has not been normalised to the internal standards due to the need for further method optimisation. The amount of compounds in winged bean pod and seed at different developmental stages in the table ± SEM (n=3), were analysed using a one-way ANOVA and Tukey's posthoc test, p-values are shown. The different letters represent significant differences among the winged bean accessions.

Plant material	Day	Catechin	Epi-catechin	Procyanidin B1	Procyanidin B2	Procyanidin C1	TOTAL
POD	**15D**	0.0022±0.0004a	0.002±0.0005a	0.0016±0.0002a	0.0022±0.0002a	0.0021±0.0003a	0.0101±0.002a
POD	**22D**	0.00067±0.015ab	0.0084±0.0032a	0.0018±0.0001a	0.0036±0.0002a	0.0035±0.0003a	0.0241±0.004a
POD	**30D**	0.0119±0.0048ab	0.0102±0.0058a	0.0017±0.0001a	0.0025±0.0001a	0.0022±0.00002a	0.0285±0.011a
POD	**37D**	0.0049±0.0016ab	0.0039±0.001a	0.0019±0.00004a	0.0026±0.0001a	0.0025±0.0001a	0.0157±0.001a
POD	**45D**	0.0074±0.002ab	0.0097±0.0055a	0.0019±0.0001a	0.003±0.0002a	0.0027±0.0002a	0.0246±0.004a
SEED	**15D**	0.0639±0.0047d	0.3559±0.275b	0.013±0.0019c	0.1051±0.0056d	0.1104±0.0047c	0.6483±0.044d
SEED	**22D**	0.0626±0.0113d	0.2972±0.0302b	0.0124±0.0026c	0.0897±0.0175cd	0.091±0.0208c	0.5529±0.081cd
SEED	**30D**	0.0494±0.0117cd	0.3225±0.0111b	0.0083±0.002bc	0.0602±0.0125bc	0.0511±0.0111b	0.4915±0.048bcd
SEED	**37D**	0.0324±0.0019bc	0.2973±0.0117b	0.0049±0.0004ab	0.0301±0.0033ab	0.0242±0.0022ab	0.3889±0.016bc
SEED	**45D**	0.025±0.0038abc	0.2984±0.0134b	0.0045±0.0001ab	0.0285±0.0007ab	0.0247±0.0008ab	0.3811±0.018b
p-Value							
Plant material		<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Days after flowering		0.010	0.300	0.005	<0.001	<0.001	0.005
Interaction		0.005	0.177	0.003	<0.001	<0.001	0.003



KEGG Pathway	Winged bean gene	Gene name
1.1.1.1219	Psote04G0261200.1	*DFR*
	Psote08G0200700.1	*DFR*
1.1420.4.	Psote01G0255900.1	*ANT17*
1.3.1.77	Psote03G0143100.1	*ANR*
1.17.1.3	Psote01G0405800.1	*LAR*
1.14.20.6	Psote09G0071300.1	*FLS*
1.14.14.81	Psote03G0208000.1	*CYP75A3*
1.14.14.82	Psote06G0265900.1	*SF3'H1*

Figure 5. 16 Heatmap of genes involved in the flavonoid biosynthesis pathway in developing pods and seeds. Transcriptomic profiles of flavonoid biosynthesis genes at Day 15, 30, 37 and 45 (from left to right) in the maturing purple pods and seeds, with the colour scale reflecting log2(FPKM+1) values. DFR: bifunctional dihydroflavonol 4-reductase/flavanone 4-reductase-like; ANT17: leucoanthocyanidin dioxygenase, ANS: anthocyanidin synthase, ANR1: anthocyanidin reductase 1; LAR1: leucoanthocyanidin reductase 1; FLS: flavonol synthase; CYP75A3: flavonoid 3', 5'-hydroxylase; SF3'H1: flavonoid 3'-monooxygenas

5.4 Discussion

From this study, parts of the results on the seed storage proteins and anthocyanins, have been published by Ho et al. 2024.

This is the first transcriptomic analysis of developing pods and seeds in winged bean, revealing differential expression of genes involved in key pathways such as seed storage proteins, fatty acid biosynthesis, and flavonoid biosynthesis. These findings provide a foundation for future research aimed at understanding the key genomic and transcriptomic mechanisms that enhance the nutritional value of winged bean seeds.

5.4.1 Seed Storage Proteins (SSP)

It has been reported that in winged bean seeds, protein and lipid bodies were not present on Day30 after flowering but accumulated at a later stage on Day45 (Saio, Nakano, and Uemoto 1983). In this study, the protein content in winged bean seeds was 34.5% on Day37 which was similar to the amount of protein in the mature seeds of the winged bean accession Ma3 (paternal accession from the cross mentioned in the previous chapter) (Figure 5. 6b). On Day45, the protein content in FP15 was increased to 37.5% which is higher than the protein of other mature winged bean accessions A6, I51 and I10 while the protein content in the mature FP15 seeds was 39% (determined in Chapter 3). A better understanding of the accumulation of seed storage proteins in winged bean seeds during pod and seed development will provide useful information for genetic improvement programs aiming to improve the protein quality of the seeds.

The previous chapters reported the amino acid composition of the seed storage proteins in winged bean seeds, comparable to soybean, as well as QTLs related to protein, oil and fatty acid contents. The winged bean gene, Psote04G0062600.1, homologous to the soybean late embryogenesis abundant (*LEA*) gene Glyma.20G147600 was upregulated in the pod and seed during the maturation process with a more significant increase observed in the seeds (Figure 5. 10). Similar results have been reported in soybean (Guo et al. 2023). In addition, Psote07G0210200.1 located in the qProt-7-1 coding for DELLA1 protein, was highly expressed

in the pods and seeds throughout the developmental stages. The homologous gene, AT5G17490 in Arabidopsis, codes for the RGL3 protein, which acts as a coactivator to ABI3 promoting seed storage protein accumulation during pod and seed development (Hu et al. 2021; Gomez et al. 2023). These results align with recent literature, indicating a clear progression in the developmental process.

In this study, 17 globulin and 14 albumin genes have been identified (Supplementary Table 5. 7). The winged bean genes, Psote03G0078700 and Psote03G0078800, are homologous to the 11S globulin gene in soybean (Glyma.08G127600) and located 6.6 Mbp upstream of the qProt-3-1 (Hu et al. 2021). Interestingly, the Psote03G0078700 gene was highly expressed in pods and seeds at the early stages of development and its expression declined over time, whereas Psote03G0078800 expression was increased on Day45 as they developed, especially in seeds by almost 3.5 folds (Figure 5. 10). The Psote04G0317000.1 gene homologous to the Glyma.02G145700 gene in soybean and AT3G22640.1 (PAP85) gene in Arabidopsis (Chen et al. 2013), which codes for vicilin (a globulin storage protein) was expressed at low levels on Day15 in both pods and seeds in winged bean. In winged bean, its expression increased as the pods and seeds matured, showing a 5-fold rise in the pods and a 14-fold surge in the seeds by Day 30, with these elevated levels sustained on Day45. A similar expression pattern was observed for the 2S albumin gene, Psote04G0281900.1. In winged bean, psophocarpin B is primarily composed of 2S proteins (albumins), contains a high amount of the limiting sulphur-containing amino acids, and includes several trypsin inhibitors, chymotrypsin inhibitors, and lectins, constituting up to 35% of the seed storage protein (Kortt 1979; 1980; 1986; 1984). Therefore, it would be important for future winged bean breeding and genetic improvement programmes to quantify the sulphur content of the major seed storage protein fractions, to improve or even maintain the protein quality of winged bean seeds. The Psote05G0126400.1 gene homologous to Glyma.08G341300 coding for a chymotrypsin inhibitor was continuously expressed in pods with the highest level of expression on Day15, while in the seeds its highest expression level was on Day30 and it was strongly downregulated by Day45. In winged bean, putative Kunitz-type chymotrypsin inhibitor genes inhibited the gut proteinases of *Helicoverpa armigera larvae* (Telang et al. 2008), suggesting that winged bean proteinase inhibitors could be a target of further studies for the development of transgenic lines resistant to *H. armigera*, a pest that affects many important crops and develops fast resistance to

pesticides (Giri et al. 2003; Telang et al. 2008). Understanding seed storage protein accumulation in winged bean seeds and the signalling pathways involved could contribute to the molecular engineering of a highly nutritious crop for the future.

5.4.2 Fatty acid biosynthesis

The first step for fatty acid biosynthesis is the carboxylation of acetyl-CoA to malonyl-CoA catalysed by ACCase. ACCase is formed of three subunits: CAC1, CAC2 and CAC3. The winged bean genes, Psote05G0053200.1, Psote03G0022300.1 and Psote05G0096100.1, previously identified on the fatty acid QTLs (Supplementary Table 5. 5), are homologous to the Arabidopsis genes, AT5G15530, AT5G35360 and AT2G3804, coding for CAC1-B, CAC2 and CAC3, respectively. The genes had similar patterns of expression as seen in the pods, where they were expressed on a high level, with Day15 showing the highest level. Psote09G0032300.1 is homologous to AT5G16390 coding for CAC1-A. Psote03G0022300.1 and Psote05G0096100.1 had a similar level of expression in the seeds. However, Psote05G0053200.1 was downregulated as the seeds matured, and Psote09G0032300.1 and Psote05G0096100.1 were upregulated. Interestingly, in Arabidopsis, two genes located in a tandem repeat, *ACC1* and *ACC2*, were continuously expressed in the wild-type seed, and the recessive *acc1-1* and *acc1-2* mutations led to embryo lethality. Both alleles seemed to disrupt embryo morphogenesis causing abnormalities such as cucumber-like structures and lower amount of triacylglycerides (Baud et al. 2003). To better understand the genes involved in the initial steps of fatty acid biosynthesis and the formation of ACCase, further research is needed to elucidate the functions of these genes as well as the next enzymes involved in the fatty acid biosynthesis pathway.

The winged bean gene Psote07G0235900.1 is homologous to AT1G67730.1 Arabidopsis gene, beta-ketoacyl reductase 1 (*KCR1*) and is one of many enzymes involved in the fatty acid elongation process. In mutant Arabidopsis plants, the loss of function of *KCR1* was embryo-lethal (Beaudoin et al. 2009). The very long fatty acids play an imperative role in the membrane lipids and the membrane's homeostasis (Batsale et al. 2021). The Psote07G0235900.1 was highly expressed throughout the pod and seed development. However, in the seeds, it was highly expressed in the early stages. In winged bean, *KCR1* could

be contributing to the production of long-chain fatty acids such as behenic (22:0), which was detected at around 20% in mature winged bean seeds (determined in Chapter 3).

An important enzyme in the biosynthesis of unsaturated fatty acids is fatty acid desaturase 2, FAD2. FAD2 catalyses the desaturation of oleic acid to linoleic acid. In winged bean, three genes were homologous to the Arabidopsis *FAD2*, AT3G12120.2. The genes Psote05G0226300.1 and Psote05G0226400.1 were located closely together on chromosome 5, with the latter showing consistent expression in seeds, whereas the former did not, indicating a tissue-specific expression pattern. The Psote04G0026200.1 was expressed at very low levels by pods and seeds on Day15. However, its expression was 8-fold increased on Day30 in seeds. The rise in the expression of *FAD2* suggests an increase in the desaturation of oleic to linoleic. Similar results have been reported showing that the fatty acid composition of winged bean seeds changes during maturation, with a decrease in saturated fatty acids and an increase in mono- and polyunsaturated fatty acids as the seeds mature (Mohanty et al. 2015). In soybean, mutations in the isoforms of *FAD2* genes increased oleic concentration at the expense of linoleic (Al Amin et al. 2019; Do et al. 2019). As reported previously in Chapter 4, the relative amount of oleic and linoleic acid are significantly and negatively correlated in the winged bean mature seeds. However, as the concentration of the oleic and linoleic acid was not quantified in the immature pods and seeds during development this remains to be determined. Understanding the genes involved in the fatty acid biosynthesis pathway will provide useful insights into the accumulation of oil and fatty acids in the winged bean seeds.

5.4.3 Flavonoid biosynthesis

Anthocyanins are responsible for the colours in plants. Proanthocyanidins, often called condensed tannins, are oligomers or polymers of the monomeric flavanols, catechin, epicatechin, gallocatechin, and epigallocatechin. The condensed tannins are antinutritional factors that form complexes with proteins, starch and digestive enzymes (Gonthier et al. 2003). FP15 has purple flowers and during maturation, the pod changes colour from green with purple specks to fully purple by Day 45 (Figure 5. 6a). As FP15 had the lowest digestibility compared to the other winged bean accessions (results reported in Chapter 3), it was selected for investigating the expression of genes involved in the flavonoid and anthocyanin pathway.

The compounds catechin, epicatechin and proanthocyanins B1, B2 and C1 were higher in the pod at the late stages whereas they were highly concentrated in seeds at the early developmental stages (Table 5. 1). However, these results, along with the anthocyanin compounds (which were not reported), provide an indication rather than actual quantification, due to the need for further optimization of the method used.

The expression of the Psote01G0405800.1 gene coding for leucoanthocyanidin reductase, involved in the synthesis of catechin and gallocatechin, was lower on Days 30 and 45 compared to Day 15 but remained high in the winged bean seeds, with low levels of expression in the pods. Anthocyanin reductase 1 (Psote03G0143100.1) catalysing the epigallocatechin and epicatechin synthesis, was expressed at low levels in pods at all time points but was highly expressed in the seeds on Day 15, with the expression levels dropping significantly by Day 45 (Figure 5. 16). The low levels of gene expression in pods, along with the progressively decreasing expression as the seeds mature and increase in size, could explain the reduction in catechin and epicatechin levels in winged bean seeds from Day 15 to Day 30 and 45 (Figure 5. 16, Table 5. 1).

The colour of legume seeds is reported to be controlled by MYB domain genes, which act as transcription factors modulating the expression of genes that code for the enzymes involved in the anthocyanin biosynthesis pathway (Herniter et al. 2018; García-Fernández, Campa, and Ferreira 2021; Zabala and Vodkin 2014). As shown by Ho et al. (2024), three tandem replications of the MYB113 transcription factors were observed in winged bean, with the *PtMYB113a*, *PtMYB113b* and *PtMYB113c* having the conserved anthocyanin-promoting (S/A)NDV motif. Sequence changes were reported between the Ma3 (green) and FP15 (purple) accessions. Amino acid substitutions with different polarities and deletions were reported in the MYB transcription factors between the two accessions. The effects of these remain to be explored. In plants, gene duplication is an important part of their evolution and adaptation to environmental stresses. It is worth considering that anthocyanins and tannins are part of the secondary metabolism of plants and act as a defence mechanism. The MYB transcription factors play a role in the defence mechanism of the plants, therefore their conservation is crucial, and their duplication might be an adaptation mechanism to the environment (Saigo et al. 2020; Ma et al. 2024; Muzquiz et al. 2012).

5.5 Conclusion

This is the first transcriptomic analysis of developing pods and seeds in winged bean. The results showed differential expression of genes coding for key enzymes in seed storage proteins, fatty acid biosynthesis, and flavonoid biosynthesis pathway. Overall, this study lays the foundation to further explore and understand the molecular mechanisms and genes involved in the regulation of protein, oil and condensed tannin contents in winged bean pods and seeds. The transcriptomic results agreed with the protein content in seeds, and proposed seed storage protein genes. Ideally, this would apply to the fatty acid composition, condensed tannins, and anthocyanins. However, due to limited sample availability, the need for method optimisation, and time constraints, these aspects will need to be investigated in future studies.

Chapter 6: Discussion

The results presented in this thesis provide valuable insights into the genome, transcriptome, and nutritional value of winged bean, addressing the key objectives outlined at the start of the thesis. The discussions in the previous chapters focused on interpreting the findings in relation to the research questions and comparing them with previous studies. This chapter will focus on how these findings contribute to the broader understanding of improving the nutritional value of winged bean and deepening our knowledge of its genetic background. Additionally, the limitations of the work and suggestions for future research will be considered.

Understanding the genetics underlying desirable quantitative and qualitative traits is crucial, as it can aid in breeding selection and contribute to the development of winged bean varieties with improved plant architecture and enhanced nutritional quality. Winged bean, an underutilised high-protein legume, is cultivated in tropical regions, which are highly dependent on soybean importation. The work included in this thesis aimed to 1) examine the nutritional profile of 20 winged bean accessions and investigate the effect of environmental factors on these profiles; 2) from these accessions, utilise the cross of a high and a low protein winged bean accession to identify QTLs related to protein, oil, and fatty acid content; 3) focusing on the high-protein parental winged bean accession from the cross, investigate the expression of genes linked to nutritional value during pod and seed development.

1)Examining the nutritional profile of 20 winged bean accessions – Chapter 3

The nutritional profile of these 20 winged bean accessions has not been studied previously, nor has the effect of environmental factors on these profiles. Fat and protein contents in winged bean seeds varied significantly among the winged bean accessions, with fat and protein being significantly and negatively correlated. This negative relationship between oil and protein content in legume seeds has been presumably driven by competition for resources, metabolic trade-offs and genetic regulation. These complex and interdependent mechanisms make it difficult to increase one component (oil or protein) without reducing the

other, particularly as the measurement is on a g/100 g basis. Protein and lipid accumulation during seed development depends on the supply of amino acids, fatty acids and sugars from maternal tissues, such as carbon assimilated in the leaves, and the metabolic activity within the developing seed, particularly in the cotyledons (Allen and Young 2013; Kambhampati et al. 2020).

In winged bean seeds, fat content appeared to be influenced by the interaction between genotype and environmental factors. Additionally, protein content and the relative amount of different fatty acids varied significantly among the winged bean accessions. In contrast, amino acid levels (mg/100 mg of protein) did not show significant variation between the accessions. Of the 20 winged bean accessions, 8 were selected for *in vitro* protein digestibility, based on their diverse protein and oil content, levels of antinutritional factors, and seed availability. The *in vitro* protein digestibility of autoclaved winged bean seeds ranged from 40% to 58%, though the results were not statistically significant due to large statistical error that could be related to replication errors, biological variances and a limited number of technical and biological replicates.

Several limitations were encountered in Chapter 3, with the most significant being the limited sample availability. Unfortunately, there was insufficient material from both locations to fully establish the nutritional profile of winged bean seeds. As a result, priority was given to the location with the greater sample availability. It would have been valuable to compare the levels of antinutritional factors (ANFs) and *in vitro* protein digestibility between both locations, to better understand the effect of environment on ANFs and protein digestibility. Both genetic and environmental factors can significantly influence the levels of antinutritional factors, which in turn affect the digestibility and bioavailability of proteins (Oluwatos 1999; Bacon et al. 1995).

Antinutritional factors are mainly part of the plant's defence mechanisms, and they are controlled by its genotype. Plants grown in different environments are exposed to various biotic and abiotic stresses. Therefore, different genes are expressed in response to these conditions and influence the levels of antinutritional factors accumulated in the seeds which could, in turn, negatively affect protein digestibility. Investigating the genetic control and the

environmental impact, as well as their interaction, for the accumulation of these antinutritional factors in winged bean seeds would provide valuable insights. Such findings could lead to more informed winged bean accession selections based on environmental stress factors, ultimately improving crop performance and nutritional quality. It's important to recognise that antinutritional factors are part of a plant's natural defence mechanism against biotic and abiotic stresses. Reducing these compounds should be done carefully, as compromising the plant's defence could lead to increased pesticide use or increased wastage. This would not only raise production costs but also negatively impact the environment by contributing to pesticide resistance and ecological harm. Therefore, any reduction in antinutritional factors must balance plant resilience with environmental and economic sustainability.

Genomic approaches could help mitigate the negative effects of antinutritional factors. Genetic manipulation of antinutritional factors must be considered carefully, as these compounds play key roles in plant defence and stress tolerance. Advances in transgenics and gene editing offer promising opportunities for modification of tissue-specific expression. Significant progress has been made in breeding common bean and soybean genotypes with reduced levels of these antinutritional factors (Valentine et al. 2017; Duraiswamy et al. 2023; Cominelli et al. 2022). However, the potentially pleiotropic impact on plant stress resilience and growth in gene-edited genotypes has yet to be fully investigated.

Various processing methods are typically applied to pulses before they are consumed by humans or incorporated into animal feed. Common processing techniques such as soaking, roasting, autoclaving, and boiling are effective in reducing heat-sensitive antinutritional factors like protease inhibitors. For heat-resistant antinutritional compounds, such as tannins, methods like dehulling have been shown to improve protein digestibility more effectively (Makkar 2003; Chang et al. 1994). Other methods that have been shown to reduce significantly the amount of antinutritional factors in legume seeds are germination and fermentation. During germination, the nutritional composition of the seeds changes due to high metabolic activity. The activation of enzymes like phytase leads to the reduction of phytic acid (Samtiya, Aluko, and Dhewa 2020; Savelkoul, Van Der Poel, and Tamminga 1992). In winged bean seeds, changes in the amino acid composition and enzymatic activity have been

reported during germination (R. King and Puwastien 1987). Winged bean requires 30h fermentation for a good quality tempeh however, fermentation on winged bean is limited (S. Sri Kantha and Erdman 1984a). Both germination and fermentation are promising processing methods that, with further exploration, could offer valuable insights into enhancing the nutritional quality and broader utilisation of winged bean seeds in the food industry.

Recent advances in biotechnology and enzyme production have introduced additional options, such as the use of exogenous enzymes. For instance, adding phytase to animal feed has proven effective in reducing phytates, leading to enhanced protein digestibility (N. Romano and Kumar 2018). Given that farming and crop production industries are keen to reduce costs, it is essential to evaluate the most cost-effective approaches, which may vary by region or country. A comparative analysis of breeding for lower phytate content versus the cost of supplementing animal feed with phytase should be considered when making such decisions.

Winged bean seeds, like many legumes, contain low levels of sulphur-containing amino acids such as methionine and cysteine. However, they are rich in lysine, which is typically the limiting amino acid in cereals. This amino acid profile makes winged bean a great complement to cereal-based diets, where lysine deficiency is common. By combining legumes and cereals in human diets and animal feed, a more balanced protein intake can be achieved, improving protein quality and mineral absorption. However, further optimisation of the cereal–legume blends might be needed to meet energy, protein quality, and fat recommendations (Suri, Tano-Debrah, and Ghosh 2014). This dietary complement can enhance nutritional outcomes, particularly in regions with protein and micronutrient deficiencies, while also promoting the sustainable use of underutilised crops like winged bean (Anitha, Govindaraj, and Kane-Potaka 2020).

Measuring the nutritional profile of winged bean accessions in different environments is crucial, as factors such as soil quality, temperature, rainfall, and exposure to biotic and abiotic stresses can significantly influence a plant's growth, development, and nutrient composition. By understanding how these environmental factors affect the nutritional profile of winged bean seeds, researchers and plant breeders can identify genotypes that consistently produce

highly nutritious seeds regardless of location or environmental stress and/or identify the best genotype for a specific environment. Identification of QTLs and genetic markers that can enhance the nutritional value and adaptability of winged bean accessions is a key step towards developing resilient accessions that can contribute to food security and sustainable agriculture.

2)Identifying QTLs related to protein, oil, and fatty acid content – Chapter 4

The study described in Chapter 4 is the first to report QTLs related to protein (Ho et al. 2024), oil, and fatty acid content in winged bean seeds. Seeds from the F_2 population of a cross between the high protein (FP15: 39% protein) and a low protein (Ma3: 34% protein) winged bean accessions, were utilised to identify QTLs. The QTL analysis identified 16 QTLs, of which three were significantly associated with linoleic and behenic acid contents; and 1 QTL consistently linked to both stearic and palmitic acid contents. Putative genes related to protein content and fatty acid synthesis were found within or near these QTL regions. Further research is needed to identify markers linked to genes associated with winged bean seed composition, as this could significantly contribute to breeding improvement efforts.

QTL analysis is a valuable tool for researchers and plant breeders that can assist targeted breeding and accelerate the development of improved winged bean accessions. However, several limitations must be considered when conducting QTL analysis. An important factor is the population size needed, with at least 100 to 200 individual plants recommended for a robust QTL analysis. In this study, the population size was 162 individual plants for the protein content and 93 for the oil and fatty acid contents. In addition, the QTL analysis was performed only on one cross that was grown in one environment. This provides limited information on the QTL regions and can both over- and under-estimate the effect of other QTLs. Additionally, the maternal genetic effects need to be taken into consideration as they can influence the gene expression and protein synthesis in seeds (Donohue 2009). Other analyses such as diallel analysis should be explored to further understand the genetic variability, identify superior parents, and optimise breeding strategies for improving complex traits.

When working on complex traits such as nutritional traits, larger populations of more than 300 individuals can increase the detection power of QTL analysis and improve the precision of the estimates. Nutritional traits may be influenced by multiple genes interacting with each other (epistasis) or affecting multiple traits (pleiotropy), complicating the identification of individual QTLs. In addition to the complexity of the nutritional traits, they can also be influenced by environmental factors, making it even more difficult to distinguish between genetic and environmental effects. Therefore, conducting experiments by growing several crosses in controlled environments could provide clearer insights into genetic effects without the confounding influence of variable environmental conditions. Additionally, performing trials across multiple locations and conditions would be essential to identify a range of environmental influences. This is essential in identifying consistent QTLs related to nutritional traits that perform well across diverse environments. In soybean, main and epistatic effect QTLs for seed protein and oil content were identified as well as their interaction with the environment. In addition, a high-density map was used to improve the accuracy and find flanking markers that could also be beneficial to breeders using marker assisted selection (MAS) (Karikari et al. 2019). A step further would be the use of expression quantitative trait loci (eQTL). The eQTL analysis associates gene expression data with QTLs. A similar study has been performed on Brassica napus for fatty acid composition, flowering and growth traits (Li et al. 2018b). This approach can help identify how environmental conditions influence gene expression related to nutritional traits.

There are more factors that need to be considered when working with complex nutritional traits, such as labour-intense and costly phenotypic measurements. In this study, which focuses on the protein of winged bean seeds, it would have been ideal to identify QTLs for amino acids. Unfortunately, as amino acid analysis is a very expensive as well as time and labour-intensive method, it was not possible to do this at the time. For future studies on winged bean seeds, QTLs related to protein and amino acid composition could provide insights on identifying genes and transcription factors that could be utilised in breeding and gene editing to increase the amount of the limiting sulphur-containing amino acids and improve the digestible indispensable amino acid score (DIAAS). The accuracy of the phenotypic measurements is important, as errors can lead to misinterpretations of QTLs.

Therefore, this increases the need for an adequate number of replicates, which increases even further the cost of an already expensive analysis. Future studies should consider these recommendations that could significantly contribute to the improvement of winged bean research and breeding selection.

Additional methods, such as genome-wide association studies (GWAS), could be used to identify genetic markers in the winged bean genome to enhance nutritional composition and resistance to biotic and abiotic stresses. In GWAS, populations are used instead of crosses to identify associations between genetic markers and traits. This method would require a large population of sequenced or genotyped winged bean genotypes and a collection of their phenotypic data. The association of the genotypic and phenotypic data would highlight correlations between genetic markers and traits. This would not require controlled crosses, and it would examine a wider range of allelic variation in natural populations. A step further could be the combination of QTLs and GWAS methods for better resolution as in soybean in the studies of Zhang et al. 2019 and Sonah et al. 2015. It is important to note that the identification of QTLs and genetic markers does not directly translate to meaningful improvements in nutritional traits within breeding programs. Further research is needed to validate and understand the effects of these QTLs in different genetic background and environments, as well as the underlying mechanisms and genes involved in the nutritional composition of seeds. A deeper understanding of the genome and transcriptome is essential to uncover the genetic mechanisms and biosynthetic pathways responsible for the accumulation of proteins, fats, and antinutritional factors present in winged bean seeds.

3)Investigate the expression of genes linked to nutritional value during pod and seed development – Chapter 5

The next step, after identifying QTLs and genes involved in the nutritional composition of winged bean seeds, was to investigate their expression levels during pod and seed development. This chapter is the first study to investigate the differential gene expression during pod and seed development in winged bean. The high-protein parental accession (FP15: 39% protein) was selected to explore the expression of genes, particularly those involved in

seed storage proteins and fatty acid biosynthesis pathways, during pod and seed development. Transcriptome sequencing of developing pods and seeds was carried out on Days 15, 30 and 45, as well as Day 37 for pod only. This analysis found that a total of 7,954 genes were differentially expressed in the pod and 10,765 genes were differentially expressed in the seed, when comparing Days 45 and 15 after flowering. Gene ontology and KEGG pathway enrichment analyses revealed 42 differentially expressed genes (DEG) involved in the fatty acid biosynthesis pathway, 29 DEG in the biosynthesis of unsaturated fatty acids, 88 DEG related to seed storage proteins and 66 DEG in the flavonoid biosynthesis pathway.

