Lupin seed proteomics for quantifying diversity and analysing food products

THE UNIVERSITY OF WESTERN AUSTRALIA

Md Shahidul Islam
BSc in Agriculture (Bangladesh Agricultural University)
MSc in Horticulture (Bangladesh Agricultural University)

This thesis is presented for the degree of
Doctor of Philosophy

School of Plant Biology
Faculty of Natural and Agricultural Sciences
The University of Western Australia

2012

ABSTRACT

The potential for using lupin grain as a food in its own right, and as an ingredient in food preparation, is increasing due to its high protein content and low level of anti-nutritional fractions. Lupin-enriched food has proven health benefits, with seed storage proteins the main contributors to these benefits. Quantification of lupin genetic diversity can increase efficiency in improvement of protein attributes through further breeding. This thesis focused on lupin seed proteomic analysis from different species, cultivars and baked products using a range of proteome technologies including Matrix-Assisted Laser Desorption Ionisation Time-Of-Flight mass spectrometry (MALDI-TOF-MS), two-dimensional gel electrophoresis and mass spectrometric peptide sequencing.

MALDI-TOF-MS is a useful high-throughput approach in producing protein profiles across lupin species and cultivars. A large number of species in the genus *Lupinus* with diversified morphological and biochemical characteristics provide genetic resources for protein improvement of commercial cultivars through breeding. This study recognised 630 polymorphic protein mass peaks across 35 genotypes (33 lupin genotypes from 19 species and two outgroup genotypes). Overall, comparison of seed protein profiles demonstrated notable diversity across the species analysed. MALDI-TOF protein profiles identified 19 species-specific protein mass peaks and 111 rare protein mass peaks common to only 2–3 species which are useful in marker-assisted breeding. The phylogenetic analysis clustered the studied species in different groups in broad agreement with their evolution and biochemical properties; two New World lupin species, *L. mutabilis* and *L. succulentus*, grouped into the cluster of smooth-seeded Old World species. The value of this diversified source for lupin breeding towards improvement of grain protein attributes is discussed.

Australia, and in particular Western Australia, is the largest lupin producer in the world. Twenty-five narrow-leafed lupin cultivars have been developed through breeding, all of which are suitable for human consumption. Quantification of grain protein diversity across these cultivars will assist breeding programs in the development of new cultivars with enhanced protein attributes. MALDI-TOF mass spectrometric fingerprinting identified 355 polymorphic peaks across the cultivars. Fifty cultivar-specific protein mass peaks and 61 rare mass peaks common to 2–3 cultivars were identified through seed protein profile analysis. The phylogenetic analysis constructed relationships among

the cultivars in accordance with their pedigree relationships. The cultivar-specific and rare protein mass peaks are useful for cultivar identification and in marker-assisted breeding. Considerable seed protein diversity across the cultivars suggests scope for proteomic improvement of the grain through further breeding.

The influence of environmental factors on seed protein expression is a concern in breeding programs targeting improvement of protein attributes. Thus this thesis investigated the environmental effect on seed protein expression of five narrow-leafed lupin cultivars across three different environmental conditions in Western Australia. Of the 133 reproducible protein mass peaks, 31 were expressed consistently across the entire set of cultivar × environment. Expression of another 20 mass peaks was influenced by cultivar, while only 6 mass peaks showed differential expression among environments irrespective of cultivar. Seventy-six mass peaks were highly variable in expression. Analysis of variance on the number of mass peaks indicated that lupin seed protein expression is significantly ($p=0.008$) influenced by cultivar, but not by environment ($p=0.131$). Multivariate analyses of mass spectrometric protein profiles supported the above analyses by demonstrating significant ($p=0.001$) influence of cultivars and no significant influence of environments ($p=0.053$) on protein expression. The overall result suggests that lupin seed protein expression is largely genetically controlled; breeding to target specific protein has the potential to improve proteome attributes.

Considerable diversity among narrow-leafed lupin cultivars lead to the detailed protomic analyses of four lupin cultivars to identify different proteins that will add value to breeding by targeting specific protein groups, in particular proteins related to health issues. A high resolution study using two-dimensional gel electrophoresis followed by MS/MS peptide sequencing revealed differential expression of allergenic and seed storage proteins. Twenty-four proteins were recognised as having differential expression within the studied cultivars; among these, 19 were identified as β-conglutins and 8 as allergenic proteins. Cultivars Tanjil and Uniharvest had more highly-expressed allergenic proteins compared to Yorrel and Coromup; this represents a new insight into the protein attributes of lupin cultivars where allergenicity is a concern. Most of the α, δ and γ conglutins showed reproducible expression among the cultivars. The proteins with differential expression, in particular cultivar-specific ones, have potential as markers for cultivar identification and for screening breeding lines for low allergenicity.

The interactions of lupin proteins with wheat proteins in lupin-wheat bread was characterised to determine the extent to which lupin proteins are extractable after the baking process. This is crucial to define the possible bioactivity of lupin proteins. Many lupin and wheat proteins remained unaffected after baking as revealed by two-dimensional gel electrophoresis while some were not extractable. Most α-conglutins were extractable from lupin–wheat bread even using milder salt (0.5 M NaCl) extraction while most β-conglutins were incorporated into the bread matrix and lost their extractability. The result indicates a significant interaction between lupin and wheat proteins with a stronger binding of β-conglutins than α-conglutins to wheat proteins that might be useful in assessing health-related issues in lupin products and to capture public interest.

This thesis has advanced our understanding of the lupin seed proteome, particularly with respect to protein diversity, expressional behaviour of proteins and their response to baking. It revealed considerable proteomic diversity across lupin species' and cultivars by characterising genotype-specific protein mass peaks and by deducing phylogenetic relationships. This thesis also demonstrated the greater genetic influence of diversity, in comparison to environments, on the expression of seed proteins. This information and in particular the methodology will be useful in lupin breeding programs. The high resolution study using two-dimensional gel electrophoresis, in combination with MS/MS peptide sequencing, demonstrated differential expression of allergenic proteins across cultivars and differential recovery of lupin conglutins from the baked product and their overall accessibility to digestion. These outcomes have potential in the food industry and will help in the understanding of the health benefits of lupin-enriched food.

THESIS DECLARATION

I declare that no experimental materials of this thesis have been previously submitted for any degree at any other university or institution. The experiments and the writing of the thesis were designed and carried out by myself in consultation with my supervisors Guijun Yan, Wujun Ma and Rudi Appels. The thesis contains published work which has been co-authored. The co-authors provided technical advice on experiments when required and editorial input into drafts.

All contributions made by other individuals have been duly acknowledged.

Md. Shahidul Islam
November, 2012

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my supervisors Associate Professor Guijun Yan, Professor Rudi Appels and Associate Professor Wujun Ma who are the mentors of my research career. I deeply appreciate your priceless suggestions, encouragement and motivation, cordial approach and enormous support that have made this journey successful. Your passion, work ethic and research commitment inspired me to endeavour for excellence in my job.

I am grateful for the financial support from the University of Western Australia, School of Plant Biology, Faculty of Natural and Agricultural Science and Centre for Food and Genomic Medicine (Murdoch Unviersity) for postgraduate scholarship support. I am also thankful to the Department of Agriculture and Food Western Australia (DAFWA) and Centre for Comparative Genomic (CCG, Murdoch University) for financially supporting my research project and the State Agricultural Biotechnology Centre for providing me with contemporary molecular level research facilities. My special thanks to the Australian Society of Plant Scientists Travel Award, CCG and School of Plant Biology for supporting me to attend the "International Botanical Congress 2011".

My earnest thanks go to Junhong Ma (wheat proteomic lab, SABC) and Liyan Gao (visiting scientist from Capital Normal University, China) for convivial support when learning some laboratory research techniques. I am very appreciative of how they helped overcome technical difficulties in laboratory experiments. I am very grateful to several visiting scientists who encouraged me including Ke Wang (China), Marie Oszvald (Hungary) and Shunli Wang (China). I am also thankful to scientists from DAFWA laboratory, SABC in particular Sharon Wescott and Dora Li for cordial assistance in learning sequencing technology.

Special thanks to Bevan J. Buirchell, lupin breeder at the Department of Agriculture and Food Western Australia for valuable suggestions in experimental design and for providing some of the seed samples. I would like to thank Sofia Sipsas (DAFWA), Dr. Jon Clements (UWA Institute of Agriculture) and Nindethana Seed Service Pty Ltd, Western Australia for providing seed material. Thanks to Wayne Hawkins for assisting me in seed grinding. Special Thanks to Dr. Jonathan Hodgson for supplying bread samples and constructive advice.

Thanks to "Proteomic International" for helping in MS/MS protein analysis. Special thanks to Bodhi's Bakery, Fremantle, Western Australia for preparing bread samples for my experiment.

I would also like to thank Dr Danica Goggin (UWA), Dr Penelope Smith (University of Sydney) and Professor Vijay Jayasena (Curtin University) for valuable advice in interpreting my results. Special thanks to Rhonda Foley and Karam Singh, from CSIRO, for supplying EST sequences for lupin protein identification.

I would like to express my thanks to Winthrop Professor Hans Lambers, Head of School of Plant Biology, and Winthrop Professor Tim Colmer for their administrative support during my candidature. Heartfelt thanks to administrative staff of the School of Plant Biology especially Renu, Alan, Pandy, Jeremy, Barbara, Kirsty, Bonnie and Natalie for their kind assistance and support. I would also like to thank the administrative staff of the SABC at Murdoch University especially Dr Dave Berryman, Bee Lay Addis,

Frances Brigg and Professor Mike Jones for administrative and technical assistance during my research work at the laboratory. Thanks to CCG staff for supporting my research project.

I would like to thank the wonderful people who shared not only office space but also my stresses and joys during the journey of my study: Chai, Marcal, Parwinder, Weihua, Sayeed, Honghua, Hamid, and many more. Special thanks to Touhid, Arpiwi and Lisinda for helping me in data analysis. I had also some fantastic colleagues and friends around me at SABC: Hollie Webster, Jingjuan Zhang, Huyen Phan, Motiur, Suren, Luo Hao, Sharif and many more.

I will always and forever be grateful to my parents who are the source of moral and emotional support in my higher studies and for showing me the right values and path in life. I am also grateful to my parents-in-law who are always supportive in my endeavours especially while staying three months with us here in Perth.

My daughter Usrat is the greatest inspiration to all of my endeavours.

Finally, and most importantly, my heartfelt thanks and indebtedness to my lovely wife Nusrat for her enormous support, understanding and patience during my candidature especially in the last few months of thesis writing. I dedicate this thesis to you as I know without your endless love and encouragement I would not have survived my PhD.

I believe God allows us to succeed.

PUBLICATIONS ARISING FROM THIS THESIS

Journal Articles

Islam S, Ma W, Appels R, Yan G. 2012. Mass spectrometric fingerprints of seed protein for defining *Lupinus spp.* relationships. *Genetic Resources and Crop Evolution* (published online: dx.doi.org/10.1007/s10722-012-9890-y) [**Chapter 3**]

Islam S, Ma W, Appels R, Buirchell BJ, Ma J, Yan G. 2011. Diversity of seed storage protein among the Australian narrow-leafed lupin (*Lupinus angustifolius* L.) cultivars. *Crop & Pasture Science* 62 (9): 765–775 [**Chapter 4**]

Islam S, Ma W, Appels R, Buirchell BJ, Yan G. Environment and genetic interaction of seed storage proteins in narrow-leafed lupin. *Crop & Pasture Science* (submitted) [**Chapter 5**]

Islam S, Yan G, Appels R, Ma W. 2012. Comparative proteome analysis of seed storage and allergenic proteins among four narrow-leafed lupin cultivars. *Food Chemistry* 135 (3): 1230-1238 [**Chapter 6**]

Islam S, Ma W, Gao L, Yan G, Appels R. 2011. Differential recovery of lupin proteins from the gluten matrix in lupin–wheat bread as revealed by mass spectrometry and two-dimensional electrophoresis. *Journal of Agricultural and Food Chemistry* 59 (12): 6696–6704 [**Chapter 7**]

Book Chapter

Islam S, Ma W, Yan G, Bekes F, Appels R. 2012. Novel approaches to modifying wheat flour processing characteristics and health attributes: from genetics to food technology. In: Cauvain, S P (ed) Bread Making. Woodhead Publishing Limited, 80 High Street, Sawston, Cambridge, CB22 3HJ UK. pp 259-296. [**Chapter 2**] (http://www.woodheadpublishing.com/en/book.aspx?bookID=2345)

Conference Proceedings

Islam S, Yan G, Appels R, Ma W. Wheat and lupin protein interaction at baking: modifying extractability from lupin-wheat bread *11th International Gluten Workshop*, Beijing, PR China, August 2012.

Islam S, Ma W, Appels R, Buirchell BJ, Ma J, Yan G. Diversity of seed storage protein among narrow-leafed lupin cultivars (*Lupinus angustifolius* L.) with reference to contributing to health. *XVIII International Botanical Congress*, Melbourne, Australia, August 2011.

Islam S, Ma W, Gao L, Yan G, Appels R. Lupin–wheat bread protein: modification of the bread matrix for improved health attributes. *XVIII International Botanical Congress*, Melbourne, Australia, August 2011.

Islam S, Ma W, Yan G, Appels R. The lupin proteins that integrate into the matrix of lupin–wheat breads. *60th Australian Cereal Chemistry Conference*, Melbourne, Sept 2010.

TABLE OF CONTENTS

Chapter 1

General introduction

1.1 Aim of thesis

The nutritional value of lupin grain is mainly attributed to its high protein content. Lupin-enriched food has the potential to provide health benefits and seed storage proteins are considered the contributor. Lupin species as a group have considerable genetic diversity that provides the scope for improvement of protein attributes through molecular-assisted breeding. The thesis is designed to quantify lupin seed protein diversity and characterise the proteins having poteinal health attributes to assist further breeding. The specific aims of this thesis are to:

1. Define the proteomic diversity among species and cultivars of lupin.
2. Investigate lupin products as healthy food at the proteome level.
3. Establish a proteome-based diagnostic tool for use in lupin breeding.

1.2 Background

Lupin has a long historical background as a cultivated crop plant for more than 3000 years (Boersma 2007; Gladstones 1970). Traditionally it is used mainly as a rotation crop for green manure or fodder, but there is evidence of human consumption since ancient Mediterranean times (Gladstones 1998). Lupin originated in the Mediterranean region and South America (Ainouche et al. 2004) and was gradually introduced to Europe (Hanelt 1960). In Australian cropping systems, lupin was first introduced in the mid-19th century (Lupin.org 2011). The modern history of lupin development in Australia is attributed to Dr John Gladstones who developed consumable modern cultivars (Cowling et al. 1998). In 2009, lupin contributed 2.7% of world pulse production (FAO 2011). Over the last three decades, Australia has been the largest lupin producer capturing approximately 70% of world production.

Due to its nutritional attributes and adaptability to diverse agronomic conditions (Dervas et al. 1999; Howieson et al. 2000), lupin cultivation increased over the last two decades (FAO 2011). The nutritional attributes of lupin grain is dominated by low fat and starch content, low anti-nutritional components and high contents of dietary fibre and protein (Sirtori et al. 2010). Functional and healthy foods are a major concern in modern society and lupin-enriched food is claimed to provide health benefits such as increasing satiety and reducing energy intake (Lee et al. 2006), decreasing blood pressure (Lee et al. 2009), glucose level (Hall *et al.* 2005) and cholesterol (Martins et al. 2005). However, allergenicity of lupin grain is an issue in the food industry although it

is reported to be similar to other legume crops (Duranti et al. 2008; Guarneri et al. 2005). Since seed protein is considered the main contributor to these claimed health attributes (Lee et al. 2009; Lee et al. 2006; Weiße et al. 2010), investigation of lupin grain using modern protein analysing technologies would assist the breeding process in improving protein attributes.

Lupins include 200–500 annual and perennial herbaceous species (Dunn and Gillett 1966) grouped as Old World species and New World species (Gladstones 1974). The New world species comprise over 90% of the total species, distributed from Alaska to South Argentina and Chile (Dunn 1984; Planchuelo-Ravello 1984). Old World species includes 13 lupin species with Mediterranean and African origins (Gladstones 1974; Amaral Franco and Pinto da Silva 1978). The most cultivated lupin species in Australia is blue lupin (*Lupinus angustifolius*) also known as narrow-leafed lupin, while white (*L. albus*) and yellow (*L. luteus*) lupins are traditionally cultivated in Europe and South America (Duranti et al. 2008). The world's largest lupin breeding program is based in Western Australia and has released 25 commercial cultivars since 1968. These cultivars are virtually free of toxic alkaloids and suitable for human consumption (French et al. 2008).

Lupin has a diverse geographical distribution, morphological variation and chromosomal polymorphisms resembling wide genetic diversity. DNA-based analyses revealed substantial variation in genetic content among lupin species (Ainouche and Bayer 1999; Christopher 2008; Hughes and Eastwood 2006; Kass and Wink 1997) and cultivars (Yuan et al. 2005), leading to suspected differences in seed protein composition. Identifying sources of protein diversity will provide scope for improving lupin grain quality, particularly in protein attributes through molecular-assisted breeding (Talhinhas et al. 2006). Considering the above facts, the overall approach for this thesis is to characterise the diversity and expression of lupin seed proteins and their response to the baking process.

1.3 Outline of thesis

The results from this study are presented and discussed in eight separate chapters, as follows:

Chapter 1: 'Introduction'

Chapter 2: The 'Literature Review' gives an overview of the knowledge and research achievements in the field of lupin proteomics. This chapter also identifies research gaps that this thesis helps to close.

[Part of this chapter is published as book chapter by Woodhead Publishing Limited, Sawston, Cambridge, CB22 3HJ UK. Pp 259-296; http://www.woodheadpublishing.com/en/book.aspx?bookID=2345]

Chapter 3: 'Mass spectrometric fingerprints of seed protein for defining *Lupinus* spp. relationships' analysed the seed protein profile of important lupin species by MALDI-TOF mass spectrometry to study the diversity of the species and to deduce relationships among the species. This study intended to establish seed protein profiling as a tool for breeders to develop cultivars enriched for specific proteins.

[This chapter has been published in *Genetic Resources and Crop Evolution* (published online); dx.doi.org/10.1007/s10722-012-9890-y]

Chapter 4: 'Diversity of seed protein among the Australian narrow-leafed lupin (*Lupinus angustifolius* L.) cultivars' investigated the feasibility of MALDI-TOF mass spectrometry for profiling lupin seed proteins in order to study the diversity at cultivar level and to validate relationships among Australian cultivars.

[This chapter has been published in *Crop & Pasture Science;* 62 (9): 765–775]

Chapter 5: 'Environment and genetic interaction of seed storage proteins in narrow-leafed lupin (*Lupinus angustifolius*)' is concerned with the environmental effect on the diversity of lupin seed storage proteins among selected cultivars.

[The manuscript from this chapter has been submitted in *Crop & Pasture Science*]

Chapter 6: 'Comparative proteome analysis of seed storage and allergenic proteins among four narrow-leafed lupin cultivars' identified differential protein expression at a higher resolution from selected narrow-leafed lupin cultivars. This study provides a new insight into protein attributes of lupin cultivars where allergenicity is a concern.

[This chapter has been published in *Food chemistry;* 135 (3): 1230-1238]

Chapter 7: 'Differential recovery of lupin proteins from the gluten matrix in lupin–wheat bread as revealed by mass spectrometry and two-dimensional electrophoresis' investigated lupin protein interactions in food preparation and their extractability from baked products.

[This chapter has been published in the *Journal of Agricultural and Food Chemistry*; *59* (12), pp 6696–6704]

Chapter 8: 'General discussion' summarises important findings from the research in this thesis and their implications for future research.

This thesis compiled individual published/submitted or potential journal articles. I have endeavoured to minimise any unnecessary repetition of text but some repetition between chapters was unavoidable.

Chapter 2

Literature Review

Part of this chapter is published as a book chapter.

Citation: Islam S, Ma W, Yan G, Bekes F, Appels R. 2012. Novel approaches to modifying wheat flour processing characteristics and health attributes: from genetics to food technology. In: Cauvain, S P (ed) Bread Making. Woodhead Publishing Limited, 80 High Street, Sawston, Cambridge, CB22 3HJ UK. pp 259-296. [http://www.woodheadpublishing.com/en/book.aspx?bookID=2345]

Lupin (*Lupinus* spp.) is a legume crop receiving attention by crop industries not only for its excellent nutritional attributes (Duranti 2006; Roccia et al. 2009) but also for its adaptability to extreme climatic conditions (Guemes-Vera et al. 2008). The grain is high in protein (comparable to soybean), low in fat and starch, and high in dietary fibre (a requirement for a low-glycaemic-indexed crop) (Hall et al. 2005; Mubarak 2001; Magni et al. 2004). A number of health benefits associated with the consumption of lupin-enriched foods have been reported, most of which are attributed to lupin seed proteins (Hall et al. 2005; Smith et al. 2008; Lee et al. 2009). This literature review outlines the flow of research findings on lupin seed proteins with emphasis on proteomic variation in terms of species, cultivars and environmental effects, health benefits and the interaction/modification at baking.

2.1 Background of lupin

Lupin has been used as a green manure since early Greek times (Gladstones 1970) and as a cultivated crop plant for more than 3000 years (Boersma 2007). There is evidence of lupin consumption in human food since early Egyptian times (Lupin.org 2011). The Mediterranean region and South America are considered the ancient homes of this crop (Ainouche et al. 2004); it was gradually introduced into Europe to improve soil health (Lupin.org 2011).

2.1.1 Evolution of *Lupinus*

In the family Fabaceae, lupin (*Lupinus* L) is a large genus consisting of herbaceous, soft-woody shrubs and small tree species (Dunn 1984; Turner 1995). Herbaceous lupins include 200–500 annual and perennial species (Dunn and Gillett 1966) that are mainly used as grain crops. The diversified species of lupin are grouped as Old World and New World species (Gladstones 1974). The New World species comprise >90% of the total species which are distributed from Alaska to South Argentina and Chile (Dunn 1984; Planchuelo-Ravelo 1984). The Old World species include 13 lupin species from Mediterranean and African regions (Gladstones 1974; Amaral Franco and Pinto da Silva 1978) which are annual, herbaceous and large seeded. Three Old World species—*L. albus* L., *L. angustifolius* L. and *L. luteus* L.—and one New World species—*L. mutabilis*—are cultivated as crops for human consumption (Petterson 1998). Some other species are used as green manure, forage, ornamentals and for land stabilisation (Gladstones 1998; Hoveland and Townsend 1985).

2.1.2 History of lupins in Australia

Lupins were first introduced into Australian cropping systems by Ferdinand Von Mueller and Richard Schomburgk in the 19th century (Lupin.org 2011). Although the potential of this crop for human consumption was well-known, it was primarily cultivated for fodder and green manure (Hanelt 1960). The modern history of lupin development in Australia is defined by the career of Western Australian plant breeder Dr John Gladstones who developed modern cultivars of narrow-leafed lupin (*L. angustifolius)* (Cowling et al. 1998). *L. angustifolius* is known as 'blue lupin' in Europe due to the blue flowers of the wild species. In Australia, cultivars of *L. angustifolius* were bred with white flowers to differentiate them from the bitter-tasting wild relatives. The lupin industry of Australia prefers the name 'Australian Sweet Lupin' for those modern cultivars of *L. angustifolius* with low levels of bitter alkaloids (French et al. 2008).

2.1.3 World Production

Lupin occupied 2.7% of world pulse production (tonnes) in 2009 (FAO 2011). In the last three decades, Australia has been the largest lupin producer capturing approximately 70% of world production (**Table 1**), which increased from about 0.26 million tonnes in 1979 to more than 2 million tonnes in 1999 (**Table 1**) mainly from increased production in Australia (FAO 2011). Blue lupin (*L. angustifolius*) is mostly cultivated in Australia, while white (*L. albus*) and yellow (*L. luteus*) lupins are traditionally cultivated in Europe and South America (Duranti et al. 2008). Cultivation of *L. angustifolius* is increasing in Europe after an outbreak of anthracnose in the 1990s (Lupin.org 2011). In African and Mediterranean countries such as Morocco, South Africa, Syria and Egypt, mainly *L. albus* and *L. luteus* are grown. Bitter *L. mutabilis* is still cultivated in some parts of Ecuador, Peru and Bolivia. Very small areas of USA and Canada are cultivated with *L. albus* (Lupin.org 2011).

Table 1: Production of lupin in various years from 1969 to 2009 by continent.

	Production (tonnes)								
	1969	**1979**	**1989**	**1999**	**2005**	**2006**	**2007**	**2008**	**2009**
Australia	1504	78 231	772 005	1 968 000	1 285 000	470 000	661 861	707 989	614 000
Africa	32 000	10 948	16 311	31 673	22 498	24008	22 162	26 141	29 797
Americas	3497	10 483	19 395	31 590	74 279	80 333	61 519	42 402	23 581
Asia	40	1100	6	1189	704	852	745	708	785
Europe	640 059	165 680	279 723	71 890	288 823	225 841	225 814	231 249	257 249
Total/ World	677 100	266 442	1 087 440	2 104 342	1 671 304	801 034	972 101	1 008 489	925 412

Source: (FAO 2011)

2.1.4 Australian production

In Australia, lupins are grown in New South Wales (NSW), Victoria, South Australia and Western Australia (ABARE 2011). The winter/spring rainfed part of south-western Western Australia produces roughly 70% of the country's total followed by South Australia, southern NSW and Victoria (**Table 2**). The lupin industries in Western Australia and South Australia are dominated by *L. angustifolius* (French et al. 2008). In contrast, a considerable proportion of *L. albus* is produced in NSW and Victoria. Moreover, small quantities of the large-seeded lupin (*L. albus)* 'lupini bean' are grown in South Australia and Tasmania (Lupin.org 2011).

Table 2: Production ('000 tonne) of lupin in Australia by state.

Year	NSW	Victoria	South Australia	Western Australia
2005–06	62	37	125	1104
2006–07	16	11	34	409
2007–08	62	25	65	210
2008–09	37	18	52	600
2009–10	77	31	75	639

Source: ABARE (Australian Bureau of Agricultural and Resource Economics) Reports (www.abareconomics.com)

2.1.5 Lupin in Western Australia

Lupin is well established in Western Australian grain-based farming systems, with *L. angustifolius* the major species grown (French et al. 2008). Western Australia is the largest lupin producer and exporter in the world, largely to the EU, Japan and Korea (Lupin.org 2011). The world's largest lupin breeding program has been based in Western Australia since the 1960s and includes a goal to improve cultivars suitable for human consumption (French et al. 2008). The Department of Agriculture and Food Western Australia (DAFWA) has released 25 commercial cultivars since 1968. These cultivars are virtually free of toxic alkaloids (French et al. 2008). There is scope to use this crop for human consumption on a large scale, which would further increase the value of the crop.

2.1.6 Consumption

Most of the global lupin production is used by stockfeed manufacturers for animal feed (Petterson 1998). Consumption by ruminants (cows and sheep) is the most significant, followed by pigs and poultry. Lupin has been increasingly used as a feed source in the aquaculture industry in recent years (Sipsas and Seymour 2008).

Although lupin is largely used as animal and fish feed, it has been used as a human food since ancient Mediterranean times (Gladstones 1998). Different communities in the world have traditional methods of producing lupin-enriched food. Most commonly, lupin seeds are first soaked in water to remove most of the alkaloids which contribute their bitter taste, followed by cooking or toasting (Hill 1998). With the improvement of nutritional science and equipment in food industries, lupin is now being considered for many food preparations such as bread, cookies, cake, pasta, noddles, muffins and drinks (Hung et al. 1990). Due to increased use of plant-derived ingredients in food preparations, lupin has attracted attention worldwide as a potential protein-rich food ingredient suitable for human consumption (Fudiyansyah et al. 1995; Johnson and Gray 1993; Petterson and Crosbie 1990). Although ~4% of global production is currently consumed as human food, it has been reported that ~500 000 tonnes of lupin-enriched food product is consumed annually in the EU (Agro products 2011), mainly through low inclusion rates of lupin flour in wheat-based baked goods.

2.2 Chemical composition of lupin seed

2.2.1 Nutritional value of lupin

Plant-derived protein for food is mainly sourced from oilseed and legume crops due to their high protein, mineral and vitamin contents (Roccia et al. 2009). Among the legumes, soybean is one of the most important for its nutritional properties and desirable functionality (Wilson et al. 2007). However, lupin is an excellent alternative to soybean or any other legume due to its high protein content, low starch and lipid contents, and high dietary fibre (non-starch polysaccharide) (Hall et al. 2005; Mubarak 2001; Magni et al. 2004). Analyses of nutritional values of sweet lupin (**Figure 1**) have shown that the bio-availability of the constituents is comparable to those of processed soybeans (Mariotti et al. 2002).

2.2.2 Amino acid composition of proteins

Lupin seeds contain approximately 40% protein (by weight) and is the largest single component (Hall et al. 2005). Seed proteins occupy a key position in the interests of nutritionists, not only for their nutritional role but also for a number of other beneficial effects for the human body. Moreover, seed proteins play important technological and functional roles, as witnessed by their increased use as food ingredients (Duranti et al. 2008).

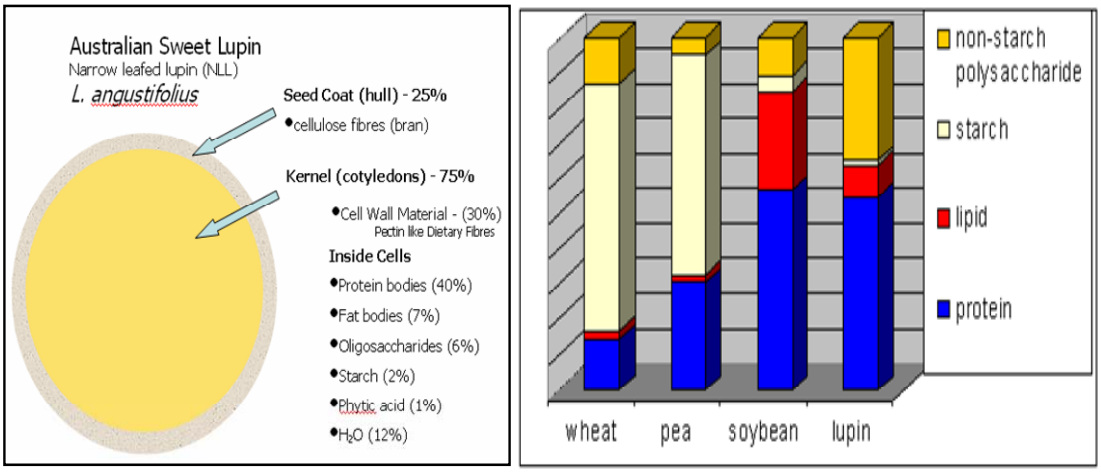


Figure 1: Comparison of major nutritional components of lupin (left) and a comparison among wheat, pea, soybean and lupin (right).

Source: Pulse Western Australia (http://www.lupins.org/feed/)

Lupin represents an important source of proteins for human beings (**Table 3**), with a biological value of 91% compared to egg proteins (Egana et al. 1992). Storage globulins, which make up the bulk of the protein of legume seeds, contain more than adequate levels of most of the essential amino acids and make this group a convenient protein source. For example, legume proteins contain high levels of lysine, an amino acid that is under-represented in cereals generally (Bressani 1973).