Performing transcriptomic analysis in pods and seeds during development is crucial for understanding the expression of genes related to nutritional traits. Genes involved in the biosynthesis of proteins, oils, carbohydrates, and antinutritional factors are often expressed at specific developmental stages, and understanding their timing is essential for improving nutritional quality. Transcriptomic analysis can help identify tissue-specific genes and uncover their functional roles. Tissue specificity is highly important for breeders aiming to improve specific tissues, such as increasing protein content or reducing antinutritional factors in seeds.

While transcriptomic analysis is a powerful tool for understanding gene expression during pod and seed development, it carries some limitations related to tissue complexity, environmental influences, data analysis, cost, as well as the interpretation of gene function. In this study, only the high-protein parental accession (FP15: 39% protein) was investigated. Comparative transcriptomic analysis between accessions of different nutritional compositions in a controlled environment would provide valuable insights into the accumulation of seed storage proteins and the genes involved; similar to studies that have been performed in soybean, oil palm and castor bean (Peng et al. 2021; Dussert et al. 2013; Yu, Li, and Liu 2020).

Sampling at 1300–1400h provides a consistent and metabolically active "snapshot"; however, it does not capture the full extent of the day-night metabolic cycles. Gene expression profiles will be biased towards processes associated with photosynthesis, active metabolism, and daytime signalling, while activities predominant at night, such as starch remobilisation, certain stress responses, and hormone signalling, will not be represented. In winged bean, examining the diel expression patterns of genes during pod and seed development could offer

additional insights, particularly regarding the influence of the circadian clock on pod and seed maturation. In addition to the clear shift from photoautotrophic metabolism during the day to heterotrophic metabolism at night, there are important interactions between these two phases, mediated by the temporal separation of storage compound synthesis and their later mobilisation. The study of Gauthier et al. (2010) showed that the carboxylic acids synthesised and stored during the night were the main source of carbon skeletons for the nitrogen assimilation the next day; highlighting the importance of day–night cycles in metabolic pathways.

Other experimental factors, such as nighttime temperature and whether the plants were grown in the field or a growth chamber, can also influence gene expression (Mi et al. 2025). A high number of samples and replicates are required to obtain a more comprehensive view of the transcriptome, while these must be balanced with the available resources. Since environmental factors also affect the nutritional composition during pod and seed development, conducting transcriptomic analysis not only in controlled environments but also under differing environmental conditions would be informative in understanding these effects. However, the complexity and cost of such experiments need to be considered.

Several DEGs have been reported in this study. However, differential gene expression does not establish the functional role of genes in development or nutrition. Additional experiments, such as gene knockouts, overexpression studies, or proteomics, are required to validate the function of identified genes. This study attempted to integrate the transcriptomic data with the metabolomic data from the phenolic analysis to identify genes that relate to the anthocyanin synthesis pathway as well as condensed tannins. Even though that was not achieved due to the need for further optimisation of the equipment and methodology, it was heading in the direction of integrating transcriptomic data with other ‘omics’ data to provide a more complete understanding of how gene expression impacts phenotype.

It is equally important to acknowledge the possibility of post-transcriptional and post-translational regulation and modification. High levels of mRNA expression do not always correlate with high levels of protein function. The transcriptomic analysis measured the mRNA levels but did not account for post-transcriptional regulation such as RNA stability and

degradation nor the post-translational modifications, such as protein phosphorylation. These processes may significantly impact gene function during seed and pod development, therefore combining RNA-Seq with proteomics or other assays that assess protein levels and modifications can provide a more complete picture of gene regulation. The reported findings provide a baseline for functional and comparative genomic analysis, helping to better understand the developmental process and mechanisms that contribute to and control the nutritional value of winged bean seeds and pods. However, further work is needed to overcome many of these challenges and identify the key genes and transcription factors in the synthesis and accumulation of nutrients in the winged bean pods and seeds that could provide information useful to researchers and plant breeders.

The genes, transcription factors, and markers identified in these accessions could then be utilised in marker-assisted selection (MAS) breeding programs aimed at reducing antinutritional factors, such as tannins, or enhancing the nutritional quality of seeds. By incorporating these molecular tools, breeders can more precisely select desirable traits, accelerating the development of improved winged bean seed accessions with higher nutritional value and reduced levels of compounds that negatively affect digestibility or nutrient absorption. This approach could enhance breeding efficiency and contribute to the development of highly nutritious winged bean accessions. It is worth noting that some antinutritional factors, such as protein inhibitors, are proteins. Reducing the levels of these inhibitors could potentially impact the protein content and amino acid composition of the seeds. Since processing methods such as heat treatments are typically used to deactivate antinutritional factors in winged bean seeds for human consumption or animal feed, breeding for lower levels of heat-liable protein inhibitors may not be worthwhile.

The thesis did not focus on plant morphology, physiology, or stress resilience, which are crucial aspects that also need to be explored and improved for a more holistic approach to winged bean's development. Integrating these factors with the nutritional improvements aligns with the 'One Health' concept, that focuses on the interconnectedness of human, animal, and environmental health. By considering not only the nutritional quality of seeds but also the plant's overall resilience and adaptability to the environment, we can develop more sustainable agricultural systems. Addressing both the genetic and environmental factors

ensures that breeding programmes contribute to long-term food security and ecological sustainability, which can be beneficial to both human health and the environment.

Conclusions

Winged bean is a high protein legume that has been cultivated mainly for its immature pods and tubers by indigenous communities in Asia and sold in local markets. As an underutilised crop, it has received limited research attention for the improvement of its vining plant architecture and the optimisation of its nutritional value. The studies described in this thesis explored the nutritional profile of winged bean accessions, and the QTLs related to protein, oil, and fatty acid content. In addition, the genetic regulation of seed and pod development was investigated with a focus on key biosynthetic pathways related to nutritional value. For the first time, this study has provided novel insights into the molecular markers associated with protein, oil, and fatty acid contents in winged bean seeds; as well as the molecular mechanisms and genes involved in the protein, lipid, and flavonoid biosynthesis pathways in developing pods and seeds.

This thesis has addressed a gap in the understanding of winged bean's developmental biology and the genetic basis of QTLs related to its nutritional value. These findings could assist genomic research and accelerate breeding selection through the use of genetic markers, by providing a solid foundation for future research and crop improvement efforts aimed at enhancing its nutritional properties. This research contributes to the field of underutilised legume crops and offers valuable information to plant breeders seeking to utilise winged bean for human consumption or animal feed. Winged bean could play a significant role in future food security. Therefore, more research is needed to fully explore its potential to become a new soybean for the tropics.

Bibliography

Abdullah, Siti Nor Akmar, Chai-Ling Ho, and Carol Wagstaff. 2017. *Crop Improvement : Sustainability Through Leading-Edge Technology*. *Crop Improvement*. Springer International Publishing. https://doi.org/10.1007/978-3-319-65079-1.

Abe, Jiro, and Hiroshi Nakamura. 1987. "Evaluation of Winged Bean in Okinawa."

Adachi, Motoyasu, Jiro Kanamori, Taro Masuda, Kazuhiro Yagasaki, Keisuke Kitamura, Bunzo Mikami, and Shigeru Utsumi. 2002. "Crystal Structure of Soybean 11S Globulin: Glycinin A3B4 Homohexamer." *PNAS*. www.rcsb.org.

Adebamowo, Clement A., Eunyoung Cho, Laura Sampson, Martijn B. Katan, Donna Spiegelman, Walter C. Willett, and Michelle D. Holmes. 2005. "Dietary Flavonols and Flavonol-Rich Foods Intake and the Risk of Breast Cancer." *International Journal of Cancer* 114 (4): 628–33. https://doi.org/10.1002/ijc.20741.

Adegboyega, Taofeek T., Michael T. Abberton, Abdel Aziz H. AbdelGadir, Mahamadi Dianda, Busie Maziya-Dixon, Olaniyi A. Oyatomi, Sam Ofodile, and Olubukola O. Babalola. 2019a. " Nutrient and Antinutrient Composition of Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) Seeds and Tubers ." *Journal of Food Quality* 2019 (October):1–8. https://doi.org/10.1155/2019/3075208.

———. 2019b. "Nutrient and Antinutrient Composition of Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) Seeds and Tubers." *Journal of Food Quality* 2019 (October):1–8. https://doi.org/10.1155/2019/3075208.

Adegboyega, Taofeek T., Michael T. Abberton, AbdelAziz H. AbdelGadir, Mahamadi Dianda, Busie Maziya-Dixon, Olaniyi A. Oyatomi, Sam Ofodile, and Olubukola O. Babalola. 2019c. "Nutrient and Antinutrient Composition of Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) Seeds and Tubers." *Journal of Food Quality* 2019 (October):1–8. https://doi.org/10.1155/2019/3075208.

Aguirre, Manuel, Edward Kiegle, Giulia Leo, and Ignacio Ezquer. 2018. "Carbohydrate Reserves and Seed Development: An Overview." *Plant Reproduction*. Springer Verlag. https://doi.org/10.1007/s00497-018-0336-3.

Alexander, Peter, Calum Brown, Almut Arneth, Clare Dias, John Finnigan, Dominic Moran, and Mark D.A. Rounsevell. 2017. "Could Consumption of Insects, Cultured Meat or Imitation

Meat Reduce Global Agricultural Land Use?" *Global Food Security* 15 (January): 22–32. https://doi.org/10.1016/j.gfs.2017.04.001.

Allen, Doug K., and Jamey D. Young. 2013. "Carbon and Nitrogen Provisions Alter the Metabolic Flux in Developing Soybean Embryos." *Plant Physiology* 161 (3): 1458–75. https://doi.org/10.1104/PP.112.203299.

Amin, Noor Al, Naveed Ahmad, Nan Wu, Xiumin Pu, Tong Ma, Yeyao Du, Xiaoxue Bo, Nan Wang, Rahat Sharif, and Piwu Wang. 2019. "CRISPR-Cas9 Mediated Targeted Disruption of FAD2-2 Microsomal Omega-6 Desaturase in Soybean (Glycine Max.L)." *BMC Biotechnology* 19 (1): 1–10. https://doi.org/10.1186/S12896-019-0501-2/FIGURES/6.

Amoo, I.A., O.T. Adebayo, and A.O. Oyeleye. 2006a. "Chemical Evaluation of Winged Beans. (Psophocarpus Tetragonolobus), Pitanga Cherries (Eugenia Uniflora) and Orchid Fruit (Orchid Fruit Myristica)." *African Journal of Food Agriculture Nutrition and Development* 6 (2): 1–12.

———. 2006b. "Chemical Evaluation of Winged Beans. (Psophocarpus Tetragonolobus), Pitanga Cherries (Eugenia Uniflora) and Orchid Fruit (Orchid Fruit Myristica)." *African Journal of Food Agriculture Nutrition and Development* 6 (2): 1–12.

———. 2006c. "Chemical Evaluation of Winged Beans. (Psophocarpus Tetragonolobus), Pitanga Cherries (Eugenia Uniflora) and Orchid Fruit (Orchid Fruit Myristica)." *African Journal of Food Agriculture Nutrition and Development* 6 (2): 1–12.

Anitha, Seetha, Mahalingam Govindaraj, and Joanna Kane-Potaka. 2020. "Balanced Amino Acid and Higher Micronutrients in Millets Complements Legumes for Improved Human Dietary Nutrition." *Cereal Chemistry* 97 (1): 74–84. https://doi.org/10.1002/CCHE.10227.

Assefa, Y., N. Bajjalieh, S. Archontoulis, S. Casteel, D. Davidson, P. Kovács, S. Naeve, and Ignacio A. Ciampitti. 2018. "Spatial Characterization of Soybean Yield and Quality (Amino Acids, Oil, and Protein) for United States." *Scientific Reports* 8 (1): 1–11. https://doi.org/10.1038/s41598-018-32895-0.

Bacon, J. R., N. Lambert, P. Matthews, A. E. Arthur, and C. Duchene. 1995. "Genotype, Environment and Genotype × Environment Effects on Trypsin Inhibitor in Peas." *Journal of the Science of Food and Agriculture* 67 (1): 101–8. https://doi.org/10.1002/JSFA.2740670116.

Bassal, Hussein, Othmane Merah, Aqeel M. Ali, Akram Hijazi, and Fawaz El Omar. 2020. "Psophocarpus Tetragonolobus: An Underused Species with Multiple Potential Uses." *Plants* 9 (12): 1–11. https://doi.org/10.3390/plants9121730.

Batsale, Marguerite, Delphine Bahammou, Laetitia Fouillen, Sébastien Mongrand, Jérôme Joubès, and Frédéric Domergue. 2021. "Biosynthesis and Functions of Very-Long-Chain Fatty Acids in the Responses of Plants to Abiotic and Biotic Stresses." *Cells* 10 (6). https://doi.org/10.3390/CELLS10061284.

Baud, Sébastien, Virginie Guyon, Jocelyne Kronenberger, Sylvie Wuillème, Martine Miquel, Michel Caboche, Loïc Lepiniec, and Christine Rochat. 2003. "Multifunctional Acetyl-CoA Carboxylase 1 Is Essential for Very Long Chain Fatty Acid Elongation and Embryo Development in Arabidopsis." *The Plant Journal* 33 (1): 75–86. https://doi.org/10.1046/J.1365-313X.2003.016010.X.

Beaudoin, Frédéric, Xianzhong Wu, Fengling Li, Richard P. Haslam, Jonathan E. Markham, Huanquan Zheng, Johnathan A. Napier, and Ljerka Kunst. 2009. "Functional Characterization of the Arabidopsis Beta-Ketoacyl-Coenzyme A Reductase Candidates of the Fatty Acid Elongase." *Plant Physiology* 150 (3): 1174–91. https://doi.org/10.1104/PP.109.137497.

Beninger, Clifford W., and George L. Hosfield. 2003. "Antioxidant Activity of Extracts, Condensed Tannin Fractions, and Pure Flavonoids from Phaseolus Vulgaris L. Seed Coat Color Genotypes." *Journal of Agricultural and Food Chemistry* 51 (27): 7879–83. https://doi.org/10.1021/jf0304324.

Bergler, Helmut, Sandra Fuchsbichler, Gregor Högenauer, and Friederike Turnowsky. 1996. "The Enoyl-[Acyl-Carrier-Protein] Reductase (FabI) of Escherichia Coli, Which Catalyzes a Key Regulatory Step in Fatty Acid Biosynthesis, Accepts NADH and NADPH as Cofactors and Is Inhibited by Palmitoyl-CoA." *European Journal of Biochemistry* 242 (3): 689–94. https://doi.org/10.1111/j.1432-1033.1996.0689r.x.

Bessada, Sílvia M.F., João C.M. Barreira, and M. Beatriz P.P. Oliveira. 2019a. "Pulses and Food Security: Dietary Protein, Digestibility, Bioactive and Functional Properties." *Trends in Food Science & Technology* 93 (November):53–68. https://doi.org/10.1016/j.tifs.2019.08.022.

———. 2019b. “Pulses and Food Security: Dietary Protein, Digestibility, Bioactive and Functional Properties.” *Trends in Food Science & Technology* 93 (November):53–68. https://doi.org/10.1016/j.tifs.2019.08.022.

Bilder, Patrick, Sandra Lightle, Graeme Bainbridge, Jeffrey Ohren, Barry Finzel, Fang Sun, Susan Holley, et al. 2006. “The Structure of the Carboxyltransferase Component of Acetyl-CoA Carboxylase Reveals a Zinc-Binding Motif Unique to the Bacterial Enzyme.” *Biochemistry* 45 (6): 1712–22. https://doi.org/10.1021/bi0520479.

Birk, Yehudith. 1985. “The Bowman-Birk Inhibitor. Trypsin- and Chymotrypsin-inhibitor from Soybeans.” *International Journal of Peptide and Protein Research*. https://doi.org/10.1111/j.1399-3011.1985.tb02155.x.

Blagrove, R J, and J M Gillespie. 1978. “Variability of the Seed Globulins of Winged Bean, Psophocarpus Tetragonolobus (L.) DC.” *Aust. J. Plant Physiol*. Vol. 5.

Bolon, Yung-Tsi, Bindu Joseph, Steven B Cannon, Michelle A Graham, Brian W Diers, Andrew D Farmer, Gregory D May, et al. 2010. “Complementary Genetic and Genomic Approaches Help Characterize the Linkage Group I Seed Protein QTL in Soybean.” *BMC Plant Biology* 10 (1): 41. https://doi.org/10.1186/1471-2229-10-41.

Brodkorb, André, Lotti Egger, Marie Alminger, Paula Alvito, Ricardo Assunção, Simon Ballance, Torsten Bohn, et al. 2019. “INFOGEST Static in Vitro Simulation of Gastrointestinal Food Digestion.” *Nature Protocols 2019 14:4* 14 (4): 991–1014. https://doi.org/10.1038/s41596-018-0119-1.

Calvert, Duncan, Tristan Dew, Arthur Gadon, Jacques Gros, and David Cook. 2024. “Valorisation of Hop Leaves for Their Bioactive Compounds: Identification and Quantification of Phenolics across Different Varieties, Crop Years and Stages of Development.” https://ssrn.com/abstract=4912262.

Carbonaro, Marina, Fabio Virgili, and Emilia Carnovale. 1996. “Evidence for Protein–Tannin Interaction in Legumes: Implications in the Antioxidant Properties of Faba Bean Tannins.” *LWT - Food Science and Technology* 29 (8): 743–50. https://doi.org/10.1006/FSTL.1996.0116.

Carrera, Constanza, María José Martínez, Julio Dardanelli, and Mónica Balzarini. 2011. “Environmental Variation and Correlation of Seed Components in Nontransgenic Soybeans: Protein, Oil, Unsaturated Fatty Acids, Tocopherols, and Isoflavones.” *Crop Science* 51 (2): 800–809. https://doi.org/10.2135/cropsci2010.06.0314.

Cater, Nilo B., and Margo A. Denke. 2001. “Behenic Acid Is a Cholesterol-Raising Saturated Fatty Acid in Humans.” *American Journal of Clinical Nutrition* 73 (1): 41–44. https://doi.org/10.1093/ajcn/73.1.41.

Ceballos, Laura Sanz, Eva Ramos Morales, Gloria de la Torre Adarve, Javier Díaz Castro, Luís Pérez Martínez, and María Remedios Sanz Sampelayo. 2009. “Composition of Goat and Cow Milk Produced under Similar Conditions and Analyzed by Identical Methodology.” *Journal of Food Composition and Analysis* 22 (4): 322–29. https://doi.org/10.1016/j.jfca.2008.10.020.

Černý, K., Maud Kordylas, F. Pospíšil, O. Švábenský, and B. Zajíc. 1971a. “ Nutritive Value of the Winged Bean (Psophocarpus Palustris Desv.) .” *British Journal of Nutrition* 26 (2): 293–99. https://doi.org/10.1079/BJN19710035.

———. 1971b. “ Nutritive Value of the Winged Bean (Psophocarpus Palustris Desv.) .” *British Journal of Nutrition* 26 (2): 293–99. https://doi.org/10.1079/bjn19710035.

CHANG, M. -J, J. L. COLLINS, J. W. BAILEY, and D. L. COFFEY. 1994. “Cowpeas Tannins Related to Cultivar, Maturity, Dehulling and Heating.” *Journal of Food Science* 59 (5): 1034–36. https://doi.org/10.1111/J.1365-2621.1994.TB08183.X.

Chankaew, Sompong, Sasiprapa Sriwichai, Teppratan Rakvong, Tidarat Monkham, Jirawat Sanitchon, Sithichoke Tangphatsornruang, Wasitthee Kongkachana, et al. 2022. “The First Genetic Linkage Map of Winged Bean [Psophocarpus Tetragonolobus (L.) DC.] and QTL Mapping for Flower-, Pod-, and Seed-Related Traits.” *Plants* 11 (4): 500. https://doi.org/10.3390/PLANTS11040500.

Chapman, Mark A. 2015a. “Transcriptome Sequencing and Marker Development for Four Underutilized Legumes.” *Applications in Plant Sciences* 3 (2): 1–5. https://doi.org/10.3732/apps.1400111.

———. 2015b. “Transcriptome Sequencing and Marker Development for Four Underutilized Legumes.” *Applications in Plant Sciences* 3 (2): 1–5. https://doi.org/10.3732/apps.1400111.

Chen, Cheng-En, Kuo-Chen Yeh, Shu-Hsing Wu, Hsiang-Iu Wang, and Hsin-Hung Yeh. 2013. “A Vicilin-like Seed Storage Protein, PAP85, Is Involved in Tobacco Mosaic Virus Replication.” *Journal of Virology* 87 (12): 6888–6900. https://doi.org/10.1128/JVI.00268-13.

Chen, Dongming, Xing Yi, Hui Yang, Hui Zhou, Ye Yu, Yun Tian, and Xiangyang Lu. 2015. "Genetic Diversity Evaluation of Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) Using Inter-Simple Sequence Repeat (ISSR)." *Genetic Resources and Crop Evolution* 62 (6): 823–28. https://doi.org/10.1007/s10722-015-0261-3.

Chen, Huan, Fa Wei Wang, Yuan Yuan Dong, Nan Wang, Ye Peng Sun, Xiao Yan Li, Liang Liu, et al. 2012. "Sequence Mining and Transcript Profiling to Explore Differentially Expressed Genes Associated with Lipid Biosynthesis during Soybean Seed Development." *BMC Plant Biology* 12 (1): 122. https://doi.org/10.1186/1471-2229-12-122.

Cheng, Acga, Hui Hui Chai, Wai Kuan Ho, Aliyu Siise, Abdullah Bamba, Aryo Feldman, Presidor Kendabie, Razlin Azman Halim, and Alberto Tanzi. 2017. "Crop Improvement." In *Crop Improvement*, 47–70. https://doi.org/10.1007/978-3-319-65079-1.

Cheng, Acga, Murthazar Naim Raai, Nurul Amalina Mohd Zain, Festo Massawe, Ajit Singh, and Wan Abd Al Qadr Imad Wan-Mohtar. 2019. "In Search of Alternative Proteins: Unlocking the Potential of Underutilized Tropical Legumes." *Food Security*. Springer Netherlands. https://doi.org/10.1007/s12571-019-00977-0.

Chiba, T. 2003. *Encyclopedia of Food Sciences and Nutrition*. Edited by Caballero Benjamin, Finglas Paul, and Trugo Luiz. *Encyclopedia of Food Sciences and Nutrition*. Second. Academic Press. https://doi.org/10.1016/b0-12-227055-x/01121-4.

Chong, Yuet Tian. 2024. "Phenotypic and Genetic Investigation of Plant Phenological Development, Architecture Adn Yield in Winged Bean (Psophocarpus Tetragonolobus (L.) DC.)."

Chrispeels, Maarten J., and Natasha V. Raikhel. 1991. "Lectins, Lectin Genes, and Their Role in Plant Defense." *Plant Cell*. American Society of Plant Biologists. https://doi.org/10.1105/tpc.3.1.1.

Claydon, A. 1975. "A Review of the Nutritional Value of the Winged Bean *Psophocarpus Tetragonolobus* (L.) DC. with Special Reference to Papua New Guinea." *Science in New Guinea*.

Cominelli, Eleonora, Francesca Sparvoli, Silvia Lisciani, Chiara Forti, Emanuela Camilli, Marika Ferrari, Cinzia Le Donne, et al. 2022. "Antinutritional Factors, Nutritional Improvement, and Future Food Use of Common Beans: A Perspective." *Frontiers in Plant Science* 13 (August). https://doi.org/10.3389/FPLS.2022.992169/FULL.

D'Agostino, Nunzio, and Pasquale Tripodi. 2017. "NGS-Based Genotyping, High-Throughput Phenotyping and Genome-Wide Association Studies Laid the Foundations for next-Generation Breeding in Horticultural Crops." *Diversity* 9 (3): 1–20. https://doi.org/10.3390/d9030038.

Das, Abhijit K., Munegowda M. Gowda, Vignesh Muthusamy, Rajkumar U. Zunjare, Hema S. Chauhan, Aanchal Baveja, Vinay Bhatt, et al. 2021. "Development of Maize Hybrids With Enhanced Vitamin-E, Vitamin-A, Lysine, and Tryptophan Through Molecular Breeding." *Frontiers in Plant Science* 0 (July):1427. https://doi.org/10.3389/FPLS.2021.659381.

DeClerck, Y A, and S Imren. 1994. "Protease Inhibitors: Role and Potential Therapeutic Use in Human Cancer." *European Journal of Cancer (Oxford, England : 1990)* 30A (14): 2170–80. https://doi.org/10.1016/0959-8049(94)00460-m.

Do, Phat T., Cuong X. Nguyen, Hien T. Bui, Ly T.N. Tran, Gary Stacey, Jason D. Gillman, Zhanyuan J. Zhang, and Minviluz G. Stacey. 2019. "Demonstration of Highly Efficient Dual GRNA CRISPR/Cas9 Editing of the Homeologous GmFAD2-1A and GmFAD2-1B Genes to Yield a High Oleic, Low Linoleic and α-Linolenic Acid Phenotype in Soybean." *BMC Plant Biology* 19 (1): 1–14. https://doi.org/10.1186/S12870-019-1906-8/FIGURES/6.

Donald, C. M. 1968. "The Breeding of Crop Ideotypes." *Euphytica* 17 (3): 385–403. https://doi.org/10.1007/BF00056241.

Dong, Jia Yi, and Li Qiang Qin. 2011. "Soy Isoflavones Consumption and Risk of Breast Cancer Incidence or Recurrence: A Meta-Analysis of Prospective Studies." *Breast Cancer Research and Treatment* 125 (2): 315–23. https://doi.org/10.1007/s10549-010-1270-8.

Donohue, Kathleen. 2009. "Completing the Cycle: Maternal Effects as the Missing Link in Plant Life Histories." *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1520): 1059. https://doi.org/10.1098/RSTB.2008.0291.

Dumas, J B A. 1831. "Procédés de l'analyse Organique." *Ann. Chim. Phys* 47:198–205.

Duraiswamy, Aishwarya, Nancy Mano Sneha A, Sherina Jebakani K, Sellakumar Selvaraj, Lydia Pramitha J, Ramchander Selvaraj, Indira Petchiammal K, et al. 2023. "Genetic Manipulation of Anti-Nutritional Factors in Major Crops for a Sustainable Diet in Future." *Frontiers in Plant Science* 13. https://doi.org/10.3389/FPLS.2022.1070398/FULL.

Duranti, Marcello. 2006. "Grain Legume Proteins and Nutraceutical Properties." *Fitoterapia*. Elsevier. https://doi.org/10.1016/j.fitote.2005.11.008.

Dussert, Stéphane, Chloé Guerin, Mariette Andersson, Thierry Joët, Timothy J. Tranbarger, Maxime Pizot, Gautier Sarah, Alphonse Omore, Tristan Durand-Gasselin, and Fabienne Morcillo. 2013. "Comparative Transcriptome Analysis of Three Oil Palm Fruit and Seed Tissues That Differ in Oil Content and Fatty Acid Composition." *Plant Physiology* 162 (3): 1337–58. https://doi.org/10.1104/PP.113.220525.

Eagleton, Graham E. 2019. "Prospects for Developing an Early Maturing Variety of Winged Bean (Psophocarpus Tetragonolobus) in Bogor, Indonesia." *Biodiversitas* 20 (11): 3142–52. https://doi.org/10.13057/biodiv/d201106.

———. 2020. "Review: Winged Bean (Psophocarpus Tetragonolobus) Cropping Systems." *Biodiversitas*. Society for Indonesian Biodiversity. https://doi.org/10.13057/biodiv/d211258.

Ebert, Andreas W. 2014. "Potential of Underutilized Traditional Vegetables and Legume Crops to Contribute to Food and Nutritional Security, Income and More Sustainable Production Systems." *Sustainability (Switzerland)* 6 (1): 319–35. https://doi.org/10.3390/su6010319.

Eckardt, Nancy A. 2007. "GA Signaling: Direct Targets of DELLA Proteins." *The Plant Cell* 19 (10): 2970. https://doi.org/10.1105/TPC.107.191010.

Ekpenyong, T. E., and R. L. Borchers. 1980a. "The Fatty Acid Composition of the Oil of the Winged Bean (Psophocarpus Tetragonolobus L.) Seeds." *Journal of the American Oil Chemists' Society* 57 (5): 147–49. https://doi.org/10.1007/BF02673105.

Ekpenyong, T E., and R L. Borchers. 1980b. "The Fatty Acid Composition of the Oil of the Winged Bean (Psophocarpus Tetragonolobus L.) Seeds." *Journal of the American Oil Chemists' Society* 57 (5): 147–49. https://doi.org/10.1007/BF02673105.

Ekpenyong, T. E., and R. L. Borchers. 1982. "Amino Acid Profile of the Seed and Other Parts of the Winged Bean." *Food Chemistry* 9 (3): 175–82. https://doi.org/10.1016/0308-8146(82)90095-4.

Erdman, J. W., Jr. Weingartner, K. E. Mustakas, G. C. Schmutz, R. D. Park, H. M. Parker., and R. M. Forbes. 1980. "Zinc and Magnesium Bioavailability From Acid-Precipitated and Neutralized Soybean Protein Products." *Journal of Food Science* 45 (5): 1193–99. https://doi.org/10.1111/j.1365-2621.1980.tb06519.x.

Erskine, W. 1979. "The Exploitation of Genetic Diversity in the Winged Bean (Psophocarpus Tetragonolobus (L. DC.) for Grain Yield."

———. 1980. “Measurements of the Cross-Pollination of Winged Bean in Papua New Guinea.” *SABRAO Journal* 12 (1): 11–13. https://www.cabdirect.org/cabdirect/abstract/19840214803.

Erskine, W., and T. N. Khan. 1977. “Inheritance of Pigmentation and Pod Shape in Winged Bean.” *Euphytica* 26 (3): 829–31. https://doi.org/10.1007/BF00021714.

———. 1980. “Variation within and between Land Races of Winged Bean (Psophocarpus Tetragonolobus (L.) DC.).” *Field Crops Research* 3 (C): 359–64. https://doi.org/10.1016/0378-4290(80)90041-6.

Esaka, Muneharu, Kanichi Suzuki, and Kioshi Kubota. 1987. “Effects of Microwave Heating on Lipoxygenase and Trypsin Inhibitor Activities, and Water Absorption of Winged Bean Seeds.” *Journal of Food Science* 52 (6): 1738–39. https://doi.org/10.1111/j.1365-2621.1987.tb05920.x.

FAO. 2010. *The Second Report on the State of the World’s Animal Genetic Resources for Food and Agriculture*. *The Second Report on the State of the World’s Animal Genetic Resources for Food and Agriculture*. https://doi.org/10.4060/i4787e.

FAO/WHO/UNU expert consultation. 2007. “Protein and Amino Acid Requirements in Human Nutrition.” *WHO Technical Report Series 935* . www.who.int/bookorders.

Fatihah, H. N.Nur, N. Maxted, and L. Rico Arce. 2012. “Cladistic Analysis of Psophocarpus Neck. Ex DC. (Leguminosae, Papilionoideae) Based on Morphological Characters.” *South African Journal of Botany* 83 (November):78–88. https://doi.org/10.1016/j.sajb.2012.07.010.

Frolov, A. V., V.G. Blinov, G.S. Golitsyn, V.P. Dymnikov, Yu.A. Izrael, V. M. Kattsov, V .M. Kotlyakov, V.I. Osipov, and S.M. Semenov. 2014. *Second Roshydromet Assessment Report on Climate Change and Its Consequences in the Russian Federation*.

Gangurde, Sunil S., Hui Wang, Shasidhar Yaduru, Manish K. Pandey, Jake C. Fountain, Ye Chu, Thomas Isleib, et al. 2020. “Nested-Association Mapping (NAM)-Based Genetic Dissection Uncovers Candidate Genes for Seed and Pod Weights in Peanut (Arachis Hypogaea).” *Plant Biotechnology Journal* 18 (6): 1457–71. https://doi.org/10.1111/pbi.13311.

Garcia, V. V., and J. K. Palmer. 1980a. “Proximate Analysis of Five Varieties of Winged Beans, Psophocarpus Tetragonolobus (L.) DC.” *International Journal of Food Science & Technology* 15 (5): 469–76. https://doi.org/10.1111/j.1365-2621.1980.tb00965.x.

———. 1980b. “Proximate Analysis of Five Varieties of Winged Beans, Psophocarpus Tetragonolobus (L.) DC.” *International Journal of Food Science & Technology* 15 (5): 469–76. https://doi.org/10.1111/j.1365-2621.1980.tb00965.x.

Garcia, V. V., J. K. Palmer, and R. W. Young. 1979. “Fatty Acid Composition of the Oil of Winged Beans, Psophocarpus Tetragonolobus (L.) DC.” *Journal of the American Oil Chemists' Society* 56 (12): 931–32. https://doi.org/10.1007/BF02674137.

Gauthier, Paul P.G., Richard Bligny, Elizabeth Gout, Aline Mahé, Salvador Nogués, Michael Hodges, and Guillaume G.B. Tcherkez. 2010. “In Folio Isotopic Tracing Demonstrates That Nitrogen Assimilation into Glutamate Is Mostly Independent from Current CO2 Assimilation in Illuminated Leaves of Brassica Napus.” *New Phytologist* 185 (4): 988–99. https://doi.org/10.1111/j.1469-8137.2009.03130.x.

GBIF. 2021. “GBIF: Global Biodiversity Information Facility.” 2021. https://www.gbif.org/occurrence/search?taxon_key=2944683.

Gibbs, Peter E M, Kara B Strongin, and Alexander Mcpherson. 1989. “Evolution of Legume Seed Storage Proteins-A Domain Common to Legumins and Vicilins Is Duplicated in Vicilins ’.” https://academic.oup.com/mbe/article/6/6/614/981427.

Gilani, G. Sarwar, Kevin A. Cockell, and Estatira Sepehr. 2005. “Effects of Antinutritional Factors on Protein Digestibility and Amino Acid Availability in Foods.” *Journal of AOAC International* 88 (3): 967–87. https://doi.org/10.1093/jaoac/88.3.967.

Gilani, G. Sarwar, Chao Wu Xiao, and Kevin A. Cockell. 2012. “Impact of Antinutritional Factors in Food Proteins on the Digestibility of Protein and the Bioavailability of Amino Acids and on Protein Quality.” *British Journal of Nutrition* 108 (SUPPL. 2). https://doi.org/10.1017/S0007114512002371.

Gillespie, J M, and R J Blagrove. 1978. “Isolation and Composition of the Seed Globulins of Winged Bean, Psophocarpus Tetragonolobus (L.) DC.” *Aust. J. Plant Physiol*. Vol. 5.

Giri, Ashok P., Abhay M. Harsulkar, Maurice S.B. Ku, Vidya S. Gupta, Vasanti V. Deshpande, Prabhakar K. Ranjekar, and Vincent R. Franceschi. 2003. “Identification of Potent Inhibitors of Helicoverpa Armigera Gut Proteinases from Winged Bean Seeds.” *Phytochemistry* 63 (5): 523–32. https://doi.org/10.1016/S0031-9422(03)00181-X.

Gomez, Maria Dolores, Isabel Cored, Daniela Barro-Trastoy, Joaquin Sanchez-Matilla, Pablo Tornero, and Miguel A. Perez-Amador. 2023. “DELLA Proteins Positively Regulate Seed

Size in Arabidopsis." *Development (Cambridge, England)* 150 (15). https://doi.org/10.1242/DEV.201853.

Gonthier, M.-P., J. L. Donovan, O. Texier, C. Felgines, C. C. Remesy, and A. Scalbert. 2003. "Metabolism of Dietary Procyanidins in Rats." *Free Radical Biology & Medicine* 35 (8): 837–844. https://doi.org/10.1016/S0891-5849(03)00394-0.

González-Thuillier, Irene, Mónica Venegas-Calerón, Rosario Sánchez, Rafael Garcés, Penny von Wettstein-Knowles, and Enrique Martínez-Force. 2016. "Sunflower (Helianthus Annuus) Fatty Acid Synthase Complex: β-Hydroxyacyl-[Acyl Carrier Protein] Dehydratase Genes." *Planta* 243 (2): 397–410. https://doi.org/10.1007/s00425-015-2410-5.

GRAMI, BAHRAM, B R. STEFANSSON, and R J. BAKER. 1977. "GENETICS OF PROTEIN AND OIL CONTENT IN SUMMER RAPE: HERITABILITY, NUMBER OF EFFECTIVE FACTORS, AND CORRELATIONS." *Canadian Journal of Plant Science* 57 (3): 937–43. https://doi.org/10.4141/cjps77-134.

Gross Rainer. 1983. "Composition and Protein Quality of Winged Bean (Psophocarpus Tetragonolobus)." *Plant Foods for Human Nutrition* 32:117–24.

Guo, B ;, J ; Zhang, C ; Yang, Binhui Guo, Jianhua Zhang, Chunhong Yang, Lu Dong, et al. 2023. "The Late Embryogenesis Abundant Proteins in Soybean: Identification, Expression Analysis, and the Roles of GmLEA4_19 in Drought Stress." https://doi.org/10.3390/ijms241914834.

Gupta, Raj Kishor, Shivraj Singh Gangoliya, and Nand Kumar Singh. 2015. "Reduction of Phytic Acid and Enhancement of Bioavailable Micronutrients in Food Grains." *Food Sci Technol (February* 52 (5): 676–87. https://doi.org/10.1007/s13197-013-0978-y.

Gurfinkel, D. M., and A. V. Rao. 2003. "Soyasaponins: The Relationship between Chemical Structure and Colon Anticarcinogenic Activity." *Nutrition and Cancer* 47 (1): 24–33. https://doi.org/10.1207/s15327914nc4701_3.

Ha, Jungmin, Myoyeon Kim, Moon Young Kim, Taeyoung Lee, Min Young Yoon, Jayern Lee, Yeong Ho Lee, et al. 2018. "Transcriptomic Variation in Proanthocyanidin Biosynthesis Pathway Genes in Soybean (Glycine Spp.)." *Journal of the Science of Food and Agriculture* 98 (6): 2138–46. https://doi.org/10.1002/JSFA.8698.

Hanson, W. D., R. C. Leffel, and Robert W. Howell. 1961. " Genetic Analysis of Energy Production in the Soybean 1 ." *Crop Science* 1 (2): 121–26. https://doi.org/10.2135/cropsci1961.0011183x000100020011x.

Heath, Richard J, and Charles O Rock. 1996. “Roles of the FabA and FabZ-Hydroxyacyl-Acyl Carrier Protein Dehydratases in Escherichia Coli Fatty Acid Biosynthesis* Downloaded From.” *THE JOURNAL OF BIOLOGICAL CHEMISTRY*. Vol. 271. http://www.jbc.org/.

Henchion, Maeve, Maria Hayes, Anne Mullen, Mark Fenelon, and Brijesh Tiwari. 2017. “Future Protein Supply and Demand: Strategies and Factors Influencing a Sustainable Equilibrium.” *Foods* 6 (7): 53. https://doi.org/10.3390/foods6070053.

Henry, C. J.K., P. A. Donachie, and J. P.W. Rivers. 1985. “The Winged Bean. Will the Wonder Crop Be Another Flop?” *Ecology of Food and Nutrition* 16 (4): 331–38. https://doi.org/10.1080/03670244.1985.9990872.

Herath, H. M.W., and D. P. Ormrod. 1979. “Effects of Temperature and Photoperiod on Winged Beans [Psophocarpus Tetragonolobus (L.) D.C.].” *Annals of Botany* 43 (6): 729–36. https://doi.org/10.1093/oxfordjournals.aob.a085686.

Hertel, Thomas W. 2016. “Food Security under Climate Change.” *Nature Climate Change*. Vol. 6. https://doi.org/10.1038/nclimate2834.

Heuzé, V, G Tran, and S Kaushik. 2020. “Soybean Meal. Feedipedia.” 2020. https://feedipedia.org/node/674.

Higuchi, Masako, Yasufumi Fukumoto, and Kazuo Iwai. 1988. “Appearance of Lectin in Winged Bean Pods during Seed Development after Flowering.” *J. Agrlc. Food Chem*. Vol. 36. https://pubs.acs.org/sharingguidelines.

Higuchi, Masako, Masako Suga, and Kazuo Iwai. 1983. “Participation of Lectin in Biological Effects of Raw Winged Bean Seeds on Rats.” *Agricultural and Biological Chemistry* 47 (8): 1879–86. https://doi.org/10.1080/00021369.1983.10865872.

Higuchi, Masako, Junji Terao, and Kazuo Iwai. 1982. “Gas Chromatography-Mass Spectrometric Determination of Fatty Acids in Seed Oil of Winged Bean(Psophocarpus Tetragonolobus (L.) DC)1.” *Journal of Nutritional Science and Vitaminology* 28 (5): 511–18. https://doi.org/10.3177/jnsv.28.511.

Ho, Wai Kuan, Alberto Stefano Tanzi, Fei Sang, Niki Tsoutsoura, Niraj Shah, Christopher Moore, Rahul Bhosale, Victoria Wright, Festo Massawe, and Sean Mayes. 2024. “A Genomic Toolkit for Winged Bean Psophocarpus Tetragonolobus.” *Nature Communications* 15 (1). https://doi.org/10.1038/s41467-024-45048-x.

Hou, Guangyun, Gary R. Ablett, K. Peter Pauls, and Istvan Rajcan. 2006. “Environmental Effects on Fatty Acid Levels in Soybean Seed Oil.” *JAOCS* 83 (9).

Hu, Wei, Xiaoxue Liu, Yajun Xiong, Tingxuan Liu, Zhan Li, Jian Song, Jun Wang, Xianzhi Wang, and Xiaofang Li. 2021. “Nutritional Evaluation and Transcriptome Analyses of Short-Time Germinated Seeds in Soybean (Glycine Max L. Merri.).” *Scientific Reports 2021 11:1* 11 (1): 1–15. https://doi.org/10.1038/s41598-021-02132-2.

Hu, Yilong, Limeng Zhou, Yuhua Yang, Wenbin Zhang, Zhonghui Chen, Xiaoming Li, Qian Qian, et al. 2021. “The Gibberellin Signaling Negative Regulator RGA-LIKE3 Promotes Seed Storage Protein Accumulation.” *Plant Physiology* 185 (4): 1697. https://doi.org/10.1093/PLPHYS/KIAA114.

Huynh, Bao Lam, Jeffrey D. Ehlers, Bevan Emma Huang, María Muñoz-Amatriaín, Stefano Lonardi, Jansen R.P. Santos, Arsenio Ndeve, et al. 2018. “A Multi-Parent Advanced Generation Inter-Cross (MAGIC) Population for Genetic Analysis and Improvement of Cowpea (Vigna Unguiculata L. Walp.).” *Plant Journal* 93 (6): 1129–42. https://doi.org/10.1111/tpj.13827.

Hwang, Eun-Young, Qijian Song, Gaofeng Jia, James E Specht, David L Hyten, Jose Costa, and Perry B Cregan. 2014. “A Genome-Wide Association Study of Seed Protein and Oil Content in Soybean.” *BMC Genomics* 15 (1): 1. https://doi.org/10.1186/1471-2164-15-1.

International Board for Plant Genetic (IBPGR). 1982. *Descriptors_winged_bean_revised.Pdf*. Rome, (Italy).

Jenkins, David J.A., Cyril W.C. Kendall, Livia S.A. Augustin, Sandra Mitchell, Sandhya Sahye-Pudaruth, Sonia Blanco Mejia, Laura Chiavaroli, et al. 2012. “Effect of Legumes as Part of a Low Glycemic Index Diet on Glycemic Control and Cardiovascular Risk Factors in Type 2 Diabetes Mellitus: A Randomized Controlled Trial.” *Archives of Internal Medicine* 172 (21): 1653–60. https://doi.org/10.1001/2013.jamainternmed.70.

Jesú Carrero, Juan, Juristo Fonollá, José Luis Marti, Jesú Jiménez, Julio J Boza, and Eduardo Ló pez-Huertas. 2007. “The Journal of Nutrition Nutrient Requirements and Optimal Nutrition Intake of Fish Oil, Oleic Acid, Folic Acid, and Vitamins B-6 and E for 1 Year Decreases Plasma C-Reactive Protein and Reduces Coronary Heart Disease Risk Factors in Male Patients in a Cardiac Rehabilitation Program 1.” *J. Nutr*. Vol. 137.

Jones, S I, and L O Vodkin. 2013. “Using RNA-Seq to Profile Soybean Seed Development from Fertilization to Maturity.” *PLoS ONE* 8 (3): 59270. https://doi.org/10.1371/journal.pone.0059270.

Jugran, H.M., P. Nath, B.K. Banerji, and S.K. Datta. 1986. “Gamma Ray Induced Dwarf Mutant of Winged Bean.” *Journal of Nuclear Agriculture and Biology* 15 (3): 175–78. https://inis.iaea.org/search/search.aspx?orig_q=RN:19050128.

Kadam, S. S. 1984a. “Winged Bean in Human Nutrition.” *C R C Critical Reviews in Food Science and Nutrition* 21 (1): 1–40. https://doi.org/10.1080/10408398409527395.

———. 1984b. “Winged Bean in Human Nutrition.” *C R C Critical Reviews in Food Science and Nutrition* 21 (1): 1–40. https://doi.org/10.1080/10408398409527395.

Kadam, S. S., L. S. Kute, K. M. Lawande, and D. K. Salunkhe. 1982. “Changes in Chemical Composition of Winged Bean (Psophocarpus Tetragonolobus L.) during Seed Development.” *Journal of Food Science* 47 (6): 2051–53. https://doi.org/10.1111/j.1365-2621.1982.tb12943.x.

Kadam, S. S., and R. R. Smithard. 1987. “Effects of Heat Treatments on Trypsin Inhibitor and Hemagglutinating Activities in Winged Bean.” *Qualitas Plantarum Plant Foods for Human Nutrition* 37 (2): 151–59. https://doi.org/10.1007/BF01092051.

Kadam, Santram S., Ronald R. Smithard, Michael D. Eyre, and David G. Armstrong. 1987. “Effects of Heat Treatments of Antinutritional Factors and Quality of Proteins in Winged Bean.” *Journal of the Science of Food and Agriculture* 39 (3): 267–75. https://doi.org/10.1002/JSFA.2740390310.

Kambhampati, Shrikaar, Jose A. Aznar-Moreno, Cooper Hostetler, Tara Caso, Sally R. Bailey, Allen H. Hubbard, Timothy P. Durrett, and Doug K. Allen. 2020. “On the Inverse Correlation of Protein and Oil: Examining the Effects of Altered Central Carbon Metabolism on Seed Composition Using Soybean Fast Neutron Mutants.” *Metabolites* 10 (1). https://doi.org/10.3390/METABO10010018.

Kantha, S. Sri, and John W. Erdman. 1984a. “The Winged Bean as an Oil and Protein Source: A Review.” *Journal of the American Oil Chemists’ Society*. Vol. 61. https://doi.org/10.1007/BF02677021.

———. 1984b. “The Winged Bean as an Oil and Protein Source: A Review.” *Journal of the American Oil Chemists’ Society* 61 (3): 515–25. https://doi.org/10.1007/BF02677021.

———. 1984c. “The Winged Bean as an Oil and Protein Source: A Review.” *Journal of the American Oil Chemists’ Society* 61 (3): 515–25. https://doi.org/10.1007/BF02677021.

Kantha, Sachi Sri, and John W Erdman. 1986. “Nutrient, Anti-Nutrient Contents and Solubility Profiles of Nitrogen, Phytic Acid and Selected Minerals in Winged Bean Flour.” *Cereal Chemistry* 63 (1): 9–13. https://www.researchgate.net/publication/235727074.

Kantha, Sachi. Sri, and John W. Erdman. 1986. “Winged Bean as a Source of Protein: Recent Advances.” In , 206–19. https://doi.org/10.1021/bk-1986-0312.ch017.

Karikari. 1972. “Pollination Requirements of Winged Beans (Psophocarpus Spp. Neck) in Ghana.” *Ghana Journal of Agricultural Science*.

Karikari, Benjamin, Shuguang Li, Javaid Akhter Bhat, Yongce Cao, Jiejie Kong, Jiayin Yang, Junyi Gai, and Tuanjie Zhao. 2019. “Genome-Wide Detection of Major and Epistatic Effect QTLs for Seed Protein and Oil Content in Soybean Under Multiple Environments Using High-Density Bin Map.” *Int. J. Mol. Sci* 20:979. https://doi.org/10.3390/ijms20040979.

Kesavan, V., and T.N. Khan. 1978. “Induced Mutations in Winged Bean.” In *Seed Protein Improvement Programme*. FAO, Vienna (Austria). Joint FAO/IAEA Div. of Atomic Energy in Food and Agriculture eng; Research Co-ordination Meeting of the Seed Protein Improvement Programme eng 28 Mar 1977 4 Baden (Austria). https://agris.fao.org/agris-search/search.do?recordID=XF2016051592.

Khan, T. N. 1976. “Papua New Guinea: A Centre of Genetic Diversity in Winged Bean (Psophocarpus Tetragonologus (L.) D c.).” *Euphytica* 25 (1): 693–705. https://doi.org/10.1007/BF00041608.

Khan, T. N., and W. Erskine. 1978. “The Adaptation of Winged Bean (Psophocavpm Tetragonolobus (L.) DC.) in P a p a New Guinea.” *Aust. J. Agric. Res*. Vol. 29.

Khor, Hun-Teik, Nget-Hong Tan, and Kai-Choo Wong. 1982a. “The Protein, Trypsin Inhibitor and Lipid of the Winged Bean [Psophocarpus Tetragonolobus (L) DC] Seeds.” *Journal of the Science of Food and Agriculture* 33 (10): 996–1000. https://doi.org/10.1002/jsfa.2740331009.

———. 1982b. “The Protein, Trypsin Inhibitor and Lipid of the Winged Bean [Psophocarpus Tetragonolobus (L) DC] Seeds.” *Journal of the Science of Food and Agriculture* 33 (10): 996–1000. https://doi.org/10.1002/jsfa.2740331009.

Khoury, Colin K., Anne D. Bjorkman, Hannes Dempewolf, Julian Ramirez-Villegas, Luigi Guarino, Andy Jarvis, Loren H. Rieseberg, and Paul C. Struik. 2014. “Increasing Homogeneity in Global Food Supplies and the Implications for Food Security.”

Proceedings of the National Academy of Sciences of the United States of America 111 (11): 4001–6. https://doi.org/10.1073/pnas.1313490111.

Kies, Arie K., Leon H. De Jonge, Paul A. Kemme, and Age W. Jongbloed. 2006. "Interaction between Protein, Phytate, and Microbial Phytase. In Vitro Studies." *Journal of Agricultural and Food Chemistry* 54 (5): 1753–58. https://doi.org/10.1021/jf0518554.

King, Richard D., and Prapasri Puwastein. 1987. "Effects of Germination on the Proximate Composition and Nutritional Quality of Winged Bean (Psophocarpus Tetragonolobus) Seeds." *Journal of Food Science* 52 (1): 106–8. https://doi.org/10.1111/j.1365-2621.1987.tb13982.x.

King, Richard, and Prapasri Puwastien. 1987. "Effects of Germination on the Proximate Composition and Nutritional Quality of Winged Bean (Psophocarpus Tetragonolobus) Seeds." *Journal of Food Science* 52 (1): 106–8. https://doi.org/10.1111/j.1365-2621.1987.tb13982.x.

Kishor Gupta, Raj, Shivraj Singh Gangoliya, Nand Kumar Singh, Raj Kishor Gupta, Shivraj Singh Gangoliya, and Nand Kumar Singh. 2013. "Reduction of Phytic Acid and Enhancement of Bioavailable Micronutrients in Food Grains." *Journal of Food Science and Technology* 52 (2): 676–84. https://doi.org/10.1007/s13197-013-0978-y.

Klu, G. Y.P., E. Jacobsen, and A. M. Van Harten. 1997. "Induced Mutations in Winged Bean (Psophocarpus Tetragonolobus L. DC) with Low Tannin Content." *Euphytica* 98 (1–2): 99–107. https://doi.org/10.1023/a:1003032408885.

Kortt, Alexander A. 1979. "Isolation and Characterization of the Trypsin Inhibitors from Winged Bean Seed (Psophocarpus Tetragonolobus (L) Dc.)." *Biochimica et Biophysica Acta* 577 (2): 271–82. https://doi.org/10.1016/0005-2795(79)90040-0.

———. 1980. *Isolation and Properties of a Chymotrypsin Inhibitor from Winged Bean Seed (Psophocarpus Tetragonolobus (L) Dc.). Biochimica et Biophysica Acta*. Vol. 624. Elsevier. https://doi.org/10.1016/0005-2795(80)90243-3.

———. 1981. "Specificity and Stability of the Chymotrypsin Inhibitor from Winged Bean Seed (Psophocarpus Tetragonolobus (L) Dc.)." *BBA - Enzymology*. Vol. 657. https://doi.org/10.1016/0005-2744(81)90145-5.

———. 1983. "Comparative Studies on the Storage Proteins and Anti-Nutritional Factors from Seeds of Psophocarpus Tetragonolobus (L.) DC from Five South-East Asian Countries."

Qualitas Plantarum Plant Foods for Human Nutrition 33 (1): 29–40. https://doi.org/10.1007/BF01093735.

———. 1984. "Purification and Properties of the Basic Lectins from Winged Bean Seed [Psophocarpus Tetragonolobus(L.) DC]." *European Journal of Biochemistry* 138 (3): 519–25. https://doi.org/10.1111/j.1432-1033.1984.tb07946.x.

———. 1986. "Isolation and Characterisation of a Major Seed Albumin." *International Journal of Peptide and Protein Research* 28 (6): 613–19. https://doi.org/10.1111/j.1399-3011.1986.tb03299.x.

———. n.d. "Isolation and Properties of a Chymotrypsin Inhibitor from Winged Bean Seed (Psophocarpus Tetragonolobus(L) Dc.) - ScienceDirect." Accessed April 13, 2020. https://www-sciencedirect-com.ezproxy.nottingham.ac.uk/science/article/abs/pii/0005279580902433.

Kortt, Alexander A., and J. Bruce Caldwell. 1984. "Characteristics of the Proteins of the Tubers of Winged Bean (Psophocarpus Tetragonolobus (L.) DC)." *Journal of the Science of Food and Agriculture*. Vol. 35. https://doi.org/10.1002/jsfa.2740350310.

Kotaru, Makoto, Tsuneo Ikeuchi, Hideki Yoshikawa, and Fumio Ibuki. 1987. "Investigations of Antinutritional Factors of the Winged Bean (Psophocarpus Tetragonolobus)." *Food Chemistry*. Vol. 24.

Kurpad, A. V. 2013. "Protein: Quality and Sources." *Encyclopedia of Human Nutrition* 4–4 (January):123–30. https://doi.org/10.1016/B978-0-12-375083-9.00241-5.

Kwan Lam, Sze, and Tzi Bun Ng. 2011. "Lectins: Production and Practical Applications." *Appl Microbiol Biotechnol*, no. 89, 45–55. https://doi.org/10.1007/s00253-010-2892-9.

Lagarda-Diaz, Irlanda, Ana Maria Guzman-Partida, and Luz Vazquez-Moreno. 2017. "Molecular Sciences Legume Lectins: Proteins with Diverse Applications." *International Journal of Molecular Sciences*, 2–18. https://doi.org/10.3390/ijms18061242.

Lepcha, Patrush, Ashley N. Egan, Jeff J. Doyle, and N. Sathyanarayana. 2017a. "A Review on Current Status and Future Prospects of Winged Bean (Psophocarpus Tetragonolobus) in Tropical Agriculture." *Plant Foods for Human Nutrition*. https://doi.org/10.1007/s11130-017-0627-0.

———. 2017b. "A Review on Current Status and Future Prospects of Winged Bean (Psophocarpus Tetragonolobus) in Tropical Agriculture." *Plant Foods for Human Nutrition*. Springer New York LLC. https://doi.org/10.1007/s11130-017-0627-0.

———. 2017c. “A Review on Current Status and Future Prospects of Winged Bean (Psophocarpus Tetragonolobus) in Tropical Agriculture.” *Plant Foods for Human Nutrition* 72 (3): 225–35. https://doi.org/10.1007/s11130-017-0627-0.

Li, Maoteng, Sarah Veronica Schiessl, Liezhao Liu, Shinje Kim, Ruijuan Li, Kwangju Jeong, John T Davis, et al. 2018a. “Integrated QTL and EQTL Mapping Provides Insights and Candidate Genes for Fatty Acid Composition, Flowering Time, and Growth Traits in a F 2 Population of a Novel Synthetic Allopolyploid Brassica Napus.” https://doi.org/10.3389/fpls.2018.01632.

———. 2018b. “Integrated QTL and EQTL Mapping Provides Insights and Candidate Genes for Fatty Acid Composition, Flowering Time, and Growth Traits in a F 2 Population of a Novel Synthetic Allopolyploid Brassica Napus.” https://doi.org/10.3389/fpls.2018.01632.

Lis, H, and N Sharon. 1986. “Lectins as Molecules and as Tools.” *Annual Review of Biochemistry* 55 (1): 35–67. https://doi.org/10.1146/annurev.bi.55.070186.000343.