Table 3: Essential amino acid content of Australian sweet lupin (g/16 g N).

Amino acids	Mean	Minimum	Maximum	Number of samples tested
Arginine	11.62	10.6	12.9	116
Histidine	2.57	2.12	2.86	118
Isoleucine	3.91	3.57	4.41	118
Leucine	6.9	5.91	7.4	118
Lysine	4.75	4.22	5.19	118
Methionine	0.66	0.49	0.82	113
Phenylalanine	3.85	3.17	4.17	113
Threonine	3.54	3	3.82	118
Tryptophan	1	0.56	1.16	17
Valine	3.78	3.17	4.23	118
Cys + Met	2.01	1.5	2.45	113

Source: DAFWA (Department of Agriculture and Food Western Australia)
(http://www.agric.wa.gov.au/PC_92144.html#pro)

2.2.3 Fibre

A significant nutritional component of lupin seed is its high fibre content which has health benefits for human consumption (Lee et al. 2006). Lupin kernel fibre (LKF) is a novel insoluble non-digestible carbohydrate that also has value as a fat replacer (Archer et al. 2004). It is beneficial for bowel health and the digestive system. LKF is primarily insoluble but is similar to pectin, a soluble fibre, in its chemical structure (Johnson and Gray 1993). LKF also appears to have an unusually high water-binding capacity and viscosity compared with other insoluble fibres (Archer et al. 2004), indicating the potential to induce physiological effects more characteristic of soluble fibres, such as lowering glycaemic response (Jenkins et al. 1977; Johnson and Gray 1993; Lee et al. 2009).

Other than the direct consumption of LKF in lupin flour, kernel fibre can also be manufactured from de-hulled lupin seeds after extraction of protein, lipid and soluble carbohydrate fractions (Johnson and Gray 1993). This fibre, in the form of a pale-coloured powder with bland flavour, can be used to boost the dietary fibre content of baked goods whilst maintaining high levels of consumer acceptability (Clark and Johnson 2002).

2.2.4 Minerals

In addition to protein and fibre, lupin seed contains many different minerals (**Table 4**) which are significant for health.

Table 4: Mineral content of Australian sweet lupin (per kg as received).

Mineral	Mean	Minimum	Maximum	Number of samples tested
Calcium (g)	2.2	1.5	2.9	623
Magnesium (g)	1.6	1.1	2	524
Phosphorus (g)	3	2.1	4.3	846
Potassium (g)	8	6.6	9.1	614
Sodium (g)	0.4	0.3	1.1	422
Sulphur (g)	2.3	1.5	2.9	611
Copper (mg)	4.7	2.5	6.8	547
Iron (mg)	68.5	31	150	535
Manganese (mg)	19	6.7	76	584
Molybdenum (mg)	1.6	0.7	2.9	41
Zinc (mg)	34.1	24.7	45	551
Cobalt (g)	78	10	260	56
Selenium (g)	89	18	240	138

Source: DAFWA (Department of Agriculture and Food Western Australia) (http://www.agric.wa.gov.au/PC_92144.html#pro)

2.2.5 Status of anti-nutritional factors

Like some other grain legumes, lupin has anti-nutritional components (**Table 5**). Some of these, such as alkaloids, have an adverse effect on taste. In the developed/modern varieties of lupin these components have been reduced to tolerable levels (French et al. 2008). For example, Australian sweet lupin is considered suitable for human consumption since it has low levels of bitter tasting and potentially toxic alkaloids (Petterson 1998) compared with wild species. Most of the cultivars of narrow leafed lupin have alkaloid between 0.004% and 0.02% at dry weight basis when 0.012 is considered as the standard for sweet lupin (Sipsas and Seymour 2008). Furthermore, lectins and protease inhibitors, that can reduce protein digestibility, are found at lower levels in lupin than in many other legumes (Petterson et al. 1997). However, some of the protease inhibitors of lupin have now been positively re-evaluated considering the impending utilization of their biological behaviour in pharmacological and medical applications. For example, an inhibitor, belongs to the Bowman–Birk Inhibitor (BBI) family, has been isolated from *Lupinus albus* L. having ability to prevent or suppress carcinogen-induced transformation (Scarafoni et. al. 2008).

Table 5: Anti-nutritional factors in narrow-leafed lupin

Factor	Units	Mean	Minimum	Maximum	Number of samples tested
Alkaloids	%	0.02	<0.01	0.04	1316
Oligosaccharides*	%	4.07	2.9	5.2	202
Phytate	%	0.5	0.32	0.84	254
Saponins	mg/kg	573	442	740	11
Tannins(total)	%	0.29	0.19	0.51	208
Tannins(cond)	%	0.01	n.d.	0.11	236
TIA	mg/kg	0.12	0.05	0.24	148
CTIA	mg/kg	0.08	n.d.	0.59	8
Lectins	dilute		n.d.		28

*Sum of raffinose, stachyose and verbascose ; n.d.: not detected

Source: DAFWA (Department of Agriculture and Food Western Australia)
(http://www.agric.wa.gov.au/PC_92144.html#pro)

2.3 Overview of lupin seed protein

Generally, the proteins of legumes seeds are the storage type, mostly located in storage vacuoles of cotyledonary tissues, and their main role is to supply nitrogen and carbon to the emerging plantlet (Duranti et al. 2008). With the advancement of proteomic and food technologies, seed proteins are considered significant components for the human

diet for their nutritional value and health benefits including hypolipidemic, hypoglycaemic, hypotensive, anti-carcinogenic and anti-obesity effects (Hall et al. 2005; Pilvi et al. 2006; Lee et al. 2009; Lee et al. 2006; Martins et al. 2005; Sirtori et al. 2004; Yoshie-Stark et al. 2004). At the same time, they have been investigated for adverse effects on human health such as food intolerance, allergies and inhibition of endogenous hydrolytic enzymes (Guillamon et al. 2010; Kenny et al. 2000; Klos et al. 2010; Rojas-Hijazo et al. 2006; Goggin et al. 2008; Magni et al. 2004).

2.3.1 Types of lupin seed proteins

According to Osborne's classification, lupin seeds have two types of proteins, namely albumin and globulins. Their quantitative ratio is approximately 1:9 (Duranti et al. 1981; Blagrove and Gillespie 1975). Other minor fractions of proteins include prolamins (Cerletti et al. 1978). Changes in this protein profile at a sub-unit level have been reported among different lupin varieties (Vaz et al. 2004).

Albumins are the functional proteins of lupin seed which are largely metabolic enzymes, linked to the storage role of the seed cotyledon (Duranti et al. 2008). Globulins comprise a number of protein molecules which are the salt-soluble storage proteins of seeds. The first separation of lupin seed protein was carried out by Blagrove and Gillespie (1975) who separated lupin protein into four main fractions in *L. albus* and *L. angustifolius* and named them as α, β, γ and δ-conglutins on the basis of their electrophoretic mobility. Further separations of lupin protein by isoelectric focusing and ion exchange chromatography showed the extreme heterogeneity of the different protein fractions (Casero et al. 1983; Duranti et al. 1981; Restani et al. 1981).

Lupin globulins contain two major types of proteins (Blagrove and Gillespie 1975). The first type contains proteins with sedimentation coefficients around 11S, named as α-conglutin and fit into the 'legumin-like' or 11S globulin family. This protein group consists of hexamers of disulphide-linked subunits which include acidic subunits of greater molecular weight, and basic subunits of lower molecular weight (Magni et al. 2007). The acidic subunits of α-conglutin, different to most other 11S globulins, are glycosylated (Duranti et al. 1995). The second major type of proteins with sedimentation coefficients of approximately 7S are known as β-conglutin and are referred to as 'vicilin-like' or 7S globulins. These are trimeric proteins with monomers consisting of polypeptides ranging in size from 16 to >70 kDa with no disulphide bridge

linking (Magni et al. 2007). The pattern of endogenous cleavage of β-conglutin in lupin is similar to that of vicilin proteins in pea, although in lupins this is greatly enhanced (Duranti et al. 1992).

In addition to these main protein types, lupin seed proteins contain considerable amounts (5–6%) of γ-conglutin, a basic 7S protein (Duranti et al. 1981) consisting of two heterogenous disulphide-linked subunits of 17 and 30 kDa (Scarafoni et al. 2001). A monomeric low molecular weight protein similar to other 2S albumins, named δ-conglutin (Duranti et al. 1981; Salmanowicz and Weder 1997), which consists of two disulphide-linked chains of about 4 and 9 kDa, has also been described. The subunits of all four abovementioned classes of proteins arise from the proteolytic cleavage of precursor molecules (Derbyshire et al. 1976; Muntz et al. 2002).

2.3.2 Seed protein variation among species

Lupins represent a group of species characterised by a diverse geographical distribution, and morphological and chromosomal polymorphisms (Ainouche and Bayer 1999; Christopher 2008; Hughes and Eastwood 2006) which potentially have large variation in seed protein levels. During the last 50 years, systematic and taxonomic relationship studies among lupin species were dominated by morphological features such as life history, leaf morphology, cotyledon form and seed coat surface pattern (Eastwood et al. 2008), with particular emphasis on Old World species (Ainouche and Bayer 1999) mainly because of the limited number of species and economic interest as fodder, food and N-fixing crops (Gladstones 1974; 1970). Only 13 lupin species are reported as Old World, all being annual, herbaceous with large fruits and seeds, and digitated leaves (Ainouche and Bayer 1999; Gladstones 1974; Heyn and Herrnstadt 1977). Based on seed coat texture, Old World species had been categorised as either smooth- or rough-seeded species (Gladstones 1974; Heyn and Herrnstadt 1977) which generally agrees with chemotaxonomic and morphological relationships among the species (Cristofolini 1989; Gladstones 1974, 1984; Nowacki 1963; Williams et al. 1983; Wink et al. 1995; Wolko and Weeden 1990a, 1990b). The five smooth-seeded species are further categorised in four sections: *Albi, Micranthi, Angustifoli* and *Lutei* (Gladstones 1984). Six or seven species have been identified as rough-seeded with higher interspecies resemblance and their typical scabrous-tuberculate testa, which is unique to the genus (Ainouche and Bayer 1999). The rough-seeded species are mainly distributed in North Africa and in the Eastern part of the Mediterranean region (Ainouche and Bayer 1999).

In contrast, New World lupins comprise hundreds of annual and perennial species which make phylogenetic relationships very complex within this genus (Ainouche and Bayer 1999). Taxonomic confusion exists in the literature, where numerous taxa or groups are distinguished based on only a few minor and inconsistent morphological characters (Ainouche and Bayer 1999). New World lupin species are mainly distributed in South and North America. For North American species, 18 groups (Smith 1944) have been categorised comprising ~200 species, although >1700 names have been proposed for New World *Lupinus* (Dunn 1984). New World lupins are reported as largely polymorphic species by floristic investigations (Welsh et al. 1987; Riggins and Sholars 1993). The complexity of this genus may come from its high morphological, breeding and eco-geographical diversity and the lack of clear diagnostic characters to distinguish species (Dunn 1984; Dunn and Gillett 1966; Planchuelo-Ravelo 1984). Although a number of groups and complexes are already recognised, some groups are roughly circumscribed which may be valuable for further modern analyses (Dunn 1984; Planchuelo-Ravelo 1984).

As genomic variation is the primary basis of proteomic diversity, understanding genetic variation within lupin species is critical for proteomic differentiation. Differences in basic genetic components among lupin species were demonstrated by several cytological studies. The 2C nuclear DNA content of 18 species and botanical forms of the genus *Lupinus*, using propidium iodide as a fluorescent dye, showed distinct infra-generic taxonomic groups with differing somatic chromosome numbers (Naganowska et al. 2003). The New World lupins' predominant chromosome number is 2n=48 with a basic chromosome number of x=6 (Dunn 1984). In contrast, Old World lupins have a series of basic chromosome numbers from x=5, 6 (or 9), 7, 8, 13 at different ploidy levels (Pazy et al. 1977). Intense selection/evolution in changeable environments is considered as the reason for recent changes in chromosome number of lupin species (Gladstones 1998). Among the smooth-seeded Old World lupins, *L. angustiflolius* has lower chromosome numbers (2n=40) compared to *L. albus* (2n=50), *L. micranthus*, *L. luteus* and *L. hispanicus* (2n=52) (Conterato and Schifino-Wittmann 2006).

Recent development and use of molecular data have significantly increased the understanding of plant systematics at various taxonomic levels (Soltis and Soltis 1995). DNA-based analyses has revealed substantial genetic variation among lupin species

(Ainouche and Bayer 1999; Kass and Wink 1997). Studies using chloroplast DNA data—obtained from both sequencing of the *rbcL* gene (Kass and Wink 1997) or restriction site mapping (Badr et al. 2000)—supported a common phylogenetic origin for Old and New World lupins. An amplified fragment length polymorphism (AFLP), inter-simple sequence repeat (ISSR) and random amplification of polymorphic DNA (RAPD) marker-based study reported extensive genetic diversity among different lupin species comprising both Old and New World species (Talhinhas et al. 2003). Though very low levels of genetic similarity were found (range 0.205–0.432), this study reported consistently more genetic similarity between *L. pilosus* and *L. cosentinii, L. mutabilis* and *L. polyphyllus*, and *L. luteus, L. hispanicus* and *L. angustifolius,* clearly separating the Old World from New World species (Talhinhas et al. 2003).

Cytological and DNA-based studies over the last four decades reported significant genetic variation among lupin species leading to expectations of considerable difference in seed protein levels. Seed globulin studies based on HPLC (Salmanowicz 1999) reported substantial variation among Old World lupin species and found that heterogeneous smooth-seeded lupins were distantly related to rather homogeneous rough-seeded lupins. One-dimensional gel electrophoresis of seed albumin (Salmanowicz and Przybylska 1994) and globulin (Przybylska and Zimniak-Przybylska 1995) also revealed significant differences among Old World lupin species but did not show any phylogenetic relationships among the species. A similar study also reported differences among the 19 species of New World lupins (Zimniak-Przybylska and Przybylska 1997). More heterogeneity was found in smooth-seeded lupins than rough-seeded species (Salmanowicz 1999). None of these studies included both Old and New World lupin species indicating scope for further investigation combining species from both groups. Moreover, since that time, proteomic technologies have improved extensively with high throughput results available within a short time (Alberghina et al. 2005; Chen et al. 2007; Cunsolo et al. 2004; Liu et al. 2009; Muccilli et al. 2005). Thus, characterising protein variation among different lupin species using modern proteomic tools would be significant for further breeding work to improve protein quality attributes of lupin seed.

2.3.3 Seed protein variation among cultivars

There has been considerable effort to increase the genetic diversity of lupins that current cultivars of narrow-leafed lupin are based (Gladstones 1984; 1994), by crossing

breeding lines, arising from mutation programs, with diverse wild germplasm (Cowling 1999). Considerable genetic diversity among different accessions/cultivars within the same species has been reported (Yang et al. 2004; Yuan et al. 2005) indicating possible proteomic differences among cultivars.

A comparative proteomic study of different lupin cultivars is essential to further assist molecular breeding to improve protein quality since protein is the main nutritional component of lupin grain (Duranti et al. 2008). In addition, proteomic analysis of cultivars is useful for studying storage protein deposition behaviours, effects of growing conditions, technological treatments and other factors of cultivar-specific proteins, variety identification, monitoring of desirable or undesirable specific protein components, and assessment of many other parameters affecting overall seed or protein quality (Magni et al. 2007). Thus, research in this area is important to improve the quality of lupin cultivars for human consumption. Until now, cultivar-specific research on lupin protein composition using modern technology has not been reported. This is particularly true for narrow-leafed lupin varieties released in recent years.

Narrow-leafed lupin was domesticated in recent times by developing a non-shattering, sweet, permeable seeded plant through mutation breeding (Cowling et al. 1998). The genetic diversity of *L. angustifolius* appears to be larger than *L. luteus* (yellow lupin) and *L. albus* due to domestication of this crop in places far from the naturally-distributed areas (Talhinhas et al. 2006). Characterisation of morphological and molecular diversity among germplasm of *L. angustifolious* revealed considerable genetic diversity among cultivars (Talhinhas et al. 2006). DAFWA has released 25 cultivars of *L. angustifolius* since the 1960s through diverse crossing. Large numbers of wild and external accessions were introduced into the crossing program that resulted in ample molecular diversification among cultivars (Cowling 1999). DNA-based fingerprinting using RAMP markers by Yuan et al. (2005) found substantial genetic variation among released cultivars. Two-dimensional gel electrophoresis revealed seed globulin differences among three varieties each of *L. albus* and *L. angustifolious* (Sirtori et al. 2008).

Reports on proteome analysis of narrow-leafed lupin were mainly on quantitative variation of total seed protein among genotypes. Variation in grain protein content among wild accessions has been reported from under 32% (dry matter basis) to almost

43% (Gladstones and Crosbie 1979). In contrast, grain protein content of cultivars ranged from 32% (Belara) to 35% (Coromup). Thus the lupin breeding program at DAFWA has set the target to increase protein content by initially crossing wild accessions with current varieties (Buirchell 2008). The qualitative variation of proteins among cultivars has not been studied; genetic variation within cultivars indicates that this should be possible.

2.3.4 Environmental effect on lupin seed protein

Environmental conditions, particularly soil nutrients and temperature, affect the amount, composition and/or polymerisation of seed proteins (Johansson et al. 2001; Lou et al. 2010; Panozzo and Eagles 2000; Wieser et al. 1998). Responses of seed storage proteins to soil nutrients are related to their amino acid composition. The mechanisms behind the balance regulation of carbon, nitrogen and sulphur in seed that influence protein composition are complex (Coruzzi and Zhou 2001; Galili et al. 2001; Shewry et al. 2001; Tabe et al. 2002). Thus adequate supply and uptake of plant nutrient components are crucial for seed protein content and quality.

Seed storage protein accumulation and their post-translation modification in particular would be expected to be influenced by temperature (Bourgeois et al. 2009). However, the effects of temperature on storage proteins are not straight forward and may vary between genotypes (DuPont and Altenbach 2003). In the case of wheat, it has been demonstrated that the influence of temperature on grain protein is dependent on soil nutrition availability. Under low levels of fertiliser application, protein content increased with increasing temperature, whereas with adequate fertiliser application, temperature variation did not have any significant influence on protein content (DuPont and Altenbach 2003). Protein content of grain is affected by temperature and may or may not lead to differences in grain protein composition and flour quality (Daniel and Triboi 2000; Randall and Moss 1990). For instance, high temperature increases the gliadin/gluten ratio and decreases the large polymer in wheat (Altenbach et al. 2002; Blumenthal et al. 1995; Corbellini et al. 1997; Panozzo and Eagles 2000). Water availability is crucial for seed storage protein quality since both vegetative and productive stages for grain development are affected by water availability (Boutraa and Sanders 2001; Carvalho et al. 2005).

Although seed storage protein attributes of grain are largely genetically inherent, storage protein gene expressions might be influenced by environmental factors (DuPont and Altenbach 2003). Wheat gluten gene expression was influenced by temperature variation during grain fill (Altenbach et al. 2002). Likewise, seed storage protein gene expression was influenced by soil nutrients in barley (Muller and Knudsen 1993) and in maize (Muller et al. 1997). In most cases, transcriptional regulation by environmental factors plays a major role in proteome variation of grain.

By comparing profiles of gene expression and protein accumulation of seeds grown under different environmental conditions, it is possible to reveal the molecular mechanisms underlying the environment influence on protein quantity and quality. In the case of lupin, protein content (Cowling and Tarr 2004) and other nutritional attributes differ according to genotypes and environmental conditions (Eggum et al. 1993; Gdala et al. 1996; Van Barneveld 1999). Effects of environment × genotype on seed protein quality have not been major considerations in lupin breeding.

2.3.5 Allergenicity of lupin proteins

Considering the increasing and potential use of lupin in food preparation, allegenicity of this grain is a significant issue that needs to be addressed. Since many legume crops, particularly peanut and soybean, have some allergenic effect, it is not unexpected that this might also be the case for lupin. Allergenic reactions to lupin have been reported either as primary lupin allergy (Hefle et al. 1994; Matheu et al. 1999; Peeters et al. 2007) or as a result of cross-reactivity to other legumes (Magni et al. 2005), especially peanut (Lindvik et al. 2008; Radcliffe et al. 2005; Shaw et al. 2008; Sirtori et al. 2011). Furthermore, lupin flour is a cause of occupational IgE-mediated allergy (Crespo et al. 2001; Moreno-Ancillo et al. 2005; Parisot et al. 2001; Prieto et al. 2010). There has been little systematic epidemiology study of lupin allergenicity due to the limited use of lupin seed and its derivatives (Duranti et al. 2008). However, a number of studies have reported on lupin globulins that comprise the bulk of lupin proteins as IgE-binding molecules (Dooper et al. 2007; Holden et al. 2008; Magni et al. 2005). The overall pattern of lupin allergenicity is similar to that of soybean, peanut and other edible legume seeds (Duranti et al. 2008; Guarneri et al. 2005). Like other legumes, the allergenic reaction to lupin is a concern (Hefle et al. 1994; Smith et al. 2008), particularly with regard to healthy lupin-enriched bread. A small proportion of the world population (Lindvik et al. 2008; Shaw et al. 2008) is allergenic to lupin proteins.

Although available data are still insufficient to identify all major lupin allergens, the basic subunits of α-conglutin (legumin-like, 11 S), β-conglutin (vicilin-like, 7 S), γ-conglutin (7 S) and δ-conglutin (2 S-albumin) seem to be responsible for the allergenic reactivity (Duranti et al. 2008; Sirtori et al. 2011). However, the candidate proteins responsible for lupin allergenic reaction have not been fully characterised. Characterisation of the lupin allergenic proteins may elucidate their role as cross-reactive and unique lupin allergens. β-conglutin was reported as the major allergen of *L. angustifolius* named as 'Lup an 1' by Goggin et al. (2008) and is in the International Union of Immunological Societies database (http://www.allergen.org). In another study, α-conglutin was reported as the predominant allergen in *L. albus* and *L. angustifolius* (Sirtori et al. 2011). On the other hand, γ-conglutin was suggested as the IgE binding component of lupin protein in *L. angustifolius* (Klos et al. 2010) and *L. albus* (Magni et al. 2005). Both α-conglutin and β-conglutin were identified as allergenic in *L. albus* by Guillamon et al. (2010) and Peeters et al. (2007). These differences in lupin allergenic reaction might be due to the genotypic variation of proteins.

2.4 Use of lupin proteins in food

As lupin has been used in different communities as a human food for a long time, there are unique processing and food preparation methods for different product purposes (Hill 1998). In modern food technologies, the most common use of lupin in human food is as flour in confectionary and baked products such as bread, crackers, biscuits, nodules and pasta (Fudiyansyah et al. 1995; Johnson and Gray 1993; Petterson and Crosbie 1990).

In addition to its direct use as flour, lupin seed derivatives including protein isolates and protein concentrates are used for various food preparations such as drinks, yogurt, sauces, infant formula etc (Hung et al. 1990; Lqari et al. 2002). To remove the bitter-tasting alkaloids, which limit the use of lupin in the human diet, protein isolates or concentrates may be an alternative choice for food preparation since alkaloids are water soluble and removed during the preparation of protein isolates (Lqari et al. 2002; Millan et al. 1995). The use of plant protein isolates in foods as functional ingredients, to improve the nutritional quality of products or for economic reasons, is increasing in modern food production (Sanchez-Vioque et al. 1999). Lupin protein isolate is useful as a food ingredient for changing the emulsifying capacity and improving the texture of food (Yoshie-Stark et al. 2004).

Due to incomplete recovery of proteins by isoelectric precipitation, a technology commonly used in industry, the protein quality of isoelectrically-precipitated lupin protein isolate is lower than that of lupin flour from which it was extracted (Sgarbieri and Galeazzi 1978). Protein isolates can also result in loss of solubility (KeShun 1997) which can impact negatively on other important physical and functional properties. It is worth noting that, for bread preparation, gluten plays a vital role in baking by forming long chains of gluten networks which help to retain gas (Butow et al. 2003; Roccia et al. 2009; Zhang et al. 2010). Unfortunately lupin does not contain gluten, which is normally present in cereals such as wheat, rye, barley and oats. For this reason, lupin is mixed with wheat in the baking process. Moreover, most legume seeds are deficient in sulphur-containing amino acids (Bressani 1973), which are more prominent in wheat. This suggests that lupin and wheat combinations would provide a more nutritionally-balanced product.

2.4.1 Effect of adding lupin flour to wheat on food properties

The nutritional attributes of food, as they relate to the impact on biochemical and physiological processes of the human body is now considered a key quality feature of grains (Duranti 2006). For example, wheat flour has moderate protein content (11.4 g/100 g) and is low in lysine, an essential amino acid (2.4 g/100 g). One way of modifying quality shortcomings of wheat flours is to add flours from other grains to increase protein and lysine contents (Duranti et al. 2008; Guemes-Vera et al. 2004). In addition, considering that a proportion of the human population is sensitive to gluten, preparation of bread with lower or even no gluten is a priority in the modern food industry. The incorporation of rice, corn, soybean and lupin flours have been of particular interest.

The health attributes of lupin and soybean are generally accepted as being desirable and development of commercially-viable food products that capture these health benefits is now a significant research activity (Guemes-Vera et al. 2004). The most common way of using lupin protein is by adding lupin flour to wheat flour in bread preparation. When lupin-enriched bread is compared with normal (only wheat) bread, changes in physical properties and nutritional properties have been reported which are briefly described below.

2.4.1.1 Physical properties

Since the physical characteristics of foods can govern consumer acceptance (Morr 1979), it is considered an important commercial issue in food industries. Physical properties of dough and bread were noticeably different when a considerable amount of wheat flour was substituted with lupin flour. Addition of lupin flour tends to improve baked products' textural properties, flavour and often colour and also increases its shelf life (Dervas et al. 1999). Bread firmness, product volume, and water absorption significantly increased with the addition of lupin flour (Guemes-Vera et al. 2008; Mubarak 2001), which can be considered a positive change for bread quality. In a sensory evaluation (appearance, flavour etc.) by trained and untrained panels, the value decreased after addition of lupin flour, but the evaluated products were considered acceptable (Clark and Johnson 2002). On a negative note, dough development time increased (Mubarak 2001) and dough strength decreased, as proteins from lupin obstruct the interconnection of the wheat gluten matrix (Guemes-Vera et al. 2008; Mubarak 2001).

The particular bread-making properties of wheat flour are recognised mainly by the ability of its gluten proteins to provide viscoelasticity by forming a gluten network when mixed with water (Butow et al. 2003; Roccia et al. 2009; Zhang et al. 2010). Polymeric glutenins and monomeric gliadins are the major storage proteins which together are known as gluten protein in wheat flour. In dough preparations and baking, these gluten proteins form large networks by linking with each other and thus help to retain gas (Cozzolino et al. 2001; Roccia et al. 2009; Rosell and Collar 2009). When wheat flour is supplemented by other flours rich in proteins, the added protein molecules interact with the gluten network in both direct and indirect ways: directly by cross-linking with gluten proteins and indirectly through the water availability of wheat gluten proteins. Reduction of gluten networks leads to weakening of dough elasticity (Bonet et al. 2006; Doxastakis et al. 2002; Mubarak 2001; Ribotta et al. 2005a; 2005b; Roccia et al. 2009; Ryan et al. 2002). A molecular level study of this protein interaction effect on the gluten network is important to improve the quality of final product.

2.4.1.2 Excellence of protein quality

Most significantly, lupin-enriched bread enhances protein quantity and quality of bread. Lupin seed contains approximately 40% protein which is 3–4 times that of wheat seed. Wheat flour has about 11.4 g protein/100 g and 2.4 g lysine/100 g protein; in contrast,

L. angustifolius contains 30–40 g protein/100 g and 5.47 g lysine /100 g protein (Brand et al. 2004; Guemes-Vera et al. 2004). Thus, addition of lupin flour increases the protein content in bread and provides a more balanced proportion of essential amino acids.

2.4.1.3 Enriched health attributes

The high fibre content of lupin seed makes it particularly significant for the human diet (Guemes-Vera et al. 2008). Lupin–wheat bread is rich in fibre so improves bowel health and keeps the digestion system smooth. It also helps to reduce fat and energy intake (Johnson et al. 2003).

The inhibitory effect of foods on appetite—the satiating power of foods—has recently been acknowledged by nutritionists to be vital for planning low-energy/weight-loss diets (Tremblay et al. 2002). High-protein diets are more satiating than high-carbohydrate diets (Anderson et al. 2004; Nickols-Richardson et al. 2005; Vandewater and Vickers 1996) and high-fibre diets are more satiating than low-fibre diets (Holt et al. 1999; Lee et al. 2006; Pereira and Ludwig 2001). As a result, foods enriched in protein, fibre or both have the potential to increase satiety and reduce energy intake. From this point of view, lupin-enriched bread is attracting attention due to the increased protein and fibre levels, and the associated positive contribution to reducing obesity (Lee et al. 2006).

Available evidence also suggests that substitution of refined carbohydrates in the diet with protein and fibre may benefit blood pressure (He et al. 2005; He and Whelton 1999; Hodgson et al. 2006; Lee et al. 2009). This means that lupin-enriched bread has the potential to control blood pressure and cardiovascular disease (Hall et al. 2005; Lee et al. 2009). It was reported that lupin-enriched bread also decreased blood glucose and cholesterol levels which benefits controlling diabetes (Hall et al. 2005; Magni et al. 2004; Sirtori et al. 2004). Since proteins play key roles in these health benefits, for example, glucose-lowering effect of lupin γ-conglutin (Bertoglio et. al 2011), it is essential to conduct proteomic research on lupin–wheat bread in order to understand the molecular basis of these health benefits.

2.4.2 Incorporation of other grain flours with wheat in baking

The incorporation of rice and corn in bread wheat flour up to a level of 10% (flour basis) and in durum wheat flour up to 20% has been achieved without any negative

effect in quality attributes such as colour, texture and flavour, offering promising nutritious and healthy alternatives for consumers (Sabanis and Tzia 2009). Increasing levels of substitution (30% and 50%) resulted in decreased dough strength, extensibility, and loaf volume, due to the replacement of gluten by the added protein. Durum flour can be substituted with non-gluten flours by up to 10% more than bread wheat flour because of its stronger gluten matrix and better dough rheological characteristics.

In the case of triticale substitution, the volume of the breads increased as the level of triticale flour increased due to the fortification of the gluten structure by the gluten added (Doxastakis et al. 2002). Nevertheless, substitution at 5 or 10% gives parameter values at least as good as the control sample and produces acceptable bread, in terms of weight, volume, and texture and crumb structure.

Wheat–oat mixes containing 5 and 10% oat products, recorded average baking values using standard assay procedures (Czubaszek and Karolini-Skaradzin´ska 2005). The same percentage of oat products brought about improved loaf volumes baked with commercial wheat flour (CWF), whereas baked products from strong flour such as lab-milled flour (LMF) enriched with oats, demonstrated less volume and poor quality crumb structure compared to control wheat breads. Wheat–oat dough and bread containing 5 and 10% oat products have been characterised as fairly good quality, with oat flakes and bran exerting a more beneficial effect on product quality, compared to oat flour.

Adding chickpea flour increased protein, fibre, ash and fat content in the blends without causing severe effects on quality even at 15% substitution. Blends increased in maltose content, W value and bread specific volume (Figuerola et al. 1987) and the final bread products were scored as good quality even without the use of maturing agents.

Among legumes, soybean protein has a proven desirable functionality, high nutritional value and relatively low price (Wilson et al. 2007). The main nutritional attribute of soybean is its high protein content, especially lysine (Olguin et al. 2003) as well as its high essential fatty acid and fibre contents, and low starch levels. In addition, its mineral content raises the food value of the grain. The types of consumable soybean protein products include flours and grits, protein concentrates and protein isolates (Ribotta et al. 2005a). Soybean-enriched bread can be satisfactorily made with 13–14% protein

content, instead of the normal 8–9% in white bread (Ribotta et al. 2005a). The high fibre content of soybean has made it of particular interest for human diets (Guemes-Vera et al. 2008). Adding soybean flour to wheat flour at relatively low levels (approximately 5% by weight) has negative effects on bread quality attributes such as extensibility properties and gas retention, thus decreasing consumer acceptability (Roccia et al. 2009). It has been reported that adding soybean flour to wheat flour in bread making decreases the solubility of soybean protein as measured by its decreased extraction from soy–wheat bread (Ribotta et al. 2005b; Ryan et al. 2002).

2.4.3 Wheat seed proteins and interaction with lupin proteins in mixture

Since wheat is one of the most important crops in the world (Yahata et al. 2005), considerable research has been conducted on wheat seed proteins. Consequently, a vast amount of information has accumulated on this highly complicated and very special group of proteins. Understanding all wheat seed protein components will assist in understanding the interaction between lupin and wheat proteins during baking.

Wheat gluten is a protein–lipid–carbohydrate complex formed as a result of specific covalent and non-covalent interactions from flour components during dough making as the components are hydrated and energy from mechanical input from mixing is provided. The usual chemical composition of untreated gluten formed during this process is approximately 75% protein, 6% fat, 15% carbohydrate and 0.85% ash (Sarkki 1980). The main composition of gluten includes high molecular weight glutenins (HMW-GS) and low molecular weight glutenins (LMW-GS), which are located on the long and short arm of group 1 chromosomes, respectively. Extensive glutenin variation has been revealed in terms of chemical composition and processing qualities. Apart from glutenins, gluten also contains a small proportion (8–10%) of soluble proteins (albumins and globulins). Removal of gluten from bread formulations often results in a liquid batter, rather than a dough system during pre-baking, and results in baked bread with crumbly texture, poor colour and other quality defects (Gallagher et al. 2003). As the main structure-forming protein-complex present in wheat dough, gluten plays a major role in bread-making functionality by providing viscoelasticity to the dough, good gas-holding properties, and good crumb structure of many baked products (Gallagher et al. 2003; Moore et al. 2004).

It has long been known that the bread-making quality of wheat flour depends on both

the quantity and quality of its gluten proteins. Gluten proteins are largely insoluble in water or dilute in salt solutions. Two functionally-distinct groups of gluten proteins are the monomeric gliadins and polymeric (extractable and unextractable) glutenins (Lindsay and Skerritt 1999). Prolamins have a unique amino acid composition, with the amino acids glutamine (Gln)/glutamic acid (Glu) and proline (Pro), accounting for more than 50% of amino acid residues (Eliasson and Larsson 1993). The low water solubility of gluten is attributable to its low content of lysine (Lys), arginine (Arg) and aspartic acid (Asp) residues, which together amount to <10% of total amino acid residues. About 30% of amino acid residues in gluten are hydrophobic, and contribute greatly to its ability to form protein aggregates by hydrophobic interactions, as well as bind lipids and other non-polar substances. The high contents of Gln and hydroxyl amino acids (~10%) in gluten are responsible for its water-binding properties. In addition, hydrogen bonding between Gln and hydroxyl residues of gluten polypeptides contributes to its cohesion–adhesion properties. Cysteine (Cys) and cysteine residues account for 2–3% of total amino acid residues and, during formation of dough, these residues undergo sulfhydryl-disulfide interchange reactions, resulting in extensive polymerisation of gluten proteins (Ma et al. 2009).

Gliadins and glutenins are usually found in more or less equal amounts in wheat. However, variation in the relative proportions of gliadin and glutenin occurs between different wheat varieties and largely depends on the growing conditions of wheat. This ratio of monomeric to polymeric proteins affects the physical properties of dough, with higher relative proportions of glutenin imparting greater dough strength (MacRitchie 1987). The gliadin fraction has been reported to contribute to the viscous properties and dough extensibility of wheat dough (Don et al. 2003a; Don et al. 2003b; Pomeranz 1988). The polymeric glutenin fraction of wheat gluten has long been considered to have a prominent role in the elasticity and strengthening of dough (Hoseney 1986; MacRitchie 1980; Xu et al. 2007).

The high level of polymorphism in wheat prolamins results in a special effect in relation to the overall functional properties of wheat dough. During dough formation, when prolamin proteins are hydrated and form the gluten network, the numerous structurally-similar but slightly different proteins produce a mass in which several characteristics (such as size, polarity, charge distribution, solubility and viscosity) are continuously distributed over a relatively large interval. This structural feature is unique among

protein complexes.

The structural complexity of the gluten complex means that relating protein composition to quality has to be investigated at different levels of protein content, and different ratios of polymeric to monomeric protein, HMW to LMW glutenin subunits, and x- and y-type HMW glutenin subunits (Bushuk 1998; Bushuk and Bekes 2002; Oliver and Allen 1992; Shewry et al. 1999; Wrigley et al. 2006). Studies of this type have demonstrated that polymeric glutenin is mostly responsible for the elasticity of dough, whereas monomeric gliadins are extensibility-related characters in the system (Hoseney 1986). Thus, the ratio of polymeric to monomeric proteins (glutenin-to-gliadin ratio) can be directly related to the balance of dough strength and extensibility of the sample.

Beyond the essential role of defining functional properties of wheat doughs, the unusual amino acid composition and primary structure of prolamins have two other nutrition- and health-related consequences. The nutritional value of gluten proteins is rather poor and this results in continued activity in the food industry to improve it by supplementing cereal products with other protein sources such as legumes. This is discussed in detail in sections 1.4.1 and 1.4.2. Consumption of gluten-containing food—those containing wheat, rye, barley and oats—causes health problems for a significant number of people. Most people in Western countries are now aware of the potential adverse effects of cereals containing gluten with reports appearing in the media (Braly and Hogganm 2002; Ford 2008; Wangen 2009) promoting gluten-free diets. Many of these reports fail to draw attention to the importance of appropriate diagnosis or defining the nature of the gluten ‘intolerance’ an individual may have, and hence pose a significant public relations threat to the grain industry.

As discussed earlier, the frequency of protein cross-linking and interactions during the process of dough preparation and baking is an important variable (Singh 2005). During dough preparation, different ingredients such as oxidising agents, salts and water are added which promote and affect the formation of protein cross-links (Gerrard et al. 2001; Lefebvre et al. 2007; Rosell and Collar 2009). Mechanical energy inputs supplied during mixing and increased temperature during baking also significantly affect protein cross-linking (Rosell and Collar 2009). Interactions of lupin and wheat proteins during dough preparation and baking of lupin–wheat bread (Dervas et al. 1999; Doxastakis et

al. 2002) occurs which can generate new proteins or protein cross-links not present in the original flour. Under this process, any modification of protein responsible for bioactivities (Duranti et al. 2008; Doxastakis et al. 2002) would be critical in relation to the claimed health benefits (Lee et al. 2009; Lee et al. 2006) of the baked products.

A number of studies on soy–wheat flour interactions have been conducted (Ribotta et al. 2005b; Roccia et al. 2009; Ryan et al. 2002) but few were on protein interactions. An electrophoresis study revealed a strong association between wheat and soy proteins in baked products (Ribotta et al. 2005b). The study also reported that the soy and wheat proteins interacted by means of non-covalent and covalent (disulfide) bonds and the extent of the interactions depended on the soy protein state. However, research on lupin–wheat flour interactions is limited to physical and chemical properties and sensory evaluation (Guemes-Vera et al. 2004; Guemes-Vera et al. 2008; Mubarak 2001; Torres et al. 2007). Thus research on lupin–wheat protein interactions at the molecular level would have the potential to promote the use of such a promising crop for human food. Moreover, it would contribute to the better understanding of health-related issues of lupin–wheat bread and have a significant practical role in food and health industries.

2.5 New technologies for proteomic study

2.5.1 Approaches to studying grain seed storage proteins

Proteomic analysis is becoming an effective way to understand plant biology in parallel to genomics and phenomics (Magni et al. 2007). In general, seed proteins are large in number and this has generated research focusing on more efficient techniques for proteome analysis. Among the methods used for characterising proteins—sodium dodecyl sulphate polyacrylamide gel electrophoresis [SDS-PAGE] (Ng and Bushuk 1987), acid PAGE (Yan et al. 2003) and reversed-phase high-performance liquid chromatography [RP-HPLC] (Margiotta et al. 1993; Wieser et al. 1998)—are used traditionally. While these techniques are popular and accurate to some extent, they have limitations in resolving complex mixtures of proteins (Cozzolino et al. 2001; Zhang et al. 2008). Proteome analysis with two-dimensional gel electrophoresis (2-DGE), where more than a thousand protein spots can be visualised, is the most powerful tool for identifying polymorphism of proteins in flours (Yahata et al. 2005). 2-DGE detected almost 1300 proteins spots in wheat endosperm, and supplied information concerning differences in protein composition to environmental influence (Skylas et al. 2001; Yahata et al. 2005). Furthermore, two-dimensional gel electrophoresis has been used

successfully for the assessment of industrially processed lupin products (Capraro et. al. 2008).

In addition, new techniques based on ionisation, such as fast-atom bombardment (FAB), electrospray iononisation (ESI) and matrix-assisted laser desorption/ionisation (MALDI) has improved protein mass spectrometry significantly (Cunsolo et al. 2004). In particular, matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) has become a high throughput tool for storage protein analysis (Alberghina et al. 2005; Chen et al. 2007; Cunsolo et al. 2004; Muccilli et al. 2005). The MALDI-TOF-MS technique provides precise results; very small amounts of sample are required (normally less than 1 pmol), and is relatively fast (requiring only a few minutes per sample) compared to the common separation methods. Moreover, it may facilitate analysis of proteins or peptides from complex mixtures without purification and separation (Kussmann et al. 1997). Storage proteins of lupin, particularly *L. angustifolius*, have not been analysed using this technique. Research in this thesis was designed to analyse seed protein profiles across lupin species and cultivars and to study lupin protein interactions in baking by using MALDI-TOF and two-dimensional gel electrophoresis technology which will contribute to proteome research of lupin and lupin–wheat interactions.

2.5.2 Protein identification by peptide sequencing

Protein identification is an important part of the gel-based proteomic studies particularly of two-dimensional gel electrophoresis. The approaches used for protein identification typically accomplished by either peptide mass fingerprinting (PMF) or amino acid sequencing of the peptide using tandem mass spectrometry (MS/MS) (Damodaran et al. 2007). In PMF, the protein is broken down into peptides by proteolytic cleavages using an enzyme such as trypsin. The resultant peptides are analysed using a MALDI-TOF mass spectrometer to determine the mass-to-charge (*m/z*) ratio that produces a fingerprint of peptides masses. Matching these masses with peptide masses of proteins in a database (for example NCBI) allows identification of the corresponding protein (Barrett et al. 2005; Damodaran et al. 2007). There is a high risk of ambiguity in protein identification due to peptide mass redundancy in PMF (Barrett et al. 2005). For example, the peptide sequence GIFKS has the same mass as SKFIG, FIGKS and so on. Consequently PMF has a low success rate for identifying proteins accurately when there are mixtures of proteins.

Recently, modifications to the PMF approach have been introduced using various sequential electrophoretic or chromatographic separations to achieve adequate peak capability to resolve complex samples (Hu et al. 2004; Wang et al. 2005). This approach of peptide sequencing involves tandem mass spectrometric (MS/MS) analysis of digested peptides using collision-induced dissociation (CID) to obtain fragment ions. Subsequently, sequence ions and intact peptide masses are matched against protein databases to identify the unknown protein (Chen and Pramanik 2009; Baldwin 2004). This approach provides more specific identification of proteins than PMF (Lasonder et al. 2002). Moreover, MS/MS can identify proteins from a mixture of proteins (Barrett et al. 2005). In spite of major limitations, such as post-translational modifications not likely to be maintained during CID at the peptide level, this is the current method of choice in protein identification in proteome studies because of the advancement in LC/MS/MS technology and compatible software development (Chen and Pramanik 2009).

Several systems can be used for MS/MS studies including: ion trap/linear ion trap (IT-LIT), quadrupole time-of-flight (Q-Q-TOF), hybrid time-of-flight (TOF TOF), Fourier transform–ion cyclotron resonance (FT-ICR), triple quadruple (Q-Q-Q) and quadrupole–linear ion trap (QQ-LIT) (Domon and Aebersold 2006). They differ in terms of performance standards and their ability to support specific analytical strategies. A basic comparison of these systems was undertaken by Domon and Aebersold (2006) and is presented in **Table 6**.

Table 6: Characteristics and performances of commonly-used types of mass spectrometers. ✓ available; (✓) optional; + possible or moderate; ++ good or high; +++ excellent or very high; Seq., sequential (source: Domon and Aebersold 2006).

	IT-LIT	Q-Q-ToF	ToF-ToF	FT-ICR	Q-Q-Q	QQ-LIT
Mass accuracy	Low	Good	Good	Excellent	Medium	Medium
Resolving power	Low	Good	High	Very high	Low	Low
Sensitivity (LOD)	Good		High	Medium	High	High
Dynamic range	Low	Medium	Medium	Medium	High	High
ESI	✓	✓		✓	✓	✓
MALDI	(✓)	(✓)	✓			
MS/MS capabilities	✓	✓	✓	✓	✓	✓
Additional capabilities	Seq. MS/MS			Precursor, Neutral loss, MRM		
Identification	++	++	++	+++	+	+
Quantification	+	+++	++	++	+++	+++
Throughput	+++	++	+++	++	++	++
Detection of modifications	+	+	+	+		+++

Time of flight (ToF) mass analyser is the basic platform for analysing with both ESI and MALDI technology. The Q-Q ToF instruments can provide high resolution and mass accuracy in MS/MS analysis (Domon and Aebersold 2006). The introduction of ion trap (IT) facilitated high-throughput and fast data acquisition. The development of linear ion trap (LIT) technology providing higher ion-trapping capacities has extended the dynamic range and overall sensitivity of this technique. In particular, LIT instruments have the option for slow scanning to boost the resolution. Moreover, these instruments have the potential for analysis of post-translational modifications (Olsen and Mann 2004). Sequentially, LIT devices have been successfully coupled to triple quadrupole–type technique (i.e. the second analyser is substituted by a LIT) to offer exclusive performance (Hager 2002). For example, the instrument offers the scanning capabilities of a triple quadrupole instrument, including precursor ion and neutral loss scanning and increased sensitivity (Domon and Aebersold 2006). These attributes enable these instruments to offer the analysis of modifications and multiple reactions monitoring (MRM) capability. Two stages analysing in combination with the high duty cycle results in quantitative analyses with unmatched sensitivity. On the other hand, introduction of robust ion cyclotron resonance (ICR) mass spectrometry with external ion sources is a breakthrough in terms of resolving power and mass accuracy (Senko et al. 1997). However, the technology has drawbacks of relatively slow acquisition rate and limited dynamic range of IT devices (Domon and Aebersold 2006).

2.5.3 Proteomics to genomics technologies

Functional genomic and proteomic studies of seed promise to reveal patterns of gene expression associated with key developmental events. Detailed identification and characterisation of lupin seed proteins are useful in terms of quality improvement of lupin grain (Duranti et al. 2008; Magni et al. 2007). Since protein is the expression of genome or gene product, proteomic analyses need to be linked with genomic studies, and with the innovation of modern technologies it becomes feasible (Foley et al. 2011). Molecular breeding to enrich lupin grain protein quality would be more efficient once we have the genomic sequence information of target regions encoding the specific seed proteins.

As described earlier, seed storage proteins are classified into four groups: α, β, γ and δ-conglutins (Blagrove and Gillespie 1975; Duranti et al. 2008). Although a number of sequences have been reported for *Lupinus* seed storage proteins, the sequence of genes

encoding the major *Lupinus* storage globulins α and β-conglutins are only in some databases. Until early 2011 there was only one sequence of α-conglutin and 3 sequences of legumin-like proteins (*L. albus*) in the database. Thus the identification and characterisation of the large number of α-conglutins is difficult (Sirtori et al. 2010) based on mass spectrometric peptide sequencing; many seed proteins belonging to this group were not identified with existing sequences of lupin seed proteins in the database (Goggin et al. 2008; Magni et al. 2007; Sirtori et al. 2010). The three sequences of β-conglutin were determined by Goggin et al. (2008) and 1 vicilin protein sequence was deposited by Magni et al. (2007). Another sequence of the gene encoding β-conglutin precursor (1791 nucleotides) was reported by Monteiro et al. (2010).

Recently, Foley et al. (2011) reported gene sequences encoding three unique α, seven β, two γ and four δ-conglutins through sequencing of ESTs of narrow-leafed lupin. Of these 16 sequences, 11 were new (two α, five β, one γ and three δ-conglutins). The sequences of many of the genes encoding conglutins and other seed storage proteins are yet to be revealed; available sequences of lupin seed proteins are less than related species such as soybean and peanut. Thus there is scope to develop new sequences of lupin seed proteins that would contribute to the whole genome sequence of lupin. With the advancement of mass spectrometry, it might be possible to generate the sequence of the target gene of individual proteins as specified by the electrophoretic mobility in 2-DGE gels based on mass spectrometric peptide sequencing.

Chapter 3

Mass spectrometric fingerprints of seed protein for defining *Lupinus spp.* relationships

Citation: Islam S, Ma W, Appels R, Yan G. 2012. Mass spectrometric fingerprints of seed protein for defining *Lupinus spp.* relationships. *Genetic Resources and Crop Evolution* (published online: dx.doi.org/10.1007/s10722-012-9890-y).

Abstract

The *Lupinus* genus consists of 200 to 600 species with diversified morphological and biochemical characteristics that provide scope for increasing the value of protein attributes in commercial lupin species through genetic improvement. This study characterised protein polymorphism among 33 selected lupin genotypes representing 19 species from Old World and New World using Matrix-Assisted Laser Desorption Ionization Time-Of-Flight (MALDI-TOF) mass spectrometry with the aim of developing a high throughput screen for assessing variation among samples within a genetic resource. MALDI-TOF mass spectrometry analysis recognized 630 polymorphic protein mass peaks. A total of 19 protein mass peaks were identified as species specific and 23 were recognized to be generally common (observed in more than 10 species). A set of 111 rare protein mass peaks common to 2-3 species were identified to uniquely distinguish the respective species in combination with the species specific mass peaks. Overall, the seed protein profiles showed extensive diversity among the species analysed. Phylogenic analysis based on the protein mass peaks clustered the Old World smooth and rough-seeded and New World lupin species separately, in broad agreement with clustering based on morphological characters, life history, cytogenetics, DNA polymorphism and electrophoreses of seed proteins. Two New World lupin species, *L. mutabilis* Lindl. and *L. succulentus* Douglas ex K. Koch were grouped into the smooth-seeded Old World group.

Keywords: Diversity, *Lupinus*, MALDI-TOF, mass peaks, phylogeny, seed protein.

Introduction

Lupin (*Lupinus* spp) is a legume crop has gained more attention in the past two decades in many countries (FAO 2011) because of its excellent nutritional attributes and adaptability to extreme climatic conditions (Dervas et al. 1999; Guemes-Vera et al. 2008). The nutritional quality of lupin is attributed to its high protein content, low fat and starch content, and high dietary fibre content that are well suited to low glycaemic indexed food products (Islam et al. 2011b; Sirtori et al. 2010). Due to the high protein content, lupin grain is widely used as animal feed. Its use as food or food ingredient for human consumption is increasing due to its nutritional properties and lower levels of anti-nutritional factors (Petterson et al. 1997). Lupin enriched food has been claimed to provide health benefits such as increased satiety and reduced energy intake (Lee et al.

2006), decreased blood pressure (Lee et al. 2009), decreased blood glucose level (Hall et al. 2005) and cholesterol-lowering effect (Martins et al. 2005).

The diverse genetic resources are the foundation for crop improvement through breeding (Talhinhas et al. 2006). Lupin (*Lupinus* L) is a large genus within the family Fabaceae consisting of herbaceous, soft-woody shrubs and small trees (Dunn 1984; Turner 1995) of more than 200 annual and perennial species (Dunn and Gillett 1966). The species of lupin are grouped as Old World species and New World species. The New World species comprise over 90% of the total species, mostly distributed from Alaska to South Argentina and Chile (Dunn 1984). The Old World species include 13 lupin species of the Mediterranean and African origin (Gladstones 1974) and are annual, herbaceous and large seeded. Three Old World species, namely, *L. albus* L., *L. angustifolius* L. and *L. luteus* L. and one New World specie *L. mutabilis* Lindl. are generally cultivated as crops (Petterson 1998). Some of the other species are used as green manure, forage, ornamentals and for land stabilization (Gladstones 1998).

Lupin species are characterized by a diverse geographical distribution, morphological variation and chromosomal polymorphisms that lead to a large variation in seed protein levels. The life history, leaf morphology, cotyledon form, seed coat surface pattern have been the main focus of detailed comparative morphological studies across the genus (Eastwood et al. 2008). Cytological studies of lupins suggested a significant difference among the species consistent with the observed morphological diversity. The New World lupins have $2n$=48 as their predominant chromosome number and their basic chromosome number is considered to be x=6 (Dunn 1984). In contrast, the Old World lupins have a series of basic chromosome numbers from x=5, 6 (or 9), 7, 8, 13 at different ploidy levels (Pazy et al. 1977). Relatively recent changes in chromosome number are believed to have occurred as a result of intense selection or evolution in changeable environments (Gladstones 1998). Among the smooth-seeded Old World lupin species, *L. angustiflolius* has lower chromosome number, $2n$= 40, compared to *L. albus* $2n$=50, *L. micranthus* Douglas ex Lindl., *L. luteus* and *L. hispanicus* Boiss. & Reut. $2n$= 52.

Genome size variations among species of major legumes including Soybean (Rayburn et. al. 2004) and *Pisum spp.* (Baranyi et. al. 1996) have been reported. Likewise, in the case of lupin, DNA based analyses have revealed substantial variation in genetic

contents among species (Ainouche and Bayer 1999; Christopher 2008; Kass and Wink 1997), leading to expectations of considerable differences in seed protein compositions. Seed globulins studies based on HPLC (Salmanowicz 1999) and one dimensional electrophoresis (Przybylska and Zimniak-Przybylska 1995; Salmanowicz and Przybylska 1994) confirmed these expectations and revealed significant differences among the Old World rough-seeded and smooth-seeded lupin species. A similar study also reported differences among the 19 species of New World lupins (Zimniak-Przybylska and Przybylska 1997). More heterogeneity was found in smooth-seeded lupins than the rough-seeded lupins (Salmanowicz 1999).

Matrix-Assisted Laser Desorption Ionization Time-Of-Flight mass spectrometry (MALDI-TOF- MS) has been used for seed protein analysis (Chen et al. 2007; Liu et al. 2009) due to the high throughput capability of this technology. The MALDI-TOF-MS technique provides accurate molecular weight data, requires very small amount of sample (normally less than 1 pmol). Moreover, it is able to analyse proteins from complex mixtures without purification and separation (Kussmann et al. 1997). Compared with the traditional one and two dimensional gel electrophoreses for seed protein identification (Islam et al. 2012; Magni et al. 2007; Yahata et al. 2005), this approach is more time and cost efficient (Barakat 2004). The current study analysed the seed protein profile of important lupin species by MALDI-TOF mass spectrometry to study the diversity and deduce the relationships among the species.

Materials and Methods

In total, 35 genotypes comprising 11 Old World and 8 New World lupin species, and two out group species from the *Crotalaria* genus (Ainouche and Bayer 2000; Ainouche et al. 2004; Ainouche and Bayer 1999; Sanz Penella et al. 2008) were included in the studies. The list of the genotypes is presented in **Table 1**. The lupin samples were supplied by the Department of Agriculture & Food, Western Australia.

Protein was extracted from lupin flour based on previously established method (Duranti et al. 2008; Islam et al. 2011a) which included grinding the seeds into fine powder manually in the laboratory. Flour samples were defatted by Hexane at 20:1 ratio (Santos et al. 1997). Protein was extracted in extraction buffer (0.5 M NaCl) at the ratio of 15 ml/g by stirring at 4 ^{o}C for 4 h and the supernatant was collected by centrifugation at 10,000 g for 10 mins. The extract was mixed with matrix (Sinapinic acid dissolved in

0.05% trifluoroacetic acid and 50% acetonitrile) at 1:9 ratio and 1 µl of mixture was spotted on MALDI-TOF plate and left at room temperature to dry. Spotting was repeated once when the previous spots were completely dry. The sinapinic acid was purchased from Sigma-Aldrich, St. Louis, MO, USA. Three separate extractions were made for MALDI-TOF protein analysis to make sure reproducibility.

Table 1: List of the species and genotypes used in this study

Group	Species name	Genotype identity	Country of origin
Old World smooth-seeded	*L. angustifolius*	Mandelup	Australia
		P27255	Morocco
		P22724	Spain
	L. luteus	Wodjil	Australia
		P26835	Morocco
	L. albus	Kiev Mutant	Russia
		subsp graecus, P23037	Greece
		subsp albus, P27715	Egypt
	L. micranthus	P22944	Morocco
		P26799	Greece
	L. hispanicus	*subsp hispanicus,* P23008	Spain
		subsp bicolour, P23006	Portugal
Old Word rough-seeded	*L. cosentinii*	P20747	Algeria
		P22911	Tunisia
	L. digitatus	P26877	Egypt
		P27045	Egypt
	L. princei	P26878	Kenya
		P26879	Kenya
	L. pilosus	P20955	Hungary
		P23029	Syria
	L. palaestinus	P20956	Israel
		P26940	Israel
	L. atlanticus	27219	Morocco
		27248	Morocco
New World	*L. mutabilis*	P27808	Ecuador
		P26961	Netherlands
	L. polyphyllus	P23299	Sudan
	L. succulentus	P27927	USA
	L. subcarnosus	P26969	USA
	L. ornatus	P26038	India
	L. hirsutissimus	P 28006	USA
	L. truncetus	P 27413	USA
	L. arizonicus		USA
Outgroup species	*Crotalaria cunninghamii*		Australia
	Crotalaria eremaea		Australia

The experiments were carried out on a Voyager DE PRO Biospectrometry Workstation from PerSpective Biosystem, Framingham, MA, USA, operated in linear mode (Lou et al. 2010). Final mass spectrum for each sample was obtained by averaging 500 shots on

a protein spot over random locations. The machine was calibrated by using "Sequazyme Peptide Mass Standards Kit" from Applied Biosystem, Foster City, USA following Sinapinic acid matrix-calibration mixture 3 as suggested by the supplier. To get the best resolution, the molecular weight range of 2,000-32,000 Dalton was split into a 3000 Dalton intervals. High molecular weight proteins of 30,000-75,000 Dalton were also analysed.

Data analysis

The results from MALDI-TOF were analysed using the Voyage machine companion software, Data Explorer, to produce the protein mass peak profiles (Liu et al. 2009). The mass spectrometric data were then analysed using software "Progenesis PG 600" from Nonlinear Dynamics Durham, NC, USA. The mass peak profiles were manually checked and the identified polymorphic mass peaks were scored visually for absence and presence. A binary dataset was constructed for multivariate analysis using the software PAUP (Phylogenetic Analysis Using Parsimony) (Swofford 1998). A distance matrix based on total character difference was constructed and the Neighbour Joining procedure was followed to produce a dendrogram setting the two species of *Crotalaria* as out group. Bootstrap analysis was carried out with 10,000 replications to assess the reliability of groupings.

Results

MALDI-TOF mass spectrometry analysis recognized 630 polymorphic mass peaks among the 35 genotypes studied. The mass peaks were clear (**Figure 1**) and generally very reproducible among the replications. The average number of mass peaks corresponding to each species (**Table 2**) varied from 65 to 177, demonstrating a high level of proteomic diversity. *Lupinus luteus* showed the highest average number of mass peaks followed by *L. albus* and *L. angustifolius* (**Table 2**). A total of 23 protein mass peaks were recognized as very common based on being observed in more than ten species (**Table 3**) which accounted for only 3.6% of the total number of mass peaks. No protein peaks were found to be common to all of the genotypes.

A total of 19 mass peaks were recognized as species specific (**Table 4**). Twelve lupin species out of nineteen had species specific proteins. *Lupinus succulentus* had the highest species specific mass peaks (3) followed by *L. angustifolius, L. subcarnosus* Hook*., L. atlanticus* Gladstones*, L. albus and L. cosentinii* Guss., each of them had 2

species specific mass peaks. The species *L. pilosus* Murray*, L. digitatus* Forssk.*, L. luteus, L. hirsutissimus* Benth.*, L. truncatus* Nutt. ex Hook. & Arn.*, L. arizonicus* S. Watson each had one species specific mass peak.