Liu, Nian, Jing Liu, Shihang Fan, Hongfang Liu, Xue-Rong Zhou, Wei Hua, and Ming Zheng. 2022. “An Integrated Omics Analysis Reveals the Gene Expression Profiles of Maize, Castor Bean, and Rapeseed for Seed Oil Biosynthesis | Enhanced Reader.” *BMC Plant Biology*. https://doi.org/https://doi.org/10.1186/s12870-022-03495-y.

Lopez-Corona, Alejandra Vanessa, Illeen Valencia-Espinosa, Fabio Antonio González-Sánchez, Angélica Lizeth Sánchez-López, Luis Eduardo Garcia-Amezquita, and Rebeca Garcia-Varela. 2022. “Antioxidant, Anti-Inflammatory and Cytotoxic Activity of Phenolic Compound Family Extracted from Raspberries (Rubus Idaeus): A General Review.” *Antioxidants*. MDPI. https://doi.org/10.3390/antiox11061192.

Lu, Nan, Xiaolan Rao, Ying Li, Ji Hyung Jun, and Richard A Dixon. 2021. “Dissecting the Transcriptional Regulation of Proanthocyanidin and Anthocyanin Biosynthesis in Soybean (Glycine Max).” *Plant Biotechnology*. https://doi.org/10.1111/pbi.13562.

Lumen, Benito O. de, Seham Fiad, and Seham Fiad. 1982. “Tocopherols of Winged Bean (Psophocarpus Tetragonolobus) Oil.” *Journal of Agricultural and Food Chemistry* 30 (1): 50–53. https://doi.org/10.1021/jf00109a010.

Lumen, Benito O. De, and Lourdes A. Salamat. 1980. “Trypsin Inhibitor Activity in Winged Bean (Psophocarpus Tetragonolobus) and the Possible Role of Tannin.” *Journal of Agricultural and Food Chemistry* 28 (3): 533–36. https://doi.org/10.1021/jf60229a042.

Lumen, Benitto O. De, Amelia L. Gerpacio, and Pran Vohra. 1982. “Effects of Winged Bean (Psophocarpus Tetragonolobus) Meal on Broiler Performance.” *Poultry Science* 61 (6): 1099–1106. https://doi.org/10.3382/ps.0611099.

Luo, Lilan, Yajun He, Yannan Zhao, Qian Xu, Jian Wu, Haiyan Ma, Hongyan Guo, et al. 2019. “Regulation of Mitochondrial NAD Pool via NAD+ Transporter 2 Is Essential for Matrix NADH Homeostasis and ROS Production in Arabidopsis.” *Science China Life Sciences* 62 (8): 991–1002. https://doi.org/10.1007/s11427-019-9563-y.

Luo, Ping, Guogui Ning, Zhen Wang, Yuxiao Shen, Huanan Jin, Penghui Li, Shasha Huang, Jian Zhao, and Manzhu Bao. 2016. “Disequilibrium of Flavonol Synthase and Dihydroflavonol-4-Reductase Expression Associated Tightly to White vs. Red Color Flower Formation in Plants.” *Frontiers in Plant Science* 6 (JAN2016). https://doi.org/10.3389/fpls.2015.01257.

Ma, Ruirui, Wenxuan Huang, Quan Hu, Guo Tian, Jie An, Ting Fang, Jia Liu, Jingjing Hou, Meixia Zhao, and Lianjun Sun. 2024. “Tandemly Duplicated MYB Genes Are Functionally Diverged in the Regulation of Anthocyanin Biosynthesis in Soybean.” *Plant Physiology* 194 (4): 2549–63. https://doi.org/10.1093/PLPHYS/KIAE019.

Makeri, Mohammad Usman, Sabo Abdulkarim Mohamed, Roselina Karim, Yogeshini Ramakrishnan, and Kharidah Muhammad. 2017. “Fractionation, Physicochemical, and Structural Characterization of Winged Bean Seed Protein Fractions with Reference to Soybean.” *International Journal of Food Properties* 20 (December):2220–36. https://doi.org/10.1080/10942912.2017.1369101.

Makkar, H. P. S. 2003a. “Measurement of Total Phenolics and Tannins Using Folin-Ciocalteu Method.” https://doi.org/10.1007/978-94-017-0273-7_3.

Makkar, H. P.S. 2003b. “Effects and Fate of Tannins in Ruminant Animals, Adaptation to Tannins, and Strategies to Overcome Detrimental Effects of Feeding Tannin-Rich Feeds.” In *Small Ruminant Research*, 49:241–56. Elsevier. https://doi.org/10.1016/S0921-4488(03)00142-1.

Maphosa, Yvonne, and Victoria A Jideani. 2017. “The Role of Legumes in Human Nutrition.” https://doi.org/10.5772/intechopen.69127.

Martineau-Côté, Delphine, Allaoua Achouri, Mélanie Pitre, Janitha Wanasundara, Salwa Karboune, and Lamia L’Hocine. 2023. “Investigation of the Nutritional Quality of Raw and Processed Canadian Faba Bean (Vicia Faba L.) Flours in Comparison to Pea and Soy Using

a Human in Vitro Gastrointestinal Digestion Model." *Food Research International* 173 (November):113264. https://doi.org/10.1016/J.FOODRES.2023.113264.

Massawe, Festo, Sean Mayes, and Acga Cheng. 2016. "Crop Diversity: An Unexploited Treasure Trove for Food Security." *Trends in Plant Science* 21 (5): 365–68. https://doi.org/10.1016/j.tplants.2016.02.006.

Maxted, Nigel. 1990. "A Phenetic Investigation of Psophocarpus Neck ." *Botanical Journal Ofthe Linnean Society*, no. 102, 103–22.

Mayes, S., F. J. Massawe, P. G. Alderson, J. A. Roberts, S. N. Azam-Ali, and M. Hermann. 2012. "The Potential for Underutilized Crops to Improve Security of Food Production." *Journal of Experimental Botany* 63 (3): 1075–79. https://doi.org/10.1093/jxb/err396.

Megazyme. 2019. "Phytic Acid (Phytate), Total Phosphorus Measured as Phosphorus Released by Phytase and Alkaline Phosphatase. Assay Procedure." www.megazyme.com.

Mensink, Ronald P, and Martijn B Katan. 1989. "Effect of a Diet Enriched with Monounsaturated or Polyunsaturated Fatty Acids on Levels of Low-Density and High-Density Lipoprotein Cholesterol in Healthy Women and Men." *New England Journal of Medicine* 321 (7): 436–41. https://doi.org/10.1056/NEJM198908173210705.

Messadi, D V, P Billings, G Shklar, and A R Kennedy. 1986. "Inhibition of Oral Carcinogenesis by a Protease Inhibitor." *Journal of the National Cancer Institute* 76 (3): 447–52. http://www.ncbi.nlm.nih.gov/pubmed/3081747.

Mi, Chao, Yanning Zhao, Xuetao Yang, Liangbin Lin, and Jinxiong Wang. 2025. "Effect of Low Nighttime Temperature on Oil Accumulation of Rapeseed Seeds (Brassica Napus L.) Based on RNA-Seq of Silique Wall Tissue." *Agriculture (Switzerland)* 15 (6). https://doi.org/10.3390/agriculture15060576.

Mnembuka, B V, and B O Eggum. 1995. "Comparative Nutritive Value of Winged Bean (Psophocarpus Tetragonolobus (L) DC) and Other Legumes Grown in Tanzania." *Plant Foods for Human Nutrition*. Vol. 47.

Mohanty, Chandra Sekhar, Rama Chandra Pradhan, Vinayak Singh, Neha Singh, Rojalin Pattanayak, Om Prakash, Chandan Singh Chanotiya, and Prasant Kumar Rout. 2015a. "Physicochemical Analysis of Psophocarpus Tetragonolobus (L.) DC Seeds with Fatty Acids and Total Lipids Compositions." *Journal of Food Science and Technology* 52 (6): 3660–70. https://doi.org/10.1007/s13197-014-1436-1.

———. 2015b. “Physicochemical Analysis of Psophocarpus Tetragonolobus (L.) DC Seeds with Fatty Acids and Total Lipids Compositions.” *Journal of Food Science and Technology* 52 (6): 3660–70. https://doi.org/10.1007/s13197-014-1436-1.

Mohanty, Chandra Sekhar, Vinayak Singh, and Mark A. Chapman. 2020. “Winged Bean: An Underutilized Tropical Legume on the Path of Improvement, to Help Mitigate Food and Nutrition Security.” *Scientia Horticulturae*. Elsevier B.V. https://doi.org/10.1016/j.scienta.2019.108789.

Mohanty, Chandra Sekhar, Naziya Syed, Deepak Kumar, Sonam Khare, Sagar Prasad Nayak, Kunwar Sarvendra, Rojalin Pattanayak, Anirban Pal, Chandan Singh Chanotiya, and Prasant Kumar Rout. 2021. “Chemical Characterization of Winged Bean (Psophocarpus Tetragonolobus (L.) DC. Seeds and Safety Evaluation of Its Fatty Oil.” *Journal of Food Measurement and Characterization* 15 (1): 807–16. https://doi.org/10.1007/s11694-020-00680-1.

Mohanty, Sekhar Chandra, Vinayak Singh, Priya Gupta, Anukool Srivsatava, Rojalin Pattanayak, Vandana Jaiswal, and Nilamani Dikshit. 2019. “Estimation of Genetic Diversity, Structure and Trait Association of Winged Bean (Psophocarpus Tetragonolobus (L.) DC.), Genotypes through AFLP and ITS Markers.” *Indian Journal of Biotechnology*. Vol. 18. www.biodiversityinternational.org.

Mohanty, Sekhar Chandra, Sushma Verma, Vinayak Singh, Shahina Khan, Priyanka Gaur, Priya Gupta, M Abdul Nizar, et al. 2014. “Characterization of Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) Based on Molecular, Chemical and Physiological Parameters.” *American Journal of Molecular Biology* 3:187–97. https://doi.org/10.4236/ajmb.2013.34025.

Molina, Antonio, Eva Miedes, Laura Bacete, Tinguaro Rodríguez, Hugo Mélida, Nicolas Denancé, Andrea Sánchez-Vallet, et al. 2021. “Arabidopsis Cell Wall Composition Determines Disease Resistance Specificity and Fitness.” *Proceedings of the National Academy of Sciences of the United States of America* 118 (5). https://doi.org/10.1073/PNAS.2010243118/-/DCSUPPLEMENTAL.

Montesinos-López, Osval Antonio, Abelardo Montesinos-López, Paulino Pérez-Rodríguez, José Alberto Barrón-López, Johannes W.R. Martini, Silvia Berenice Fajardo-Flores, Laura S. Gaytan-Lugo, Pedro C. Santana-Mancilla, and José Crossa. 2021. “A Review of Deep

Learning Applications for Genomic Selection." *BMC Genomics*. BioMed Central Ltd. https://doi.org/10.1186/s12864-020-07319-x.

Mosse, J, and J Baudet. 1983. "Crude Protein Content and Aminoacid Composition of Seeds: Variability and Correlations ." *Qual Plant Plant Foods Hum Nutr*. Vol. 32.

Mossé, Jacques. 1990. "Nitrogen to Protein Conversion Factor for Ten Cereals and Six Legumes or Oilseeds. A Reappraisal of Its Definition and Determination. Variation According to Species and to Seed Protein Content." *Journal of Agricultural and Food Chemistry* 38 (1): 18–24. https://doi.org/10.1021/jf00091a004.

Muleya, Molly, Dongfang Li, Gabriella Chiutsi-Phiri, Lester Botoman, John M. Brameld, and Andrew M. Salter. 2023. "In Vitro Determination of the Protein Quality of Maize Varieties Cultivated in Malawi Using the INFOGEST Digestion Method." *Heliyon* 9 (9). https://doi.org/10.1016/j.heliyon.2023.e19797.

Muzquiz, Mercedes, Alejandro Varela, Carmen Burbano, Carmen Cuadrado, Eva Guillamón, and Mercedes M. Pedrosa. 2012. "Bioactive Compounds in Legumes: Pronutritive and Antinutritive Actions. Implications for Nutrition and Health." *Phytochemistry Reviews* 11 (2–3): 227–44. https://doi.org/10.1007/s11101-012-9233-9.

Nasi, Antonella, Gianluca Picariello, and Pasquale Ferranti. 2009. "Proteomic Approaches to Study Structure, Functions and Toxicity of Legume Seeds Lectins. Perspectives for the Assessment of Food Quality and Safety." *Journal of Proteomics* 72 (3): 527–38. https://doi.org/10.1016/j.jprot.2009.02.001.

National Academy of Science. 1981. "The Winged Bean: High-Protein Crop for the Humid Tropics (BOSTID, 1981, 41 p.): 2 Agronomy." 1981. http://www.nzdl.org/gsdlmod?e=d-00000-00---off-0cdl--00-0----0-10-0---0---0direct-10---4-------0-1l--11-en-50---20-about---00-0-1-00-0--4----0-0-11-10-0utfZz-8-00&a=d&c=cdl&cl=CL2.23&d=HASH01dc8a536327ee6eb7e15196.5#HASH01dc8a536327ee6eb7e15196.5.

National Research Council. 1981. *Winged Bean: A High-Protein Crop for the Tropics. Winged Bean*. Washington, DC: The National Academies Press. https://doi.org/10.17226/19754.

Ne Triboi, Eugè, and Anne-Marie Triboi-Blondel. 2002. "Productivity and Grain or Seed Composition: A New Approach to an Old Problem-Invited Paper." *European Journal of Agronomy*. Vol. 16. www.elsevier.com/locate/eja.

O'Fallon, J. V., J. R. Busboom, M. L. Nelson, and C. T. Gaskins. 2007. "A Direct Method for Fatty Acid Methyl Ester Synthesis: Application to Wet Meat Tissues, Oils, and Feedstuffs." *Journal of Animal Science* 85 (6): 1511–21. https://doi.org/10.2527/jas.2006-491.

Ojuederie, Omena Bernard, Nkoyo Ani Nkang, Kafilat Abiodun Odesola, and David Okeh Igwe. 2020. "Genetic Diversity Assessment of Winged Bean (Pso-Phorcarpus Tetragonolobus) Accessions Revealed by Inter-Simple Sequence Repeat (ISSR) Markers." *Journal of Plant Biology and Crop Research* 3 (1): 1014. http://meddocsonline.org/.

Okezie, B. O., and F. W Martin. 1980. "Chemical Composition of Dry Seeds and Fresh Leaves of Winged Bean Varieties Grown in the U.S. and Puerto Rico." *Journal of Food Science* 45 (4): 1045–51. https://doi.org/10.1111/j.1365-2621.1980.tb07509.x.

Okuley, John, Jonathan Lightner, Kenneth Feldmann, Narendra Yadav, Ellen Lark, and John Browseai'. 1994. "Arabidopsis FAD2 Gene Encodes the Enzyme That 1s Essential for Polyunsaturated Lipid Synthesis." *The Plant Cell*. Vol. 6. https://academic.oup.com/plcell/article/6/1/147/5984849.

Oluwatos, Olusola Bayo. 1999. "Genotype × Environment Influence on Cowpea (Vigna Unguiculata (L) Walp) Antinutritional Factors: 1 – Trypsin Inhibitors, Tannins, Phytic Acid and Haemagglutinin." Journal of Science of Food and Agriculture. 1999. https://scijournals.onlinelibrary.wiley.com/doi/epdf/10.1002/%28SICI%291097-0010%28199902%2979%3A2%3C265%3A%3AAID-JSFA190%3E3.0.CO%3B2-Z?saml_referrer.

Omode, Ayo A, Olalekan S Fatoki, K A 01aogun+, and K A Olaogun. 1995. "Physicochemical Properties of Some Underexploited and Nonconventional Oilseeds." *Food Chem*. Vol. 43. https://pubs.acs.org/sharingguidelines.

Ooijen, J W Van. 2009. "MapQTL 6." *Genome*, no. April.

Oomah, B Dave, François Caspar, Linda J Malcolmson, and Anne-Sophie Bellido. 2010. "Phenolics and Antioxidant Activity of Lentil and Pea Hulls." https://doi.org/10.1016/j.foodres.2010.09.027.

Pallares, Pallares Andrea, Shannon Gwala, Katharina Pälchen, Dorine Duijsens, Marc Hendrickx, and Tara Grauwet. 2021. "Pulse Seeds as Promising and Sustainable Source of Ingredients with Naturally Bioencapsulated Nutrients: Literature Review and Outlook." *Comprehensive Reviews in Food Science and Food Safety* 20 (2): 1524–53. https://doi.org/10.1111/1541-4337.12692.

Patil, Gunvant, Rouf Mian, · Tri Vuong, Vince Pantalone, Qijian Song, · Pengyin Chen, · Grover, et al. 2017. “Molecular Mapping and Genomics of Soybean Seed Protein: A Review and Perspective for the Future.” *Theoretical and Applied Genetics* 130:1975–91. https://doi.org/10.1007/s00122-017-2955-8.

Peng, Li, Linlin Qian, Meinan Wang, Wei Liu, Xiangting Song, Hao Cheng, Fengjie Yuan, and Man Zhao. 2021. “Comparative Transcriptome Analysis during Seeds Development between Two Soybean Cultivars.” *PeerJ* 9 (March):e10772. https://doi.org/10.7717/PEERJ.10772/SUPP-3.

Peumans, Willy J, and Eis J M Van Damme. 1995. “Lectins as PIant Defense Proteins’.” *Plant Physiol*. Vol. 1.

PFAF. 2021. “PFAF: Plants For A Future.” 2021. https://pfaf.org/User/Plant.aspx?LatinName=Psophocarpus+tetragonolobus.

Phokas, Alexandros, and Juliet C. Coates. 2021. “Evolution of DELLA Function and Signaling in Land Plants.” *Evolution & Development* 23 (3): 137. https://doi.org/10.1111/EDE.12365.

“Policy Brief Changing Policy Concepts of Food Security.” 2006. http://www.foodsecinfoaction.org/.

Poole, M. M. 1978. “Pollen Morphology of Psophocarpus (Leguminosae) in Relation to Its Taxonomy.” *Kew Bulletin* 34 (2): 211–20.

Porter, John R., Liyong Xie, Andrew J. Challinor, Kevern Cochrane, S. Mark Howden, Muhammad Mohsin Iqbal, David B. Lobell, et al. 2015. “Food Security and Food Production Systems.” *Climate Change 2014 Impacts, Adaptation and Vulnerability: Part A: Global and Sectoral Aspects*. https://doi.org/10.1017/CBO9781107415379.012.

Prakash, Dhan, P N Misra, and P S Misra. 1987a. “Amino Acid Profile of Winged Bean (Psophocarpus Tetragonolobus (L.) DC): A Rich Source of Vegetable Protein.” *Plant Foods for Human Nutrition*. Vol. 37.

———. 1987b. “Amino Acid Profile of Winged Bean (Psophocarpus Tetragonolobus (L.) DC): A Rich Source of Vegetable Protein.” *Plant Foods for Human Nutrition*. Vol. 37.

Prohens, Jaime, Pietro Gramazio, Mariola Plazas, Hannes Dempewolf, Benjamin Kilian, María J. Díez, Ana Fita, et al. 2017. “Introgressiomics: A New Approach for Using Crop Wild Relatives in Breeding for Adaptation to Climate Change.” *Euphytica* 213 (7): 158. https://doi.org/10.1007/s10681-017-1938-9.

Qi, Zhaoming, Zhanguo Zhang, Zhongyu Wang, Jingyao Yu, Hongtao Qin, Xinrui Mao, Hongwei Jiang, et al. 2018. “Meta-Analysis and Transcriptome Profiling Reveal Hub Genes for Soybean Seed Storage Composition during Seed Development.” https://doi.org/10.1111/pce.13175.

Quan, Miao-hua, Xiang-hui, Jiang, Dong-ming Chen, Xiang-yang Lu, and Yun Tian. 2011. “Identification and Analysis of A New Dwarf Mutant of Winged Bean.” 2011. https://en.cnki.com.cn/Article_en/CJFDTotal-ZGSC201108011.htm.

Raai, Murthazar Naim, Nurul Amalina Mohd Zain, Normaniza Osman, Nur Ardiyana Rejab, Nurul Amylia Sahruzaini, and Acga Cheng. 2020. “Effects of Shading on the Growth, Development and Yield of Winged Bean (Psophocarpus Tetragonolobus).” *Ciencia Rural* 50 (2): 1–7. https://doi.org/10.1590/0103-8478cr20190570.

Raboy, Victor, Kevin A Young, John A Dorsch, and Allen Cook. 2001. “Genetics and Breeding of Seed Phosphorus and Phytic Acid.” *J. Plant Physiol*. Vol. 158. http://www.urbanfischer.de/journals/jpp.

Rahman, M. M., M. A. Islam, M. S. Azirun, and A. N. Boyce. 2014. “Agronomic and Nitrogen Recovery Efficiency of Rice under Tropical Conditions as Affected by Nitrogen Fertilizer and Legume Crop Rotation.” *Journal of Animal and Plant Sciences* 24 (3): 891–96.

Rambaran, Theresa F. 2020. “Nanopolyphenols: A Review of Their Encapsulation and Anti-Diabetic Effects.” *SN Applied Sciences*. Springer Nature. https://doi.org/10.1007/s42452-020-3110-8.

Ray, Deepak K., Paul C. West, Michael Clark, James S. Gerber, Alexander V. Prishchepov, and Snigdhansu Chatterjee. 2019. “Climate Change Has Likely Already Affected Global Food Production.” *PLoS ONE*. https://doi.org/10.1371/journal.pone.0217148.

Reddy, P. Parvatha, and P. Parvatha Reddy. 2015. “Winged Bean, Psophocarpus Tetragonolobus.” In *Plant Protection in Tropical Root and Tuber Crops*, 293–303. Springer India. https://doi.org/10.1007/978-81-322-2389-4_10.

Reverdatto, Sergei, Vadim Beilinson, and Niels C Nielsen. 1999. “A Multisubunit Acetyl Coenzyme A Carboxylase from Soybean 1.” https://academic.oup.com/plphys/article/119/3/961/6087527.

Rodriguez, Jesse, Nitin Gupta, Richard D. Smith, and Pavel A. Pevzner. 2008. “Does Trypsin Cut before Proline?” *Journal of Proteome Research* 7 (1): 300–305. https://doi.org/10.1021/PR0705035/SUPPL_FILE/PR0705035-FILE002.PDF.

Romano, Nicholas, and Vikas Kumar. 2018. “Phytase in Animal Feed.” *Enzymes in Human and Animal Nutrition: Principles and Perspectives*, January, 73–88. https://doi.org/10.1016/B978-0-12-805419-2.00004-6.

Romano, Raffaele, Gioacchino Filosa, Fabiana Pizzolongo, Alessandra Durazzo, Massimo Lucarini, Patricia Severino, Eliana B. Souto, and Antonello Santini. 2021. “Oxidative Stability of High Oleic Sunflower Oil during Deep-Frying Process of Purple Potato Purple Majesty.” *Heliyon* 7 (3). https://doi.org/10.1016/j.heliyon.2021.e06294.

Romano, Raffaele, Nadia Manzo, Laura Le Grottaglie, Anella Giordano, Annalisa Romano, and Paolo Masi. 2013. “Comparison of the Frying Performance of High Oleic Oils Subjected to Discontinuous and Prolonged Thermal Treatment.” *JAOCS, Journal of the American Oil Chemists’ Society* 90 (7): 965–75. https://doi.org/10.1007/s11746-013-2238-8.

Rombolá-Caldentey, Belén, Paloma Rueda-Romero, Raquel Iglesias-Fernández, Pilar Carbonero, and Luis Oñate-Sánchez. 2014. “Arabidopsis DELLA and Two HD-ZIP Transcription Factors Regulate GA Signaling in the Epidermis through the L1 Box Cis-Element.” *The Plant Cell* 26 (7): 2905–19. https://doi.org/10.1105/TPC.114.127647.

Rotundo, José L., and Mark E. Westgate. 2009. “Meta-Analysis of Environmental Effects on Soybean Seed Composition.” *Field Crops Research* 110 (2): 147–56. https://doi.org/10.1016/j.fcr.2008.07.012.

Ruiz-Aquino, Faustino, Rossy Feria-Reyes, José Guadalupe Rutiaga-Quiñones, Luis Humberto Robledo-Taboada, and Rosalío Gabriel-Parra. 2023. “Characterization of Tannin Extracts Derived from the Bark of Four Tree Species by HPLC and FTIR.” *Forest Science and Technology* 19 (1): 38–46. https://doi.org/10.1080/21580103.2023.2166593.

Saadi, Sami, Nazamid Saari, Hasanah Mohd Ghazali, and Mohammed Sabo Abdulkarim. 2022. “Mitigation of Antinutritional Factors and Protease Inhibitors of Defatted Winged Bean-Seed Proteins Using Thermal and Hydrothermal Treatments: Denaturation/Unfolding Coupled Hydrolysis Mechanism.” *Current Research in Food Science* 5 (January):207–21. https://doi.org/10.1016/j.crfs.2022.01.011.

Sacks, Frank M., Alice H. Lichtenstein, Jason H.Y. Wu, Lawrence J. Appel, Mark A. Creager, Penny M. Kris-Etherton, Michael Miller, et al. 2017. “Dietary Fats and Cardiovascular Disease: A Presidential Advisory from the American Heart Association.” *Circulation* 136 (3): e1–23. https://doi.org/10.1161/CIR.0000000000000510.

Saigo, Tomoki, Tong Wang, Mutsumi Watanabe, and Takayuki Tohge. 2020. “Diversity of Anthocyanin and Proanthocyanin Biosynthesis in Land Plants.” *Current Opinion in Plant Biology* 55 (June):93–99. https://doi.org/10.1016/J.PBI.2020.04.001.

Saio, K., Y. Nakano, and S Uemoto. 1983. “Microstructure of Winged Beans.” *Food Structure* 2 (2): 175–76.

Salunkhe, D. K., and J. K. Chavan. 1990. “Dietary Tannins: Consequences and Remedies - D. K. Salunkhe, J. K. Chavan - Google Books.” 1990. https://books.google.co.uk/books?printsec=frontcover&vid=ISBN0849368111&redir_esc=y#v=onepage&q=condensed tannin&f=false.

Samtiya, Mrinal, Rotimi E. Aluko, and Tejpal Dhewa. 2020a. “Plant Food Anti-Nutritional Factors and Their Reduction Strategies: An Overview.” *Food Production, Processing and Nutrition* 2 (1): 1–14. https://doi.org/10.1186/s43014-020-0020-5.

Samtiya, Mrinal, Rotimi E Aluko, and Tejpal Dhewa. 2020b. “Plant Food Anti-Nutritional Factors and Their Reduction Strategies: An Overview.” *Food Production, Processing and Nutrition*. https://doi.org/10.1186/s43014-020-0020-5.

Sánchez-Chino, Xariss, Cristian Jiménez-Martínez, Gloria Dávila-Ortiz, Isela Álvarez-González, and Eduardo Madrigal-Bujaidar. 2015. “Nutrient and Nonnutrient Components of Legumes, and Its Chemopreventive Activity: A Review.” *Nutrition and Cancer* 67 (3): 401–10. https://doi.org/10.1080/01635581.2015.1004729.

Savelkoul, F H M G, A F B Van Der Poel, and S Tamminga. 1992. “The Presence and Inactivation of Trypsin Inhibitors, Tannins, Lectins and Amylase Inhibitors in Legume Seeds during Germination. A Review.” *Plant Foods for Human Nutrition*. Vol. 42. Kluwer Academic Publishers.

Schlueter, Jessica A., Iryna F. Vasylenko-Sanders, Shweta Deshpande, Jing Yi, Mojesta Siegfried, Bruce A. Roe, Shannon D. Schlueter, Brian E. Scheffler, and Randy C. Shoemaker. 2007. “The FAD2 Gene Family of Soybean: Insights into the Structural and Functional Divergence of a Paleopolyploid Genome.” *Crop Science* 47 (SUPPL. 1). https://doi.org/10.2135/cropsci2006.06.0382tpg.

Schmid, Markus, Timothy S. Davison, Stefan R. Henz, Utz J. Pape, Monika Demar, Martin Vingron, Bernhard Schölkopf, Detlef Weigel, and Jan U. Lohmann. 2005. “A Gene Expression Map of Arabidopsis Thaliana Development.” *Nature Genetics* 37 (5): 501–6. https://doi.org/10.1038/ng1543.