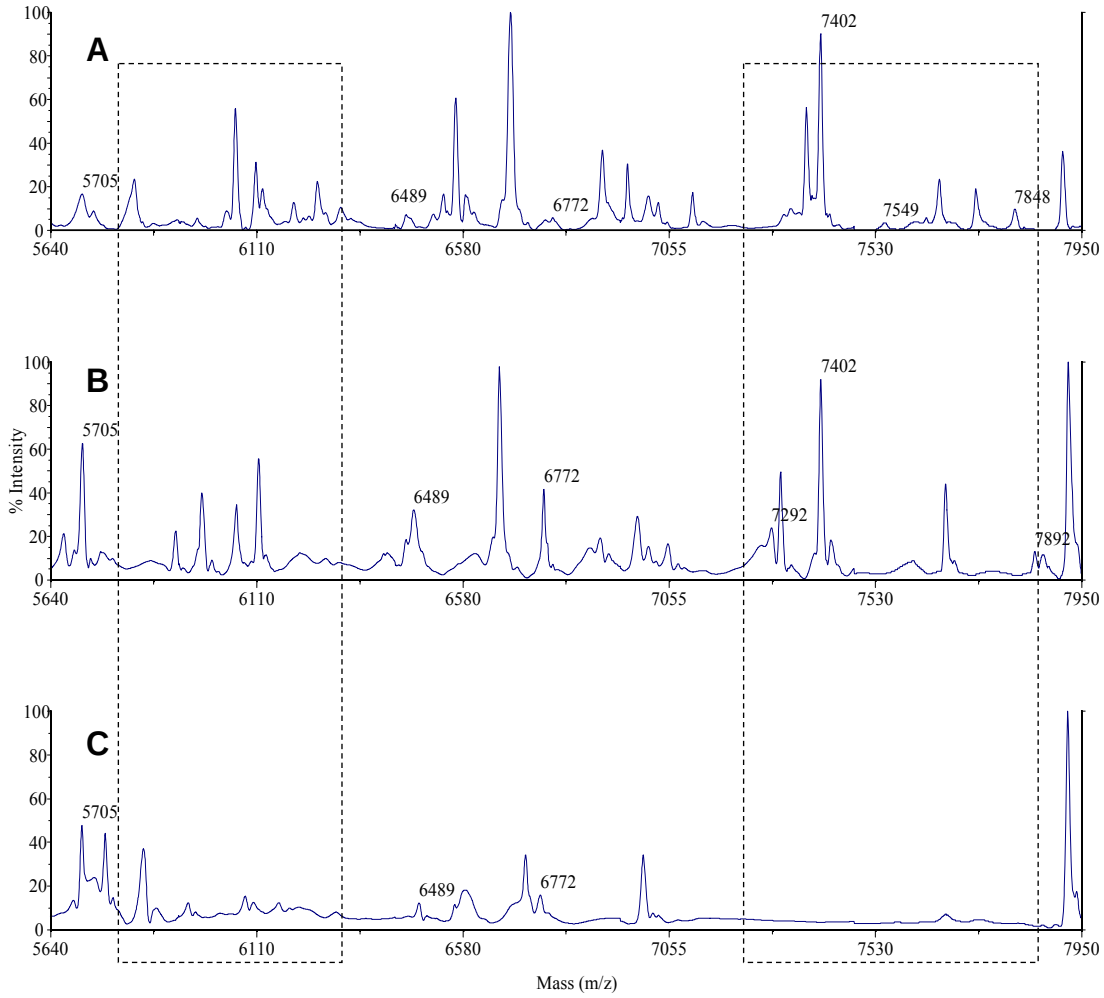


Figure 1: MALDI–TOF outputs of seed protein profiles demonstrating easily visible and identifiable polymorphism of protein mass peaks among different lupin species. Plate A, B and C represents smooth seeded old world species *L. luteus* (P26835); rough seeded old world species *L. digitatus* (P27045) and the new world species *L. hirsutissimus,* respectively. The rectangles are highlighting the regions with polymorphic peaks. The numbers on the protein peaks indicate the molecular weight of the corresponding protein in Daltons

The mass peaks found in only 2-3 species were defined as rare mass peaks and a set of 111 rare mass peaks from all species has been compiled based on the mass spectrometric protein profiles (**Table 5**). *Lupinus angustifolius* had the highest number

(34) of rare mass peaks followed by *L. luteus* (26) and *L. albus* (20) and the species *L. ornatus* Douglas ex Lindl. had the lowest number (4) of rare mass peaks. Pairwise differences in mass peaks among the genotypes were analysed by the distance matrix calculated by PAUP (**Table 6**). The pairwise difference ranged from 81 to 247 mass peaks.

Table 2: Average number of mass peaks in the species seed protein profiles by MALDI-TOF mass spectrometry

Group	Species name	Average mass peaks identified ±(SE)
Old World-smooth seeded	*L. angustifolius*	148 (±4.3)
	L. luteus	177 (±24.0)
	L. albus	150 (±0.6)
	L. micranthus	94 (±15.5)
	L. hispanicus	105 (±13.5)
	L. cosentinii	108 (±12.5)
Old World-rough seeded	*L. digitatus*	122 (±8.0)
	L. princei	83 (±6.0)
	L. pilosus	103 (±7.0)
	L. palaestinus	77 (±18.0)
	L. atlanticus	82 (±4.0)
	L. mutabilis	133 (±0.5)
New World	*L. polyphyllus*	85 (-)
	L. succulentus	116 (-)
	L. subcarnosus	85 (-)
	L. ornatus	71 (-)
	L. hirsutissimus	88 (-)
	L. truncetus	65 (-)
	L. arizonicus	86 (-)
Outgroup genus	*Crotalaria*	86 (±1.5)

The dendrogram produced from the distance matrix showed that there is a considerable level of diversity among the species and close relationship among the genotypes within the species (**Figure 2**). The dendrogram separated the 19 species into two major groups. The largest group consisted of fifteen genotypes that included all the 12 genotypes of five Old World smooth-seeded lupin species studied. Remarkably, all the genotypes studied of New World species *L. mutabilis* (two genotypes) and *L. succulentus* (one genotype) were clustered in this group of Old World smooth-seeded lupins. Within this group *L. angustifolius, L. luteus* and *L. albus* formed a subgroup with 67% bootstrap supports. The new world species *L. mutabilis* placed with this sub-group with 77% bootstrap supports. The genotypes within the species of Old World smooth-seeded lupin grouped together with a high bootstrap (>99%) value (**Figure 2**).

Table 3: List of the very common mass peaks* of lupin seed protein as identified by mass spectrometry

Protein mass peaks as molecular weight (Dalton)	Present in Number of lupin species out of 19 studied	Missing species
2,121	18	*L. digitatus*
3,136	12	*L. digitatus, L. Polyphyllus, L. subcarnosus, L. ornatus, L. arizonicus, L. truncetus,, L. hirsutissimus*
3,726	12	*L. princei, L. atlanticus, L. succulentus, L. subcarnosus, L. ornatus, L. truncetus, L. hirsutissimus*
4,147	11	*L. cosentinii, L. princei, L. palaestinus, L. atlanticus, L. Polyphyllus, L. subcarnosus, L. ornatus, L. hirsutissimus,*
5,198	11	*L. princei, L. pilosus , L. palaestinus , L. atlanticus, L. polyphyllus, L. succulentus, L. subcarnosus, L. ornatus*
5,224	17	*L. truncetus, L. arizonicus*
5,522	13	*L. palaestinus, L. ornatus, L. subcarnosus, L. succulentus, L. polyphyllus, L. truncetus*
5,705	12	*L. albus, L. micranthus, L. cosentinii, L. princei, L. atlanticus, L. succulentus, L. ornatus,*
6,489	13	*L. cosentini, L. princei, L. atlanticus, L. polyphyllu, L. succulentus, L. ornatus*
6,772	12	*L. micranthus,L. hispanicus, L. cosentinii, L. polyphyllus, L. succulentus, L. subcarnosus, L. ornatus*
7,402	15	*L. subcarnosus, L. hirsutissimus, L. truncetus, L. arizonicus*
7,965	16	*L. subcarnosus, L. albus , L. truncetus*
7,990	15	*L. ornatus, L. subcarnosus, L. succulentus, L. truncetus*
8,013	12	*L. pilosus, L. palaestinus, L. atlanticusL, Polyphyllus, L. succulentus, L. subcarnosus, L. ornatus,*
8,139	11	*L. albus, L. princei, L. palaestinus, L. atlanticus, L. Polyphyllus, L. subcarnosus, L. arizonicus, L. hirsutissimus*
8,195	12	*L. luteus, L. hispanicus, L. palaestinus, L. subcarnosus, L. arizonicus, L. truncetus, L. hirsutissimus*
8,275	12	*L. mutabilis, L. cosentinii, L. palaestinus, L. atlanticus, L. succulentus, L. subcarnosus, L. ornatus,*
8,522	13	*L. palaestinus, L. cosentinii, L. micranthus, L. hirsutissimus, L. truncetus, L. arizonicus*
8,681	13	*L. micranthus, L. atlanticus, L. polyphyllus, L. succulentus, L. truncetus, L. hirsutissimus*
8,795	11	*L. micranthus, L. hispanicus, L. atlanticus, L. polyphyllus, L. subcarnosus, L. ornatus, L. truncetus, L. hirsutissimus*
8,817	14	*L. subcarnosus, L. micranthus, L. hirsutissimus, L. truncetus, L. arizonicus*
9,039	12	*L. ornatus, L. subcarnosus, L. succulentus, L. polyphyllus, L. hirsutissimus, L. truncetus, L. arizonicus*
16813	12	*L. cosentinii, L. Polyphyllus, L. subcarnosus, L. ornatus, L. arizonicus, L. truncetus, L. hirsutissimus*

* Very common mass peaks are those found in more than ten lupin species out of 19 studied

Table 4: List of species-specific mass peaks* of lupin seed protein as revealed by mass spectrometry

Group	Species name	Molecular weight of the protein mass peaks (Dalton)
Old World-smooth seeded	*L. angustifolius*	10545**, 19193
	L. luteus	21857
	L. albus	3646**, 14879**
	L. micranthus	-
	L. hispanicus	-
Old World-rough seeded	*L. cosentinii*	12506, 64088
	L. digitatus	12913
	L. princei	-
	L. pilosus	12835
	L. palaestinus	-
	L. atlanticus	9847, 10608
New World	*L. mutabilis*	-
	L. polyphyllus	-
	L. succulentus	9289, 10659, 19108
	L. subcarnosus	3452, 16906
	L. ornatus	-
	L. hirsutissimus	23813
	L. truncetus	23882
	L. arizonicus	14379

*Species-specific mass peaks are those specific to a single species

**Protein present in 2 genotypes of the species out of 3 studied

The second largest group of the dendrogram consisting of twelve genotypes and includes all the genotypes of six species of Old World rough-seeded lupin species. The species *L. cosentinii* and *L. atlanticus*; *L. digitatus* and *L. princei* Harms appeared as sister species respectively and formed a sub-group together. In this sub-group the genotypes within the species *L. cosentinii, L. atlanticus, L. digitatus* and *L. princei* positioned together with the bootstrap supports 94%, 84%, 60% and 67% respectively. The four genotypes of *L. pilosus* and *L. palaestinus* Boiss. grouped together with 85% bootstrap. Besides these two large groups there were two other small groups in the dendrogram comprising 3 New World lupin species in each. The species *L. arizonicus, L. truncatus* and *L. hirsutissimus* formed one small group with 96% bootstrap support. The other small group supported by 78% bootstrap comprises *L. polyphyllus* Lindl., *L. subcarnosus* and *L. ornatus*. The out-group species (*C. cunninghamii and C. eremaea*) were separated from the lupin species with 100% bootstrap support.

Table 5: List of rare mass peaks* of lupin seed protein as revealed by mass spectrometry

Group	Species Name	Rare proteins (identified as molecular weight in Dalton)	Total
Old World smooth-seeded	*L. angustifolius*	2306, 2396, 3157, 3171, 3576, 3907, 4751, 7171, 7651, 7848, 10565, 11048, 11213, 11980, 14811, 16669, 17036, 17111, 17149, 17199, 17237, 19659, 20843, 20910, 21410, 21485, 21982, 23612, 24203, 24457, 24896, 28668, 30642, 49726	34
	L. luteus	2663, 2677, 3157, 3576, 3609, 3907, 4751, 7549, 7640, 7848, 9473, 9667, 10275, 10296, 16669, 17237, 20390, 20910, 21982, 23000, 24542, 24619, 24933, 26243, 26392, 28668	26
	L. albus	2306, 3664, 4257, 4511, 4720, 7472, 7640, 7651, 8508, 11980, 16403, 17464, 18371, 21919, 22386, 24203, 24457, 30642, 40000, 49726	20
	L. micranthus	7292, 10275, 11213, 11446, 13223, 13490, 17199, 19342, 19659, 24152, 26392, 31161, 64730	13
	L. hispanicus	3171, 3397, 3609, 4257, 7171, 9166, 10939, 17855, 18258, 21806, 21919, 25976, 30080	13
Old World rough-seeded	*L. cosentinii*	7549, 9243, 9667, 11643, 13490, 22731, 23000, 23410, 26500, 26628, 64730	11
	L. digitatus	7292, 7893, 10365, 10939, 13223, 15650, 15807, 19002, 22731, 23410, 23612	11
	L. princei	17464, 17546, 17855, 21410, 26500, 31161	6
	L. pilosus	2396, 3397, 4720, 5473, 7472, 9318, 11297, 11643, 15650, 17464, 19843	11
	L. palaestinus	2663, 4511, 8508, 10365, 13223, 14683, 17111, 17855, 19843, 22386	10
	L. atlanticus	8502, 9243, 11213, 11446, 11917, 13615, 15807, 20390, 23000, 24542, 26500	11
New World	*L. mutabilis*	3664, 5473, 7292, 7651, 9166, 9435, 10565, 16507, 17199, 18371, 20843, 21806, 24152	13
	L. polyphyllus	7893, 8222, 10565, 16507, 19342, 26243, 26628	7
	L. succulentus	2807, 4751, 7549, 16403, 16571, 17149, 19002, 25495, 25976, 26243, 40000	11
	L. subcarnosus	2807, 3576, 3664, 8487, 9473, 11048, 11917, 14811, 16753, 17036, 47478	11
	L. ornatus	8487, 17036, 17149, 47478	4
	L. arizonicus	7171, 10296, 11048, 11297, 15093, 18621	6
	L. hirsutissimus	5165, 8222, 9435, 9777, 9929, 10365, 15093, 18258, 25495	9
	L. truncetus	3157, 3609, 5165, 8222, 13615, 18621, 24896, 24933	8
Outgroup species	*Crotalaria*	2677, 9777, 9929, 14683, 16571, 16753, 17546, 21485, 24619	9

* Rare mass peaks are those specific to 2-3 species

Table 6: Pairwise distances between different genotypes as analysed by PAUP. The total character differences are shown among the 630 protein mass peaks examined.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	1	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1 *L. angustifolius,* Mandelup	-																																		
2 *L. angustifolius,* P27255	161	-																																	
3 *L. angustifolius,* P22724	154	125	-																																
4 *L. luteus,* Wodjil	213	216	213	-																															
5 *L. luteus,* P26835	187	190	187	134	-																														
6 *L. mutabilis,* P27808	198	181	196	223	179	-																													
7 *L. mutabilis,* P26961	197	196	195	218	182	141	-																												
8 *L. albus,* Kiev Mutant	207	186	187	224	202	189	188	-																											
9 *L. albus subsp albus,*	207	192	197	216	178	193	196	146	-																										
10 *L. albus subsp graecus,*	204	189	192	221	173	192	185	139	153	-																									
11 *L. micranthus,* P22944	177	182	185	238	196	173	170	178	178	195	-																								
12 *L. micranthus,* P26799	184	191	182	233	181	174	167	185	181	192	105	-																							
13 *L. hispanicus subsp hispanicus*	181	182	173	224	176	167	172	192	198	177	128	145	-																						
14 *L. hispanicus subsp bicolour*	208	195	202	227	185	180	177	197	203	196	161	170	115	-																					
15 *L. cosentinii,* P20747	205	200	201	234	210	197	180	200	204	199	150	159	158	183	-																				
16 *L. cosentinii,* P22911	190	181	194	203	181	192	185	185	173	184	137	144	135	168	107	-																			
17 *L. digitatus,* P26877	198	187	196	221	179	172	191	189	189	184	153	168	153	180	153	112	-																		
18 *L. digitatus,* P27045	194	191	204	207	175	174	189	191	191	194	167	172	169	194	181	164	134	-																	
19 *L. princei,* P26878	181	178	173	222	178	165	174	170	180	171	116	135	132	153	136	121	117	135	-																
20 *L. princei,* P26879	191	174	187	226	176	169	176	180	190	177	142	161	140	159	146	117	111	143	86	-															
21 *L. pilosus,* P20955	194	183	188	229	173	176	171	193	183	182	131	160	147	162	151	130	150	162	115	129	-														
22 *L. pilosus,* P23029	200	185	184	215	167	172	175	179	177	184	155	156	145	166	165	142	150	144	135	133	116	-													
23 *L. palaestinus,* P20956	163	162	173	212	158	155	160	174	166	167	106	141	124	141	140	127	125	141	108	110	97	111	-												
24 *L. palaestinus,* P26940	185	174	185	210	158	175	168	190	182	181	136	163	142	165	148	133	143	151	132	130	113	111	88	-											
25 *L. atlanticus,* 27219	188	187	194	235	195	178	175	183	197	192	125	150	133	162	137	122	140	174	111	129	134	152	113	143	-										
26 *L. atlanticus,* 27248	184	181	184	225	191	172	163	169	177	170	127	156	145	162	133	116	132	164	81	103	110	142	109	127	88	-									
27 *L. polyphyllus,* P23299	185	180	195	226	190	163	170	178	184	197	136	149	154	177	164	157	169	175	130	152	145	155	130	158	139	141	-								
28 *L. succulentus,* P27927	194	193	190	247	191	186	181	183	213	198	147	168	151	174	173	158	168	182	141	159	152	164	149	161	162	144	159	-							
29 *L. subcarnosus,* P26969	183	174	187	240	192	177	172	182	188	189	148	157	142	169	174	151	165	175	132	140	153	157	130	154	143	135	118	159	-						
30 *L. ornatus,* P26038	189	166	173	230	188	165	158	166	182	175	130	147	122	153	146	127	149	163	118	122	135	117	106	132	127	119	116	129	98	-					
31 L. hirsutissimus	164	165	176	219	185	164	169	181	189	194	141	156	137	158	161	144	152	162	131	141	148	154	119	145	142	140	137	166	139	139	-				
32 *L. truncetus*	167	174	185	212	180	157	156	168	188	179	124	139	128	149	160	147	147	159	122	138	151	145	104	140	129	129	122	151	132	118	101	-			
33 *L. arizonicus*	168	179	174	231	185	168	163	179	193	180	137	156	135	168	165	148	148	160	129	143	152	146	115	143	146	142	143	160	133	121	118	103	-		
34 *Crotalaria cunninghamii*	194	203	210	245	205	188	185	193	207	196	135	166	143	182	161	150	170	192	129	147	148	162	125	145	138	130	151	170	151	133	144	137	140	-	
35 *Crotalaria eremaea*	191	194	209	230	194	187	172	180	196	187	138	163	148	189	172	147	163	185	132	148	143	157	124	142	137	133	134	163	134	134	147	126	137	111	-

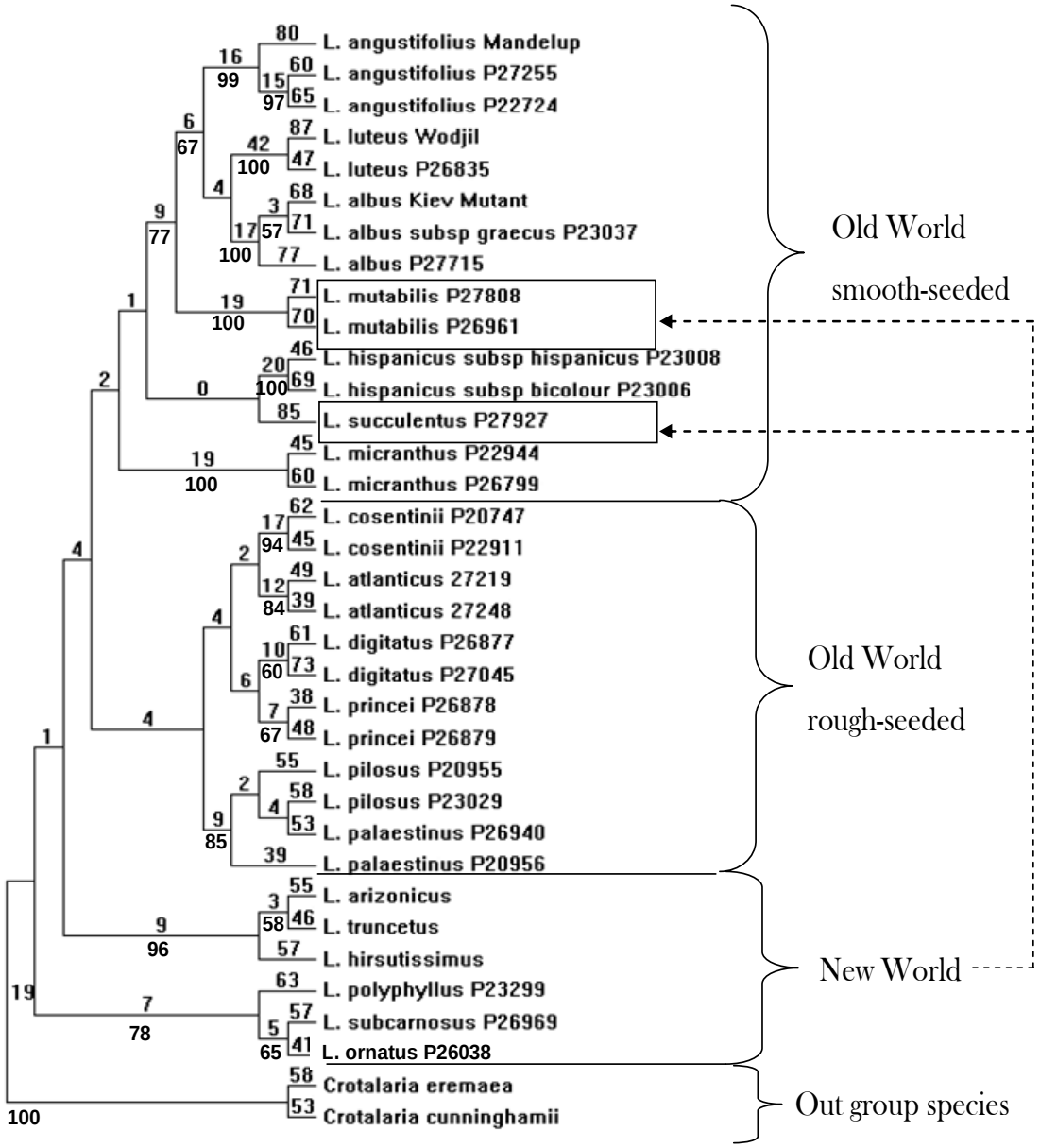


Figure 2: Dendrogram showing relationship among 33 lupin genotypes from 19 species along with two outgroup species the genus *Crotalaria* by unweighted pair-group method with arithmetic averages based on total character differences. Numbers above branches are branch length and those under the branches are bootstrap values based on 10,000 reiterations

Discussion

MALDI-TOF as a proteomic fingerprinting technology

MALDI-TOF- MS produced protein mass peak profiles of 33 lupin genotypes representing 19 species and two *Crotalaria* species as outgroup (Ainouche and Bayer 2000; Ainouche et al. 2004; Ainouche and Bayer 1999). In total, 630 protein mass peaks

have been revealed in this study. The differences among the protein mass peak profiles were clearly visible that made the scoring easy and accurate (**Figure 1**).

Previous studies on seed protein analysis of lupin species were mainly based on limited number of species and by using HPLC or 1D gel electrophoresis (Przybylska and Zimniak-Przybylska 1995; Salmanowicz 1999; Salmanowicz and Przybylska 1994; Zimniak-Przybylska and Przybylska 1997). Our current study included species from both New World and Old World groups (**Table 1**). All the Old World lupin species as recognized by Gladstones (1994) were included in the study except the one rough-seeded species *L. somaliensis* BAKER, which is only known from its type specimen at Kew (Gladstones 1984; Przybylska and Zimniak-Przybylska 1995). On the other hand, eight species including annual and perennial species from New World lupin were included. Multiple genotypes were studied from each of the Old World species and from one New World species *L. mutabilis* (**Table 1**) considering their traditional use and importance as crop for consumption (Gladstones 1998; Petterson 1998). This allowed us to compare the interspecies and intra-species seed protein mass peak diversity in a single frame.

The mass peaks recognized in MALDI-TOF protein profiles may not all represent intact proteins since the molecular weights of intact proteins of lupin are relatively high, ranging from 13 to 430 KDa (Duranti et al. 2008). Mass peaks above 10 KDa might representing subunits of seed proteins (conglutins) or polypeptides (fragment of proteins) (Islam et al. 2011b; Duranti et al. 2008). These subunits generally come up with the proteolytic cleavages of indigenous protein molecules (Muntz et al. 2002). Likewise the mass peaks having lower molecular weights are assumed to represent polypeptides derived from post-translational biochemical processes that lead to proteolysis of lupin seed proteins (Cerletti et al. 1978). To avoid any confusion, we used the term ‘mass peak’ to represent either intact protein or a fragment of protein.

Assessment of seed protein diversity among lupin species

MALDI-TOF based mass spectrometry revealed that the seed protein mass peaks of lupin species are largely polymorphic. All of the recognized 630 mass peaks were polymorphic and approximately 96% mass peaks were highly polymorphic. Only 3.6 % mass peaks were classed as common, appearing in more than ten species (**Table 3**). The comparison of this finding with that of a similar study on different cultivars of *L.*

angustifolius (Islam et al. 2011a) indicates higher mass peak polymorphism among the species (96% highly polymorphic peaks) than within the species (84% highly polymorphic peaks). This polymorphism is attributed to heterogeneity of polypeptides due to the multigenic origin of seed proteins and very distinct post-translational proteolysis of the core protein molecules (Cerletti et al. 1978).

A study on seed globulin and albumin composition of Old World species revealed low occurrence of species specific eceltrophoretic polypeptide bands, particularly among rough-seeded species where only *L. atlanticus* could be distinguished from other (Przybylska and Zimniak-Przybylska 1995; Salmanowicz 1999; Salmanowicz and Przybylska 1994; Zimniak-Przybylska and Przybylska 1997). This study revealed a total of 19 species specific mass peaks, distributed among the12 lupin species (**Table 4**) including four rough-seeded species namely *L. cosentinii, L. digitatus, L. pilosus* and *L. atlanticus*.

To identify the seven species without any species specific mass peak, a set of mass peaks common only to a very few species (2-3) were identified (**Table 5**) and are referred to as rare protein mass peaks in this study. Generally, the Old World smooth-seeded lupin species had more rare protein mass peaks compared to the Old World-rough-seeded species, indicating a higher diversification of seed proteins in the first group. Morphological differentiation and genetic barriers (Gladstones 1974, 1984), alakaloid composition (Nowacki 1963), chemotaxonomic investigations (Williams et al. 1983), isoenzyme phenotypes (Wolko and Weeden 1990a, b) and seed globulins pattern (Salmanowicz 1999) also suggest a higher level of diversity in the smooth-seeded species over the rough-seeded species. A considerable number of rare protein mass peaks (**Table 5**) were common among the species having close relation as defined by morphological and molecular characters (Ainouche et al. 2004; Eastwood et al. 2008; Gladstones 1974, 1984; Naganowska et al. 2003; Talhinhas et al. 2003).

The pairwise distance analysis showed the highest 247 mass peaks difference between the New World lupin species *L. succulentus* (var P27927) and Old World lupin species *L. luteus (*var Wodjil) (**Table 6**) in agreement with large morphological, chemotaxonic and molecular differences between Old World and New World groups (Dunn 1984; Gladstones 1974; Christopher 2008). The lowest 81 character difference was recorded between two Old World rough-seeded species *L. atlanticus* (var 27248) and *L. princei*

(var P26878) in accordance with reported homogeneity of the rough-seeded Old World lupin species based on seed coat texture (Heyn and Herrnstadt 1977), alkaloids (Ainouche et al. 1996; Nowacki 1963; Wink et al. 1995), flavonoids (Williams et al. 1983), seed globulins (Przybylska and Zimniak-Przybylska 1995), protein serology (Cristofolini 1989), and isozymes (Wolko and Weeden 1990a). However, in general 25% to 35% pairwise distance (**Table 6**) suggesting a considerable seed protein mass peak variation among the species. This might be attributed by significant genetic diversity among the different lupin species as revealed by DNA marker based studies (Christopher 2008; Eastwood et al. 2008; Talhinhas et al. 2003). The result also demonstrated that the pairwise distance among the different genotypes within the same species are generally lower (**Table 6**) suggesting less proteomic difference within the species.

MALDI-TOF fingerprinting for phylogenies

The phylogenetic analysis of seed protein firstly clustered the Old World and New World species (with the exception of two New World species) in separate group and then the Old World species were grouped into smooth-seeded and rough-seeded clusters in agreement with the generally accepted classification of lupin species (Dunn 1984; Gladstones 1974, 1984) and chemotaxonomic studies (Williams et al. 1983; Nowacki 1963; Wink et al. 1995; Cristofolini 1989). However, this study showed exceptional clustering of two New World lupin species namely *L. mutabilis* and *L. succulentus* with smooth-seeded Old World species. In a study on nuclear DNA content of lupin species, Naganowska et al. (2003) reported that *L. mutabilis* is closely linked with *L. angustifolius* and positioned together with the other smooth-seeded species. Likewise, a DNA based study (Talhinhas et al. 2003) using AFLP marker (individual DNA) and AFLP+ISSR+RAPD marker suggested a close relation of *L. mutabilis* and *L. albus*.

Seed proteomic data placed species *L. micranthus* at a transitional position between smooth-seeded and rough-seeded group which is in agreement with the analyses based on nuclear DNA content (Naganowska et al. 2003), flavonoids (Williams et al. 1983), protein serology (Cristofolini 1989), and isozymes (Wolko and Weeden 1990a). These biochemical analyses and some morphological characters (Gladstones 1974) indicate the similarity of this species to both smooth and rough-seeded species (Salmanowicz 1999). The intermediate position of *L. micranthus* among Old World lupin species might be due to hybridization and introgression between smooth and rough-seeded

species or reflect a remainder of the transitional lineage from which the smooth-seeded species were derived (Naganowska et al. 2003).

Clustering of all of the six rough-seeded Old World lupin species in a group indicated apparent homogeneity of this group in accordance with series of studies on seed coat texture (Heyn and Herrnstadt 1977), alkaloids (Ainouche et al. 1996; Nowacki 1963; Wink et al. 1995), flavonoids (Williams et al. 1983), seed globulins (Przybylska and Zimniak-Przybylska 1995), protein serology (Cristofolini 1989), and isozymes (Wolko and Weeden 1990a, b). Within the group further clustering of *Lupinus pilosus* and *L. palaestinus* with 85% bootstrap support is supported by cross-compatibilty (Gladstones 1974; Plitmann et al. 1980), internal transcribed spacers (ITS), nrDNA analysis (Ainouche and Bayer 1999), and seed globulin and albumin electrophoreses (Salmanowicz 1999; Salmanowicz and Przybylska 1994). Most of those studies and the chromosome investigation (Gupta et al. 1996) identified *L. princei* as an isolated species of rough-seeded lupin species. However, our study revealed the close relationship of this species with *L. digitatus*. This result is in agreement with the observation by (Kass and Wink 1997) who reported that *L. princei* has ITS sequence identical to that of *L. digitatus*.

Three New World annual lupin species *L. arizonicus*, *L. truncatus* and *L. hirsutissimus* clustered together with 96% bootstrap support in broad agreement with DNA based studies (Eastwood et al. 2008; Christopher 2008). The other annual species *L. subcarnosus* clustered with two perennial species *L. polyphyllus* and *L. ornatus* with 78% bootstrap support. Electrophoreses of seed globulin (Zimniak-Przybylska and Przybylska 1997) showed notable similarity between *L. polyphyllus* and *L. ornatus*. Nevertheless, the lack of phylogenetic data on these species limited the discussion of the relationship revealed by seed protein mass peaks and suggests further investigation would add value.

The results of this study establish the MALDI-TOF based analysis of seed protein as a valuable technology for investigating diversity in lupin species. Generally speaking, the MALDI-TOF based analysis of protein can be more useful when a precise problem can be defined.