Sekhar Mohanty, Chandra, Naziya Syed, · Deepak Kumar, · Sonam Khare, · Sagar, Prasad Nayak, Kunwar Sarvendra, et al. 2021. “Chemical Characterization of Winged Bean (Psophocarpus Tetragonolobus (L.) DC. Seeds and Safety Evaluation of Its Fatty Oil” 15:807–16. https://doi.org/10.1007/s11694-020-00680-1.

Serrano, José, Riitta Puupponen-Pimiä, Andreas Dauer, Anna Marja Aura, and Fulgencio Saura-Calixto. 2009. “Tannins: Current Knowledge of Food Sources, Intake, Bioavailability and Biological Effects.” *Molecular Nutrition and Food Research* 53 (SUPPL. 2): S310–29. https://doi.org/10.1002/mnfr.200900039.

Shafii, Bahman, Karen A. Mahler, William J. Price, and Dick L. Auld. 1992. “Genotype × Environment Interaction Effects on Winter Rapeseed Yield and Oil Content.” *Crop Science* 32 (4): 922–27. https://doi.org/10.2135/cropsci1992.0011183x003200040017x.

Sharon, Nathan, and Halina Lis. 1990. “Legume Lectins — a Large Family of Homologous Proteins.” *The FASEB Journal* 4 (14): 3198–3208. https://doi.org/10.1096/fasebj.4.14.2227211.

———. 2004. “History of Lectins: From Hemagglutinins to Biological Recognition Molecules.” *Glycobiology* 14 (11): 53–62. https://doi.org/10.1093/glycob/cwh122.

Shewry, Peter R, Johnathan A Napier, and Arthur S Tatham. 1995. “Seed Storage Proteins: Structures ’and Biosynthesis.” *The Plant Cell*. Vol. 7. American Society of Plant Physiologists.

Shibata, Hiroshi, Saburo Hara, Tokuji Ikenaka, and Jiro Abe. 1986. “Purification and Characterization of Proteinase Inhibitors from Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) Seeds.” *Journal of Biochemistry* 99 (4): 1147–55. https://doi.org/10.1093/oxfordjournals.jbchem.a135578.

Silva, Luiz Cláudio Costa, Loreta Buuda da Matta, Guilherme Ramos Pereira, Rafael Delmond Bueno, Newton Deniz Piovesan, Andrea J. Cardinal, Pedro Ivo Vieira Good God, Cleberson Ribeiro, and Maximiller Dal-Bianco. 2021. “Association Studies and QTL Mapping for Soybean Oil Content and Composition.” *Euphytica* 217 (2). https://doi.org/10.1007/s10681-020-02755-y.

Simons, P. C. M., H. A. J. Versteegh, A. W. Jongbloed, P. A. Kemme, P. Slump, K. D. Bos, M. G. E. Wolters, R. F. Beudeker, and G. J. Verschoor. 1990. “Improvement of Phosphorus

Availability by Microbial Phytase in Broilers and Pigs." *British Journal of Nutrition* 64 (2): 525–40. https://doi.org/10.1079/bjn19900052.

Singh, Balwinder, Jatinder Pal Singh, Amritpal Kaur, and Narpinder Singh. 2016. "Bioactive Compounds in Banana and Their Associated Health Benefits - A Review." *Food Chemistry*. Elsevier Ltd. https://doi.org/10.1016/j.foodchem.2016.03.033.

Singh, Balwinder, Jatinder Pal Singh, Khetan Shevkani, Narpinder Singh, and Amritpal Kaur. 2017. "Bioactive Constituents in Pulses and Their Health Benefits." *Journal of Food Science and Technology*. Springer. https://doi.org/10.1007/s13197-016-2391-9.

Singh, M., R. K. Dubey, T. K. Koley, A. Maurya, P. M. Singh, and B. Singh. 2019a. "Valorization of Winged Bean (Psophocarpus Tetragonolobus (L) DC) by Evaluation of Its Antioxidant Activity through Chemometric Analysis." *South African Journal of Botany* 121 (March):114–20. https://doi.org/10.1016/j.sajb.2018.10.026.

———. 2019b. "Valorization of Winged Bean (Psophocarpus Tetragonolobus (L) DC) by Evaluation of Its Antioxidant Activity through Chemometric Analysis." *South African Journal of Botany* 121 (March):114–20. https://doi.org/10.1016/j.sajb.2018.10.026.

Singh, Vinayak, Ridhi Goel, Veena Pande, Mehar Hasan Asif, and Chandra Sekhar Mohanty. 2017. "De Novo Sequencing and Comparative Analysis of Leaf Transcriptomes of Diverse Condensed Tannin-Containing Lines of Underutilized Psophocarpus Tetragonolobus (L.) DC." *Scientific Reports* 7 (March). https://doi.org/10.1038/srep44733.

Smartt, J. 1980. "Some Observations on the Origin and Evolution of the Winged Bean (Psophocarpus Tetragonolobus)." *Euphytica* 29 (1): 121–23. https://doi.org/10.1007/BF00037256.

Smith, Margaret E., and Dion G. Morton. 2010. "DIGESTION AND ABSORPTION." In *The Digestive System*, 129–52. Elsevier. https://doi.org/10.1016/b978-0-7020-3367-4.00008-6.

Smulikowska, S., B. Pastuszewska, E. Święch, A. Ochtabińska, A. Mieczkowska, V. C. Nguyen, and L. Buraczewska. 2001. "Tannin Content Affects Negatively Nutritive Value of pea for Monogastrics." *Journal of Animal and Feed Sciences* 10 (3): 511–23. https://doi.org/10.22358/jafs/68004/2001.

Sonah, Humira, Louise O'donoughue, Elroy Cober, Istvan Rajcan, and Franc ßois Belzile. 2015. "Identification of Loci Governing Eight Agronomic Traits Using a GBS-GWAS Approach

and Validation by QTL Mapping in Soya Bean." *Plant Biotechnology Journal* 13:211–21. https://doi.org/10.1111/pbi.12249.

Song, Chao, Da Bing Xiang, Lin Yan, Yue Song, Gang Zhao, Yue Hua Wang, and Bao Lin Zhangb. 2016. "Changes in Seed Growth, Levels and Distribution of Flavonoids during Tartary Buckwheat Seed Development." *Plant Production Science* 19 (4): 518–27. https://doi.org/10.1080/1343943X.2016.1207485.

Srinivas, R., S. S. Udikeri, S. K. Jayalakshmi, and K. Sreeramulu. 2004. "Identification of Factors Responsible for Insecticide Resistance in Helicoverpa Armigera." *Comparative Biochemistry and Physiology - C Toxicology and Pharmacology* 137 (3): 261–69. https://doi.org/10.1016/j.cca.2004.02.002.

Sriwichai, Sasiprapa, Kularb Laosatit, Tidarat Monkham, Jirawat Sanitchon, Sanun Jogloy, and Sompong Chankaew. 2022. "Genetic Diversity of Domestic (Thai) and Imported Winged Bean [Psophocarpus Tetragonolobus (L.) DC.] Cultivars Assessed by Morphological Traits and Microsatellite Markers." *Annals of Agricultural Sciences* 67 (1): 34–41. https://doi.org/10.1016/j.aoas.2022.04.002.

Sriwichai, Sasiprapa, Tidarat Monkham, Jirawat Sanitchon, Sanun Jogloy, Sompong Chankaew, Chirakkuzhyil Abhilash, and Rodomiro Ortiz. 2021. "Dual-Purpose of the Winged Bean (Psophocarpus Tetragonolobus (L.) DC.), the Neglected Tropical Legume, Based on Pod and Tuber Yields." https://doi.org/10.3390/plants10081746.

Suri, Devika J., Kwaku Tano-Debrah, and Shibani A. Ghosh. 2014. "Optimization of the Nutrient Content and Protein Quality of Cereal—Legume Blends for Use as Complementary Foods in Ghana." *Http://Dx.Doi.Org/10.1177/156482651403500309* 35 (3): 372–81. https://doi.org/10.1177/156482651403500309.

Suwareh, Ousmane, David Causeur, Julien Jardin, Valérie Briard-Bion, Steven Le Feunteun, Stéphane Pezennec, and Françoise Nau. 2021. "Statistical Modeling of in Vitro Pepsin Specificity." *Food Chemistry* 362 (November):130098. https://doi.org/10.1016/J.FOODCHEM.2021.130098.

Tan, Nget Hong, Zubaidah Hj A. Rahim, Hun Teik Khor, and Kai Choo Wong. 1983. "Winged Bean (Psophocarpus Tetragonolobus) Tannin Level, Phytate Content, and Hemagglutinating Activity." *Journal of Agricultural and Food Chemistry* 31 (4): 916–17. https://doi.org/10.1021/jf00118a063.

Tan, Nget Hong, Kai Choo Wong, and Benito O. de Lumen. 1984. “Relationship of Tannin Levels and Trypsin Inhibitor Activity with the in Vitro Protein Digestibilities of Raw and Heat-Treated Winged Bean (Psophocarpus Tetragonolobus).” *Journal of Agricultural and Food Chemistry* 32 (4): 819–22. https://doi.org/10.1021/JF00124A030.

Tan, Nget-Hong, Kai-Choo Wong, and Benito O de Lumen. 1984. “Relationship of Tannin Levels and Trypsin Inhibitor Activity with the in Vitro Protein Digestibilities of Raw and Heat-Treated Winged Bean (Psophocarpus Tetragonolobus).” *J. Agric. Food Chem.* Vol. 32. Wiley. https://pubs.acs.org/sharingguidelines.

Tanzi, Alberto Stefano. 2018. “Genetic Investigation of Plant Architecture, Yield, and Diversity in Winged Bean (Psophocarpus Tetragonolobus (L.) Dc.) for Breeding Programmes.”

Tanzi, Alberto Stefano, Graham Ewen Eagleton, Wai Kuan Ho, Quin Nee Wong, Sean Mayes, and Festo Massawe. 2019a. “Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) for Food and Nutritional Security: Synthesis of Past Research and Future Direction.” *Planta* 250 (3): 911–31. https://doi.org/10.1007/s00425-019-03141-2.

———. 2019b. “Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) for Food and Nutritional Security: Synthesis of Past Research and Future Direction.” *Planta*. Springer Verlag. https://doi.org/10.1007/s00425-019-03141-2.

Tanzi, Alberto Stefano, Wai Kuan Ho, Festo Massawe, and Sean Mayes. 2019a. “Development and Interaction between Plant Architecture and Yield-Related Traits in Winged Bean (Psophocarpus Tetragonolobus (L.) DC.).” *Euphytica* 215 (2): 1–14. https://doi.org/10.1007/s10681-019-2359-8.

———. 2019b. “Development and Interaction between Plant Architecture and Yield-Related Traits in Winged Bean (Psophocarpus Tetragonolobus (L.) DC.).” *Euphytica* 215 (2). https://doi.org/10.1007/s10681-019-2359-8.

Telang, Manasi A., Ashok P. Giri, Prashant S. Pyati, Vidya S. Gupta, Mechthild Tegeder, and Vincent R. Franceschi. 2008. “Winged Bean Chymotrypsin Inhibitors Retard Growth of Helicoverpa Armigera ☆.” *Gene* 431 (1–2): 80–85. https://doi.org/10.1016/j.gene.2008.10.026.

Ting, Ngoot Chin, Zulkifli Yaakub, Katialisa Kamaruddin, Sean Mayes, Festo Massawe, Ravigadevi Sambanthamurthi, Johannes Jansen, et al. 2016. “Fine-Mapping and Cross-Validation of QTLs Linked to Fatty Acid Composition in Multiple Independent

Interspecific Crosses of Oil Palm." *BMC Genomics* 17 (1). https://doi.org/10.1186/s12864-016-2607-4.

Tsay, Jiu Tsair, Won Oh, Timothy J. Larson, Suzanne Jackowski, and Charles O. Rock. 1992. "Isolation and Characterization of the β-Ketoacyl-Acyl Carrier Protein Synthase III Gene (FabH) from Escherichia Coli K-12." *Journal of Biological Chemistry* 267 (10): 6807–14. https://doi.org/10.1016/s0021-9258(19)50498-7.

Tyagi, Parul, Deeksha Singh, Shivangi Mathur, Ayushi Singh, and Rajiv Ranjan. 2022. "Upcoming Progress of Transcriptomics Studies on Plants: An Overview." *Frontiers in Plant Science* 13 (December):1030890. https://doi.org/10.3389/FPLS.2022.1030890/BIBTEX.

Urbano, G., M. López-Jurado, P. Aranda, C. Vidal-Valverde, E. Tenorio, and J. Porres. 2000. "The Role of Phytic Acid in Legumes: Antinutrient or Beneficial Function?" *Journal of Physiology and Biochemistry* 56 (3): 283–94. https://doi.org/10.1007/BF03179796.

US Department of Agriculture, Agricultural Research Service. 2019. Nutrient Data Laboratory. USDA FoodData Central (FDC). 2019. "FoodData Central." 2019. https://fdc.nal.usda.gov/fdc-app.html#/food-details/172436/nutrients.

Valentine, Michelle F., Joann R. De Tar, Muruganantham Mookkan, Jeffre D. Firman, and Zhanyuan J. Zhang. 2017. "Silencing of Soybean Raffinose Synthase Gene Reduced Raffinose Family Oligosaccharides and Increased True Metabolizable Energy of Poultry Feed." *Frontiers in Plant Science* 8 (May). https://doi.org/10.3389/FPLS.2017.00692.

Vatanparast, Mohammad, Prateek Shetty, Ratan Chopra, Jeff J. Doyle, N. Sathyanarayana, and Ashley N. Egan. 2016a. "Transcriptome Sequencing and Marker Development in Winged Bean (Psophocarpus Tetragonolobus; Leguminosae)." *Scientific Reports* 6 (June):1–14. https://doi.org/10.1038/srep29070.

———. 2016b. "Transcriptome Sequencing and Marker Development in Winged Bean (Psophocarpus Tetragonolobus; Leguminosae)." *Scientific Reports* 6 (June):1–14. https://doi.org/10.1038/srep29070.

Verdcourt, B., and P. Halliday. 1978. "A Revision of Psophocarpus (Leguminosae-Papilionoideae-Phaseoleae)." *Kew Bulletin* 33 (2): 191. https://doi.org/10.2307/4109575.

Vlasova, Anna, Salvador Capella-Gutiérrez, Martha Rendón-Anaya, Miguel Hernández-Oñate, André E. Minoche, Ionas Erb, Francisco Câmara, et al. 2016. "Genome and Transcriptome

Analysis of the Mesoamerican Common Bean and the Role of Gene Duplications in Establishing Tissue and Temporal Specialization of Genes." *Genome Biology* 17 (1). https://doi.org/10.1186/s13059-016-0883-6.

Voss-Fels, Kai P., Andreas Stahl, and Lee T. Hickey. 2019. "Q&A: Modern Crop Breeding for Future Food Security." *BMC Biology*. BioMed Central Ltd. https://doi.org/10.1186/s12915-019-0638-4.

Wainaina, Irene, Elizabeth Wafula, Daniel Sila, Clare Kyomugasho, Tara Grauwet, Ann Van Loey, and Marc Hendrickx. 2021. "Thermal Treatment of Common Beans (Phaseolus Vulgaris L.): Factors Determining Cooking Time and Its Consequences for Sensory and Nutritional Quality." *Comprehensive Reviews in Food Science and Food Safety*. Blackwell Publishing Inc. https://doi.org/10.1111/1541-4337.12770.

Wan Mohtar, Wan Abd Al Qadr Imad, Azizah Abdul Hamid, Suraini Abd-Aziz, Sharifah Kharidah Syed Muhamad, and Nazamid Saari. 2014a. "Preparation of Bioactive Peptides with High Angiotensin Converting Enzyme Inhibitory Activity from Winged Bean [Psophocarpus Tetragonolobus (L.) DC.] Seed." *Journal of Food Science and Technology* 51 (12): 3658–68. https://doi.org/10.1007/s13197-012-0919-1.

———. 2014b. "Preparation of Bioactive Peptides with High Angiotensin Converting Enzyme Inhibitory Activity from Winged Bean [Psophocarpus Tetragonolobus (L.) DC.] Seed." *Journal of Food Science and Technology* 51 (12): 3658–68. https://doi.org/10.1007/s13197-012-0919-1.

Wang, Ming Li, Pawan Khera, Manish K Pandey, Hui Wang, Lixian Qiao, Suping Feng, Brandon Tonnis, et al. 2015. "Genetic Mapping of QTLs Controlling Fatty Acids Provided Insights into the Genetic Control of Fatty Acid Synthesis Pathway in Peanut (Arachis Hypogaea L.)." *PLOS One*.

Wang ML, Khera P, Pandey MK, Wang H, Qiao L, Feng S, and Tonnis B. 2015. "Genetic Mapping of QTLs Controlling Fatty Acids Provided Insights into the Genetic Control of Fatty Acid Synthesis Pathway in Peanut (Arachis Hypogaea L.)." *PLoS ONE* 10 (4): 119454. https://doi.org/10.1371/journal.pone.0119454.

Wang, Xianzhi, Guo Liang Jiang, Marci Green, Roy A. Scott, David L. Hyten, and Perry B. Cregan. 2012. "Quantitative Trait Locus Analysis of Saturated Fatty Acids in a Population of Recombinant Inbred Lines of Soybean." *Molecular Breeding* 30 (2): 1163–79. https://doi.org/10.1007/s11032-012-9704-0.

Wang, Zhihui, Ke Cheng, Liyun Wan, Liying Yan, Huifang Jiang, Shengyi Liu, Yong Lei, and Boshou Liao. 2015. “Genome-Wide Analysis of the Basic Leucine Zipper (BZIP) Transcription Factor Gene Family in Six Legume Genomes.” *BMC Genomics* 16 (1). https://doi.org/10.1186/s12864-015-2258-x.

Warrington, C V, · H Abdel-Haleem, D L Hyten, P B Cregan, J H Orf, · A S Killam, · N Bajjalieh, et al. 2015. “QTL for Seed Protein and Amino Acids in the Benning × Danbaekkong Soybean Population.” *Theor Appl Genet* 128:839–50. https://doi.org/10.1007/s00122-015-2474-4.

Wati, Richa Kusuma, Theerapong Theppakorn, Soottawat Benjakul, and Saroat Rawdkuen. 2010. “Trypsin Inhibitor from 3 Legume Seeds: Fractionation and Proteolytic Inhibition Study.” *Journal of Food Science* 75 (3): C223–28. https://doi.org/10.1111/j.1750-3841.2010.01515.x.

Whelan, Jay. 2008. “The Health Implications of Changing Linoleic Acid Intakes.” *Prostaglandins Leukotrienes and Essential Fatty Acids* 79 (3–5): 165–67. https://doi.org/10.1016/j.plefa.2008.09.013.

Wong, Q. N., F. Massawe, K. Mayes, M. Blythe, and S. Mayes. 2016. “Molecular Genetic Tools to Support Genetic Improvement of Winged Bean (Psophocarpus Tetragonolobus) for Food and Nutrition Security.” In *Acta Horticulturae*, 1110:1–6. International Society for Horticultural Science. https://doi.org/10.17660/ActaHortic.2016.1110.1.

Wong, Q. N., F. Massawe, and S. Mayes. 2015. “Improving Winged Bean (Psophocarpus Tetragonolobus) Productivity: An Analysis of the Determinants of Productivity.” *Acta Horticulturae* 1102 (May 2019): 83–88. https://doi.org/10.17660/ActaHortic.2015.1102.9.

Wong, Quin Nee, Alberto Stefano Tanzi, Wai Kuan Ho, Sunir Malla, Martin Blythe, Asha Karunaratne, Festo Massawe, and Sean Mayes. 2017. “Development of Gene-Based SSR Markers in Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) for Diversity Assessment.” *Genes* 8 (3). https://doi.org/10.3390/genes8030100.

Wong, Quin Nee, Alberto Stefano Tanzi, Wai Kuan Ho, Sunir Malla, Martin Blythe, Asha Karunaratne, Festo Massawe, and Sean Mayes. 2017. “Development of Gene-Based SSR Markers in Winged Bean (Psophocarpus Tetragonolobus (L.) DC.) for Diversity Assessment.” https://doi.org/10.3390/genes8030100.

Worthington, R. E., Ray O. Hammons, and John R. Allison. 1972. “Varietal Differences and Seasonal Effects on Fatty Acid Composition and Stability of Oil from 82 Peanut Genotypes.” *Journal of Agricultural and Food Chemistry* 20 (3): 729–30. https://doi.org/10.1021/jf60181a032.

Wyckoff S., and Vohra P. 1982. “The Nutritional Value of Autoclaved and Ammonia Treated Winged Beans [Psophocarpus Tetragonolobus (L.) Dc] for Japanese Quail,” 359–64.

Xu, B. J., S. H. Yuan, and S. K.C. Chang. 2007. “Comparative Analyses of Phenolic Composition, Antioxidant Capacity, and Color of Cool Season Legumes and Other Selected Food Legumes.” *Journal of Food Science* 72 (2). https://doi.org/10.1111/j.1750-3841.2006.00261.x.

Xu, Baojun, and Sam K.C. Chang. 2008. “Antioxidant Capacity of Seed Coat, Dehulled Bean, and Whole Black Soybeans in Relation to Their Distributions of Total Phenolics, Phenolic Acids, Anthocyanins, and Isoflavones.” *Journal of Agricultural and Food Chemistry* 56 (18): 8365–73. https://doi.org/10.1021/JF801196D.

Xu, Shizhong. 2003. “Theoretical Basis of the Beavis Effect.” *Genetics* 165 (3). https://doi.org/https://doi.org/10.1093/genetics/165.4.2259.

Yanagi, Sonoe Ochiai. 1983. “Properties of Winged Bean (Psophocarpus Tetragonolobus) Protein in Comparison with Soybean (Glycine Max) and Common Bean (Phaseolus Vulgaris) Protein.” *Agricultural and Biological Chemistry* 47 (10): 2273–80. https://doi.org/10.1080/00021369.1983.10865943.

Yang, Jinsong, and Haisheng Tan. 2011. “Study on Winged Bean Milk.” *ICAE 2011 Proceedings: 2011 International Conference on New Technology of Agricultural Engineering*, 814–17. https://doi.org/10.1109/ICAE.2011.5943916.

Yang, Shuyi, Aurélie Grall, and Mark A. Chapman. 2018. “Origin and Diversification of Winged Bean (Psophocarpus Tetragonolobus (L.) DC.), a Multipurpose Underutilized Legume.” *American Journal of Botany* 105 (5): 888–97. https://doi.org/10.1002/ajb2.1093.

Yang, Songnan, Long Miao, Jianbo He, Kai Zhang, Yan Li, and Junyi Gai. 2019. “Dynamic Transcriptome Changes Related to Oil Accumulation in Developing Soybean Seeds.” *International Journal of Molecular Sciences Article*. https://doi.org/10.3390/ijms20092202.

Yoneyama, Tadakatsu, Kounosuke Fujita, Tomio Yoshida, Tetsuo Matsumoto, Iwao Kambayashi, and Jinya Yazaki. 1986. “Variation in Natural Abundance of 15 N among

Plant Parts and in 15 N/ 14 N Fractionation during N 2 Fixation in the Legume-Rhizobia Symbiotic System." *Plant CellPhysiol*. Vol. 27. https://academic.oup.com/pcp/article/27/5/791/1899683.

Yu, Anmin, Fei Li, and Aizhong Liu. 2020. "Comparative Proteomic and Transcriptomic Analyses Provide New Insight into the Formation of Seed Size in Castor Bean." *BCM Plant Biology*. https://doi.org/10.1186/s12870-020-2249-1.

Yu, Ki Seon, Hyunnho Cho, and Keum Taek Hwang. 2018. "Physicochemical Properties and Oxidative Stability of Frying Oils during Repeated Frying of Potato Chips." *Food Science and Biotechnology* 27 (3): 651–59. https://doi.org/10.1007/s10068-017-0292-y.

Zhang, Tengfei, Tingting Wu, Liwei Wang, Bingjun Jiang, Caixin Zhen, Shan Yuan, Wensheng Hou, Cunxiang Wu, Tianfu Han, and Shi Sun. 2019. "A Combined Linkage and GWAS Analysis Identifies QTLs Linked to Soybean Seed Protein and Oil Content." *International Journal of Molecular Sciences*. https://doi.org/10.3390/ijms20235915.

Zhang, Ying Hu, Mei Feng Liu, Jian Bo He, Yu Feng Wang, Guang Nan Xing, Yan Li, Shou Ping Yang, Tuan Jie Zhao, and Jun Yi Gai. 2015. "Marker-Assisted Breeding for Transgressive Seed Protein Content in Soybean [Glycine Max (L.) Merr.]." *Theoretical and Applied Genetics* 128 (6): 1061–72. https://doi.org/10.1007/s00122-015-2490-4.

Zhao, Chuang, Bing Liu, Shilong Piao, Xuhui Wang, David B. Lobell, Yao Huang, Mengtian Huang, et al. 2017. "Temperature Increase Reduces Global Yields of Major Crops in Four Independent Estimates." *Proceedings of the National Academy of Sciences of the United States of America* 114 (35): 9326–31. https://doi.org/10.1073/pnas.1701762114.

Zhou, J. R., E. J. Fordyce, V. Raboy, D. B. Dickinson, M. S. Wong, R. A. Burns, and J. W. Erdman. 1992. "Reduction of Phytic Acid in Soybean Products Improves Zinc Bioavailability in Rats." *Journal of Nutrition* 122 (12): 2466–73. https://doi.org/10.1093/jn/122.12.2466.

Appendices - Supplementary material

	2019								2021									
Month	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
max temp (°C)																		
FRC	33.4	32.7	32.8	32.1	32.9	32.4	32.6	32.7										
FF									31.9	32.8	32.2	32.0	31.6	31.7	31.5	31.2	30.8	-
min temp (°C)																		
FRC	23.6	23.5	22.9	22.5	22.5	22.7	22.8	23.1										
FF									23.5	23.5	23.7	23.8	24.5	24.2	24.1	23.4	23.3	-
RH (%)																		
FRC	90.3	94.1	91.1	93.7	88.1	91.7	90.6	81.0										
FF									83.0	80.0	81.0	82.0	82.0	79.0	78.0	78.0	79.0	82.0
total rainfall (mm)																		
FRC	62	21	49	112	59	178	51	144										
FF									-	14	107	275	111	112	170	118	196	8

Supplementary Figure 3. 1 Average temperature, relative humidity and rainfall in the CFF-FRC, UNM rainout shelter and FF from 2019 to 2021. As featured in the PhD thesis Chong, 2024 .

Location	Sampling stage	pH	Organic C (%)	N (%)	C/N	P	Exchangable cations (m.e. %)			Mechanical analysis (%)			
						(Acid fluoride soluble, ppm)	K	Ca	Mg	Clay	F silt	Sand	Texture class
FRC	before planting	4.26	0.93	0.13	7.6	205.67	9.37	0.54	0.11	37.7	8.7	16.2	sandy clay
FF	before planting	6.16	0.86	0.10	9.1	37.18	0.41	3.03	0.52	16.8	6.8	76.4	sandy clay loam - sandy loam
	after planting	5.76	0.74	0.09	8.3	21.04	0.22	2.58	0.39	23.6	6.4	70.0	sandy clay loam - sandy loam

Supplementary Figure 3. 2 The analysis of soil physicochemical properties in the CFF-FRC (2019), UNM rainout shelter (2020) and FF (2021) before and after planting. As featured in the PhD thesis Chong, 2024 .

Supplementary Table 3. 1 Relative amount of fatty acids (%). The relative amounts of fatty acids in the two locations (FF and FRC) are presented as the mean (n=3) ± SEM, analysed using a two-way ANOVA and Tukey's post-hoc test. The different letters represent significant differences among the winged bean accessions. The numbers next to the fatty acids correspond to: 1 = Saturated fatty acids, 2 = Monounsaturated fatty acids, and 3 = Polyunsaturated fatty acids.