Acknowledgements

The authors are grateful to Jon Clements and Nindethana Seed Service Pty Ltd, WA, Australia for providing part of seed materials and thankful to Junhong Ma for her technical support in laboratory works.

Chapter 4

Diversity of seed protein among the Australian narrow-leafed lupin (*Lupinus angustifolius* L.) cultivars

Citation: Islam S, Ma W, Appels R, Buirchell BJ, Ma J, Yan G. 2011. Diversity of seed storage protein among the Australian narrow-leafed lupin (*Lupinus angustifolius* L.) cultivars. *Crop & Pasture Science* 62 (9): 765–775

Abstract

Narrow-leafed lupin (NLL) is one of the major legume crops in Australian farming systems which is largely used as animal feed. Several modern cultivars have been developed through breeding making NLL feasible for use as human food. Significant health benefits have been recognised for NLL. The current study characterised protein polymorphism among 25 Australian cultivars through mass spectrometry (MALDI-TOF) with the aim of developing molecular breeding strategies to improve protein quality and content. A total of 364 seed protein mass peaks were clearly identified by MALDI-TOF and 50 protein mass peaks were cultivar-specific. In addition, 9 protein mass peaks were found present in all cultivars and 61 protein mass peaks present in 2–3 cultivars only. Phylogenic analysis based on the protein profile categorised the cultivars into 2 major groups, which are broadly supported by pedigree information. The low proportion (2.4%) of common protein mass peaks among the cultivars suggested a high level of diversity in seed protein of NLL.

Key words: Protein profiling, diversity, MALDI-TOF, phylogeny, breeding.

Introduction

Lupin is one of the major legume crops in Australian farming systems because of its high nutritional value and adaptability to marginal soils (Dervas et al. 1999; Guemes-Vera et al. 2008; Howieson et al. 2000). The lupin grain is widely used as animal feed due to its high protein content (~32%) and low anti nutritional factors (Hill 1986; Petterson 1998). In addition, the high dietary fibre, low fat content and negligible amount of starch features of lupin flour dictate its suitability as human health food. One of the major cultivated species, *Lupinus angustifolius*, known as narrow-leafed lupin (NLL), is commonly cultivated as a rotation crop especially under Mediterranean climatic environment (Gladstones 1994; Siddique and Sykes 1997). To date, a total of 25 commercial cultivars of *L. angustifolius* have been released in Australia since 1968.

Understanding genetic diversity is the foundation for crop improvement (Talhinhas et al. 2006). Many beneficial genes were discovered in the last 50 years (Cowling et al. 1998; Gladstones 1994) and several genetic diversity studies have been conducted on *L. angustifolius* based on DNA technologies (Yang et al. 2004; Yuan et al. 2005), which showed a high level of genetic variation among cultivars. These studies have assisted

lupin breeding programs in increasing yield, overcoming major diseases and contributing to the agronomic success of the crop (Buirchell 2008). Considerable research on seed proteins has been conducted in different crops (Aly et al. 2000; Ma et al. 2005) including lupin and its legume relatives such as soybean for the purpose of improving protein content and quality (Fahmy and Salama 2002; Hsieh et al. 2001; Stejskal and Griga 1995). However, proteome diversity based on seed protein of *L. angustifolius* has not been reported. Seed protein profiling will be helpful for breeders to develop cultivars enriched for specific proteins. The information can also be used for constructing phylogenetic relationships among cultivars and for fingerprinting cultivars and subsequently assist cultivar identification, breeding line selection, and quality prediction during breeding.

Matrix-assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) has been proved to be a powerful tool for seed protein analysis (Chen et al. 2007; Liu et al. 2009; Muccilli et al. 2005; Alberghina et al. 2005; Cunsolo et al. 2004) with high throughput capability. The MALDITOF- MS technique provides accurate results; requires very small amounts of sample (normally less than 1 pmol), and is relatively fast (requiring only a few minutes per sample) compared with other common separation methods. Moreover, it may facilitate the analysis of proteins from complex mixtures without purification and separation (Kussmann et al. 1997). Thus the technique is particularly suitable for protein analysis of large number of individuals within a short time. Compared with the traditional one- and two-dimensional gel electrophoreses for seed protein identification (Fahmy and Salama 2002; Magni et al. 2007; Yahata et al. 2005) this approach is more time and cost efficient (Barakat 2004). However, MALDI-TOF mass spectrometry has not been widely used to produce the protein profiles of lupin cultivars. The current study investigated the feasibility of MALDI-TOF mass spectrometry to profiling lupin seed proteins in order to study the diversity and to deduce the relationships among the Australian cultivars.

Materials and methods

In total 25 cultivars of *L. angustifolius* released in Australia were supplied by the Department of Agriculture and Food, Western Australia (DAFWA). All the cultivars were grown in the same year at the same experimental station of DAFWA (Wongan Hills, WA) under the same environmental conditions. Thirty grams of seeds containing more than 100 seeds from each cultivar were taken as a working sample and ground in

the 'Retsch 2M 200' and sieved with 750 mM. The details of the cultivars are listed in **Table 1**.

Protein was extracted from lupin flour based on Duranti et al. (2008) and Lampart-Szczapa (1996). Lupin flour samples were defatted by Hexane at 20 : 1 ratio (Santos et al. 1997) and extraction buffer (0.5M NaCl) was added at the ratio of 15ml/g. Protein was extracted by stirring at 48C for 4 h and the supernatant was collected by centrifugation at 10 000*g* for 10 min. The extract was mixed with matrix (sinapinic acid dissolved in 0.05% TFA and 50% ACN) at 1 : 9 ratio and 1 ml of mixture was spotted on MALDI-TOF plate and left at room temperature to dry. Spotting was repeated once when the previous spots were completely dry. The sinapinic acid was purchased from Sigma-Aldrich, St. Louis, MO, USA. Three separate extractions were made for MALDI-TOF protein analysis to make sure of reproducibility.

Table 1: List of narrow-leafed lupin (*Lupinus angustifolius* L) cultivars used in this study.

Series No.	Name of cultivar	Year of release	Series No.	Name of cultivar	Year of release
1	Uniwhite	1967	14	Myallie	1995
2	Uniharvest	1971	15	Kalya	1996
3	Unicrop	1973	16	Wonga	1996
4	Marri	1976	17	Belara	1997
5	Illyarrie	1979	18	Tallerack	1997
6	Yandee	1980	19	Tanjil	1998
7	Chittick	1982	20	Moonah	1998
8	Danja	1986	21	Quilinock	1999
9	Geebung	1987	22	Jindalee	2000
10	Gungurru	1988	23	Mandelup	2004
11	Yorrel	1989	24	Coromup	2006
12	Warrah	1989	25	Jenabillup	2007
13	Merrit	1991			

The experiment was carried out on a Voyager DE PRO Biospectrometry Workstation from PerSeptive Biosystems, Framingham, MA, USA, operated in linear mode (Lou et al. 2010). Final mass spectrum for each sample was obtained by averaging 500 shots on a protein spot over random locations. The machine was calibrated by using 'Sequazyme Peptide Mass Standards Kit' from Applied Biosystems, Foster City, CA, USA following sinapinic acid matrix-calibration mixture 3 as suggested by the supplier. To get the best resolution, the molecular weight range of 2000–32 000 Dalton was split into 3000-Dalton intervals. High molecular weight proteins of 30 000–75 000 Dalton were also analysed.

Data analysis

The results from MALDI-TOF were analysed using the Voyage machine companion software, Data Explorer, to produce the protein mass peak profiles (Liu et al. 2009). The mass spectrometric data were then analysed by using software 'Progenesis PG 600' from Nonlinear Dynamics, Durham, NC, USA. The mass peak profiles were manually checked and the identified polymorphic mass peaks were scored visually for absence and presence. Mass peaks clearly detected in all three replicates were scored to ensure reproducibility. A binary dataset was constructed for multivariate analysis using the software PAUP (Phylogenetic Analysis Using Parsimony) (Swofford 1998). A distance matrix based on total character difference was constructed and the UPGMA (unweighted pair-group method with arithmetic averages) procedure was followed to produce a dendrogram. Bootstrap analysis was carried out with 10 000 replications to assess the reliability of groupings.

Results

MALDI-TOF mass peaks of the seed protein of NLL were clear and easy to score (**Figure 1**). The analysis obtained 364 mass peaks including 355 polymorphic protein peaks ranging from 2 to 60 KDa among the 25 cultivars of NLL. The number of mass peaks identified for each cultivar varied from 88 to 186, demonstrating a high level of proteomic diversity. In total, 58 mass peaks were categorised as very commonly observed in more than 20 cultivars (**Table 2**), accounting for 15.9% of the total mass peaks. Nine mass peaks with molecular weight of 3058, 3103, 4030, 4427, 4575, 4673, 4975, 5850 and 14 642 Da were found to be common to all 25 cultivars, comprising 2.4% of the total profiled mass peaks observed. A total of 50 mass peaks were cultivar-specific (**Table 3**). Eighteen cultivars out of 25 had cultivar-specific mass peaks ranging from 1 to 8. The largest number (8) of cultivar-specific mass peaks was found in the cultivars Coromup and Geebung followed by the cultivars Wonga and Uniharvest. Afew (2–3) cultivar-specific mass peaks were observed in the cultivars Chittick, Gungurru, Mandelup, Tallerack, Danja, Moonah, Marri and Uniwhite. In cultivars Illyarrie, Merrit, Unicrop, Tanjil, Warrah and Yorrel, only one cultivar-specific mass peak was observed. Cultivars Yandee, Myallie, Kalya, Belara, Quilinock, Jindalee and Jenabillup did not show any cultivar-specific mass peaks.

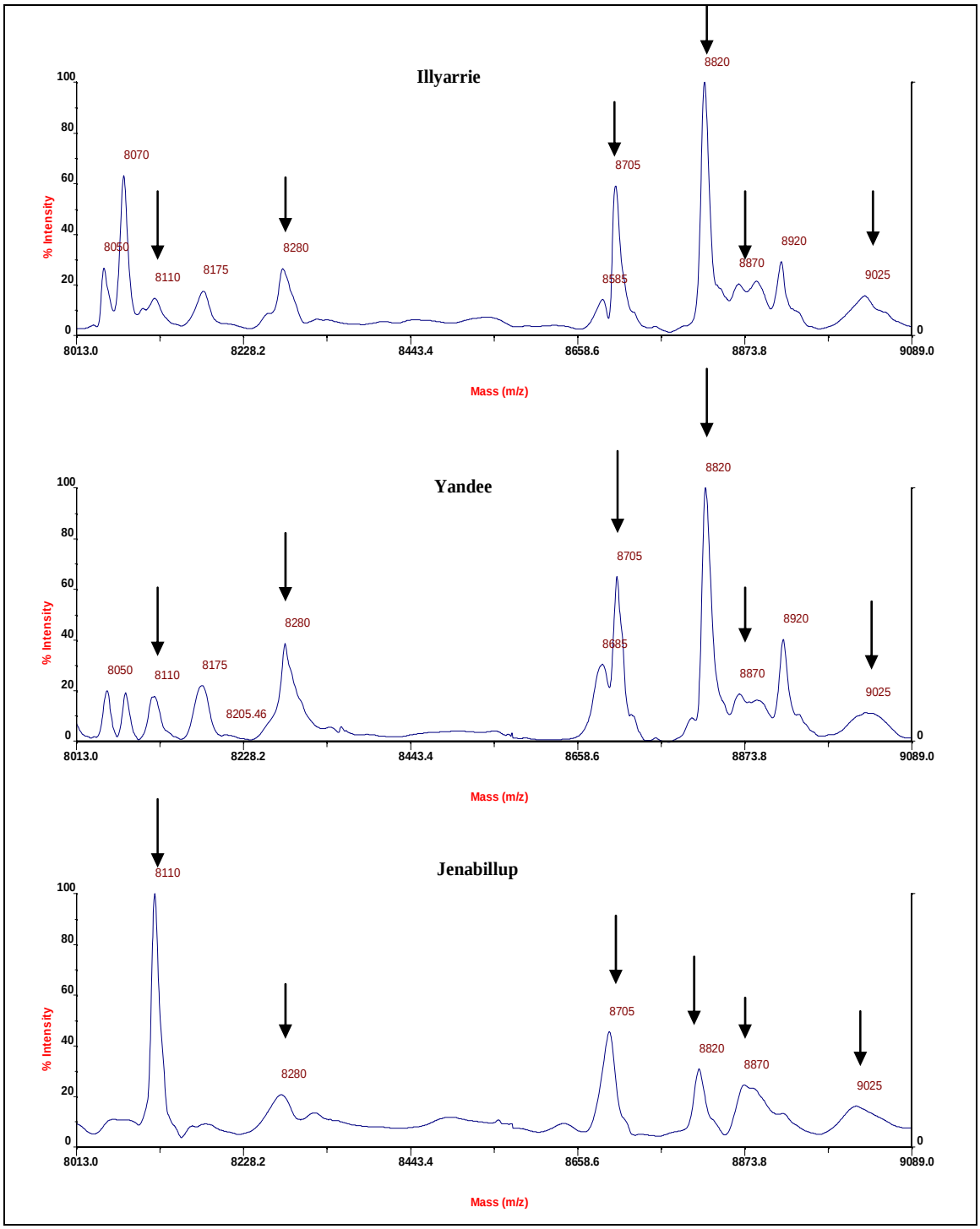


Figure 1: MALDI – TOF outputs of narrow-leafed lupin protein profiles demonstrating easily visible and identifiable polymorphism of protein mass peaks among different cultivars. The numbers on the protein peaks indicate the molecular weight of the corresponding protein in Daltons. Two sister cultivars Illyarrie and Yandee showed very similar protein profiles. The other cultivar Jenabillup showed many common proteins to them (arrowed) but apparently missing some of the proteins.

Table 2: List of the very common mass peaks* for seed protein of narrow-leafed lupin (*Lupinus angustifolius* L) cultivars as identified by mass spectrometry.

Protein Mass peaks (molecular weight in Dalton)	No of cultivars having the mass peak	Name of the cultivars missing the mass peak
2115	22	Marri, Geebung, Wonga
2635	22	Uniharvest, Marri, Coromup
2767	22	Gungurru, Myallie , Coromup
2936	20	Gungurru, Myallie , Wonga, Coromup, Jenabillup
3058	25	-
3086	22	Marri, Geebung, Coromup
3103	25	-
3349	20	Gungurru, Wonga, Jindalee
3440	21	Illyarrie, Wonga, Yorrel, Coromup
3625	23	Geebung, Coromup
3640	20	Marri, Myallie, Kalya, Tallerack, Coromup
3731	20	Marri, Illyarrie, Yandee, Yorrel, Coromup
3931	23	Kalya, Wonga
4030	25	-
4129	24	Coromup
4427	25	-
4575	25	-
4673	25	-
4821	20	Kalya, Wonga, Quilinock, Mandelup, Coromup
4975	25	-
5198	21	Marri, Chittick, Merrit, Wonga
5225	20	Uniwhite, Marri, Geebung, Mandelup, Moonah
5376	21	Myallie, Kalya, Wonga, Jindalee
5395	22	Mandelup, Coromup, Jenabillup
5414	21	Chittick, Geebung, Mandelup, Coromup
5556	20	Geebung, Yorrel, Belara, Mandelup, Coromup
5850	25	-
5917	22	Tanjil, Coromup, Geebung
6036	21	Myallie, Kalya, Wonga, Coromup
6096	23	Quilinock, Jenabillup
6199	22	Myallie, Kalya, Wonga
6484	20	Myallie, Kalya, Wonga, Belara, Tanjil
6545	24	Coromup
6667	24	Coromup
6713	20	Yorrel, Belara, Moonah, Jindalee, Coromup
6868	21	Yorrel, Myallie, Geebung, Coromup
7346	20	Warrah, Myallie, Kalya, Wonga, Tallerack
7408	23	Myallie, Kalya
7656	20	Yorrel, Kalya, Belara
8116	21	Geebung, Kalya, Mandelup, Coromup
8175	22	Kalya, Coromup, Jenabillup,
8280	21	Kalya, Tallerack, Tanjil, Coromup
8705	23	Kalya, Coromup
8820	22	Kalya, Coromup, Tanjil
9377	20	Illyarrie, Kalya, Belara, Mandelup, Coromup
13127	23	Chittick, Coromup
13926	21	Uniharvest, Illyarrie, Chittick, Geebung
14092	23	Yandee, Coromup,
14459	21	Chittick, Gungurru, Kalya, Belara
14642	25	-
15024	24	Chittick
15182	24	Chittick
15332	21	Unicrop, Marri, Illyarrie, Tallerack
15393	24	Wonga
15495	24	Illyarrie
15694	23	Uniwhite, Warrah
16242	23	Kalya, Moonah
21350	20	Gungurru, Yorrel, Warrah, Merrit, Mandelup

* Very common mass peaks are those found in 20 or more cultivars among the 25 studied.

Table 3: List of cultivar-specific mass peaks* for seed protein of narrow-leafed lupin (*Lupinus angustifolius* L) cultivars as revealed by mass spectrometry.

Name of cultivars	Number of cultivar-specific mass peaks	Molecular weight of the protein mass peaks (Dalton)
Coromup	8	8661, 11962, 12132, 12965, 13652, 16320, 18841, 19813
Geebung	8	2258, 2603, 3402, 3536, 3740, 6562, 7640, 13593
Wonga	5	5488, 10728, 16450, 16904, 28603
Uniharvest	4	2462, 40000, 53000, 60000
Chittick	3	2420, 9498, 14878
Gungurru	3	3665, 21121, 22008
Mandelup	3	7524, 13322, 14690
Tallerack	2	4993, 10682
Danja	2	4611, 28436
Moonah	2	17774, 17922
Marri	2	5932, 20002
Uniwhite	2	10632, 28639
Illyarrie	1	3200
Merrit	1	3284
Unicrop	1	24521
Tanjil	1	14055
Warrah	1	17245
Yorrel	1	10423
Yandee	0	-
Myallie	0	-
Kalya	0	-
Belara	0	-
Quilinock	0	-
Jindalee	0	-
Jenabillup	0	-

* Cultivar-specific mass peaks are those specific to a single cultivar

Another 96 mass peaks were specific to a small number (2–5) of cultivars (**Table 4**). The molecular weight of most of these mass peaks ranged from 2 to 15 KDa. Among these 96 peaks 33 were specific to only 2 cultivars. Similarly, the other 28, 25 and 10 mass peaks were specific to 3, 4 and 5 cultivars, respectively. Pairwise differences in mass peaks among the cultivars were analysed by the distance matrix calculated by PAUP (**Table 5**). The pairwise difference ranged from 56 to 186 mass peaks.

The dendrogram produced from the distance matrix showed that there is a considerable level of diversity among the cultivars (**Figure 2**). The dendrogram separated the cultivars into two major groups. The largest group, supported by 88% bootstraps, consisted of 12 cultivars at the upper side of the dendrogram and includes Uniwhite, Illyariie, Yandee, Danja, Marri, Uniharvest, Unicrop, Chittick, Gungurru, Warrah, Merrit and Yorrel. Similarly, cultivars Myallie, Wonga, Kalya, Tallerack, Jenabillup and Tanjil clustered together in the middle part of the dendrogram. Likewise, Belara, Moonah, Jindalee, Quilinock and Mandelup were placed closely. These 11 cultivars

formed a large group with 51% bootstrap supports. In contrast, cultivars Coromup and Geebung were isolated from the two major groups.

Table 4: List of rare mass peaks* for seed protein of narrow-leafed lupin (*Lupinus angustifolius* L) cultivars as observed by mass spectrometry

Protein mass peaks (Molecular weight in Dalton)	Number of cultivars with the mass peak	Name of the specific lupin cultivars with the mass peak
2193	2	Unicrop, Tanjil
2451	3	Kalya, Belara,Moonah
2535	5	Geebung, Kalya, Moonah, Jenabillup, Tanjil
2653	2	Geebung, Quilinock
2838	3	Geebung, Kalya, Moonah
2919	2	Unicrop, Illyarrie
3122	4	Yorrel, Myallie, Belara, Jindalee
3326	4	Geebung, Merrit, Gungurru, Yorrel
3412	2	Marri, Yorrel
3459	3	Jindalee, Tanjil, Jenabillup
3700	5	Danja, Gungurru, Wonga, Tanjil, Unicrop
3797	5	Uniwhite, Yorrel, Belara, Jindalee, Wonga
3849	3	Uniharvest, Tanjil, Jindalee
3869	2	Jindalee, Moonah
3893	2	Uniharvest, Jindalee
3906	5	Uniharvest, Geebung, Illyarrie, Warrah, Tallerack
3943	4	Myallie, Quilinock, Jindalee, Wonga
4015	3	Unicrop, Illyarrie, Merrit
4057	4	Illyarrie, Chittick, Merrit, Marri
4169	4	Chittick, Merrit, Kalya, Jenabillup
4195	3	Marri, Chittick, Danja
4291	2	Geebung, Mandelup
4642	3	Geebung, Kalya, Myallie
4708	3	Belara, Warrah, Merrit
4736	2	Kalya, Wonga
4798	2	Myallie, Tallerack
4891	4	Belara, Quilinock, Mandelup, Jenabillup
4918	2	Quilinock, Mandelup
5285	3	Geebung, Uniwhite, Tallerack
5535	3	Kalya, Quilinock, Jenabillup
5710	4	Geebung, Gungurru, Belara, Quilinock
5823	2	Myallie, Tanjil
5945	2	Uniharvest, Kalya
6082	2	Quilinock, Jenabillup
6348	2	Unicrop, Jenabillup
6425	2	Chittick, Tanjil
6496	4	Uniharvest, Unicrop, Wonga, Jindalee
6634	4	Illyarrie, Belara, Tanjil, Coromup
6900	3	Uniharvest, Marri, Geebung
7167	5	Mandelup, Coromup, Belara, Moonah, Geebung
7255	3	Uniwhite, Unicrop, Jenabillup
7584	2	Belara, Mandelup
7687	5	Uniwhite, Uniharvest, Tanjil, Quilinock, Jindalee
7839	2	Geebung, Myallie
8026	4	Geebung, Moonah, Quilinock, Tanjil
8087	3	Marri, Moonah, Jindalee

Protein mass peaks (Molecular weight in Dalton)	Number of cultivars with the mass peak	Name of the specific lupin cultivars with the mass peak
8165	3	Quilinock, Uniwhite, Danja
8221	3	Moonah, Geebung, Belara
8302	3	Warrah, Merrit, Tanjil
8322	2	Kalya, Myallie
8340	4	Geebung, Quilinock, Mandelup, Coromup
8453	3	Geebung, Mandelup, Coromup
8473	4	Myallie, Tanjil, Moonah, Jindalee
8732	3	Uniharvest, Unicrop, Geebung
9008	2	Uniharvest, Yorrel
9438	3	Marri, Chittick, Danja
9590	5	Belara, Jenabillup, Myallie, Wonga, Marri
10061	2	Yorrel, Belara
10521	2	Myallie, Tanjil
10533	3	Mandelup, Yorrel, Merrit
10666	4	Warrah, Belara, Moonah, Coromup
11214	2	Tanjil, Coromup
11562	4	Mandelup, Coromup, Marri, Geebung
11797	4	Uniharvest, Marri, Illyarrie, Chittick
11898	4	Illyarrie, Yorrel, Warrah, Coromup
12468	4	Danja, Moonah, Quilinock, Jenabillup
12561	5	Myallie, Tanjil, Gungurru, Jindalee, Mandelup
12834	4	Uniharvest, Tallerack, Myallie, Wonga
12858	2	Tanjil, Coromup
13106	2	Unicrop, Chittick
13794	3	Kalya, Quilinock, Coromup
13975	5	Chittick, Danja, Geebung, Gungurru, Warrah
13997	2	Jindalee, Tanjil
15068	3	Tanjil, Uniharvest, Chittick
15349	4	Illyarrie, Unicrop, Marri, Tallerack
15542	3	Unicrop, Tallerack, Uniharvest
16181	2	Jindalee, Moonah
17082	4	Geebung, Quilinock, Mandelup, Jenabillup
17213	3	Geebung, Quilinock, Mandelup
18396	4	Uniharvest, Marri, Yandee, Mandelup
18496	3	Belara, Uniharvest, Chittick
19177	4	Uniwhite, Uniharvest, Danja, Geebung
19471	2	Geebung, Mandelup
19750	4	Chittick, Wonga, Illyarrie, Belara
21010	3	Marri, Tallerack, Jenabillup
21330	4	Yorrel, Warrah, Gungurru, Mandelup
21565	4	Unicrop, Illyarrie, Yandee, Mandelup
23000	2	Uniwhite, Uniharvest
23777	2	Yorrel, Uniharvest
24289	5	Uniharvest, Unicrop, Yandee, Geebung, Kalya
28378	2	Warrah, Uniharvest
28523	3	Illyarrie, Chittick, Geebung
28698	2	Kalya, Myallie
28803	2	Kalya, Wonga
31000	3	Yandee, Chittick, Merrit
35000	2	Merrit, Yorrel

* Rare mass peaks are those specific to 2-5 cultivars.

Table 5: Pairwise distance between different cultivars of narrow-leafed lupin (*Lupinus angustifolius* L) as analysed by PAUP. The total character differences are shown among the 364 protein mass peaks examined.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	Uniwhite	-																								
2	Uniharvest	85	-																							
3	Unicrop	71	68	-																						
4	Marri	86	109	95	-																					
5	Illyarrie	75	90	80	83	-																				
6	Yandee	74	81	67	62	63	-																			
7	Chittick	88	105	99	100	85	88	-																		
8	Danja	72	89	75	84	81	56	76	-																	
9	Geebung	137	148	144	141	150	143	135	131	-																
10	Gungurru	91	104	88	107	106	85	95	81	124	-															
11	Yorrel	82	105	99	112	89	86	102	84	139	97	-														
12	Warrah	71	90	80	99	84	75	87	71	128	74	77	-													
13	Merrit	81	98	82	99	80	77	77	83	130	80	87	60	-												
14	Myallie	126	145	125	124	133	118	142	116	137	111	132	121	127	-											
15	Kalya	136	155	137	136	147	130	136	130	129	123	144	125	127	80	-										
16	Wonga	123	130	118	129	134	121	135	123	146	128	137	122	124	75	81	-									
17	Belara	117	142	134	125	120	109	123	111	138	120	99	126	116	119	123	112	-								
18	Tallerack	93	118	106	101	112	97	109	97	132	104	117	96	114	83	85	86	106	-							
19	Tanjil	126	145	131	138	137	122	132	120	145	119	136	127	129	98	104	101	97	89	-						
20	Moonah	119	136	134	127	118	115	123	115	124	118	109	118	114	113	115	122	92	104	95	-					
21	Quilinock	105	124	116	125	112	105	115	103	124	104	129	104	100	109	111	108	104	102	97	80	-				
22	Jindalee	118	131	125	118	129	98	124	106	135	97	120	111	115	92	110	103	99	97	86	69	79	-			
23	Mandelup	133	150	142	133	132	115	133	121	104	112	107	116	120	131	125	140	100	114	123	90	98	87	-		
24	Coromup	153	186	180	137	148	135	161	155	130	144	133	158	160	135	123	146	108	128	125	122	136	125	104	-	
25	Jenabillup	108	125	107	114	119	104	112	94	139	105	128	101	99	88	90	87	115	81	92	105	91	98	117	147	-

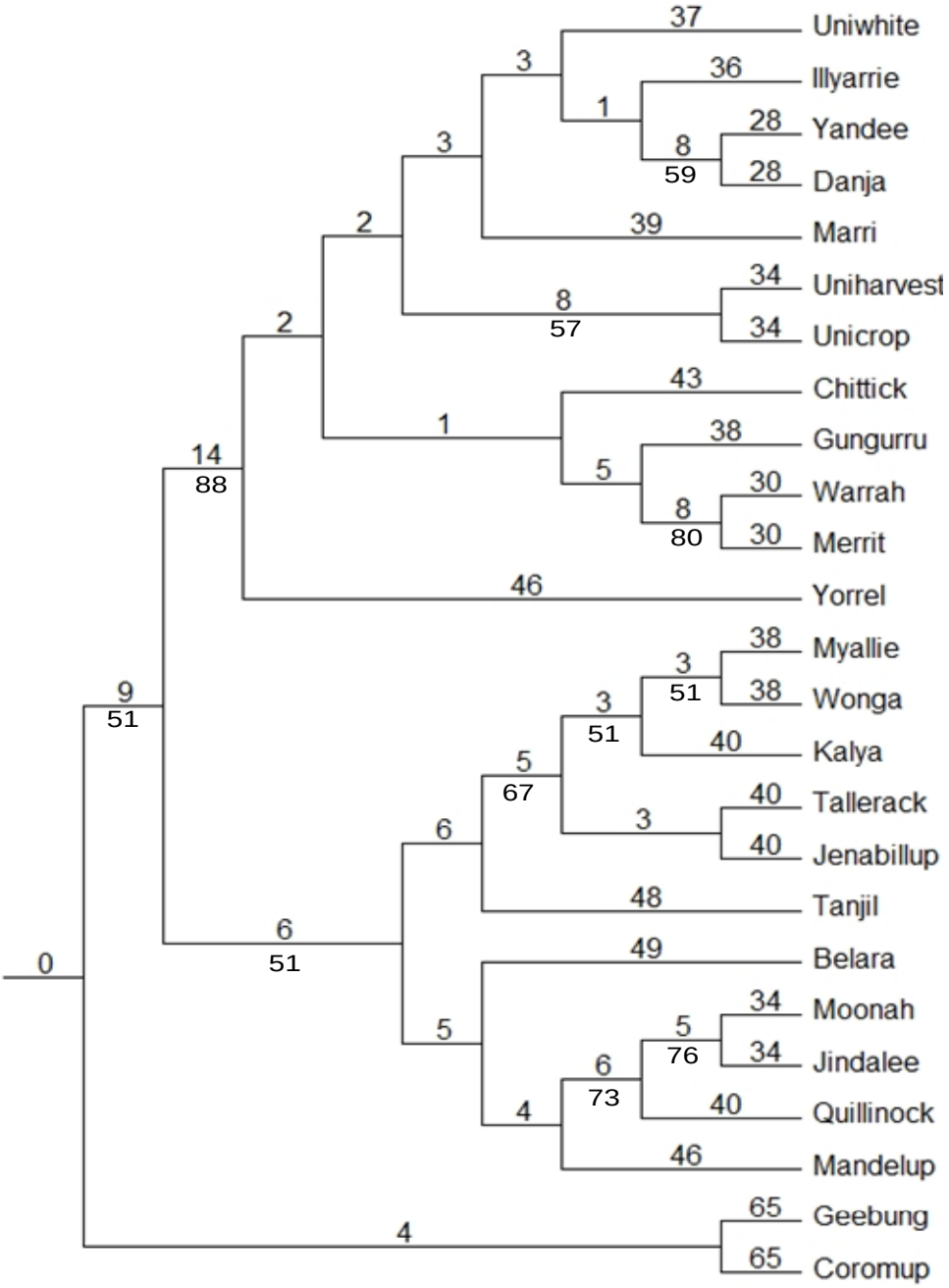


Figure 2: Dendrogram showing relationship among 25 narrow-leafed lupin (*Lupinus angustifolius*) cultivars by unweighted pair-group method with arithmetic averages based on total character differences. Numbers above branches are branch length and those under the branches are bootstrap values based on 10,000 reiterations.