Winged bean accessions	Location	Fat (%)	Palmitic[1]	Stearic[1]	Oleic[2]	Linoleic[3]	Behenic[1]
A27	**FF**	14.15±0.33ab	8.83±0.34bcdef	5.29±0.27ab	35.08±1.47abc	23.76±1.07b	18.81±0.25
	FRC	16.14±0.82abc	8.13±0.26bcd	5.94±0.14b	38.26±0.71bc	21.07±0.68ab	18.84±0.72
I17	**FF**	14.38±1.19ab	9.07±0.5def	4.77±0.1ab	33±0.4ab	25.44±1.71b	19.3±0.67
	FRC	15.41±0.61abc	8.1±0.14bcd	5.18±0.16ab	34.35±0.24abc	24.55±0.5b	19.72±0.71
A30	**FF**	15.41±0.78abc	9.65±0.33ef	5.14±0.2ab	32.42±1.07a	24.33±1.4b	19.76±1.01
	FRC	18.89±0.85cd	9.84±0.13f	5.24±0.28ab	34.42±0.54abc	21.79±0.67ab	19.65±0.9
A4	**FF**	15.79±0.55abc	8.86±0.18cdef	4.99±0.28ab	32.59±0.61ab	24.53±0.84b	20.15±0.96
	FRC	12.49±0.41a	7.59±0.04ab	4.35±0.22a	36.12±0.76abc	23.96±1.22b	19.27±0.68
I53	**FF**	15.94±0.47abc	7.89±0.23abcd	5.75±0.48b	36.99±2.5abc	22.81±1.78ab	18.21±1.38
	FRC	16.78±1.49bc	7.71±0.04abc	5.98±0.19b	35.77±0.97abc	22.23±0.42ab	19.9±0.43
I10	**FF**	16.09±0.18abc	8.45±0.13bcde	5.48±0.13ab	32.74±0.59ab	24.56±0.71b	20.49±0.47
	FRC	17.75±0.71bc	8.04±0.18bcd	5.83±0.37b	32.63±0.61ab	23.33±0.24b	21.29±0.75
Ma3	**FF**	21.78±0.89d	7.66±0.06abc	6.05±0.2b	35.7±0.98abc	20.91±0.14ab	19.97±1.34
	FRC	18.45±0.17cd	6.77±0.32a	5.78±0.19b	39.66±1.79c	17.16±1.87a	21.82±0.31
p-value							
	G	<0.001	<0.001	<0.001	<0.001	<0.001	0.106
	E	0.414	<0.001	0.380	0.005	0.006	0.228
	GxE	<0.001	0.085	0.194	0.209	0.695	0.643

Supplementary Table 3. 2 Actual amount of fatty acids (mg/100 mg dry seed weight). The actual amount of fatty acids was expressed as mg of fatty acid per 100 mg of dry seed weight. The amount of the fatty acids shown in the table ± SEM (n=3), was analysed using a two-way ANOVA and Tukey's post-hoc test. From the two-way ANOVA (n=3), the p-values are shown. The different letters represent significant differences among the winged bean accessions. The numbers next to the fatty acids correspond to: 1 = Saturated fatty acids, 2 = Monounsaturated fatty acids, and 3 = Polyunsaturated fatty acids.

Winged bean accessions	Location	Oil yield (%)	Palmitic[1]	Stearic[1]	Oleic[2]	Linoleic[3]	Behenic[1]
A27	**FF**	14.15±0.33ab	1.25±0.03ab	0.75±0.05abc	4.97±0.33ab	3.36±0.01ab	2.66±0.04ab
	FRC	16.14±0.82abc	1.31±0.03abc	0.96±0.05bc	6.18±0.36abcd	3.41±0.28ab	3.03±0.11abcd
I17	**FF**	14.38±1.19ab	1.31±0.18abc	0.68±0.05ab	4.75±0.42a	3.62±0.04abc	2.79±0.33abc
	FRC	15.41±0.61abc	1.25±0.03ab	0.8±0.06abc	5.29±0.19ab	3.78±0.13abc	3.04±0.18abcd
A30	**FF**	15.41±0.78abc	1.49±0.13bcd	0.8±0.07abc	5±0.31ab	3.73±0.09abc	3.05±0.26abcd
	FRC	18.89±0.85cd	1.86±0.11d	0.99±0.09bcd	6.5±0.34bcd	4.13±0.39bc	3.7±0.04bcde
A4	**FF**	15.79±0.55abc	1.4±0.07bc	0.79±0.07abc	5.14±0.08ab	3.86±0.07abc	3.19±0.27abcd
	FRC	12.49±0.41a	0.95±0.04a	0.54±0.02a	4.52±0.24a	2.99±0.22a	2.4±0.08a
I53	**FF**	15.94±0.47abc	1.26±0.01abc	0.92±0.08bc	5.89±0.37abc	3.64±0.26abc	2.91±0.3abc
	FRC	16.78±1.49bc	1.29±0.11abc	1±0.09bcd	5.99±0.48abc	3.74±0.21abc	3.34±0.3abcde
I10	**FF**	16.09±0.18abc	1.36±0.01abc	0.88±0.03bc	5.27±0.15ab	3.95±0.13abc	3.3±0.09abcde
	FRC	17.75±0.71bc	1.43±0.06bc	1.03±0.03cd	5.78±0.14abc	4.14±0.17bc	3.79±0.28cde
Ma3	**FF**	21.78±0.89d	1.67±0.08cd	1.32±0.09d	7.79±0.46d	4.55±0.23c	4.33±0.23e
	FRC	18.45±0.17cd	1.25±0.06abc	1.07±0.03cd	7.32±0.32cd	3.17±0.29ab	4.02±0.05de
p-value	**G**	<0.001	<0.001	<0.001	<0.001	0.039	<0.001
	E	0.414	0.193	0.288	0.028	0.108	0.181
	GxE	<0.001	<0.001	0.001	0.02	0.002	0.023

Supplementary Table 3. 3 Fatty acid of 8 winged bean accessions used in the in vitro digestion system. The actual amount of fatty acids was expressed as mg of fatty acid per 100 mg of dry seed weight. The relative and actual amount of the fatty acids shown in the table ± SEM (n=3), was analysed using a two-way ANOVA and Tukey's post-hoc test. From the two-way ANOVA (n=3), with the p-values. The different letters represent significant differences among the winged bean accessions. The numbers next to the fatty acids correspond to: 1 = Saturated fatty acids, 2 = Monounsaturated fatty acids, and 3 = Polyunsaturated fatty acids.

Winged bean		**Oil yield (%)**	**Palmitic[1]**	**Stearic[1]**	**Oleic[2]**	**Linoleic[3]**	**Behenic[1]**
Relative	A30		9.65±0.33b	5.14±0.2ab	32.42±1.07a	24.33±1.4a	19.76±1.01a
Amount	A6		7.87±0.47a	6.49±0.13c	35.15±1.22a	21.91±0.36a	19.81±0.66a
	FP15		8.27±0.2ab	6.51±0.21c	33.61±0.92a	23.21±1.24a	19.38±0.22a
	Gmya4		9.1±0.15ab	5.23±0.12ab	34.81±1.05a	26.05±0.42a	17.23±0.55a
	I10		8.45±0.13ab	5.48±0.13abc	32.74±0.59a	24.56±0.71a	20.49±0.47a
	I17		9.07±0.5ab	4.77±0.1a	33±0.4a	25.44±1.71a	19.3±0.67a
	I53		7.89±0.23a	5.75±0.48abc	36.99±2.5a	22.81±1.78a	18.21±1.38a
	Ma3		7.66±0.06a	6.05±0.2bc	35.7±0.98a	20.91±0.14a	19.97±1.34a
	p-value		0.002	<0.001	0.178	0.071	0.259
mg/100	A30	15.41±0.78a	1.49±0.13ab	0.8±0.07a	5±0.31a	3.73±0.07ab	3.05±0.26ab
mg	A6	17.18±0.52ab	1.35±0.1a	1.12±0.03bc	6.03±0.2ab	3.76±0.09ab	3.41±0.21abc
	FP15	20.61±0.96bc	1.71±0.09b	1.34±0.07c	6.92±0.33bc	4.79±0.38c	3.99±0.17bc
	Gmya4	14.54±0.18a	1.32±0.03a	0.76±0.01a	5.07±0.22a	3.79±0.04ab	2.5±0.05a
	I10	16.09±0.18a	1.36±0.01a	0.88±0.03ab	5.27±0.15a	3.95±0.09abc	3.3±0.09abc
	I17	14.38±1.19a	1.31±0.18a	0.68±0.05a	4.75±0.42a	3.62±0.06a	2.79±0.33a
	I53	15.94±0.47a	1.26±0.01a	0.92±0.08ab	5.89±0.37ab	3.64±0.3ab	2.91±0.3ab
	Ma3	21.78±0.89c	1.67±0.08b	1.32±0.09c	7.79±0.46c	4.55±0.17bc	4.33±0.23c
	p-value	<0.001	0.026	<0.001	<0.001	0.002	<0.001

Supplementary Table 3. 4 Statistical analysis of each amino acid among the 8 winged bean accessions (A30, Ma3, I17, I10, GMYA4, A6, FP15, I53). The amino acids were expressed as mg per gram of protein. The statistical test used was one-way ANOVA and no significant differences were detected.

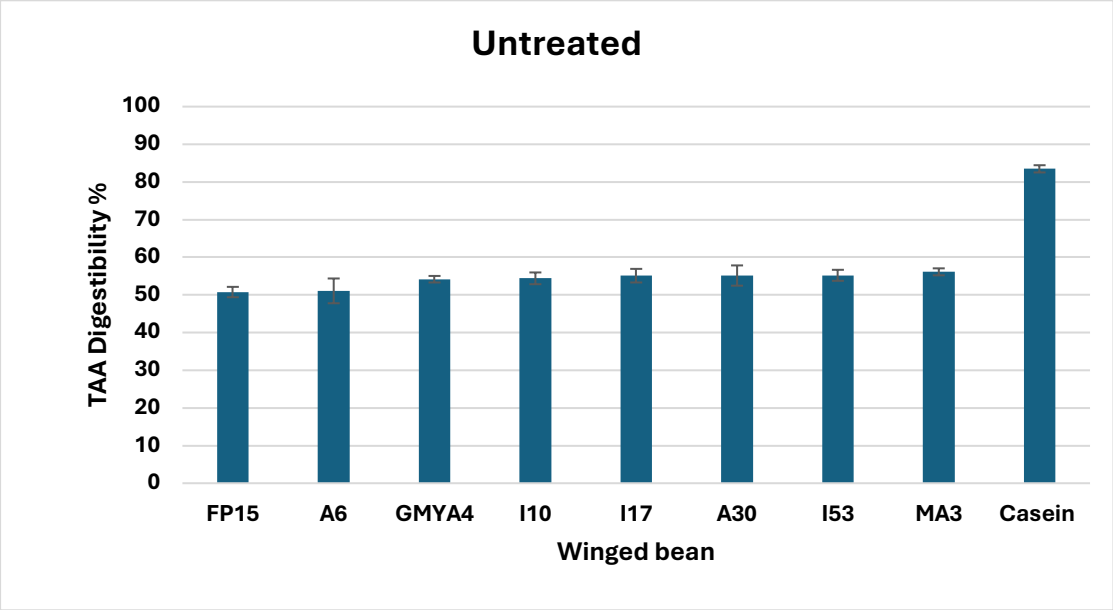
Amino acids mg.gr protein	F (DFn, DFd)	P-value
Alanine	F (7, 15) = 0.6849	P=0.6834
Arginine	F (7, 15) = 0.6652	P=0.6981
Aspartate	F (7, 15) = 0.09561	P=0.9978
Cysteine	F (7, 15) = 1.409	P=0.2722
Glutamate	F (7, 15) = 0.9701	P=0.4865
Glycine	F (7, 15) = 0.9335	P=0.5095
Histdine	F (7, 15) = 0.7992	P=0.6002
Isoleucine	F (7, 15) = 0.3859	P=0.8964
Leucine	F (7, 15) = 0.4318	P=0.8673
Lysine	F (7, 15) = 0.8075	P=0.5943
Methionine	F (7, 15) = 1.030	P=0.4506
Phenylalanine	F (7, 15) = 0.3076	P=0.9394
Proline	F (7, 15) = 0.7931	P=0.6045
Serine	F (7, 15) = 0.7948	P=0.6033
Threonine	F (7, 15) = 0.9719	P=0.4854
Tyrosine	F (7, 15) = 1.159	P=0.3807
Valine	F (7, 15) = 1.253	P=0.3359

Supplementary Table 3. 5 Amino acid g/100 g of protein.

	Soybean	A30	Ma3	I17	I10	GMYA4	A6	FP15	I53
Histidine	3.8	2.97±0.08	2.99±0.03	2.85±0.03	2.75±0.09	2.82±0.18	2.7±0.18	2.77±0.15	2.65±0.18
Isoleucine	5.2	4.63±0.21	4.51±0.1	4.55±0.09	4.49±0.06	4.34±0.17	4.45±0.38	4.34±0.16	4.21±0.28
Leucine	7.2	8.47±0.34	8.21±0.25	8.21±0.27	7.84±0.2	8.09±0.32	8.03±0.25	7.69±0.46	8.06±0.59
Lysine	5.4	7.49±0.06	7.21±0.06	7.14±0.14	6.97±0.24	6.89±0.4	6.96±0.34	6.82±0.4	6.66±0.32
Phenylalanine	4.9	4.71±0.06	4.71±0.1	4.56±0	4.58±0.12	4.44±0.19	4.5±0.36	4.39±0.24	4.53±0.29
Threonine	5.4	3.88±0.03	3.88±0.07	3.43±0.3	3.76±0.15	3.39±0.18	3.66±0.25	3.58±0.16	3.64±0.26
Valine	4.9	5.16±0.19	5.32±0.22	5.16±0.07	5.39±0.17	4.96±0.44	5.04±0.37	4.59±0.09	4.7±0.09
Methionine	1.2	1.15±0.02	1.27±0.03	1.13±0.1	1.19±0.02	1.07±0.03	1.17±0.11	1.13±0.04	1.13±0.07
Cysteine	1.5	1.02±0.06	1±0.04	0.95±0.01	0.97±0.13	1.04±0.03	0.98±0.13	0.73±0.06	0.89±0.07
Alanine	6.6	4.19±0.2	3.97±0.42	3.53±0.28	3.7±0.11	3.63±0.3	3.53±0.38	3.6±0.09	3.68±0.29
Arginine	5.3	7.37±0.1	7.68±0.06	7.43±0.03	7.3±0.15	7.34±0.25	7.11±0.33	7.54±0.32	7.3±0.18
Aspartate	10	10.39±0.1	10.53±0.22	10.48±0.08	10.21±0.65	10.29±0.63	9.96±0.6	10.25±0.37	10.35±0.98
Glutamate	9.9	18.13±0.16	17.15±0.03	17.44±0.28	16.52±0.58	17.27±0.55	16.12±1.01	17.06±0.94	16.35±0.81
Glycine	7.3	4.76±0.54	4.76±0.14	4.2±0.46	4.52±0.22	3.93±0.07	4.49±0.28	4.33±0.12	4.39±0.29
Proline	6.8	6.02±0.09	5.83±0.07	5.71±0.21	5.64±0.18	5.83±0.23	5.44±0.38	5.45±0.29	5.51±0.27
Serine	6.6	4.32±0.03	4.36±0.07	3.91±0.18	4.05±0.16	4.02±0.09	4.26±0.28	4.02±0.15	3.99±0.32
Tyrosine	2.2	3.78±0.1	3.85±0.08	3.55±0.12	3.53±0.11	3.73±0.15	3.52±0.19	3.52±0.19	3.38±0.21

Supplementary Table 3. 6 DIAAS of autoclaved winged bean seeds. The analysis was performed on two out of three biological replicates. LEU stands for leucine and SAA for methionine as sulphur-containing amino acid.

Average in vitro DIAAR	HIS	ILE	LEU	LYS	SAA	AAA	THR	VAL	TRP	DIAAS	Limiting AA
A30	0.84	0.57	0.57	0.72	0.18	0.43	0.63	0.58		0.18	SAA
A6	0.81	0.81	0.62	0.74	0.20	0.47	0.67	0.65		0.20	SAA
FP15	0.66	0.45	0.51	0.52	0.14	0.32	0.46	0.40		0.14	SAA
GMYA4	0.74	0.60	0.53	0.65	0.16	0.39	0.49	0.47		0.16	SAA
I10	0.62	0.54	0.52	0.53	0.14	0.36	0.51	0.50		0.14	SAA
I17	0.82	0.70	0.70	0.73	0.19	0.48	0.57	0.59		0.19	SAA
I53	0.85	0.74	0.75	0.73	0.21	0.52	0.72	0.59		0.21	SAA
M3	0.77	0.47	0.47	0.63	0.18	0.40	0.56	0.50		0.18	SAA
Casein	1.05	1.09	1.20	1.06	0.77	0.70	0.96	0.90	0.97	0.70	SAA



Supplementary Figure 3. 3 Total amino acid digestibility (%) of untreated winged bean accessions. The analysis was performed by a technician. Casein was used as a positive control.

Supplementary Table 3. 7 DIAAS of untreated winged bean seeds. The analysis was performed by a technician. LEU stands for leucine and SA

Average in vitro DIAAR	DIAAS	Limiting AA
A30	0.55	SAA
A6	0.35	LEU
FP15	0.68	SAA
GMYA4	0.35	LEU
I10	0.50	SAA
I17	0.38	LEU
I53	0.35	LEU
MA3	0.34	LEU

Supplementary Table 4. 1 Winged bean markers in close proximity to winged bean genes and the genes' position on the genome.

No marker	Marker	Chr	Position	WB Gene	Start	Finish
56	20122_68:C>T	Chr01.1	3,553,364	Psote01G0031000.1	3727155	3732594
136	17711_29:G>A	Chr02.1	28,305,947	Psote02G0085400.1	28549122	28563980
139	21926_8:C>G	Chr02.3	1,529,838	Psote02G0162800.1	2458974	2461322
138	15138_30:C>T	Chr02.3	4,644,907	Psote02G0175100.1	5431298	5434950
140	20773_45:T>C	Chr02.2	252,098	Psote02G0195400.1	2067527	2072525
134	8506_16:A>G	Chr02.5	21,036,687	Psote02G0269200.1	17549045	17554199
135	20984_68:G>A	Chr02.5	14,155,614	Psote02G0269200.1	17549045	17554199
66	25763_26:G>A	Chr03.1	3,964,708	Psote03G0022300.1	4231297	4247418
86	23154_46:T>C	Chr03.1	33,368,309	Psote03G0166900.1	33688946	33692586
120	22489_20:T>C	Chr03.5	26,019,388	Psote03G0381000.1	25430795	25434046
175	13546_10:C>G	Chr05.1	6,301,652	Psote05G0053200.1	7543397	7548100
173	10233_32:T>G	Chr05.1	11,795,082	Psote05G0083100.1	11990726	11995665
177	14418_66:A>T	Chr05.1	13,222,002	Psote05G0096100.1	14177221	14186716
306	18511_24:G>T	Chr06.1	2,624,815	Psote06G0023300.1	2681564	2689273
325	22828_20:C>G	Chr06.4	16,109,326	Psote06G0283600.1	16034371	16038702
327	24665_28:G>T	Chr06.4	18,115,712	Psote06G0352200.1	24076620	24083027
360	24227_50:G>C	Chr07.2	21,707,553	Psote07G0237900.1	22534654	22558592
283	11190_34:G>T	Chr08.1	17,057,522	Psote08G0128000.1	16739495	16742025
379	14162_37:C>G	Chr09.1	2,000,223	Psote09G0032300.1	4230362	4234278

Supplementary Table 4. 2 Summary table of the position of winged bean fatty acid QTLs and genes on winged bean genome. The soybean homologous to the winged bean genes are shown.

QTL	cM	Genome position	Gene	Gene Name	Gene Chr	Gene Start	Gene end	Soybean homolog	E-value	Arabidopsis homolog	E-value
qLin-2-1	25.9	Chr02.5: 14,155,614 .. 31,732,450	Psote02G0269200.1	FabZ	Chr02.5	17,549,045	17,554,199	Glyma.15G052500	2.94E-138	AT5G10160.1	1.30E-106
qStear-3-1	7.4	Chr03.1: 3,440,557 .. 6,951,816	Psote03G0022300.1	accC	Chr03.1	4231297	4247418	Glyma.05G221100	0.00E+00	AT5G35360.1	#####
qPal-3-1	7.4	Chr03.1: 3,964,708 .. 6,951,816									
qBeh-3-1	45.8	Chr03.1: 19,318,294 .. 23,039,109	Psote03G0119900.1	xylem bark cysteine peptidase	Chr03.1	22,734,739	22,736,264	Glyma.08G116300	2.36E-153	AT1G20850.1	3.43E-55
qBeh-3-2	65.7	Chr03.1: 24,596,586 .. 32,187,119	Psote03G0166900.1	SAD2/ACPD	Chr03.1	33,688,946	33,692,586	Glyma.07G207200	0.00E+00	AT2G43710.1	0.00E+00
qLin-5-1	37.1	Chr05.1: 6,164,990 .. 11,795,082	Psote05G0053200.1	BCCP-ACC	Chr05.1	7,543,397	7,548,100	Glyma.09G248900	4.97E-87	AT5G15530.1	9.46E-41
qLin-5-2	##	Chr05.1: 8,654,322 .. 15,574,958	Psote05G0083100.1	FabH - KASIII	Chr05.1	11,990,726	11,995,665	Glyma.09G277400	0.00E+00	AT1G62640.1	0.00E+00
			Psote05G0096100.1	accA	Chr05.1	14,177,221	14,186,716	Glyma.18G195700	7.13E-168	AT2G38040.2	1.35E-146
qProt-7-1	90.8	Chr07.2: 16,623,267 .. 21,707,553	Psote07G0237900.1	FabG	Chr07.2	22,534,654	22,558,592	Glyma.18G009200	0.00E+00	AT1G24360.1	2.94E-138

Supplementary Table 4. 3 Winged bean genes homologues to soybean and Arabidopsis.

WB Genes	QTL	KEGG name	Accession (E-value)	Lowest E-value	Greatest identity %
Psote03G0078700.1/ Psote03G0078800.1	qProt-3-1		Glyma.05G169200/ Glyma.05G169100	0	86.31
			AT1G07750.1/ AT2G28680.1	2.68872E-95	41.34
Psote03G0119900.1	qProt-3-1		Glyma.08G116300	2.357E-153	77.2
			AT1G20850.1	3.43E-55	40.36
Psote07G0210200.1	qProt-7-1		Glyma.11G216500	0	82.04
			AT3G03450.1	0	64.06
Psote01G0031000.1	qOl-1-1	FabI/ENR1, MOD1 At	Glyma.11G101400	0	89.14
			AT2G05990.2	4.02651E-167	70
Psote02G0175100.1		FabD	Glyma.11G164500	0	93.99
			AT2G30200.1	0	84.97
Psote06G0023300.1			Glyma.07G019100	0	87.32
			AT3G16170.1	0	63.35
Psote05G0096100.1	qLin-5-1 qLin-5-2	ACCA-3 ACCA-2	Glyma.18G195700	7.1312E-168	86.67
			AT2G38040.2	1.3508E-146	79.37
Psote05G0053200.1			Glyma.09G248900	4.96797E-87	90.14

	qPal-5-1 qLin-5-1 qLin-5-2		AT5G15530.1	9.46479E-41	85.71
Psote02G0085400.1		FAD2	Glyma.04G104900	0	93.85
			AT1G36160.1	0	81.79
Psote09G0032300.1			Glyma.13G057400	1.2553E-109	80.38
			AT5G16390.1	1.24161E-46	50.96
Psote03G0022300.1	qPal-3-1 qStear-3-1	ACCC-2	Glyma.05G221100	0	95.90
			AT5G35360.1	0	84.91
Psote05G0083100.1	qLin-5-1 qLin-5-2	FabH - KASIII	Glyma.09G277400	0	93.95
			AT1G62640.1	0	75.12
Psote06G0352200.1		FabF- KASII-B	Glyma.13G112700	0	94.21
			AT1G74960.1	0	72.13
Psote07G0237900.1		FabG	Glyma.18G009200	0	92.17
			AT1G24360.1	3.4322E-126	63.67
Psote02G0269200.1	qLin-2-1	FabZ	Glyma.15G052500	2.9388E-138	94.04
			AT5G10160.1	1.2979E-106	81.36
Psote01G0031000.1	qOl-1-1	FabI / ENR1, MOD1 At	Glyma.11G101400	0	89.14
			AT2G05990.2	4.0265E-167	70.00
Psote03G0381000.1		MERC	Glyma.13G330100	1.4091E-162	81.94
			AT3G45770.2	4.6585E-140	70.03
Psote08G0128000.1			Glyma.13G038600	0	83.91

		SACPD-C / SACPD	AT1G43800.1	0	74.57
Psote03G0166900.1		SAD2 / ACPD	Glyma.07G207200	0	96.68
			AT2G43710.1	0	81.05
Psote06G0283600.1		FAT1 / FATB for AT	Glyma.05G012300	0	94.23
			AT1G08510.1	0	75.85
Psote02G0195400.1			Glyma.06G211300	0	85.89
			AT1G08510.1	0	71.96
Psote02G0162800.1		FATB AT	Glyma.06G168100	0	88.66
			AT1G08510.1	4.2127E-108	57.14

Supplementary Table 4. 4 QTLs detected from MapQTL 6.0

Trait	Nr	Group	Position	Marker	LOD	PVE (%)	Additive	Dominance
protein		Chr3						
protein	270	Chr3	56.409	23504_38:A>G	3.51	7.8	-1.09557	0.646436
protein	273	Chr3	59.104	13560_30:G>T	4.19	9.2	-1.1532	0.800761
protein	279	Chr3	64.614	22478_40:C>T	0.03	0.1	-0.210716	-0.233147
protein	416	Chr2	0	24732_23:G>A	0.04	0.1	0.171205	-0.0735221
protein	432	Chr2	14.748	24545_60:G>A	2.42	5.2	0.803529	-0.500064
protein	435	Chr2	17.285	24406_22:C>G	0.03	0.1	-0.102797	-0.250712
protein	508	Chr2	83.605	21743_45:T>C	0.19	0.4	-0.270029	0.0266708
protein	526	Chr2	101.38	13849_61:A>C	2.98	6.5	-0.991129	0.796457
protein	535	Chr2	109.291	26618_40:C>G	0.03	0.1	-0.152851	-0.121659
protein	1237	Chr6	0	10305_25:T>C	2.69	5.8	-0.965631	0.445229
protein	1243	Chr6	5.232	11144_66:A>G	0.44	0.9	-1.60179	-0.856501
protein	1444	Chr7	85.826	12919_21:G>A	0.15	0.4	0.584826	0.472104
protein	1450	Chr7	90.798	18371_20:A>G	2.96	8.1	-1.00391	0.791486
protein	1455	Chr7	94.087	24158_45:G>A	1.03	2.7	1.74917	1.7835
fat	416	Chr2	0	24732_23:G>A	3.3	12.3	-0.551773	0.44105
fat	429	Chr2	12.615	14631_48:A>C	0.75	2.5	0.456873	0.283534

fat	859	Chr5	245.036	13186_20:T>C	0.01	0	-0.0027824	-0.0901261
fat	874	Chr5	259.438	8171_14:C>G	3.23	12	0.690793	0.054827
fat	886	Chr5	270.571	19732_29:G>A	0.15	0.5	-0.136762	0.208927
fat	1466	Chr9	8.818	13440_29:C>T	0.82	3.2	-0.26714	0.19704
fat	1474	Chr9	16.211	11154_46:G>T	3.04	11.3	-0.588509	0.223131
fat	1487	Chr9	28.534	15409_18:T>A	0.92	3.1	0.270055	-0.35342
Palmitic	212	Chr3	4.549	21324_46:T>C	1.56	4.8	-0.32839	-0.830365
Palmitic	215	Chr3	7.37	25763_26:G>A	3.38	9.9	-0.436129	-1.09414
Palmitic	219	Chr3	11.014	25702_43:A>G	0.73	1.9	0.150666	0.800627
Palmitic	351	Chr3	120.018	15330_48:T>G	0.56	1.5	-0.397775	-0.269526
Palmitic	362	Chr3	129.664	22383_53:C>T	2.45	7	0.610103	-0.161908
Palmitic	366	Chr3	132.912	8749_67:G>T	0.34	0.9	0.206503	0.491579
Palmitic	598	Chr2	160.486	14181_16:T>G	0.04	0.1	-0.0640104	-0.0846277
Palmitic	599	Chr5	0	14758_43:A>C	3.5	10.2	0.589571	-0.386369
Palmitic	615	Chr5	15.837	12217_11:A>G	0.11	0.3	-0.124178	0.176572
Palmitic	755	Chr5	147.987	18027_13:C>T	0.86	2.5	0.415702	0.0995473
Palmitic	767	Chr5	159.545	23909_49:A>G	2.18	6.2	0.702633	0.204634
Palmitic	792	Chr5	184.482	24470_67:C>T	0.19	0.5	0.491471	0.124615
Palmitic	874	Chr5	259.438	8171_14:C>G	0.41	1.3	-0.306357	-0.146074
Palmitic	886	Chr5	270.571	19732_29:G>A	3	8.7	-0.25608	-0.974231

Palmitic	890	Chr5	273.897	11489_5:T>C	0.39	1	0.850603	-0.200743
Palmitic	1318	Chr6	75.405	19053_20:G>C	0.29	0.8	-0.0055611	0.417763
Palmitic	1323	Chr6	79.263	24410_13:G>C	2.76	7.9	-0.627821	-0.196934
Palmitic	1324	Chr6	79.675	8215_65:G>T	0.04	0.1	13.1929	-13.5631
Steraric	211	Chr3	4.244	16321_22:A>G	0.32	1.3	0.652398	-0.391753
Steraric	215	Chr3	7.37	25763_26:G>A	2.27	10	-0.306562	-0.906885
Steraric	219	Chr3	11.014	25702_43:A>G	0.11	0.5	0.211029	0.30243
Steraric	1318	Chr6	75.405	19053_20:G>C	0.09	0.4	0.051521	0.235015
Steraric	1323	Chr6	79.263	24410_13:G>C	2.01	8.8	-0.557658	0.0533016
Steraric	1324	Chr6	79.675	8215_65:G>T	0.6	2.5	19.0675	-17.6514
Oleic	80	Chr1	72.035	17891_25:T>C	0.27	1	-0.688187	0.632911
Oleic	92	Chr1	83.493	13470_64:C>T	3.5	14.3	3.94874	1.55128
Oleic	96	Chr1	86.901	15891_23:T>G	0.33	1.2	-1.71845	-1.63695
Oleic	1262	Chr6	23.24	25629_22:C>A	0.34	1.2	-1.83312	1.53892
Oleic	1270	Chr6	30.869	16696_52:C>T	1.45	5.6	-0.769838	2.21133
Oleic	1284	Chr6	44.043	12155_32:G>A	0.27	1	0.46034	-0.957521
Oleic	874	Chr5	259.438	8171_14:C>G	0.34	1.3	-0.776448	-1.62683
Oleic	890	Chr5	273.897	11489_5:T>C	2.34	9.3	-1.3397	2.70804
Oleic	901	Chr5	283.422	13184_65:C>A	0.56	2	-1.31464	-2.00649
Linoleic	441	Chr2	22.039	7847_38:T>C	0.58	2.1	0.177029	-0.324137

Linoleic	445	Chr2	25.921	8506_16:A>G	2.7	9.4	0.749452	-0.190664
Linoleic	456	Chr2	36.575	20984_68:G>A	0.31	1	-0.397264	-0.375376
Linoleic	1493	Chr9	34.058	26188_63:C>T	0.07	0.2	0.0541274	-0.162517
Linoleic	1508	Chr9	48.053	27258_49:C>T	4.58	16.9	-0.912233	-0.375454
Linoleic	1511	Chr9	50.207	22666_55:A>C	0.34	1.1	-0.198507	-0.841156
Linoleic	626	Chr5	26.713	22484_42:C>T	0.55	1.7	-0.248434	-0.841324
Linoleic	637	Chr5	37.137	23128_19:A>G	4.06	14.7	-0.917411	-2.08329
Linoleic	650	Chr5	49.63	13546_10:C>G	0.17	0.5	0.337901	0.276932
Linoleic	673	Chr5	70.83	9586_63:G>C	0.51	1.6	-0.146271	0.560484
Linoleic	690	Chr5	86.797	19704_36:C>T	3.43	12.2	1.30853	1.5662
Linoleic	700	Chr5	95.836	24910_50:T>A	0.25	0.8	-0.388908	0.098834
Behenic	248	Chr3	37.38	22709_12:C>G	1.64	7.1	0.881745	0.802813
Behenic	257	Chr3	45.835	17225_8:C>G	4.19	16.9	1.23385	0.987873
Behenic	258	Chr3	46.697	14405_26:T>G	0.13	0.5	-0.161268	-0.411531
Behenic	273	Chr3	59.104	13560_30:G>T	0.4	1.4	-0.706622	0.419798
Behenic	281	Chr3	65.703	26777_47:A>C	2.66	10.3	-0.779598	-0.988813
Behenic	286	Chr3	69.733	11206_45:G>A	0.03	0.1	0.173771	-0.0880023
Behenic	536	Chr2	109.546	28539_54:G>A	0.95	3.4	-0.30215	1.71929
Behenic	539	Chr2	111.961	21034_27:G>C	2.12	8.1	0.291745	-0.987478
Behenic	542	Chr2	114.79	23158_18:C>T	0.25	0.9	0.419058	0.323928

Supplementary Table 5. 1 Clean reads, Q20 and Q30 percentages of the bases whose Q Phred values are greater than 20 and 30 respectively are shown as well as GC (%) of total bases

sample	Raw reads	Raw bases	Clean reads	Clean bases	Clean reads (%)	Error rate	Q20	Q30	GC (%)
FP15_1_D15	83019636	12.45G	81859908	12.28G	98.60	0.03	96.88	91.46	44.66
FP15_2_D15	86279460	12.94G	85033992	12.76G	98.56	0.03	96.91	91.58	44.59
FP15_3_D15	84726250	12.71G	83541046	12.53G	98.60	0.03	97.04	91.89	44.67
FP15_1_D30	81845304	12.28G	80778998	12.12G	98.70	0.03	96.68	91.14	44.57
FP15_2_D30	105874034	15.88G	104543234	15.68G	98.74	0.03	97.41	92.6	44.53
FP15_3_D30	108650604	16.3G	107347656	16.1G	98.80	0.03	96.64	91.02	44.58
FP15_1_D37	78995824	11.85G	78545950	11.78G	99.43	0.03	96.68	91.3	44.31
FP15_2_D37	97222602	14.58G	96039640	14.41G	98.78	0.03	96.68	91.01	44.39
FP15_3_D37	70670242	10.6G	69579460	10.44G	98.46	0.03	96.72	91.18	44.34
FP15_1_D45	82109052	12.32G	80966484	12.14G	98.61	0.03	97.07	92.1	43.88
FP15_2_D45	97840800	14.68G	96727148	14.51G	98.86	0.03	97.23	92.19	44.16
FP15_3_D45	80420126	12.06G	79494356	11.92G	98.85	0.03	97.46	92.7	44.2
SFP15_1_D15	84062670	12.61G	83258516	12.49G	99.04	0.03	96.31	90.33	44.42
SFP15_2_D15	70424948	10.56G	69553418	10.43G	98.76	0.03	95.98	89.63	44.48
SFP15_3_D15	76063566	11.41G	75136384	11.27G	98.78	0.03	96.51	90.7	44.44
SFP15_1_D30	100036094	15.01G	99291972	14.89G	99.26	0.03	96.27	90.01	44.59
SFP15_2_D30	86684918	13G	86084220	12.91G	99.31	0.03	96.38	90.24	44.77
SFP15_3_D30	127194732	19.08G	126106718	18.92G	99.14	0.03	96.4	90.38	44.55
SFP15_1_D45	90356160	13.55G	89681762	13.45G	99.25	0.03	96.42	90.3	45.42
SFP15_2_D45	83418114	12.51G	82777686	12.42G	99.23	0.03	96.4	90.21	45.3
SFP15_3_D45	91372542	13.71G	90720426	13.61G	99.29	0.03	96.43	90.43	45.25

Supplementary Table 5. 2 Summary of RNA-Seq read number and mapping results

Sample	Total reads	Mapped reads	Unmapped reads
FP15_1_D15	81859908	92.04%	7.96%
FP15_2_D15	85033992	92.68%	7.32%
FP15_3_D15	83541046	92.78%	7.22%
FP15_1_D30	80778998	92.32%	7.68%
FP15_2_D30	104543234	93.73%	6.27%
FP15_3_D30	107347656	92.36%	7.64%
FP15_1_D37	78545950	91.69%	8.31%
FP15_2_D37	96039640	93.59%	6.41%
FP15_3_D37	69579460	93.27%	6.73%
FP15_1_D45	80966484	93.29%	6.71%
FP15_2_D45	96727148	94.26%	5.74%
FP15_3_D45	79494356	94.28%	5.72%
SFP15_1_D15	83258516	93.43%	6.57%
SFP15_2_D15	69553418	93.44%	6.56%
SFP15_3_D15	75136384	93.88%	6.12%
SFP15_1_D30	99291972	91.22%	8.78%
SFP15_2_D30	86084220	92.51%	7.49%
SFP15_3_D30	126106718	91.16%	8.84%
SFP15_1_D45	89681762	91.14%	8.86%
SFP15_2_D45	82777686	93.03%	6.97%
SFP15_3_D45	90720426	92.86%	7.14%

Supplementary Table 5. 3 Winged bean gene homologues to LEA soybean gene

Query	Chromosome	Lowest E-value	Accession	Greatest identity %
Psote04G0062600.1	Chr04.1	2.8158E-106	Glyma.20G147600	87.19

Supplementary Table 5. 4 Winged bean genes in with high similarity to soybean genes involved in the fatty acid biosynthesis pathway

KEGG Pathway	Query	Chromo-some	Lowest E-value	Soybean ID gene	Greatest identity %	Accession
6.4.1.2	Psote05G0096100.1	Chr05.1	7.13124E-168	Glyma.18G195700	86.67	Glyma.18G196000
	Psote05G0053200.1	Chr05.1	4.968E-87	Glyma.09G248900	90.14	Glyma.18G243500
	Psote02G0085400.1	Chr02.1	0	Glyma.04G104900	93.85	Glyma.04G104900
	Psote09G0032300.1	Chr09.1	1.255E-109	Glyma.13G057400	80.38	Glyma.13G057400
	Psote03G0022300.1	Chr03.1	0	Glyma.05G221100	95.90	Glyma.08G027600
	Psote05G0032900.1	Chr05.1	1.645E-108	Glyma.18G265300	73.85	Glyma.18G265300
FabD	Psote02G0175100.1	Chr02.3	0	Glyma.11G164500	93.99	Glyma.18G057700
FabH	Psote05G0083100.1	Chr05.1	0	Glyma.09G277400	93.95	Glyma.09G277400
	Psote06G0352200.1	Chr06.4	0	Glyma.13G112700	94.21	Glyma.13G112700
FabF	Psote04G0300400.1	Chr04.3	1.393E-80	Glyma.13G128000	96.43	Glyma.10G041100
	Psote03G0026000.1	Chr03.1	0	Glyma.08G024700	91.73	Glyma.18G091100
	Psote03G0088900.1	Chr03.1	0	Glyma.08G084300	93.18	Glyma.08G084300
FabG	Psote05G0203200.1	Chr05.2	8.985E-154	Glyma.16G042000	88.54	Glyma.16G042000
	Psote07G0237900.1	Chr07.2	0	Glyma.18G009200	92.17	Glyma.11G248100
	Psote03G0105900.1	Chr03.1	2.551E-176	Glyma.08G102100	90.77	Glyma.08G102100
FabZ	Psote02G0269200.1	Chr02.5	2.939E-138	Glyma.15G052500	94.04	Glyma.08G073900
FabI	Psote05G0117000.1	Chr05.1	0	Glyma.18G156100	86.19	Glyma.08G345900
	Psote01G0031000.1	Chr01.1	0	Glyma.11G101400	89.14	Glyma.11G101400
	Psote03G0381000.1	Chr03.5	1.409E-162	Glyma.13G330100	81.94	Glyma.13G330100
MECR	Psote06G0283600.1	Chr06.4	0	Glyma.05G012300	94.23	Glyma.05G012300

3.1.2.21	Psote02G0195400.1	Chr02.5	0	Glyma.06G211300	85.89	Glyma.06G211300
	Psote02G0163100.1	Chr02.3	0	Glyma.04G197500	71.58	Glyma.04G197500
	Psote02G0162800.1	Chr02.3	0	Glyma.06G168100	88.66	Glyma.06G168100
3.1.2.14	Psote02G0195400.1	Chr02.5	0	Glyma.06G211300	85.89	Glyma.06G211300
	Psote02G0163100.1	Chr02.3	0	Glyma.04G197500	71.58	Glyma.04G197500
	Psote02G0162800.1	Chr02.3	0	Glyma.06G168100	88.66	Glyma.06G168100
	Psote04G0059800.1	Chr04.1	0	Glyma.20G143900	91.69	Glyma.20G143900
6.2.1.3	Psote02G0093600.1	Chr02.1	0	Glyma.06G112900	91.35	Glyma.06G112900
	Psote07G0243000.1	Chr07.2	0	Glyma.11G254100	93.67	Glyma.11G254100
	Psote04G0191600.1	Chr04.2	0	Glyma.10G010800	93.47	Glyma.10G010800
	Psote01G0429300.1	Chr01.7	0	Glyma.07G161900	88.77	Glyma.07G161900
	Psote04G0339800.1	Chr04.3	0	Glyma.13G010100	88.76	Glyma.20G060100
	Psote08G0127900.1	Chr08.1	0	Glyma.14G121400	84.10	Glyma.14G121400
1.14.19.2	Psote03G0166900.1	Chr03.1	0	Glyma.07G207200	96.68	Glyma.07G207200
	Psote08G0128000.1	Chr08.1	0	Glyma.13G038600	83.91	Glyma.13G038600

Supplementary Table 5. 5 Winged bean and soybean genes involved in the fatty acid biosynthesis pathway, as shown in KEGG

KEGG Path-way	Gene symbol	Soybean gene mentioned in the KEGG	Winged bean gene	Chromo-some	QTL	Lowest E-value	Soybean accession	Identity %
6.4.1.2	ACCA-3	Glyma.18G196000	Psote05G0096100.1	Chr05.1	qLin-5-2	7.1312E-168	Glyma.18G195700	86.67

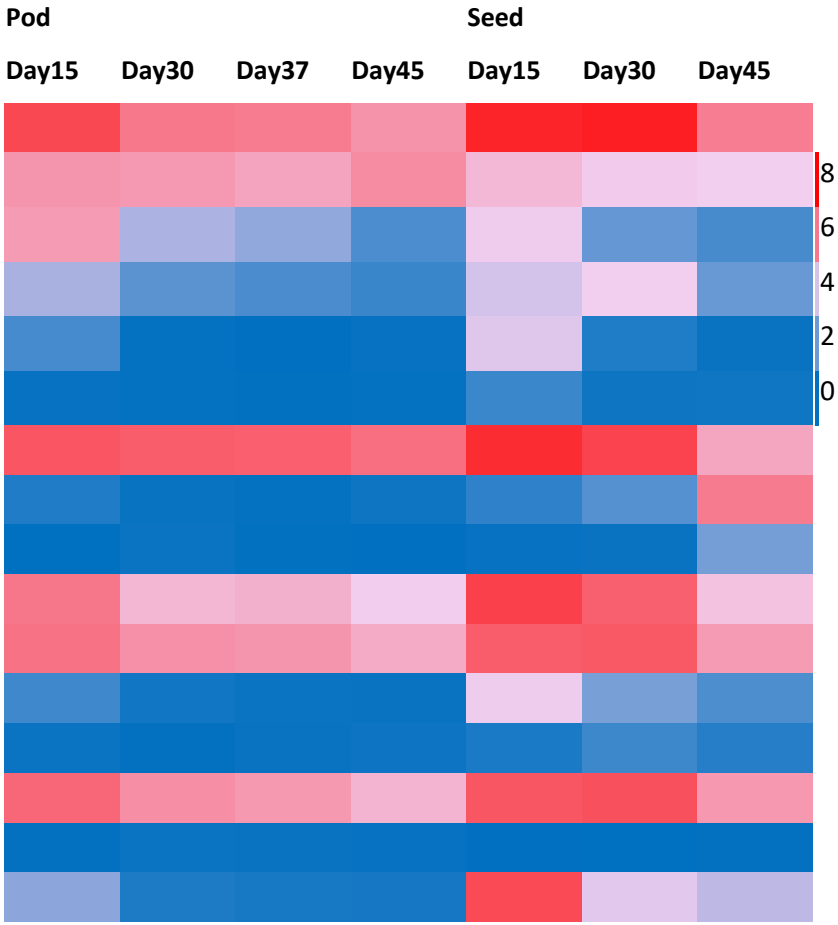
	ACCA-2	Glyma.18G195900	Psote05G0096100.1	Chr05.1	qLin-5-2	7.1312E-168	Glyma.18G195700	86.67
		Glyma.19G028800	Psote05G0032900.1	Chr05.1		1.645E-108	Glyma.18G265300	73.85
		Glyma.09G248900	Psote05G0053200.1	Chr05.1	qPal-5-1	4.96797E-87	Glyma.09G248900	90.14
	ACC1	Glyma.04G104900	Psote02G0085400.1	Chr02.1		0	Glyma.04G104900	93.85
	ACC2/BCCP	Glyma.13G057400	Psote09G0032300.1	Chr09.1	qLin-9-1	1.2553E-109	Glyma.13G057400	80.38
	ACCC-2	Glyma.05G221100	Psote03G0022300.1	Chr03.1	qPal-3-1 qStear-3-1	0	Glyma.05G221100	95.90
	ACCB-2	Glyma.18G265300	Psote05G0032900.1	Chr05.1		1.645E-108	Glyma.18G265300	73.85
	ACCA-1	Glyma.18G195700	Psote05G0096100.1	Chr05.1		7.1312E-168	Glyma.18G195700	86.67
FabD		Glyma.11G164401	Psote07G0193100.1	Chr07.2		0	Glyma.11G164200	68.40
		Glyma.11G164500	Psote02G0175100.1	Chr02.3		0	Glyma.11G164500	93.99
	MT2	Glyma.18G057700	Psote02G0175100.1	Chr02.3		0	Glyma.11G164500	93.99
FabH		Glyma.15G003100	Psote05G0083100.1	Chr05.1		0	Glyma.09G277400	93.95
	KASIII	Glyma.09G277400	Psote05G0083100.1	Chr05.1	qLin-5-1	0	Glyma.09G277400	93.95
		Glyma.18G211400	Psote05G0083100.1	Chr05.1		0	Glyma.09G277400	93.95
FabF	KASII-B	Glyma.13G112700	Psote06G0352200.1	Chr06.4		0	Glyma.13G112700	94.21
		Glyma.13G128000	Psote04G0300400.1	Chr04.3		1.39302E-80	Glyma.13G128000	96.43
		Glyma.10G041100	Psote04G0300400.1	Chr04.3		1.39302E-80	Glyma.13G128000	96.43

		Glyma.15G181500	Psote06G0352200.1	Chr06.4		0	Glyma.13G112700	94.21
		Glyma.08G024700	Psote03G0026000.1	Chr03.1		0	Glyma.08G024700	91.73
		Glyma.05G218600	Psote03G0026000.1	Chr03.1		0	Glyma.08G024700	91.73
		Glyma.18G091100	Psote03G0026000.1	Chr03.1		0	Glyma.08G024700	91.73
	KASI	Glyma.08G084300	Psote03G0088900.1	Chr03.1		0	Glyma.08G084300	93.18
	KASII-A	Glyma.17G047000	Psote06G0352200.1	Chr06.4		0	Glyma.13G112700	94.21
FabG		Glyma.11G248000	Psote07G0237900.1	Chr07.2		0	Glyma.18G009200	92.17
		Glyma.16G042000	Psote05G0203200.1	Chr05.2		8.985E-154	Glyma.16G042000	88.54
		Glyma.18G009200	Psote07G0237900.1	Chr07.2		0	Glyma.18G009200	92.17
		Glyma.08G102100	Psote03G0105900.1	Chr03.1		2.5506E-176	Glyma.08G102100	90.77
FabZ		Glyma.08G179900	Psote02G0269200.1	Chr02.5	qLin-2-1	2.9388E-138	Glyma.15G052500	94.04
		Glyma.15G052500	Psote02G0269200.1	Chr02.5	qLin-2-1	2.9388E-138	Glyma.15G052500	94.04
		Glyma.08G073900	Psote02G0269200.1	Chr02.5	qLin-2-1	2.9388E-138	Glyma.15G052500	94.04
FabI		Glyma.12G027300	Psote01G0031000.1	Chr01.1	qOl-1-1	0	Glyma.11G101400	89.14
		Glyma.18G156100	Psote05G0117000.1	Chr05.1		0	Glyma.18G156100	86.19
		Glyma.11G101400	Psote01G0031000.1	Chr01.1	qOl-1-1	0	Glyma.11G101400	89.14
		Glyma.08G345900	Psote05G0117000.1	Chr05.1		0	Glyma.18G156100	86.19
MECR	MECR	Glyma.13G330100	Psote03G0381000.1	Chr03.5	qBeh-3-2	1.4091E-162	Glyma.13G330100	81.94

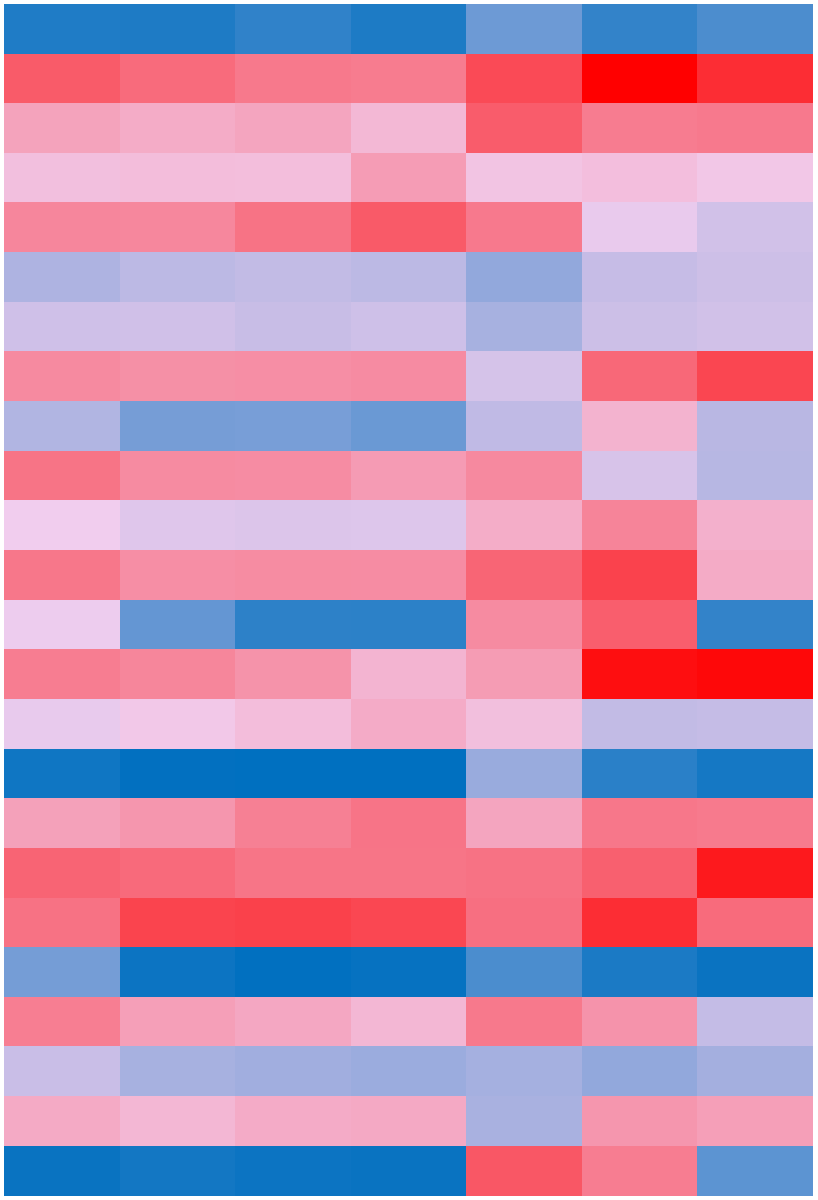
3.1.2.21	FAT1	Glyma.05G012300	Psote06G0283600.1	Chr06.4	0	Glyma.05G012300	94.23
3.1.2.14		Glyma.06G211300	Psote02G0195400.1	Chr02.5	0	Glyma.06G211300	85.89
	FATB1B	Glyma.17G120400	Psote06G0283600.1	Chr06.4	0	Glyma.05G012300	94.23
	FATB	Glyma.04G197400	Psote02G0162800.1	Chr02.3	0	Glyma.06G168100	88.66
		Glyma.04G197500	Psote02G0163100.1	Chr02.3	0	Glyma.04G197500	71.58
	FATB	Glyma.06G168000	Psote02G0162800.1	Chr02.3	0	Glyma.04G197500	71.58
		Glyma.04G151600	Psote06G0283600.1	Chr02.5	0	Glyma.06G211300	85.89
	FATB	Glyma.06G168100	Psote02G0162800.1	Chr02.3	0	Glyma.06G168100	88.66
6.2.1.3		Glyma.20G143900	Psote04G0059800.1	Chr04.1	0	Glyma.20G143900	91.69
		Glyma.06G112900	Psote02G0093600.1	Chr02.1	0	Glyma.06G112900	91.35
		Glyma.20G007900	Psote04G0339800.1	Chr01.7	0	Glyma.07G161900	88.77
		Glyma.11G254100	Psote07G0243000.1	Chr07.2	0	Glyma.11G254100	93.67
		Glyma.10G249700	Psote04G0059800.1	Chr04.1	0	Glyma.20G143900	91.69
		Glyma.12G047400	Psote04G0191600.1	Chr01.7	0	Glyma.07G161900	88.77
		Glyma.10G010800	Psote04G0191600.1	Chr04.2	0	Glyma.10G010800	93.47
		Glyma.07G161900	Psote01G0429300.1	Chr01.7	0	Glyma.07G161900	88.77
		Glyma.20G060300	Psote04G0339800.1	Chr04.3	0	Glyma.13G010100	88.76
		Glyma.13G010100	Psote04G0339800.1	Chr04.3	0	Glyma.13G010100	88.76
1.14.19.2	SACPD-C	Glyma.14G121400	Psote08G0127900.1	Chr08.1	0	Glyma.14G121400	84.10
	SAD2	Glyma.07G207200	Psote03G0166900.1	Chr03.1	0	Glyma.07G207200	96.68
	SACPD	Glyma.13G038600	Psote08G0128000.1	Chr08.1	0	Glyma.13G038600	83.91
	ACPD	Glyma.02G138100	Psote03G0166900.1	Chr03.1	0	Glyma.07G207200	96.68

Supplementary Table 5. 6 Heatmap of genes in the fatty acid biosynthesis pathway. Transcriptomic profiles of fatty acid biosynthesis genes at Day 15, 30, 37 and 45 (from left to right) in the maturing pods and seeds, with the colour scale reflecting log2(FPKM+1) values.