Discussion

MALDI-TOF-MS has been shown to be very useful in detecting protein mass peak profiles effectively in NLL. The reproducibility of repeated extraction and detection of the same sample is generally very good. From the three replicated experiments for each sample, almost all the detected peaks are present in all the replicates. Only a very small number of weak peaks were not reproducible, which were not included in the analysis. In total, 364 protein mass peaks have been revealed in this study. Visible differences among the protein mass peak profiles made the scoring easy and accurate (**Figure 1**). High throughput results suggested that the method is suitable to study a large number of genotypes within a short period of time.

All the mass peaks identified through MALDI-TOF may not represent intact proteins as the molecular weights of intact proteins of lupin are relatively high (Duranti et al. 2008). Mass peaks above 10 KDa might represent subunits of NLL seed proteins (conglutins) or polypeptides (fragment of proteins), as identified by 2D gel electrophoresis in another study by our research group (Islam et al. 2011b). Identified subunits of seed proteins (conglutins) of *Lupinus albus* also have been reported with a molecular weight of 9 KDa or above (Duranti et al. 2008). These subunits generally arise from the proteolytic cleavages of native protein molecules (Derbyshire et al. 1976; Muntz et al. 2002). Likewise the mass peaks having lower molecular weights are assumed as representing polypeptides derived from post-translational biochemical processes that lead to proteolysis of NLL seed proteins (Cerletti et al. 1978).To avoid any confusion, we used the term ‘mass peak’ that represent the intact protein or a fragment of protein, which are useful for fingerprinting (Horneffer et al. 2007). Theoretically, the detected mass peaks are confounded outcome of cultivar and cultivar-by-site interactions as all the samples were from the same site, same year and same growing condition although we envisage that most of the variation was due to difference among cultivars.

This study revealed that the seed storage protein mass peaks of NLL are highly polymorphic. Over 97% mass peaks were found to be polymorphic as only 2.4% of the observed peaks were common to every genotype. This polymorphism is attributed to heterogeneity of polypeptides due to the multigenic origin of seed proteins and very distinct post-translational proteolysis of the core protein molecules (Cerletti et al. 1978). Among the polymorphic mass peaks, 15.9% (of total) was common to a minimum of 20

genotypes, which can be considered as generally common seed protein mass peaks of NLL. Thus 81% of NLL seed mass peaks were highly polymorphic, which might be useful for cultivar identification based on protein profiling.

A total of 50 cultivar-specific mass peaks have been identified among the 25 NLL cultivars. Out of the 18 cultivars having cultivar-specific mass peaks, 12 had multiple cultivar-specific mass peaks. These cultivar-specific mass peaks can be used for the identification of corresponding cultivars (**Table 3**). To identify the 7 cultivars without any cultivar-specific mass peaks, the set of mass peaks specific to very few cultivars (2–5) can be used (**Table 4**). Among the observed 96 rare mass peaks, 61 were specific to only 2–3 cultivars. Thus the combined information is useful for the identification of all 25 NLL cultivars.

Existence of most of the rare mass peaks in the cultivars is generally in agreement with their pedigree. For example one of the early cultivars Illyarrie had four rare mass peaks common to one parent Unicrop and three common to another parent Marri. Likewise the latest cultivar Jenabillup had five mass peaks common to its one parental cultivar Quilinock. On the other hand, few mass peak profiles were not in agreement with the pedigree links suggesting the protein components came through hybridisation. Geebung showed seven rare mass peaks common to Mandelup although their pedigree showed a far distant relationship. Cultivar Illyarrie had one cultivar-specific mass peak at 3200 Da (**Table 3**) and another five rare mass peaks at 3906, 6634, 11 898, 19 750 and 28 523 Da (**Table 4**), which were not present in either of its parental cultivars Unicrop or Marri.

Mass peak profiling of NLL seed protein demonstrated extensive diversity among the studied cultivars. The number of mass peaks identified for each cultivar varied from 88 to 186 indicating noteworthy qualitative differences among the cultivars. Likewise, the pairwise distance analysis showed the highest 51% distance (mean character difference) between one of the earliest cultivars Uniharvest and one of the latest cultivars Coromup (**Table 5**). The lowest 15% distance was between the cultivars Yandee and Danja originating from the same parent Unicrop. However, most of the pairwise distances ranged from 25 to 40% (**Table 5**), suggesting a considerable seed protein variation. The significant genetic diversity among the cultivars of *L. angustifolius* has previously been

revealed by DNA marker-based studies (Talhinhas et al. 2006; Yang et al. 2004; Yang et al. 2001; Yuan et al. 2005).

The constructed phylogenetic tree (**Figure 2**) based on the polymorphic seed protein mass peaks is in broad agreement with the pedigree information and DNA marker-based studies (Talhinhas et al. 2006; Yuan et al. 2005) of the NLL cultivars. Cultivars having common parental lines were grouped together as the seed proteins are generally genetically inherited (Bolon et al. 2010; Timmerman-Vaughan et al. 2005). Cultivars Warrah and Merrit derived from the cross of Ilyarrie with wild types from Spain were placed together closely with 80% bootstrap support. Likewise, cultivars Yandee and Danja sharing the same parent Unicrop were placed together with 59% bootstrap support. One of the earliest cultivars Unicrop released by the University of Western Australia clustered with its parental cultivar Uniharvest with 57% bootstrap support, which is in agreement with the findings of Yuan et al. (2005). Including the abovementioned cultivars, a total of 12 cultivars of NLL originating from a series of early crosses (until1991) were clustered together in a group with 88% bootstrap support. This finding is generally supported by the randomly amplified microsatellite polymorphism (RAMP)-based fingerprinting (Yuan et al. 2005) and by an amplified fragment length polymorphism (AFLP) and inter-simple sequence repeat (ISSR) marker-based study (Talhinhas et al. 2006).

Likewise, cvv. Moonah and Quilinock sharing the wild type from Italy were placed in a group with 73% bootstrap support. The other cultivar of this group Jindalee paired with Moonah with 76% bootstrap support in agreement with their pedigree as shared parental line from Gungurru. Four cultivars Myallie, Tallerack, Wonga and Kalya originating from a series of complex crosses incorporating the cultivar Illyarrie and different wild types from Spain and Morocco were clustered together with 67% bootstrap support. The abovementioned 7 cultivars of NLL along with the cultivars Tanjil, Belara, Jenabillup and Mandelup formed the other cluster with 51% bootstrap support. This finding is generally in agreement with the result of DNA-based fingerprinting of NLL by Yuan et al. (2005).

Some exceptions in grouping in relation to the pedigree links suggest that the seed protein diversity among the cultivars might be introduced beyond their pedigree relationship. The latest (2007) cultivar Jenabillup was placed in the group of Myallie,

Tallerack, Wonga and Kalya with 67% bootstrap support but it shares a very different pedigree relationship with other cultivars. Geebung was not placed closely with parental cultivars Uniwhite and Uniharvest and the case was similar in the DNA-based study (Yuan et al. 2005). These exceptional placements might be due to heterogeneity of polypeptides (Cerletti et al. 1978) and also suggest that new proteins different from the parents might be formed in hybrids through hybridisation. The isolated position of the cultivar Coromup (**Figure 2**) suggests further investigation of this cultivar is warranted using alternative proteomic approaches.

For further comparison of groupings of NLL based on protein mass peak profiles with pedigree consequence, a most parsimonious tree was constructed (not presented here) taking the earliest cultivar Uniwhite as outgroup. The groupings were very similar to the UPGMA tree and also broadly support the pedigree relationships as demonstrated by (Cowling 1999). However, the limitation of this method is that it is based on qualitative approach (Dekker et al. 2005; Szajli et al. 2008). Thus where the protein quantity is a major concern, this method has to be used in combination with other quantitative methodologies such as NIR, Kjeldahl, Biuret, and Combustion etc. (Moore et al. 2010).

Acknowledgements

The authors are grateful to Sophie Sipsas for providing lupin seed materials.

Chapter 5

Environment and genetic interaction of seed storage proteins in narrow-leafed lupin (*Lupinus angustifolius*)

Citation: Islam S, Ma W, Appels R, Buirchell BJ, Yan G. Environment and genetic interaction of seed storage proteins in narrow-leafed lupin. *Crop & Pasture Science* (submitted)

Abstract

This study investigated the environmental influence on seed protein profiles of five narrow-leafed lupin cultivars grown under three different environmental conditions. High throughput MALDI-TOF mass spectrometry revealed 133 reproducible seed protein mass peaks. Thirty-one seed protein mass peaks were expressed in all 15 combinations of cultivar × environment. Expression of 20 mass peaks was influenced by cultivar irrespective of environment. Only six protein mass peaks were influenced by the environment. The expressions of 76 mass peaks were highly variable. Statistic analysis of the number of mass peaks (NMP) indicates that lupin seed protein mass peak expression is mostly genetically controlled ($p=0.008$) with no significant influence of the environment ($p=0.131$). Environment and cultivar interactions were not significant ($p=0.889$). Multivariate analyses of mass peaks supported the above analysis showing that protein mass peak expression was significantly ($p=0.001$) influenced by cultivar but not by environment ($p=0.053$). This result indicates the importance of breeding new lupin cultivars targeting specific proteins for human food and animal feed without being too concerned about environmental influences.

Keywords: Lupin, environment, cultivar, seed protein, MALDI-TOF, mass peak expression

Introduction

Lupin (*Lupinus* spp) is a legume crop which has gained increased attention in the past two decades in many countries (FAO 2011) because of its excellent nutritional attributes. Due to its high protein content, lupin grain is widely used as animal feed and food, or for food ingredients for human consumption. The nutritional properties and lower levels of anti-nutritional factors in lupin are particularly attractive for human consumption (Petterson et al. 1997). Lupin seed protein is considered as the main contributor for several health benefits (Lee et al. 2009; Lee et al. 2006) such as increasing satiety, reducing energy intake, decreasing blood pressure and blood glucose level. Lupin is well adapted to extreme climatic conditions (Dervas et al. 1999; Guemes-Vera et al. 2008; Howieson et al. 2000). However, it is possible that environmental variation may influence grain composition, especially seed storage proteins.

Grain quality, particularly protein content and the spectrum of protein types, is a vital issue to the miller and baker. Seed storage protein accumulation and their post translation modification are expected to be influenced by environmental variation (DuPont and Altenbach 2003; Panozzo and Eagles 2000). Field studies have reported the effect of environmental variation on composition and/or polymerisation of seed storage proteins (Huebner et al. 1997; Triboi et al. 2000; Zhu and Khan 2001). In particular, the nutrient composition of soil, temperature and rainfall (rate and pattern) are crucial for seed protein synthesis. The effect of temperature on seed storage protein composition is not well defined and may vary between genotypes (DuPont and Altenbach 2003). Protein content of grain is affected by temperature and may or may not lead to differences in grain protein composition and flour quality (Randall and Moss 1990; Daniel and Triboi 2000). In wheat, the protein content of grain is dependent on the fertiliser status of the soil. Under low levels of fertiliser application, protein content increased with increasing temperature, whereas with adequate fertiliser application, temperature variation did not have any significant influence on protein content (DuPont and Altenbach 2003; DuPont et al. 2000). Water availability is also crucial at both vegetative and productive stages for grain development with a possible effect on seed storage protein quality (Boutraa and Sanders 2001). It is interesting that in the case of lupin, water stress did not appear to have any significant effect on seed protein content (Carvalho et al. 2005).

Although seed storage protein attributes of grain are largely genetically inherent, storage protein gene expressions might be influenced by environmental factors (DuPont and Altenbach 2003). Wheat gluten gene expression was influenced by temperature variation at grain filling (Altenbach et al. 2002). Likewise, seed storage protein gene expression was influenced by soil nutrients in barley (Muller and Knudsen 1993) and maize (Muller et al. 1997). In most cases, transcriptional regulations by environmental factors play a major role in proteome variation of the grain. The comparisons of seed storage protein profiles under different environmental conditions would be expected to assist in revealing the basic molecular mechanisms that are influenced by the environment and that affect productivity and quality.

Seed storage protein variation across different genotypes of grain crops is well documented including lupin (Islam et al. 2011a), soybean (Murphy and Resurreccion 1984; Yaklich 2000), wheat (Skylas et al. 2001; Yahata et al. 2005) and peanut

(Kottapalli et al. 2008). Previous studies on the influence of environmental variation on lupin seed mainly focused on protein content and amino acid composition (Bhardwaj et al. 1998; Carvalho et al. 2005; Cowling and Tarr 2004). However, the qualitative variation in lupin seed storage proteins as affected by environment has not been reported. Development of mass spectrometric technology enabled more efficient and cost-effective analysis of seed proteins compare to traditional PAGE-based methods. This study used MALDI-TOF mass spectrometry for high throughput analysis of seed proteins to study genetic environment interactions of lupin seed storage proteins.

Materials and methods

Cultivars: Five cultivars of narrow-leafed lupin released in Australia were used in this study: Belara, Coromup, Jenabillup, Mandelup and Tanjil. The cultivars were selected based on their potential cultivation in Western Australia. Seeds were supplied by the Department of Agriculture and Food, Western Australia (DAFWA).

Locations/environmental variation: Each cultivar was grown at three field experimental stations in Western Australia located at (1) Lake Varley, 70 km south east of Hyden in the eastern wheat-belt with comparatively low rainfall; (2) Valentine Road, 40 km east of Geraldton on the Eradu sandplain with medium to high rainfall and warmer winters than other sites; and (3) Wongan Hills, 180 km north east of Perth in the medium rainfall area (**Figure 1**). Details of the climate of the three locations in the experimental period are in **Appendix I**. In each location, each cultivar was grown in three different plots as three replications. Thus the sample size of the experiment was: Cultivars × Locations × Replications =5 × 3 × 3 = 45.

Sample preparation: Ten randomly-selected plants were harvested from the experimental field, seeds removed and mixed thoroughly. Thirty grams of seed containing more than 100 seeds from each replication was taken as a working sample, ground in the Retsch 2M 200 and sieved at 750 µM.

Protein analysis: Protein was extracted from lupin flour based on a previously established method (Duranti et al. 2008; Islam et al. 2011a). Flour samples were defatted with hexane at 20:1 ratio (Santos et al. 1997). Protein was extracted in extraction buffer (0.5 M NaCl) at a ratio of 15 ml/g by stirring at 4^{o}C for 4 h; supernatant was collected by centrifugation at 10,000 g for 10 mins. The extract was

mixed with matrix (sinapinic acid dissolved in 0.05% trifluoroacetic acid and 50% acetonitrile) at 1:9 ratio and 1 µl of mixture was spotted on MALDI-TOF plate and left at room temperature to dry. Spotting was repeated once when the previous spots were completely dry. The sinapinic acid was purchased from Sigma-Aldrich, St. Louis, MO, USA. Three separate extractions were made for MALDI-TOF protein analysis for reproducibility.

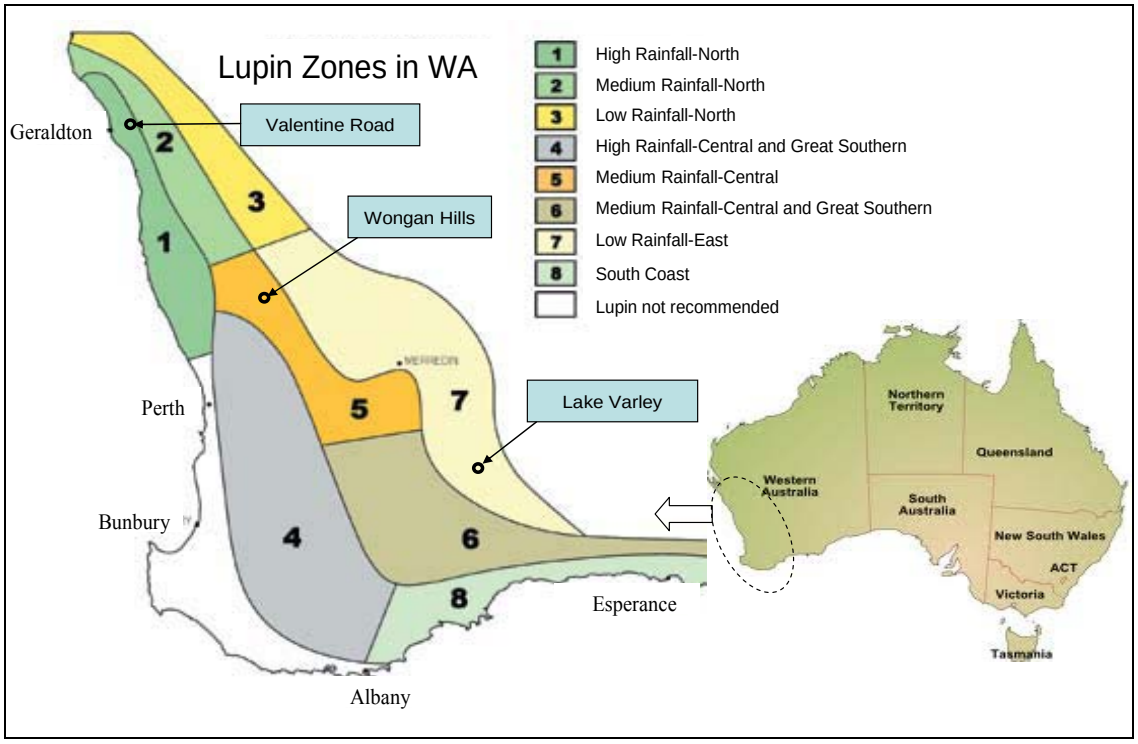


Figure 1: Location of experimental sites in Western Australia.

The experiments were carried out on a Voyager DE PRO Biospectrometry Workstation from PerSpective Biosystem, Framingham, MA, USA, operated in linear mode (Lou et al. 2010). Final mass spectrum for each sample was obtained by averaging 500 shots on a protein spot over random locations. The machine was calibrated by using 'Sequazyme Peptide Mass Standards Kit' from Applied Biosystem, Foster City, USA following sinapinic acid matrix-calibration mixture 3 as suggested by the supplier. For the best resolution, a molecular weight range of 2000–32 000 Dalton was split into 3000 Dalton intervals. High molecular weight proteins (30 000–75 000 Dalton) were also analysed.

Data analysis

The results from MALDI-TOF were analysed using the Voyage machine companion software, Data Explorer, to produce the protein mass peak profiles (Liu et al. 2009). The mass spectrometric data were then analysed using Progenesis PG 600 software from Nonlinear Dynamics, Durham, NC, USA. The mass peak profiles were manually

checked and the clearly identified polymorphic mass peaks were scored visually for absence or presence. Mass peaks were qualitatively measured based on peak intensity and peak area to differentiate them from the base line noises. Essentially, only prominent peaks were used in the analysis and non-prominent ones were ignored. A binary dataset was constructed for multivariate analysis.

Comparisons of protein mass peaks expression at cultivar and environment levels were tested using a permutation-based hypothesis testing, Analysis of Similarities (ANOSIM) was conducted in PRIMER v 6 (Clarke and Gorley 2005) from the binary data set. An overall R statistic was generated that was on a scale from 0 or negative value (identical) to 1 (dissimilar). Non-metric multidimensional scaling (MDS) plot was used to explore the relationships among cultivars and environments based on the Bray-Curtis similarity matrix. Number of mass peaks (NMP) was calculated across seed protein profiles to quantify the environment × cultivar influence. NMP was calculated as the average number of clearly visible mass peaks. ANOVA was performed using Minitab version 14 (Minitab Pty Ltd, Sydeny, NSW, Australia).

Results

MALDI-TOF mass spectrometry revealed 133 reproducible mass peaks. The peaks were clear and reproducible across the mass spectrometric replications. Mass peak similarity and NMP were analysed across cultivars and environments.

Differential expression of protein mass peaks

Protein mass peaks, reproducible between replications at every site, were considered to visualise the pattern of mass peak expression. Thirty-one protein mass peaks were expressed consistently among the entire set of cultivars and environmental conditions (**Table 1**). Expressions of 20 mass peaks were influenced by cultivar irrespective of environment. In contrast, expression of six mass peaks was changeable among environments irrespective of cultivar. The expressions of 76 mass peaks were highly variable across the combinations of cultivar and environments. Of these, 8 protein mass peaks were recognised in 11–14 of 15 environment and cultivar combinations (**Table 1**), termed as generally common mass peaks in this study. Another 17 protein mass peaks were expressed in a single combination (only in a single cultivar at a single environment) and 40 mass peaks were expressed in 2–5 combinations (**Table 1**).

Table 1: Overall expression pattern of lupin seed protein mass peaks as influenced by environment and genetic interactions

Protein mass peak category according to expression pattern across the samples	Number of combinations* showing protein expression (/15)	Lupin seed protein mass peaks (present in Dalton)	Total number of mass peaks
Mass peaks consistently expressed at entire set of locations and cultivars	15	2121, 2532, 2758, 2930, 2954, 3101, 3928, 4021, 4128, 4572, 4670, 5210, 6091, 6646, 6846, 8078, 8096, 8471, 8681, 9006, 9136, 9521, 12835, 13012, 13136, 14811, 14985, 15158, 15351, 15452, 15650	31
Mass peaks influenced by cultivar	NA	3515, 3726, 3795, 5195, 5987, 6030, 6108, 6195, 6692, 13470, 13711, 13870, 13907, 14234, 14295, 14408, 14460, 14560, 14646, 14845	20
Mass peaks influenced by environment	NA	3491, 7334, 17111, 17210, 20843, 21384	6
Mass peaks with variable expression across combinations of cultivar × environment			
a) Generally common mass peaks	11–14	3090, 5855, 9365, 9473, 9570, 9890, 10699, 19450	8
b) Irregular mass peaks	6–10	2155, 2910, 4147, 4921, 5542,.5826, 6537, 8702, 9101, 9595, 14258	11
c) Generally rare mass peaks	2–5	2342, 2396, 2418, 2663, 3245, 3351, 4240, 4902, 4963, 5425, 5522, 5552, 5684, 5884, 6131, 6175, 6267, 6303, 6410, 7674, 7696, 8310, 8558, 8625, 8654, 8795, 9610, 10626, 11297, 12440, 13421, 13755, 14600, 15807, 16620, 20976, 21622, 24255, 24291, 24457	40
d) Very rare mass peaks	1	2585, 3646, 4214, 4478, 4511, 5473, 5705, 7621, 8363, 9667, 11826, 12307, 12687, 14765, 15034, 24364, 24619	17

*Total combination means; number of locations (3) × number of cultivars (5) = 15

NA: Not applicable

Influence of cultivar on protein mass peak expression

Analysis of variance (ANOVA) on NMP showed significant (p=0.008) influence of cultivar on lupin seed protein mass peak expression (**Table 2**). Non-metric multidimensional scaling (MDS) ordination of protein mass peak profiles showed significant (1-way ANOSIM, R=0.58, p=0.001) differences among cultivars (**Figure 2 A**) based on Bray-Curtis similarity matrix. All possible pairwise comparisons had significant differences among cultivars at 1% level of significance (**Table 3**). Comparison of mass peak profiles recognised 20 protein mass peaks influenced by cultivar (**Table 1**).

Table 2: Analysis of Variance (ANOVA) showing influence of cultivar and environment on number of mass peaks (NMP) for lupin seed protein.

Source	df	Seq SS	Adj SS	Adj MS	F	P
Replication	2	116.0	116.0	58.0	0.29	0.752
Environment	2	879.2	879.2	439.6	2.18	0.131
Cultivar	4	3456.6	3456.6	864.1	4.29	0.008
Environment × cultivar	8	701.4	701.4	87.7	0.44	0.889
Error	28	5637.3	5637.3	201.3		
Total	44	10 790.6				

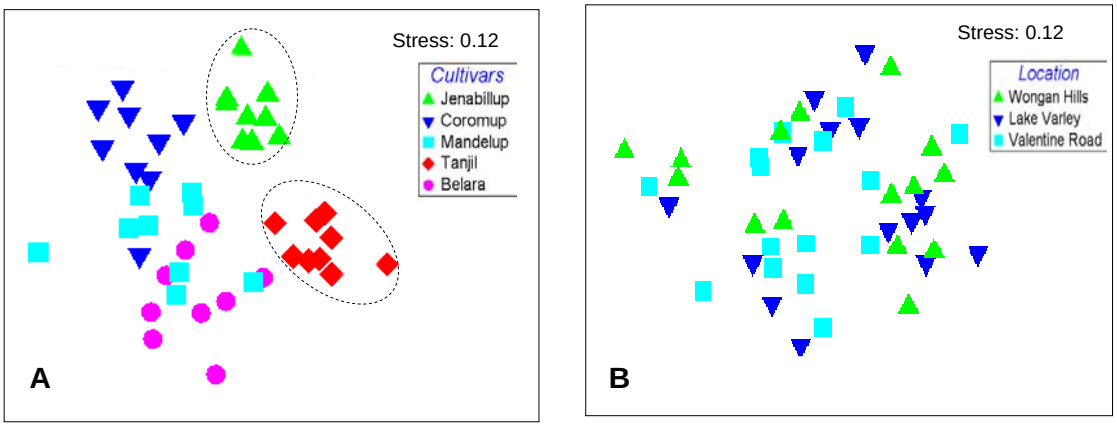


Figure 2: Non-metric multidimensional scaling (MDS) plot based on Bray-Curtis similarity matrix of protein mass peaks profiles showing relationships among samples at cultivar (Plate A) and environment (Plate B) levels as analysed by PRIMER 6. Plate A shows less similarity between cultivars (higher similarity within cultivars); circles indicate clear separation of certain cultivars. Plate B indicates higher similarities across environments.

Effect of environmental variation on protein mass peak expression

Variation in number of mass peaks (NMP) across different locations was not significant (p=0.131) as identified by Analysis of Variance (**Table 2**). Non-metric multidimensional scaling (MDS) plot based on Bray-Curtis similarity matrix of protein mass peak profiles showed that differences between environments (**Figure 2 B**) were not significant (1-way ANOSIM, R=0.07, p=0.053). All possible pairwise comparisons showed non-significant differences among environments at 1% level of significance (**Table 3**). Nevertheless, analysis of seed protein profiles across samples showed expressions of certain mass peaks within a cultivar were influenced by environmental variation. This study found six protein mass peaks (**Table 1**) influenced by environmental variations. **Figure 3** shows the expressional variation of mass peaks 17,111 and 17,210 Dalton in the cultivar Coromup due to variation in environmental conditions.

Table 3: Result of the analysis of similarities (ANOSIM) based on lupin seed protein mass peak profiles using PRIMER 6. All possible pairwise comparisons across cultivars and locations are presented sequentially. Scale of R values is from 0 or negative values (identical) to 1 (dissimilar) and p indicates level of significance.

Groups	R value	p value
Cultivars	**0.580**	**0.001**
Jenabillup–Coromup	0.678	0.001
Jenabillup–Mandelup	0.662	0.001
Jenabillup–Tanjil	0.606	0.001
Jenabillup–Belara	0.812	0.001
Coromup–Mandelup	0.401	0.002
Coromup–Tanjil	0.754	0.001
Coromup–Belara	0.637	0.001
Mandelup–Tanjil	0.519	0.004
Mandelup–Belara	0.346	0.004
Tanjil–Belara	0.496	0.003
Environments	**0.070**	**0.053**
Wongan Hills–Lake Varley	0.030	0.230
Wongan Hills–Valentine Road	0.054	0.132
Lake Varley–Valentine Road	0.129	0.017

Cultivar × environment interactions were not significant (p=0.889) for number of mass peaks (NMP). However, comparative analysis of lupin seed protein profiles recognised 76 protein mass peaks with variable expression across the combinations of cultivar and

environment apparently indicating influence of cultivar × environment interactions (**Table 1**).

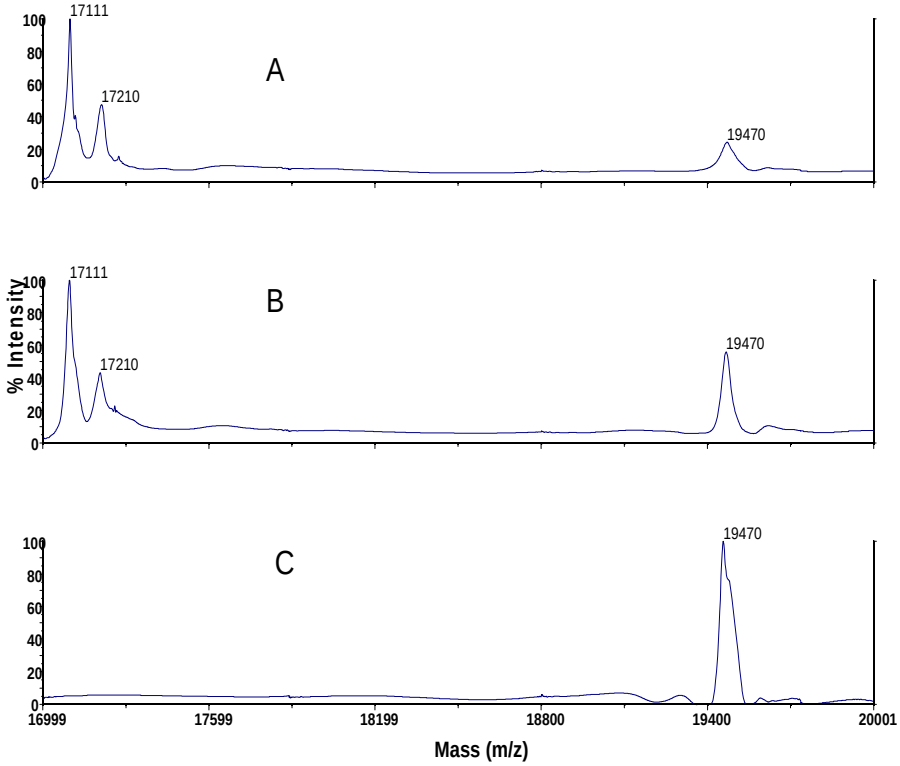


Figure 3: Some protein expressions influenced by environment. The figure shows that mass peak 17,111 and 17,210 in cultivar Coromup are expressed at A: Wongan Hills and B: Lake Varley but not at C: Valentine Road.

Discussion

This study demonstrated the use of the MALDI-TOF method to analyse environment × genetic influences on lupin seed storage proteins among 45 samples from five cultivars grown in three environmental conditions. A total of 133 reproducible peaks were identified by MALDI-TOF mass spectrometry. Only the reproducible protein mass peaks among the entire field replications were considered to attain a higher confidence level in the comparison of mass peak expression. NMP analysis using ANOVA and MDS based on mass peak similarity matrix visualises relationships at cultivar and environment levels. The results demonstrate that lupin seed protein mass peak expression was significantly influenced by cultivar variations. In contrast, seed protein expression was not significantly affected by environmental variation. However, protein

profile comparison showed that expressions of certain mass peaks were apparently influenced by cultivar × environment interactions.

Seed storage protein attributes is largely influenced by genetic variability

Analysis of the results of MDS plot and NMP found a significant influence of cultivar on protein mass peaks expression indicating lupin storage protein is largely genetically controlled. MDS plot based on Bray-Curtis similarity matrix of protein mass peaks profiles showed that the similarity within cultivars was higher than between cultivars (**Figure 2 A**). Cultivar Jenabillup and Tanjil showed a clear separation while Mandelup and Belara were positioned together with Coromup close by. This observation is in close agreement with the pedigree history as Mandelup had a parental line from Belara, and Coromup had a parental line of Belara sister (Cowling 1999). In contrast, both Jenabillup and Tanjil were developed through distant pedigree lines. This result is in accordance with the findings of proteomic relationships of 25 narrow-leafed lupin cultivars on mass spectrometric analysis (Islam et al. 2011a). Seed protein content of this species has been reported as highly variable in respect to cultivar and genotype (Bhardwaj et al. 1998). Seed protein quality attributes of soybean varied among different cultivars (Carrao-Panizz et al. 2008; Murphy and Resurreccion 1984; Yaklich 2000) which support our observation of the effect of cultivar on mass peak expression of lupin seed.

Influence of cultivar on expression of mass peaks indicates that protein diversity across different genotypes has the potential for proteome improvement in lupin through breeding. Cultivars with higher proteome diversity could be a good source of genes in lupin breeding. However, the five cultivars used in this study represent a small part of the potential genetic sources of lupin seed protein variability available to breeders. For instance, seed protein content in 234 lupin accessions ranged from 32% to 43% on a dry basis (Cowling and Tarr 2004) which may have much more variation at proteome level.

Seed storage protein composition in relation to environmental variation

ANOSIM and MDS plot (**Figure 2 B**) showed no significant variation across environments indicating lupin seed protein mass peak expression is not largely influenced by environment. Likewise, NMP was not significantly different in response to environmental variation. This result agrees with the observation that protein content of dry seed was not affected by growing environment in *L. albus* grown at two different

locations in USA (Bhardwaj et al. 1998). Likewise, water stress did not affect protein content of *L. albus* and *L. mutabilis* (Carvalho et al. 2005) indicating less effect of environmental variation on seed storage protein. Similarly, in the case of bread wheat, there was no significant difference in protein content due to temperature variation when well-watered and post-anthesis fertiliser application (DuPont et al. 1998). Nevertheless, protein content is not always correlated with protein quality attributes. The present study reports less effect of environment on mass spectrometric quality of lupin protein which agrees with the similar effects observed for seed protein content.

This study revealed that the expression of certain mass peaks in a cultivar could be influenced by environmental variation (**Table 1; Figure 3**) which may affect protein quality attributes of grain. This result is supported by the increased proportion of gliadins to glutenins and decreased proportion of large polymers in wheat due to high temperature effects (Blumenthal et al. 1995; Corbellini et al. 1997; Panozzo and Eagles 2000). Yet, very small proportions of protein mass peaks expressed an influence of environment which indicates that lupin seed protein mass peaks are largely stable in response to environmental variations.

Due to little influence of environmental variation on seed protein profile, none of the environmental factors such as temperature, rainfall and sunlight could be associated with mass peak expression in this study. This also happened in a study on narrow-leafed lupin seed protein and oil content (Cowling and Tarr 2004) where no strong association of rainfall, temperature or sunlight with quality traits tested was found in different locations of Western Australia.

Protein mass peak expression variability across cultivar × environment

Majority (76 of 133) of lupin seed protein mass peaks (**Table 1**) showed variable expression across the different combinations of cultivar × environment. Although there was lack of significant influence of cultivar × environment interactions on NMP (**Table 2**), the arbitrary expression of these mass peaks suggests an influence of cultivar × environment interactions on protein mass peaks expression in lupin. Substantial influences of environment × genotype interactions on seed composition and seed protein content has been reported in Brazilian soybean cultivars (Carrao-Panizz et al. 2008).

This study found two notable effects on mass peak expression due to environment × cultivar interactions. Firstly, some protein mass peaks appeared as rare i.e. observed in few combinations of environment and cultivar. For example, ~5% of all recognised protein mass peaks were expressed on one occasion i.e. at a single environment in a single cultivar. Secondly, some common mass peaks that did not appear in certain combinations of environment and cultivar. For instance, proteins mass peaks of molecular weights 3086, 5855, 9365, 9473, 9570, 9890, 10699, 19450 Dalton were not expressed on 2–5 occasions might be due to the interaction of environment × cultivar (**Table 1**). Overall, the results showed that the environment × cultivar influence on mass peak expression occurs in an unpredictable fashion, which is in accordance with the findings of Cowling and Tarr (2004) who reported considerable variation of seed quality and protein content in narrow-leafed lupin. The result of this study suggests that several trials are needed to be confident with the ranking of genotypes for seed protein mass peak expression. Trials over time may also add value to the findings of this study.

Possible reasons behind the mass peak variations

It is not known why individual mass peaks were unpredictable to environmental variation, but temperature variations across sites might have some influence on seed protein gene expression. Regulation of grain seed protein gene expression is complex and the mechanism is not fully understood. Studies revealed that temporal expressions of wheat gluten genes are influenced by different environmental factors (DuPont and Altenbach 2003). High temperature induces earlier accumulation of transcripts within all major gene groups of wheat seed proteins, resulting in earlier deposition of proteins during grain development (Altenbach et al. 2002). This temporal expression shift for protein genes might be due to developmental or environmental signals. Different seed protein genes like wheat endosperm genes may be involved in processing and post-translational modification of storage proteins (DuPont and Altenbach 2003). Thus alteration of the expression of these genes due to environmental factors would be responsible for seed protein quality variation. On the other hand, due to genotypic variation across cultivars, they might respond differently to the similar environment. For example, in response to temperature variation, wheat genotypes with the alleles for the HMW-GS pair 1Dx5, 1Dy10 were less variable in storage protein composition than genotypes with alleles pair 1Dx2, 1Dy12 (Blumenthal et al. 1995), which supports our observation on lupin. It is noteworthy that seed protein accumulation and post-transitional modification is a complex process which may vary even seed to seed.

However, overall similarities in mass peak profiles in replications indicate reproducibility of the method as proteome fingerprinting.

Mass peaks of the MALDI-TOF protein profile may not represent intact seed storage protein as the molecular weight range of intact lupin seed protein is 13 to 430 KDa (Duranti et al. 2008). However, subunits or polypeptides arising from proteolytic cleavages of native protein molecules (Derbyshire et al. 1976; Muntz et al. 2002; Cerletti et al. 1978) might have lowered their molecular weights (Islam et al. 2011b). Fragmentation of native proteins as a result of the ion-source decay (ISD) or post-source decay (PSD) in the MALDI instrument (Liu and Schey, 2005) may also results low molecular weight protein mass peaks.. Therefore, we consistently use the term 'mass peaks' that represent either intact proteins or protein fragments.

Conclusion

The effect of environmental conditions on lupin seed protein quality is a key concern in breeding programs, especially when designed for specific growing areas. Due to the limitation of environmental and geographical data, this study can not give a complete prediction of environment × genotype influence on seed protein mass peak expression. However, the revealed insignificant influence of environment on lupin seed protein mass peak expression provides an insight. The recognition of consistent cultivar-specific proteins across environmental variation suggests that these could be a good source for protein improvement.

Acknowledgements

I am grateful to Wayne Hawkins (DAFWA) for assisting me in seed grinding and thankful to Junhong Ma for her technical support in laboratory works.

Chapter 6

Comparative proteome analysis of seed storage and allergenic proteins among four narrow-leafed lupin cultivars

Citation: Islam S, Yan G, Appels R, Ma W. 2012. Comparative proteome analysis of seed storage and allergenic proteins among four narrow-leafed lupin cultivars. *Food Chemistry* 135 (3): 1230-1238.

Abstract

Lupin is an emerging crop worldwide due to its wide range of health benefits. In this study, a comprehensive proteome analysis has been conducted using mature seed of four narrow-leafed lupin cultivars, Uniharvest, Yorrel, Tanjil and Coromup, through two-dimensional gel electrophoresis followed by mass spectrometric protein sequencing. Two-dimensional gels recognised about 400 protein spots among the cultivars in the 10-100 KDa molecular weight and 5.0-8.5 pI ranges. The results revealed a considerable variation of protein expression patterns with a total of 24 proteins showed differential expression among the cultivars, among which 19 were identified as β-conglutin, and 8 were identified as allergenic proteins. Most of the α, δ and γ conglutins were showing similar expression among the cultivars. Overall, the differentially-expressed proteins especially the cultivar-specific proteins would be valuable markers for cultivar identification and for screening parental lines of low allergenicity in breeding process.

Keywords: Lupin, cultivars, protein, two-dimensional gel electrophoresis, differential expression, allergenic

Introduction

Recently, functional foods are attracting more attention at the consumer level due to their potential to provide health benefits. Proteins sourced from plants are considered to be valuable ingredients by food industry in the preparation functional foods (Sirtori et al. 2007). Although soybean is currently the major source of plant protein in food preparation, other grain legumes especially lupin are in rapid development as protein sources (Dijkstra et al. 2003). The use of lupin as a food or food ingredient is increasing due to its nutritional properties and lower levels of anti-nutritional factors (Petterson et al. 1997). The nutritional properties of lupin include its high protein content (Sirtori et al. 2004) along with higher content of fibre (Gorecka et al. 2000), oligosaccharides (Zdunczyk et al. 1998), and phenolic compounds content (Lampart-Szczapa et al. 2003). Due to the increased concern over the GMO issues, lupin is increasingly replacing soybean in food industries (Leduc et al. 2002).

Lupin flour is mainly used as an additive to wheat flour or a substitute for other protein rich flours in food preparations. Lupin-enriched foods provide health benefits such as

increased satiety and reduced energy intake (Lee et al. 2006), decreased blood pressure (Lee et al. 2009) , decreased blood glucose level (Hall et al. 2005) and cholesterol-lowering effect (Martins et al. 2005). The lupin seed proteins have been considered to be the main contributor to these claimed health benefits. The other reported bioactivities of lupin protein include plasma cholesterol and triglyceride lowering effects (Sirtori et al. 2004), antihypertensive properties (Pilvi et al. 2006) and angiotensin converting enzyme (ACE) inhibitory activity (Yoshie-Stark et al. 2004) . Since lupin is becoming popular as a food ingredient there have been reports concerning its allergic properties. Usually seed storage proteins are considered as the cause of allergic reactions upon ingestion (Breiteneder and Radauer 2004), suggesting that more defined information on lupin seed protein is crucial for the continued adoption of this legume grain in the food industry.

A number of studies have been carried out on seed proteins of lupin focused on storage protein compositions (Duranti et al. 2008; Magni et al. 2007; Sirtori et al. 2007), nutritional value (Brand et al. 2004; Lqari et al. 2002), protein modification due to processing (Islam et al. 2011b; Sirtori et al. 2010) and immunological and health properties of the protein fractions (Goggin et al. 2008; Guillamon et al. 2010; Klos et al. 2010). However, the mechanisms of differential accumulation of various protein components that result in differences in seed quality (taste, allergenicity) and morphology (Kottapalli et al. 2008) remain largely unknown (Ruuska et al. 2002).

Two-dimensional gel electrophoresis (2-DGE)-based proteomics approaches have been used successfully to identify and profiling proteins expressed during seed development or in mature seed of model plant species including soybean (Hajduch et al. 2005), rapeseed (Hajduch et al. 2006), *Medicago* (Gallardo et al. 2003), *Arabidopsis* (Gallardo et al. 2002), wheat (Islam et al. 2002; Majoul et al. 2003) and barley (Finnie et al. 2004). It is noteworthy that most of the previous studies on lupin proteins used the species *Lupinus albus* (Peeters et al. 2007). Very few preliminary studies using 2-DGE were reported on *Lupinus angustifolius* (Goggin et al. 2008; Islam et al. 2011b; Sirtori et al. 2010) and there is no systematic proteomics-level study on this species. Since proteomic studies on grain species (Liu et al. 2009; Yahata et al. 2005; Barakat 2004; Kottapalli et al. 2008) reported considerable variation of storage and allergenic proteins among cultivars, it is worth to study the important cultivars of narrow-leafed lupin (NLL) to understand the expression level of different seed proteins. A fingerprinting

study by our research group (Islam et al. 2011a) using direct mass spectrometry on 25 cultivars of NLL grown in Australia showed considerable seed protein variation among the cultivars. Moreover, total protein extracts from three different cultivars of blue lupin showed differential effects on plasma lipids in rats (Bettzieche et al. 2008). The latter finding in particular indicates variation of seed proteins among the cultivars might lead to differential bioactivity that is significant for food with certain health benefits.

Although white lupin is mostly used for human consumption, the use of NLL is increasing. In recent times the inadequate knowledge regarding the functional properties of proteins of NLL grain has limited its use in food stuff (Sirtori et al. 2010) . In order to characterise the proteome of NLL cultivars, we have carried out a 2-DGE based study for high resolution protein profiling of NLL cultivars. The selection of four NLL cultivars for detailed analysis was based on the phylogenetic relationship of the NLL cultivars following mass spectrometric seed protein analysis (Islam et al. 2011a). Cultivars from each of the major groups comprising pre-wild crosses, primary wild crosses and complex wild crosses were selected for this study. The information defines the variation among the cultivars with respect to allergenicity and nutritional aspects for utilisation in the breeding of lupin.

Materials and Methods

Materials: The four cultivars were selected from 25 Australian NLL cultivars based on proteomic phylogenetic relationship by direct mass spectrometry that broadly supported the pedigree relationship (Islam et al. 2011a). Cultivars Uniwhite, Yorrel and Tanjil were selected from the group of Pre-wild, Primary wild and complex wild crosses, respectively. The other cultivar Coromup had an isolated position in mass spectrometric study though its pedigree is in complex wild crossing. Seeds of the cultivars were supplied by the Department of Agriculture and Food, Western Australia (DAFWA). All the cultivars were grown in the same year at the same experimental station of DAFWA (Wongan Hills, WA). Thirty gram of seeds containing more than 100 seeds from each cultivar were taken as a working sample and ground in the “Retsch 2 M 200” and sieved with 750 µm.

Protein extraction: Flour samples were defatted using Hexane at 20:1 ratio (Santos et al. 1997) . The extraction buffer (8 M urea, 4% CHAPS, 60 mM DTT and 2% (v/v) IPG

buffer) was added to the defatted flour in the proportion of 20 ml/g to extract the protein at room temperature for 3 h (Goggin et al. 2008; Islam et al. 2011b). The protein extract (supernatant) was collected by centrifugation at 12,000g for 30 min and was precipitated by incubating with ice cold acetone at -20 ^{o}C for 16 h followed by centrifugation. The protein pellet was then washed with 10% ethanol and then with acetone containing β-mercaptoethanol (0.07%) to remove the additional salts. Ten milligrams of dried protein was dissolved in rehydration buffer containing 7 M urea, 2 M thiourea, 2% CHAPS, 65 mM DTT and 2% IPG buffer for 5-6 h at room temperature. Protein concentration was determined by using RC DC protein assay kit (Bio-Rad, Herculles, CA) and Lambda 25 UV-vis spectrometer (PerkinElmer). For each sample, 1100 µg of protein was loaded onto IPG strips (Bio-Rad, Hercules, CA) to optimise resolution and to ensure the adequate loading of minor components for MS/MS analysis (Goggin et al. 2008; Islam et al. 2011b).

Two-dimensional gel electrophoresis: Iso-electric focusing (IEF) was conducted on 17 cm IPG strips with pH 3-10. The strips were rehydrated with the buffer (7 M urea, 2 M thiourea, 2% CHAPS, 65 mM DTT and 2% IPG buffer) containing 1100 µg of protein for 12 h. Strips were focussed at 60,000 Vh, with a maximum of 10,000 V, at 20 ^{o}C using Protein IEF cell (BioRad). Before running SDS-PAGE, the strips were equilibrated with 50 mM Tris-HCl (pH 8.8), 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS and 0.002% bromophenol blue, containing 65 mM DTT for 15 min and another 10 min by substituting DTT with 135 mM iodoacetamide in the same buffer.

Protein separation was carried out on 12% acrylamide/bis (31.5:1) gels, using Protean II Xi cell (Bio-Rad). The running buffer consisted of 2.5 mM Tris-Base, 19.2 mM glycine and 0.01% SDS. The gels were stained by Coomassie Brilliant Blue (CBB). Protein standards (Bio-Rad) were used to estimate the molecular size of the proteins. To minimise experimental variability, all samples were run three times with individual extraction and IEF.

The gels were analysed by a 2-D Proteomic Imaging Systems (PerkinElmer) using ProScan 4.0 software. The digital gel maps of different samples were analysed and compared by using Progenesis Same Spots software (Nonlinear Dynamics). Master gels were generated for each sample by matching all of the available gels. Normalisation was carried out by determining the gain factor for each sample which can be modelled as

$y_i/y'_i = 1/\alpha_k$ where y_i is the measured abundance of feature i on sample k, $1/\alpha_k$ is the gain factor for sample k and y'_i is the normalised abundance of feature i on sample k.

Protein identification by MS/MS: Protein spots of interest were excised from Coomassie Brilliant Blue stained two-dimensional gels and analysed further by mass spectrometric peptide sequencing. To avoid the overlapping parts of closely related spots, the centre portions of each spot was sampled. The spots were analysed by Proteomics International Ltd Pty, UWA, Perth, Australia. Protein samples were trypsin digested and the resulting peptides were extracted according to standard techniques (Bringans et al. 2008). Peptides were analysed by electrospray ionisation mass spectrometry using the Ultimate 3000 nano HPLC system (Dionex) coupled to a 4000 Q TRAP mass spectrometer (Applied Biosystems). Tryptic peptides were loaded onto a C18 PepMap100, 3 μm (LC Packings) and separated with a linear gradient of water/acetonitrile/0.1% formic acid (v/v).

Spectra were analysed to identify proteins of interest using Mascot sequence matching software (Matrix Science) with taxonomy set to Viridiplantae (Green Plants). All searches used the Ludwig NR. The software was set to allow 1 missed cleavage, a mass tolerance of ± 1.2 Da for peptides and ± 0.6 for fragment ions. The peptide charges were set at 1+, 2+ and 3+, and the significance threshold at $P<0.05$. Generally a match was accepted where two or more peptides from the same protein were present in a protein entry in the Viridiplantae database.

Results

Two-dimensional gel electrophoresis

Protein extracted from seeds of each of the four cultivars (Uniharvest, Yorrel, Tanjil and Coromup) was analysed by two-dimensional gel electrophoresis and produced high resolution protein profiles as showed in the **Figure 1**. The result indicates successful standardisation of 2 DGE procedures to study the expression profile and comparative proteomic analysis of seed protein of NLL (*L. angustifolius*) cultivars. The results showed considerable differences in the protein profiles among the cultivars (**Figure 1**) although in general the protein patterns were similar. About 400 spots were revealed in the respective gels of each cultivar by 2DGE software (Progenesis Same Spots, Nonlinear Dynamics) and 97 protein spots showed some difference in their expression

levels among the cultivars. However, A total of 24 protein spots were found to be either present or absent, or showing markedly differential expression among the cultivars when the difference threshold was set to 2.5-fold. Most of the proteins were located in the 10-100 KDa and 5.0-8.5 pI ranges.

The differential proteins among the cultivars were positioned largely in 3 specific areas of the gels as showed in **Figures 1** and **2**(A-C). Noticeably some of the differentiating proteins were as a chain form in the gels with similar molecular weights but different pI values. The most striking region that differentiated cultivars had proteins in the 32-35 KDa range with 5.5-8.0 pI range. In this region, 13 proteins (spot numbers 1, 2, 3, 5, 7, 9, 10, 11, 12,13, 24, 26 and 27) showed either fully present versus absent or different level of expression among the cultivars (**Figure 2B**, **Table 1**). Six proteins (spots number 14, 15, 16, 17, 18 and 19) from the higher molecular weight range (65-70 KDa with 5.5 to 6.5 pI; **Figure 2A**) showed differential expressions. Likewise, five proteins (spot numbers 20, 21, 22, 23 and 25) from relatively low molecular weight range (10-20 KDa range) were found (**Figure 2C**) as differentiating among the cultivars.

Identifying cultivar-specific proteins

Proteins specific to single cultivars: Eight proteins (spot numbers 18, 19, 20, 21, 23, 25, 26 and 27) were found present in only one of the cultivars (**Table 1**, **Figure 2A, B** and **C**). Cultivar Tanjil had the highest with five cultivar-specific proteins (spot numbers 20, 21, 23, 26 and 27) and a cultivar Coromup had two cultivar-specific proteins (spot numbers 18 and 19) at the higher molecular weight range (75KDa). Cultivar Yorrel possessed a single cultivar-specific protein (spot number 25) while cultivar Uniharvest did not have any cultivar-specific proteins.

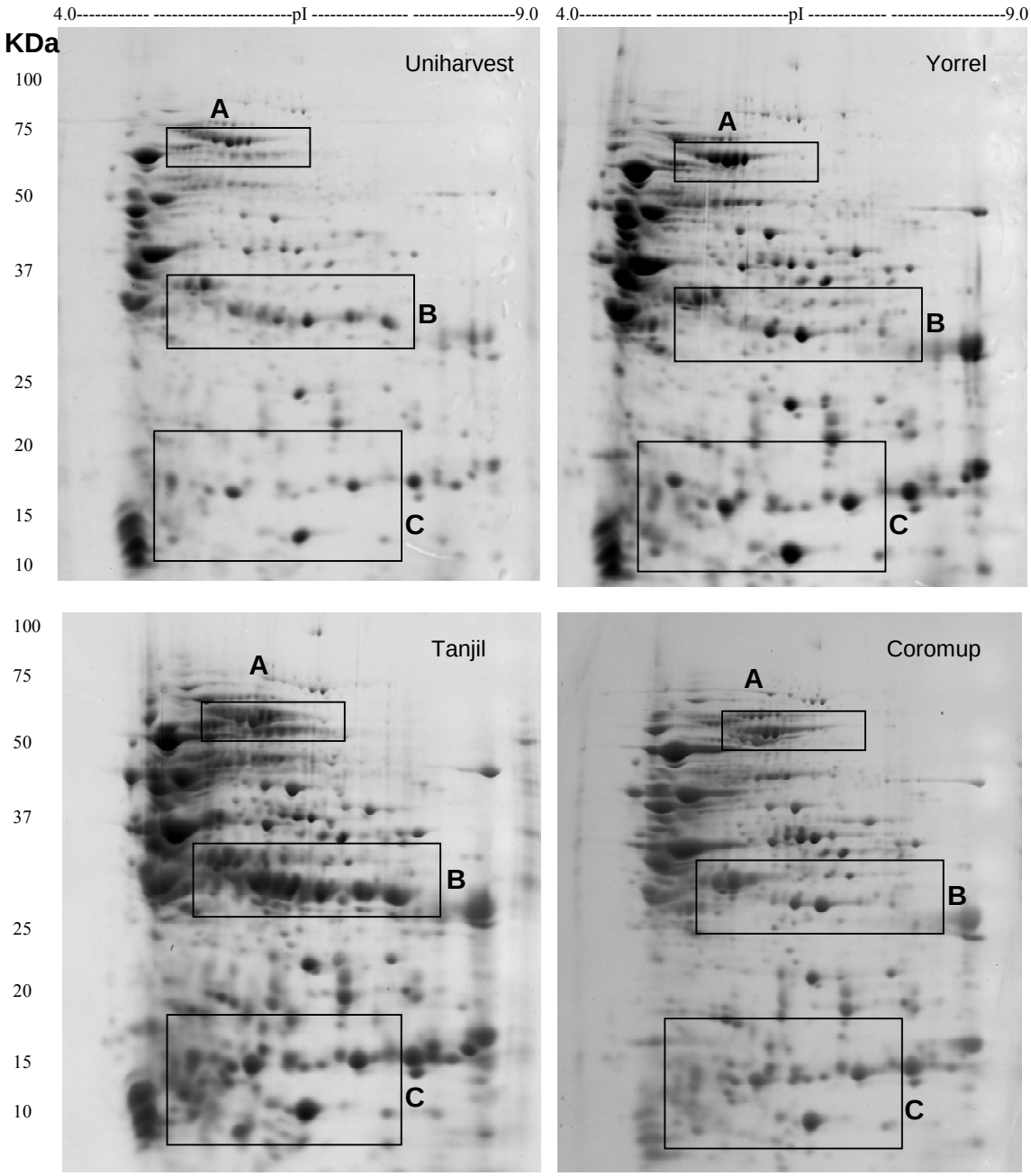
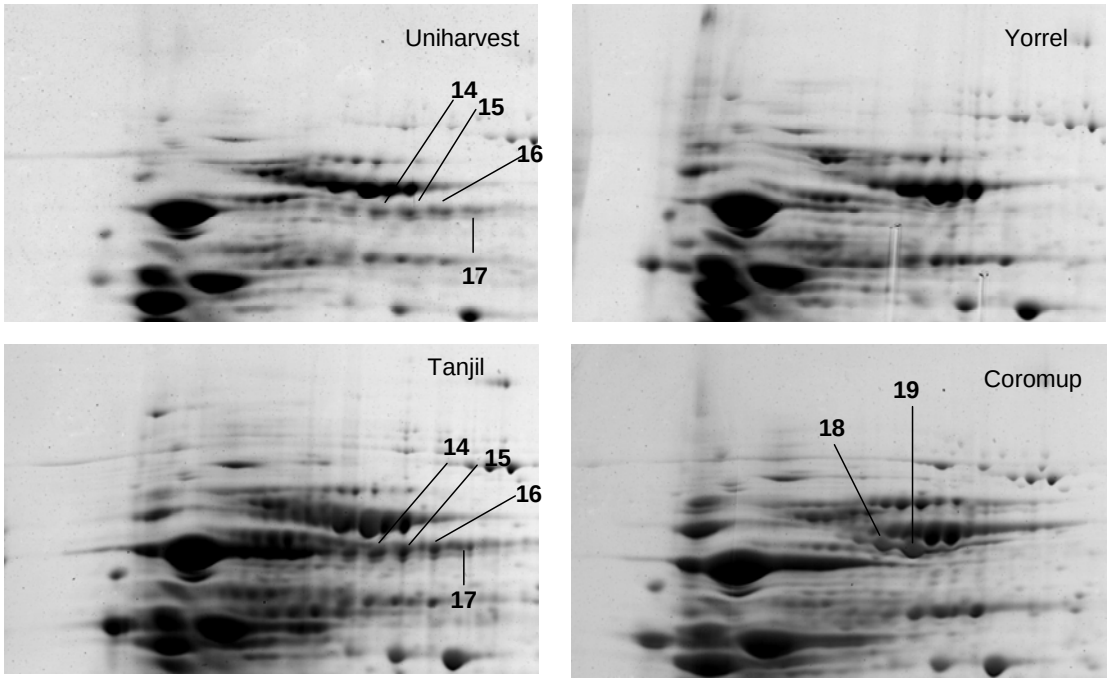
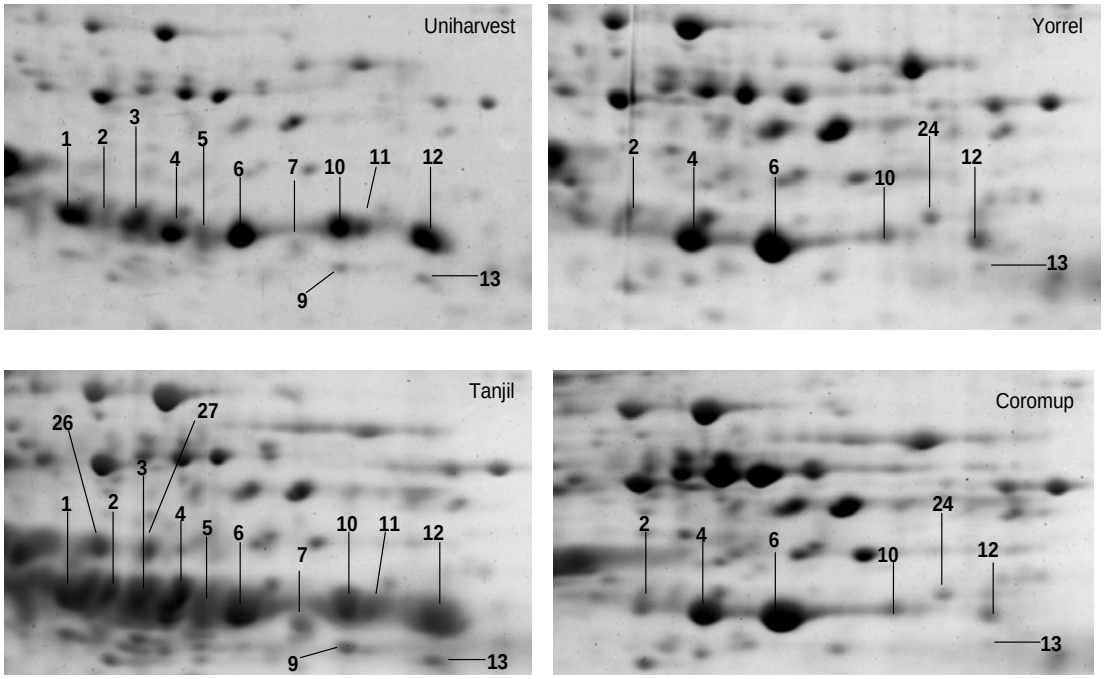


Figure 1: Seed protein profile of four cultivars of *Lupinus angustifolius* as revealed by two-dimensional gel electrophoresis of total protein indicating overall variation of proteins. The rectangles indicate the regions with differentiating proteins. The elaboration of those specific areas is showed in **Figure 2**.