Winged bean gene	Gene name
Psote01G0031000.1	MOD1
Psote01G0242100.1	LACS4
Psote01G0429300.1	CCL6
Psote02G0093600.1	LACS9
Psote02G0162800.1	FATB
Psote02G0163100.1	FATB1
Psote02G0175100.1	Mcat
Psote02G0195400.1	FATB1
Psote02G0195500.1	Psote02G0195500.1
Psote02G0269200.1	fabZ
Psote03G0022300.1	POPTRDRAFT_831870
Psote03G0026000.1	KAS1
Psote03G0029000.1	LACS4
Psote03G0088900.1	KAS1
Psote03G0105900.1	CLKR27
Psote03G0112600.1	AAE16

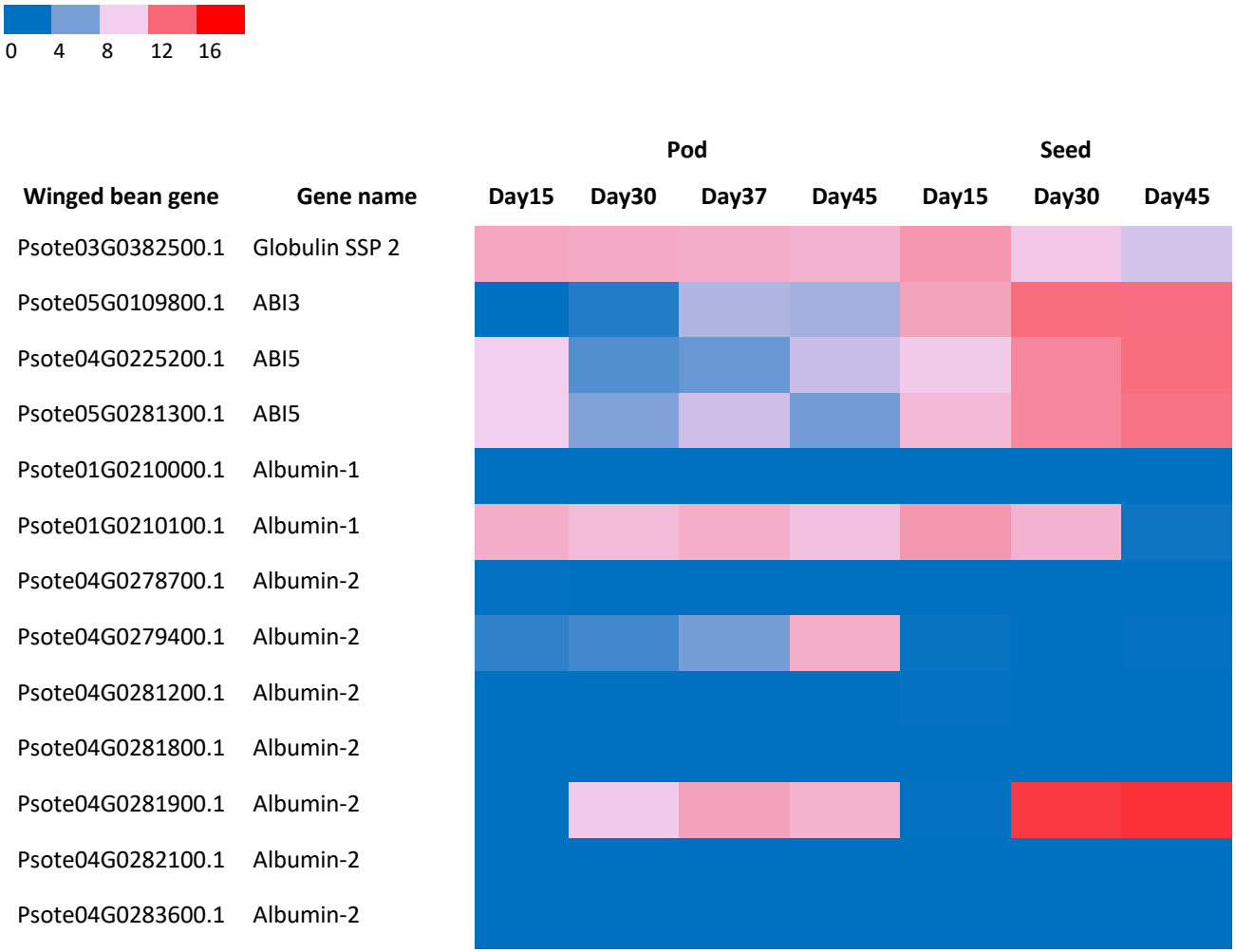


Psote03G0131400.1	fabZ
Psote03G0166900.1	ACPD
Psote03G0381000.1	At3g45770
Psote04G0059800.1	LACS7
Psote04G0191600.1	LACS1
Psote04G0300200.1	KAS
Psote04G0300400.1	KAS
Psote04G0339800.1	LACS8
Psote05G0032900.1	BCCP2
Psote05G0053200.1	ACCB-1
Psote05G0083100.1	KAS3A
Psote05G0096100.1	CAC3
Psote05G0117000.1	MOD1
Psote05G0119500.1	FATA2
Psote05G0203200.1	At3g03980
Psote05G0308600.1	LACS1
Psote06G0023300.1	CCL8
Psote06G0283600.1	Psote06G0283600.1
Psote06G0352200.1	Psote06G0352200.1
Psote07G0192800.1	Psote07G0192800.1
Psote07G0237900.1	CLKR27
Psote07G0243000.1	AAE16
Psote08G0117600.1	LACS9
Psote08G0127900.1	S-ACP-DES6

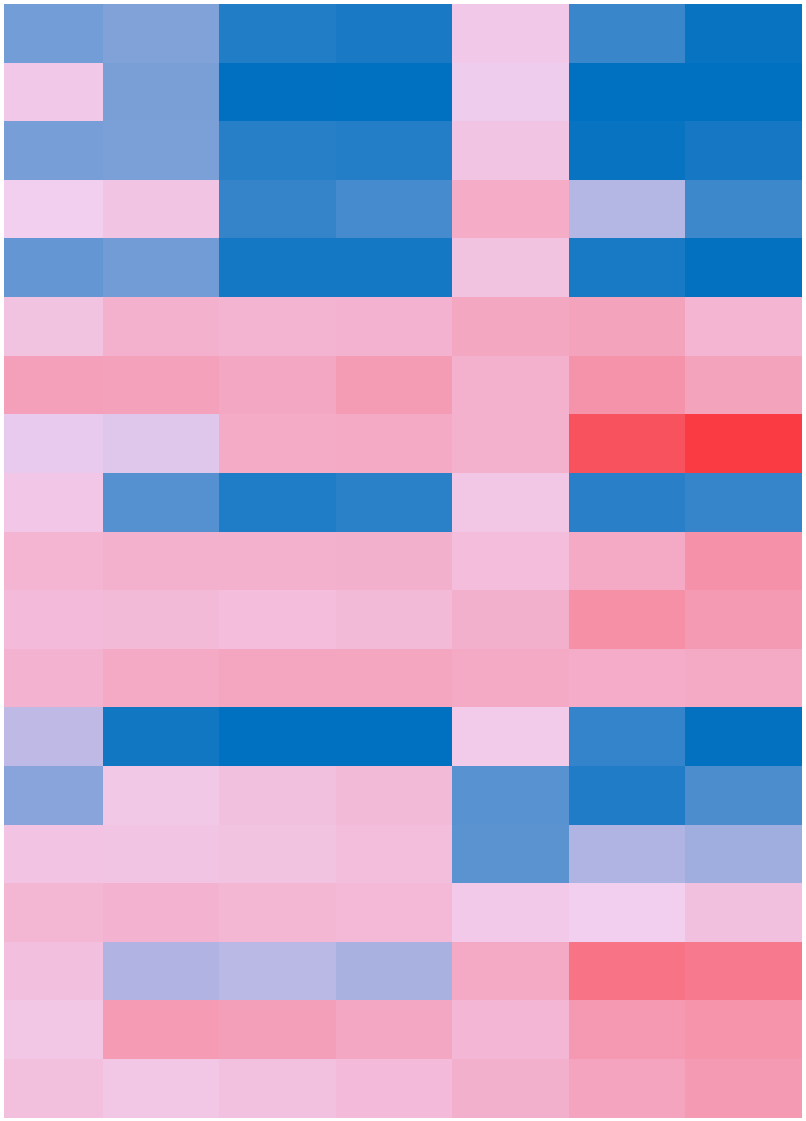


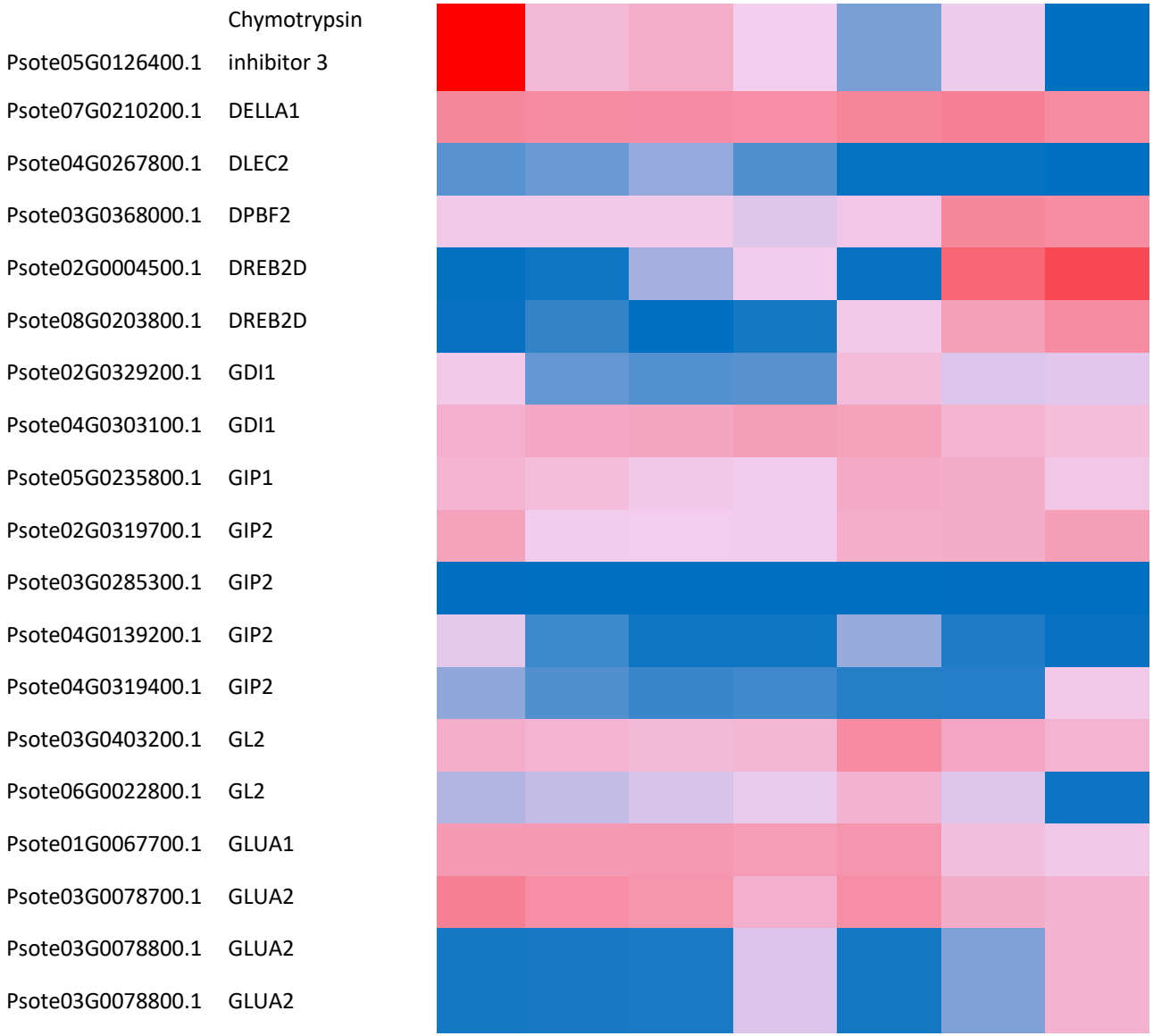
Psote08G0128000.1	S-ACP-DES6	
Psote09G0032300.1	BCCP2	

Supplementary Table 5. 7 Heatmap of the expression profile of seed storage protein-related genes in developing pod and seed tissues. Expression levels [log2 (FPKM+1)] of albumin, globulin, and related genes.

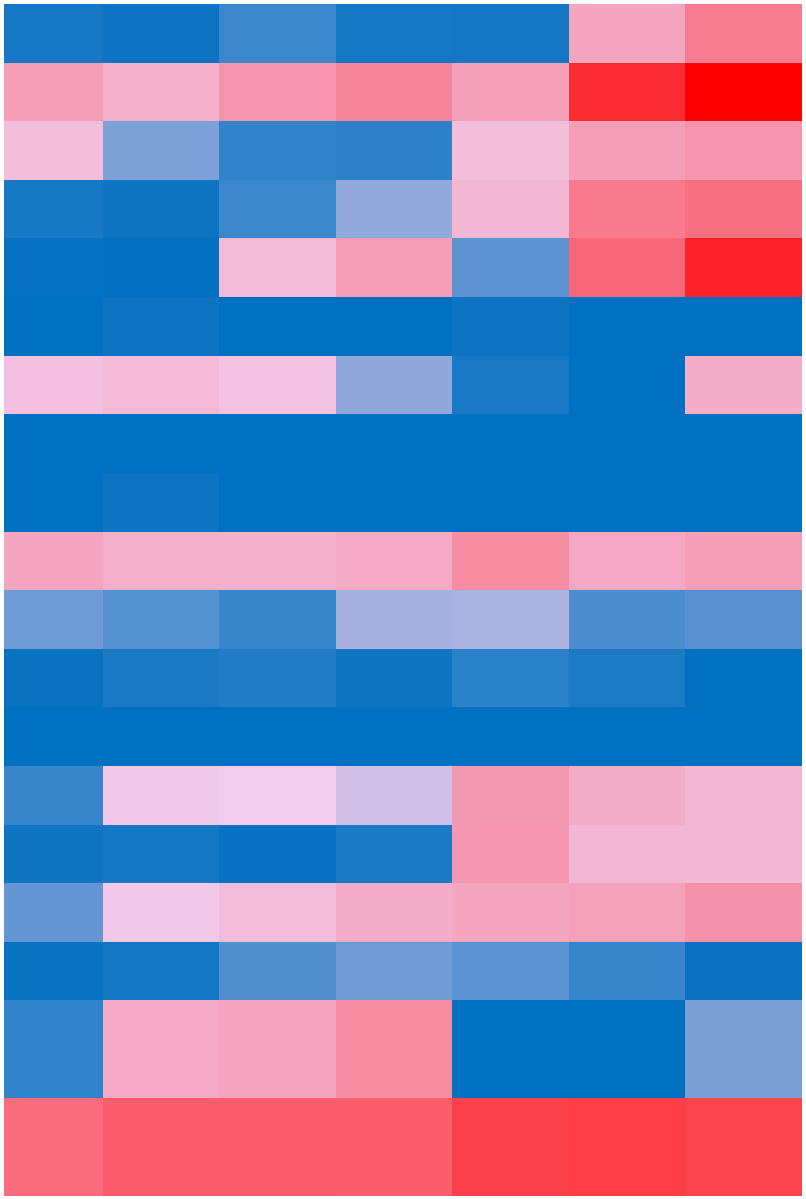


Psote04G0284200.1	Albumin-2
Psote04G0285200.1	Albumin-2
Psote04G0285500.1	Albumin-2
Psote04G0285600.1	Albumin-2
Psote04G0288400.1	Albumin-2
Psote01G0270900.1	APRR2
Psote07G0043000.1	APRR2
Psote01G0296200.1	At2g18540
Psote04G0249100.1	At2g41710
Psote05G0215500.1	At2g41710
Psote09G0131200.1	At5g08430
Psote02G0028400.1	ATH1
Psote08G0176100.1	ATH1
Psote04G0168800.1	ATL6
Psote01G0274500.1	BHLH25
Psote07G0046600.1	BHLH25
Psote07G0046700.1	BHLH25
Psote01G0294000.1	BLH2
Psote07G0069300.1	BLH2

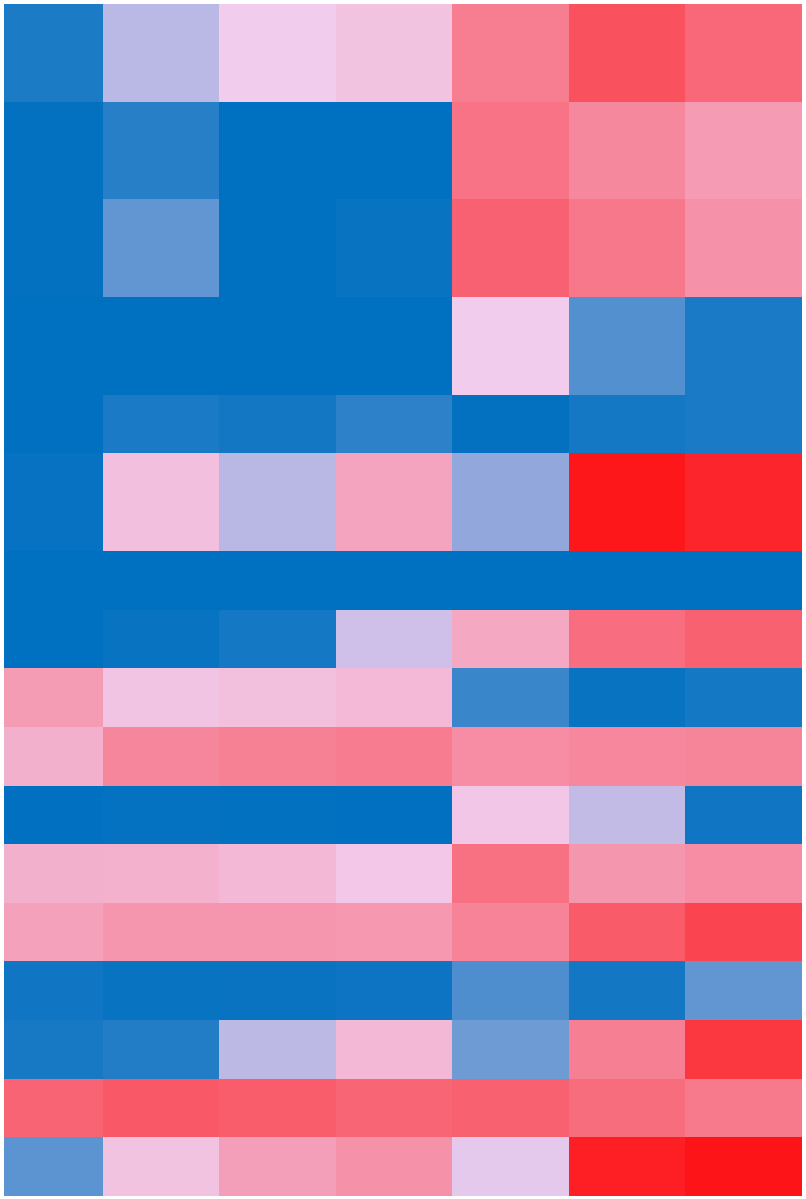


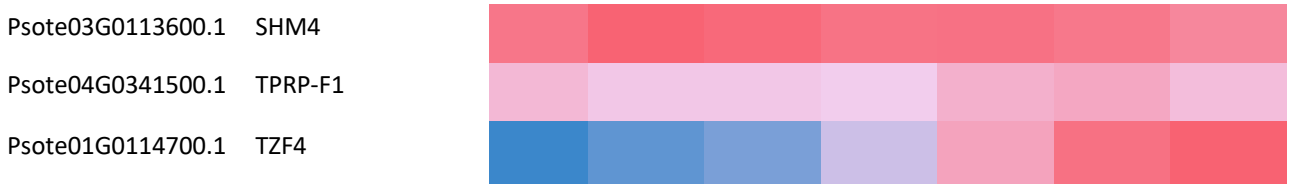

Chymotrypsin inhibitor 3
Psote05G0126400.1
Psote07G0210200.1 DELLA1
Psote04G0267800.1 DLEC2
Psote03G0368000.1 DPBF2
Psote02G0004500.1 DREB2D
Psote08G0203800.1 DREB2D
Psote02G0329200.1 GDI1
Psote04G0303100.1 GDI1
Psote05G0235800.1 GIP1
Psote02G0319700.1 GIP2
Psote03G0285300.1 GIP2
Psote04G0139200.1 GIP2
Psote04G0319400.1 GIP2
Psote03G0403200.1 GL2
Psote06G0022800.1 GL2
Psote01G0067700.1 GLUA1
Psote03G0078700.1 GLUA2
Psote03G0078800.1 GLUA2
Psote03G0078800.1 GLUA2

Psote04G0240700.1	GMPM1
Psote05G0226200.1	GMPM1
Psote02G0243900.1	GRF3
Psote06G0041700.1	GRF5
Psote04G0062600.1	LEA D-34
Psote04G0267900.1	LE1
Psote04G0268000.1	LE1
Psote04G0279000.1	Lectin
Psote04G0279100.1	Lectin
Psote03G0217000.1	LTPG2
Psote01G0228200.1	LTPG5
Psote06G0041300.1	MYB1
Psote02G0244100.1	MYB2
Psote06G0041200.1	MYB2
Psote03G0379100.1	NAC029
Psote02G0211300.1	NAC100
Psote05G0037400.1	NAC100
Psote07G0197200.1	Non-specific lipid-transfer protein
Psote06G0099600.1	Non-specific lipid-transfer protein1

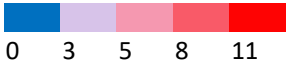


Psote05G0081300.1	Non-specific lipid-transfer protein3
Psote08G0141300.1	Non-specific lipid-transfer protein A
Psote08G0141400.1	Non-specific lipid-transfer protein A
Psote08G0141500.1	Non-specific lipid-transfer protein A
Psote03G0123600.1	NPF4.6
Psote03G0119900.1	P34 probable thiol-protease
Psote03G0245700.1	Psote03G0245700.1
Psote06G0196300.1	Psote06G0196300.1
Psote06G0196400.1	Psote06G0196400.1
Psote06G0259600.1	Psote06G0259600.1
Psote06G0359300.1	Psote06G0359300.1
Psote06G0359600.1	Psote06G0359600.1
Psote06G0400900.1	Psote06G0400900.1
Psote06G0444200.1	Psote06G0444200.1
Psote04G0317500.1	RAB28
Psote03G0113400.1	SAHH
Psote04G0317000.1	SBP

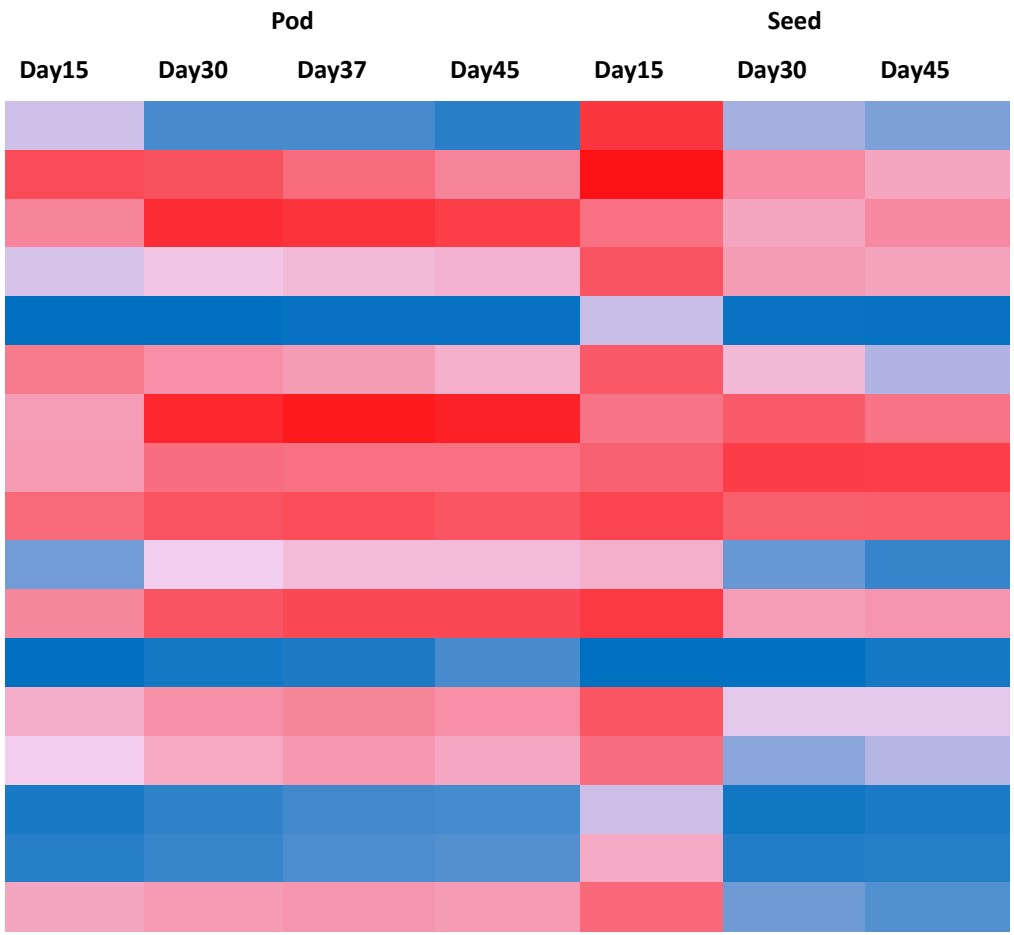



Psote03G0113600.1 SHM4
Psote04G0341500.1 TPRP-F1
Psote01G0114700.1 TZF4

Supplementary Table 5. 8 Heatmap of genes in the flavonoid biosynthesis pathway. Transcriptomic profiles of flavonoid biosynthesis genes at Day 15, 30, 37 and 45 (from left to right) in the maturing pods and seeds, with the colour scale reflecting log2(FPKM+1) values.



Winged bean gene	Gene name
Psote03G0143100.1	ANR
Psote01G0255900.1	ANT17
Psote03G0183300.1	CCOMT
Psote07G0060300.1	CCOMT
Psote07G0060400.1	CCOMT
Psote04G0009400.1	CHI
Psote04G0009600.1	CHI1A
Psote04G0009500.1	CHI1B2
Psote02G0128700.1	CHIL2
Psote01G0233000.1	CHS
Psote01G0233200.1	CHS
Psote02G0100100.1	CHS
Psote01G0232400.1	CHS
Psote01G0232500.1	CHS
Psote01G0232200.1	CHS
Psote01G0232700.1	CHS
Psote01G0232100.1	CHS



Psote01G0232300.1	CHS
Psote01G0207000.1	CHS
Psote01G0233300.1	CHS
Psote05G0176800.1	CHS
Psote07G0244500.1	CHS-1A
Psote03G0114700.1	CHS1
Psote08G0082200.1	CYP73A11
Psote03G0204600.1	CYP75A1
Psote03G0208000.1	CYP75A3
Psote02G0204600.1	CYP75B1
Psote02G0204700.1	CYP75B1
Psote09G0091400.1	CYP93B16
Psote05G0186000.1	CYP98A2
Psote04G0029100.1	Cytochrome P450 CYP73A100
Psote04G0261200.1	DFR
Psote08G0200700.1	DFR
Psote03G0189000.1	F3H
Psote07G0080300.1	FHT
Psote08G0114200.1	FL
Psote08G0114300.1	FL
Psote09G0071300.1	FLS
Psote05G0179900.1	GLC
Psote09G0083100.1	GLC
Psote06G0007900.1	HST

Psote03G0414400.1	HST
Psote03G0416800.1	HST
Psote01G0405800.1	LAR
Psote02G0223400.1	NAD(P)H-dependent
Psote01G0412700.1	6'-deoxychalcone synthase
Psote08G0005500.1	NAD(P)H-dependent
Psote02G0223300.1	6'-deoxychalcone synthase
	NAD(P)H-dependent
Psote04G0247200.1	6'-deoxychalcone synthase
Psote01G0233100.1	Psote01G0233100.1
Psote03G0196200.1	Psote03G0196200.1
Psote05G0015600.1	Psote05G0015600.1
Psote06G0160400.1	Psote06G0160400.1
Psote06G0265900.1	Psote06G0265900.1
Psote08G0082300.1	Psote08G0082300.1
Psote06G0000500.1	SAT
Psote01G0306100.1	SAT
Psote05G0200600.1	SHT
Psote07G0150700.1	SHT
Psote05G0200700.1	SHT
Psote07G0159700.1	SHT
Psote07G0159900.1	SHT
Psote09G0080700.1	SHT
Psote07G0160100.1	SHT

Psote03G0108300.1	TSM1	
Psote03G0273200.1	UGT88F4	

Supplementary Table 5. 9 Phenolic compounds µg per 1 g of plant material. This is an estimate as the data has not been normalised to the internal standards due to the need for further method optimisation. The amount of compounds in winged bean pod and seed at different developmental stages in the table ± SEM (n=3), were analysed using a one-way ANOVA and Tukey's posthoc test, p-values are shown. The different letters represent significant differences among the winged bean accessions.

Plant material	Day	Catechin	Epi-catechin	Procyanidin B1	Procyanidin B2	Procyanidin C1	TOTAL
POD	**15D**	0.022±0.004a	0.02±0.005a	0.016±0.002a	0.022±0.002a	0.021±0.003a	0.101±0.02a
POD	**22D**	0.067±0.015ab	0.084±0.032a	0.018±0.001a	0.036±0.002a	0.035±0.003a	0.241±0.04a
POD	**30D**	0.119±0.048ab	0.102±0.058a	0.017±0.001a	0.025±0.001a	0.022±0.0002a	0.285±0.11a
POD	**37D**	0.049±0.016ab	0.039±0.01a	0.019±0.0004a	0.026±0.001a	0.025±0.001a	0.157±0.01a
POD	**45D**	0.074±0.02ab	0.097±0.055a	0.019±0.001a	0.03±0.002a	0.027±0.002a	0.246±0.04a
SEED	**15D**	0.639±0.047d	3.559±0.275b	0.13±0.019c	1.051±0.056d	1.104±0.047c	6.483±0.44d
SEED	**22D**	0.626±0.113d	2.972±0.302b	0.124±0.026c	0.897±0.175cd	0.91±0.208c	5.529±0.81cd
SEED	**30D**	0.494±0.117cd	3.225±0.111b	0.083±0.02bc	0.602±0.125bc	0.511±0.111b	4.915±0.48bcd
SEED	**37D**	0.324±0.019bc	2.973±0.117b	0.049±0.004ab	0.301±0.033ab	0.242±0.022ab	3.889±0.16bc
SEED	**45D**	0.25±0.038abc	2.984±0.134b	0.045±0.001ab	0.285±0.007ab	0.247±0.008ab	3.811±0.18b
p-Value							
Plant material		<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Days after flowering		0.010	0.300	0.005	<0.001	<0.001	0.005
Interaction		0.005	0.177	0.003	<0.001	<0.001	0.003