A


Uniharvest
14
15
16
17
Yorrel
Tanjil
14
15
16
17
Coromup
19
18

B


Uniharvest
1
2
3
4
5
6
7
10
11
12
9
13
Yorrel
2
4
6
10
24
12
13
Tanjil
26
27
3
1
2
4
5
6
10
11
12
7
9
13
Coromup
2
4
6
10
24
12
13

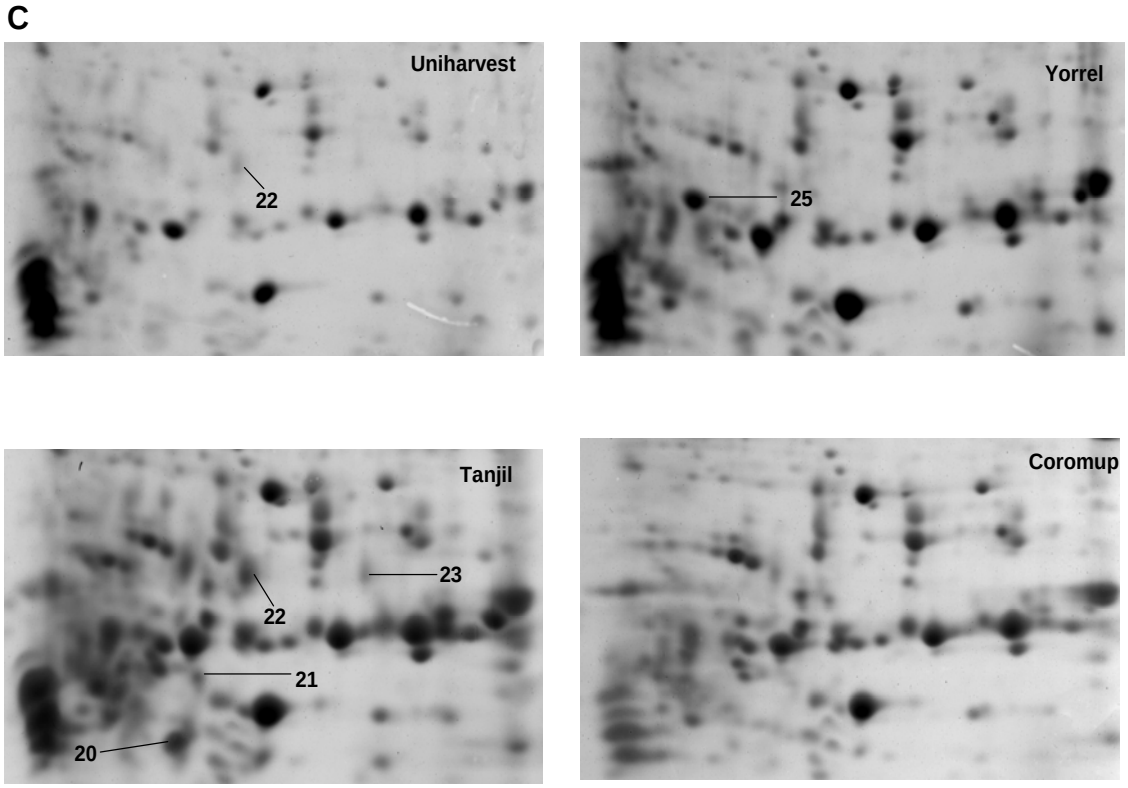


Figure 2: Comparison of specific regions on the 2-D gel demonstrating the expression differentiating proteins among the four cultivars examined. The letters **A, B** and **C** indicate the regions showed in **Figure 1**.

Differentiating proteins present in two cultivars: Twelve proteins (spot numbers 1, 3, 5, 7, 9, 11, 14, 15, 16, 17, 22 and 24) were recognised clearly in two of the four cultivars and absent in the other two cultivars (**Table 1**, **Figure 2**). Eleven spots out of the twelve (except 24) were expressed in cultivar Uniharvest and Tanjil and were absent in cultivars Yorrel and Coromup. In contrast, the protein corresponding to spot number 24 was clearly expressed in cultivars Yorrel and Coromup whereas absent in Uniharvest and Tanjil.

Proteins present in all of the four cultivars with differential expression level: Two-dimensional gel analysis by the software "progenesis same spots" recognised 77 proteins those are present in all of the four cultivars with different level of expression among the cultivars. However, only four of these proteins (spot numbers 2, 10, 12 and 13; **Figure 2B**) met our stringy threshold (2.5-fold) (**Table 1**). All of these proteins

showed relatively higher expression in the cultivars Uniharvest and Tanjil compared to Yorell and Coromup.

Protein Identification by mass spectrometry

A total of 58 different protein spots were analysed through mass spectrometry excised from the two-dimensional gels. These included the 24 differential proteins (**Figure 2**) and 34 common proteins (**Figure 3**) of the cultivars, making a total of 184 proteins samples analysed. Corresponding protein spots from each of the cultivars (where available) were analysed separately and matched together and gave a very good homology of the corresponding protein spots from different cultivars. Of the 58 individual proteins analysed, 52 proteins were identified as one of the major lupin seed protein groups i.e. conglutins. Four proteins were identified with proteins from other species and two proteins could not be identified.

Table 1: MS/MS identification of differentiating proteins among the cultivars. Matching has been achieved using Mascot sequence matching software (Matrix Science) with the taxonomy set to Viridiplanate (Green Plants). β-Conglutin and α-Conglutin are sometimes referred to as vicilin-like proteins and legumin-like proteins, respectively, but we have used the term β-Conglutin and α-Conglutin for all closely matching cases to avoid confusion.

Spot no.	Spot relative abundance Protein expression as fold ratio: Uniharvest/Yorell/ Coromup/Tanjil	Anova (p)	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
Cultivar-specific proteins									
18	--/--/2.8/--	0.008	β-conglutin [*Lupinus angustifolius]*	gi|149208401	61490/5.34	60000/5.7	13%	180	GKPSESGPFNLR QAYNLEYGDALR TNRLENLQNYR HSDADYILVVLNGR AIFIVVVDEGEGNYELVGIR
19	--/--/2.6/--	0.002	β-conglutin [*Lupinus angustifolius*]	gi|149208401	61490/5.34	60000/5.8	10%	167	QAYNLEYGDALR TNRLENLQNYR HSDADYILVVLNGR AIFIVVVDEGEGNYELVGIR
20	--/--/--/3.2	0.007	Putative uncharacterised protein [*Zea mays*]	gi|223947879	42658/8.13	10000/5.9	3%	97	REEEEEATPAAR
21	--/--/--/3.1	2.39E-04	Not Identified	-	-	13000/6.0	-	-	-
23	--/--/--/3.8	0.006	β-conglutin [*Lupinus albus*]	gi|89994190	61994/6.08	18000/7.2	8%	225	YEEIQR LENLQNYR NPYHFNSQR DQQSYFSGFSR NTLEATFNTRYEEIQR
25	--/8.6/--/--	1.08E-04	β-conglutin [*Lupinus angustifolius*]	gi|149208403	54267/6.27	17500/5.0	6%	432	KQIQELR HAQSSSGEGKPSESGPFNLR KHAQSSSGEGKPSESGPFNLR
26	--/--/--/2.8	0.008	BLAD (Fragment) Tax_Id=3870 [*Lupinus albus*]	gi|77994351	20393/9.66	33000/6.2	15%	134	DQQSYFSGFSR NTLEATFNTRYEEIQR
27	--/--/--/2.5	0.002	Not identified			33000/6.3			

Spot no.	Spot relative abundance Protein expression as fold ratio: Uniharvest/Yorell/ Coromup/Tanjil	Anova (p)	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
Proteins present in two cultivars									
1	3.6/--/--/2.1	2.45E-04	β-conglutin [*Lupinus albus*]	gi|89994190	61994/6.08	29000/6.0	9%	216	NPYHFNSQR NFLAGSEDNVIR TNRLENLQNYR NTLEATFNTRYEEIQR
3	3.6/--/--/2.2	1.99E-05	β-conglutin [*Lupinus angustifolius*]	gi|149208401	61490/5.34	28500/6.1	8%	256	TNRLENLQNYR NTLEATFNTRYEEIQR NQQQSYFANAQPQQKQQR
5	2.5/--/--/2.3	1.42E-04	β-conglutin [*Lupinus angustifolius*]	gi|149208401	61490/5.34	28000/6.5	5%	160	TNRLENLQNYR NQQQSYFANAQPQQKQQR
7	4.6/--/--/7.6	1.67E-05	β-conglutin [*Lupinus angustifolius*]	gi|149208403	54267/6.27	27000/6.7	22%	349	NPYHFSSNR DQQSYFSGFSK QQEEQEREHR TNRLENLQNYR EREQEQQPRPQR QEEEEEEEEWQPR RQEEEEEEEEWQPR NTLEATFNTRYEEIER LPAGTTSYILNPDDNQNLR
9	4.2/--/--/5.0	9.03E-05	β-conglutin [*Lupinus angustifolius*]	gi|328684565	71853/5.82	25500/7.0	4%	160	TNRLENLQNYR NQQQSYFANAQPQQQQQR
11	3.5/--/--/2.2	1.97E-05	β-conglutin [*Lupinus angustifolius*]	gi|149208401	61490/5.34	28000/7.1	3%	127	NQQQSYFANAQPQQKQQR
14	5.4/--/--/4.5	0.001	β-conglutin [*Lupinus albus*]	gi|89994190	61994/6.08	68000/5.8	9%	242	YEEIQR NPYHFNSQR DQQSYFSGFSR NFLAGSEDNVIR NTLEATFNTRYEEIQR
15	5.4/--/--/4.3	0.005	β-conglutin [*Lupinus angustifolius*]	gi|149208401	61490/5.34	68000/5.9	12%	296	EQEQQPQHGR NTLEATFNTRYEEIQR

Spot no.	Spot relative abundance Protein expression as fold ratio: Uniharvest/Yorell/ Coromup/Tanjil	Anova (p)	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
16	4.4/--/--/4.2	0.002	β-conglutin [*Lupinus albus*]	gi|89994190	61994/6.08	68000/6.0	9%	277	LPAGTTSYILNPDDNQNLR NQQQSYFANAQPQQKQQR NPYHFNSQR DQQSYFSGFSR NFLAGSEDNVIR NTLEATFNTRYEEIQR
17	3.9/--/--/4.5	6.89E-04	β-conglutin [*Lupinus albus*]	gi|89994190	61994/6.08	68000/6.1	9%	254	NPYHFNSQR DQQSYFSGFSR NFLAGSEDNVIR NTLEATFNTRYEEIQR
22	3.0/--/--/4.1	6.47E-06	β-conglutin [*Lupinus angustifolius*]	gi|149208401	61490/5.34	18000/6.4	8%	394	LENLQNYR TNRLENLQNYR NTLEATFNTRYEEIQR LPAGTTSYILNPDDNQNLR
24	--/4.0/4.3/--	2.51E-07	Putative uncharacterised protein [*Oryza sativa* subsp. *indica*]	gi|125539365	67761/7.12	29000/7.2	2%	56	NVSEEIICLAAALR
Proteins present in all cultivars with different expression level									
2	2.5/1.0/1.1/2.5	0.006	β-conglutin [*Lupinus angustifolius*]	gi|89994190	61490/5.34	28000/6.3	9%	250	NPYHFNSQR NFLAGSEDNVIR TNRLENLQNYR NTLEATFNTRYEEIQR
10	3.8/1.3/1.0/2.2	3.68E-05	β-conglutin [*Lupinus angustifolius*]	gi|149208401	61490/5.34	28000/7.2	3%	147	NQQQSYFANAQPQQKQQR
12	3.8/1.2/1.0/3.0	9.99E-05	β-conglutin [*Lupinus albus*]	gi|89994190	61994/6.08	28000/7.5	2%	126	NFLAGSEDNVIR
13	3.8/2.0/1.0/3.6	4.15E-05	β-conglutin [*Lupinus angustifolius*]	gi|328684565	71853/5.82	26000/7.6	6%	177	NPYHFSSNR TNRLENLQNYR NQQQSYFANAQPQQQQQR

Out of the 24 differentially-expressed proteins, 19 proteins (spots numbers 1-3, 5, 7, 9-19, 22, 23, 25; **Figure 2**, **Table 1**) were identified as β-conglutin, the major seed storage protein of lupin. Of these 19 proteins 7 (spot numbers 1, 2, 12,14,16,17, 23; **Figure 2**, **Table 1**) were identified as allergenic protein by matching with the accession number gi|89994190, an allergenic protein of lupin (Goggin et al. 2008; Guillamon et al. 2010; Peeters et al. 2007). Protein spot 26, a cultivar-specific protein of Tanjil, was also identified as allergenic protein as matched with BLAD of *L. albus,* a previously reported allergenic protein of lupin (Guillamon et al. 2010). Protein spot numbers 20 and 24 were matched closely with an uncharacterised protein of the species *Zea mays* and *Oryza sativa* subsp. *Indica* respectively. Peptides generated from the protein spots 21 and 27 present in cultivar Tanjil only did not give any perfect match with sequences present at the database.

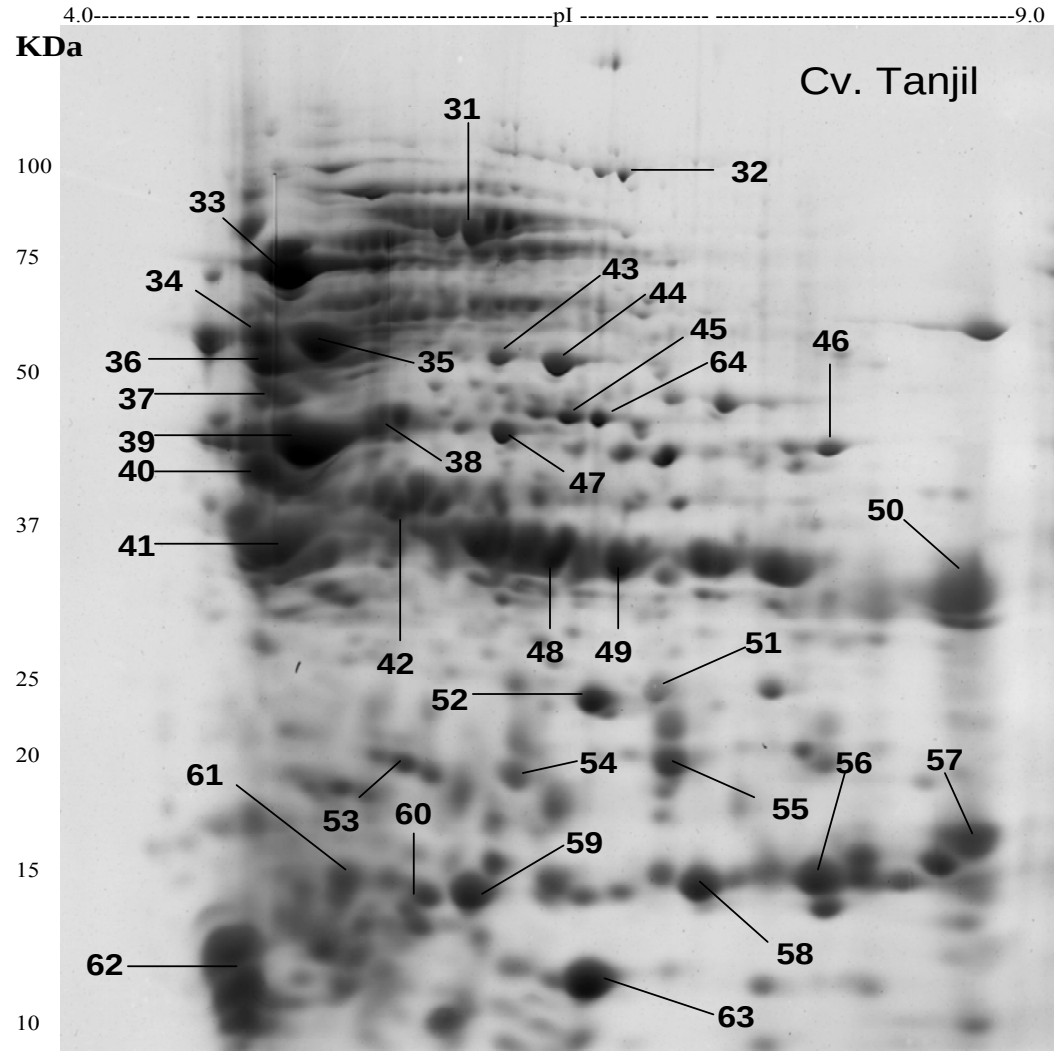


Figure 3: Identical (common) lupin seed proteins with similar expressions among the four cultivars as identified by MS/MS. The spots are shown on a 'Tanjil' background and the identification of the spots is presented in **Table 2**.

With the 34 common proteins among the four cultivars that were analysed through MS/MS (**Table 2** and **Figure 3**), 33 were identified with a close match with the sequences of conglutin groups (α, β, γ and δ conglutins) in the databases. Spot number 46 was identified as Glyceraldehyde-3 phospahte dehydrogenase (see the supporting document for details).

Table 2: MS/MS protein identification results of 2-DGE gel spots. (Protein spot numbers are indicated in **Figure 2** and **3**)

Spot number	Present in Cultivars	Identified protein
23	Tanjil	β-conglutin
20,21,26	Tanjil	Peptides did not match with any known lupin protein sequence
27	Tanjil	BLAD [*Lupinus albus*]
18,19	Coromup	β-conglutin
25	Yorrel	β-conglutin
24	Yorrel and Coromup	Peptides did not match with any known lupin protein sequence
1,3,5,7,9,11,14,15, 16,17, 22	Uniharvest and Tanjil	β-conglutin
2,10,12,13	All cultivars (with differential expression)	β-conglutin
31, 36-37, 41-45, 47-49, 52, 57,64	All cultivars (with similar expression)	β-conglutin
33-35, 38-40, 51, 53-56, 58-61	All cultivars (with similar expression)	α-conglutin
32, 50, 63	All cultivars (with similar expression)	γ-conglutin
62	All cultivars (with similar expression)	δ-conglutin
46	All cultivars (with similar expression)	Glyceraldehyde-3-phosphate-dehydrogenase [*Lupinus albus*]

Discussion

In the current study, we have focused on variation in the lupin seed storage protein among different cultivars. In the past, traditional protein analysis methods failed to reveal the vast variations among germplasms especially cultivars. The current study demonstrated that 2-DGE technology coupled with mass spectrometry peptide sequencing is a powerful and high resolution approach to reveal the extent of variations among cultivars.

Allergenic variation among cultivars in other crops including peanut have been reported (Kottapalli et al. 2008). Protein isolates from different cultivars of blue lupin showing differential effects on plasma lipid regulation on rat (Bettzieche et al. 2008) indicates variation of bioactive proteins among cultivars. Moreover, dissimilar groups of conglutins have been claimed as the major allergenic protein of lupin while using different species and cultivars (Goggin et al. 2008; Guillamon et al. 2010; Klos et al. 2010) suggesting variation in expression of allergenic proteins may occur among cultivars. The study based on direct MALDI-TOF protein profiling suggested considerable variation of seed proteins among the NLL cultivars (Islam et al. 2011a). The current study applied a high resolution 2-DGE based proteomic analysis of selected four NLL cultivars to search for the differentiating seed proteins patterns with an emphasis on the allergenic and bioactive proteins.

Detection of approximately 400 spots using the 2DGE based proteomic approach indicated that in general the patterns of proteins among the cultivars were similar. Homology of many common protein spots among the cultivars allowed the identity of the differentiating proteins to be assigned on the basis of their electrophoretic mobility. A total of 97 protein spots showed some difference in their expression levels among the cultivars but we only considered the spots having more than 2.5-fold variation as qualifying for a differentiating protein. The 2.5-fold variation requirement minimised the effects of any experimental error. The analysis identified 24 differentiating proteins among the cultivars. Seed protein composition is generally genetically controlled (Bolon et al. 2010) although some environmental affects would be expected. All the samples studied in this study were grown at the same experimental conditions, suggesting that the proteomic variations revealed by the current study was due to the genetic variation. The complex crossing systems used in the breeding of cultivars (Cowling 1999) has led to diverse genetic variation. DNA based study (Yuan et al. 2005) suggested considerable genetic variation among the NLL cultivars released in Australia. Thus the information on variation at proteomic level might be useful in selecting appropriate germplasm as parental lines to breed cultivars with low or even no allergenicity.

Most of the differentiating proteins among cultivars were in the β-conglutin group, the largest seed protein family of lupin (Duranti et al. 2008). Many of the differential proteins as well as some common ones appeared as a chain in the gels at the same

molecular weight with different pI values. Protein spot numbers 1- 6, 10-12, 14-17 and 18-19 (**Figure 2A** and **B**) were placed in the gels as a form of chain indicating the existence of different isoforms of the same protein. Matching of protein spots within a chain with the sequence of same protein accession also indicated their homology. (Liang et al. 2006) identified differences in isoforms of basic arachin (iso-Ara h3) among peanut cultivars and suggested the cause was variation in post-translation modification. Most of the α, γ and δ conglutins including both the high and low molecular weight appeared as consistent representatives in the cultivars, indicating their consistency and stability. β-Conglutin and α-conglutin are sometimes described as vicilin-like proteins and legumin-like proteins, respectively, but we have used β-conglutin and α-conglutin consistently to avoid confusion.

The results indicate considerable variation of allergenic proteins among cultivars that provides an insight about the significance of cultivar-specific lupin proteomics. All of the 8 differentiating allergenic proteins are highly expressed in the cultivar Tanjil while Cultivar Uniharvest has 6 highly expressed. However, cultivar Yorrel and Coromup have only two differentially-expressed allergenic proteins at very low expression level. This predicts that cultivar Yorrel and Coromup may have low allergenic effect than the other two cultivars.

The differential expression of seed proteins among cultivars might have the potential to relate the cultivars to bioactivities of lupin proteins. Most of the differentiating proteins were identified as β-conglutin, the major seed storage protein of lupin similar to 7s glubulins of soybean which has been investigated for some biological activity (Maruyama et al. 2003; Prak et al. 2006). In contrast, γ-conglutins having blood glucose lowering effect is the only lupin seed protein has been reported for individual bioactivity (Duranti et al. 2008) that showed similar expression among the studied cultivars. However, the lack of information regarding the biological activity of individual proteins or individual protein groups (α, β, γ and δ) of lupin (Duranti et al. 2008) has limited the discussion and suggested the necessity of more detailed studies of individual protein groups in terms of functionality and bioactivity.

The expression of differentiating proteins suggested two distinct groups among the cultivars. Cultivars Uniharvest and Tanjil showed similar patterns of differentiating protein expression except for the cultivar-specific proteins in Tanjil whereas cultivars

Yorrel and Cormup formed a separate group with similar patterns. At the 32-35 KDa molecular weight range, 11 β-conglutins were present or highly expressed in the cultivars Uniharvest and Tanjil (**Figure 2B**). In contrast, these proteins were absent or poorly expressed in Yorrel and Coromup. Likewise a group of high molecular weight proteins (spots 14-17) were found in cultivars Uniharvest and Tanjil and absent in Yorrel and Coromup. On the other hand one uncharacterised protein (spot number 24, **Figure 2B** and **Table 1**) was present in cultivar Yorrel and Coromup but absent at cultivars Uniharvest and Tanjil. The pedigree history suggested the cultivar Coromup has one parental line from Yorrel. It is noted that DNA based studies suggest cultivar Tanjil is closer to the Uniharvest than the other cultivars (Yuan et al. 2005).

Cultivar-specific proteins have been used for cultivar identification in some species (Kottapalli et al. 2008; Yahata et al. 2005). The eight cultivar-specific proteins detected in this study will be useful for lupin cultivar identification. Cultivar Tanjil possesses the highest number (6) of cultivar-specific proteins. This may be due to its complex crossing process during breeding (Cowling 1999). Coromup, the other cultivar with complex wild crossing pedigree history had 2 cultivar-specific proteins. A comparatively simpler crossing system that comprises both primary and secondary crosses (Cowling 1999; Islam et al. 2011a) made Yorrel to have just one cultivar-specific proteins. On the other hand, one of the oldest cultivar, Uniharvest, bred from only primary crosses did not have any cultivar-specific protein. These cultivar-specific proteins are certainly useful for cultivar identification and for proteomic improvement of lupin through further breeding once the function of those proteins are known.

MS/MS peptide sequences of 52 different proteins out of 58 gave very good matching with the lupin protein sequences in the databases including all the conglutin groups, indicating successful identification of lupin proteins. In all analysed protein samples, the corresponding protein spots from all four cultivars were analysed separately for a better confirmation of the identification. In all cases (with few exceptions), highly similar peptide sequencing and matching with the similar accessions indicate uniformity and homogeneity of the proteins among the cultivars. Three proteins were successfully identified as proteins from other species and 3 proteins were not identified at all, suggesting lacking of sequence information in the databases.

Visible differences in the expression of important seed proteins among the four cultivars signify a valuable tool for cultivar identification for further molecular breeding. The reported differential expression of allergenic protein suggests further studies including more cultivars could lead to a targeted selection of lupin cultivars for food industries.

Acknowledgements

The authors are grateful to Sophie Sipsas for providing lupin seed materials and thankful to Junhong Ma for her technical support in laboratory works.

Supporting information

Matching of the mass spectrometric peptide sequences to identify the common proteins among the cultivars is presented at **Appendix II**.

Chapter 7

Differential recovery of lupin proteins from the gluten matrix in lupin–wheat bread as revealed by mass spectrometry and two-dimensional electrophoresis

Citation: Islam S, Ma W, Gao L, Yan G, Appels R. 2011. Differential recovery of lupin proteins from the gluten matrix in lupin–wheat bread as revealed by mass spectrometry and two-dimensional electrophoresis. *Journal of Agricultural and Food Chemistry* 59 (12): 6696–6704

Abstract

Bread made from a mixture of wheat and lupin flour possesses a number of health benefits. The addition of lupin flour to wheat flour during breadmaking has major effects on bread properties. The present study investigated the lupin and wheat flour protein interactions during breadmaking process including dough formation and baking by using proteomics research technologies including MS/MS to identify the proteins. Results revealed that qualitatively most proteins from both lupin and wheat flour remained unchanged after baking as per electrophoretic behaviour, while some were incorporated into the bread gluten matrix and became unextractable. Most of the lupin α-conglutins could be readily extracted from the lupin-wheat bread even at low salt and non-reducing/non-denaturing extraction conditions. In contrast, most of the β-conglutins lost extractability suggesting that they were trapped in the bread gluten matrix. The higher thermal stability of α-conglutins compare to β-conglutins is speculated to account for this difference.

Keywords: Lupin; wheat; bread; protein; interaction; extractability

Introduction

Use of plant-derived proteins is widely accepted as a way of meeting the demand for health requirements in food products. Lupin (*Lupinus* spp.) is a legume crop well adapted for a wide range of climates and soils and has been investigated for its health benefits in relation to its high protein and fibre contents (Lee et al. 2009; Lee et al. 2006) . It represents a relatively new health-food source and its consumer awareness and acceptance are still low. One efficient way to bring lupin-associated health benefits to consumers is to add lupin flour to a widely consumed food product such as bread.

The main attribute of lupin is its high protein and dietary fibre as well as negligible amounts of starch content. It has been proven that bread enriched by lupin flour has the potential to provide health benefits, such as increased satiety and reduced energy intake (Lee et al. 2006), decrease the blood pressure (Lee et al. 2009), and decrease the blood glucose level (Hall et al. 2005). The mixing of flour from different sources such as lupin and wheat has major effects on the properties of the final product (Doxastakis et al. 2002). Wheat gluten is required for breadmaking because it provides the basis for forming a network that has viscoelasticity and gas retention properties as dough is

formed (Roccia et al. 2009; Zhang et al. 2008). When wheat flour is supplemented with other flour-rich in proteins, the added protein molecules interact with the wheat protein (gluten) network in both direct and indirect ways: the direct way is an interaction between added proteins and gluten proteins by cross-linking and the indirect way is related to competition for water availability required by wheat gluten proteins (Bonet et al. 2006; Ribotta et al. 2005b; Ryan et al. 2002). During dough preparation and baking, different ingredients such as oxidizing agents, salts, water that promote and affect the formation of protein cross-links (Gerrard et al. 2001) could also be affected by lupin flour addition. The interactions between soybean and wheat proteins in bread baking have been studied (Roccia et al. 2009; Ryan et al. 2002). Results demonstrated that the addition of soybean flours to wheat flour at relatively low levels (approximately 5% by weight) has negative effects on bread quality attributes such as extensibility properties and gas retention, thus decreasing the consumer acceptability (Roccia et al. 2009). Similarly, the addition of lupin flour to wheat flour breaks the connectivity of the wheat gluten matrix (Guemes-Vera et al. 2004), resulting in a reduced consumer-level quality (Dervas et al. 1999) . Some of these negative effects can be overcome by the addition of extra gluten during processing (Lee et al. 2006).

It has been reported that adding soybean flour to wheat flour in bread-making decreases the solubility of soybean protein as measured by its decreased extraction from soy-wheat bread (Ribotta et al. 2005b; Ryan et al. 2002). The possible reason for this solubility change is that covalent cross-links (different disulfide bonds) are formed during baking. Some soybean proteins were also suggested to be bound into the bread matrix by non-covalent bonds. In a study on wheat protein changes during bread-making, the decrease of protein extractability in the bread has been suggested to be due to protein cross-linking and/or aggregation (Singh 2005). It also suggested that disulfide bonds were mainly responsible for protein insolubility as a result of formation of very high molecular weight protein matrix in bread.

The aim of the current study was to monitor the lupin and wheat flour protein complexes formed in bread-making by using a combination of protein analysis techniques including intact protein analysis by MALDI-TOF and 2-D protein separation followed by MS/MS protein identification. A detailed analysis of protein from lupin-wheat bread was carried out in order to define the extractability of a wide range of proteins. The relationships of these properties to possible health attributes are discussed.

Materials and Methods

Flour and Bread samples: Lupin-wheat bread was prepared by substituting 40% of the wheat flour present in the normal wheat bread with lupin kernel flour. Extra gluten was added to the lupin-wheat bread to match the amount found in wheat flour (Lee et al. 2006). Breads were baked as single batch at Bodhi's Bakery, Fremantle, Western Australia, following the procedure used to prepare breads for studying health attributes of lupin-enriched bread (Lee et al. 2006). The same batch of lupin and wheat flour samples were used in this study. The lupin flour was from one single cultivar "Kalya" of *Lupinus angustifolius* and the wheat flour was a commercially available mix generally used to prepare normal wheat bread.

Protein extraction: Under non-reducing and non-denaturing condition, protein was extracted by using 0.5 M NaCl (pH 7.0) based on the techniques outlined in Duranti et al. (2008) and Lampart-Szczapa (1996). During the extraction process, both the bread and flour samples were defatted by hexane at 20:1 ratio (Santos et al. 1997). The extraction buffer (0.5M NaCl) was added to the flour and bread sample at the ratio of 15 mL/g and the protein was extracted by stirring at 4 ^{o}C for 4 h. The extraction buffer for protein under reduced and denatured condition contained 8 M urea, 4% CHAPS, 60 mM DTT and 2% (v/v) IPG buffer. The extraction buffer was added to the defatted flour and bread sample at the ratio of 20 mL/g and protein was extracted at room temperature for 3 h (Goggin et al. 2008). In both the cases the protein extract (supernatant) was collected by centrifugation at 12000g for 30 min.

For 2-D electrophoresis, all protein extracts were precipitated by incubating with ice cold acetone at -20 ^{o}C for 16 h followed by centrifugation. The protein pellet was then washed with 10% ethanol and then with acetone containing β-mercaptoethanol (0.07%) to remove the additional salts. Ten mg of dried protein was dissolved in rehydration buffer containing 7 M urea, 2 M thiourea, 2% CHAPS, 65 mM DTT and 2% IPG buffer for 5-6 h at room temperature. Protein concentration was determined by using RC DC protein assay kit (Bio-Rad, Herculles, CA) and Lambda 25 UV/Vis spectrometer (PerkinElmer). For each sample, 1100 µg of protein was loaded onto IPG strips (Bio-Rad).

Intact protein analysis by MALDI-TOF: An overall survey of proteins from lupin-wheat bread, normal wheat bread and lupin and wheat flours was conducted by using

matrix-assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI-TOF) (Chen et al. 2007). Protein extracted by 0.5 M NaCl was directly mixed with matrix (sinapic acid dissolved in 0.05% trifluoroacetic acid and 50% acetonitrile) at 1:9 ratio. A 1 µl sample from the mixture was spotted on a MALDI-TOF plate and left at room temperature to dry. Spotting on the plate was repeated once and the analysis was carried out on a Voyager DE PRO Biospectrometry Workstation (PerSpective Biosystem, USA), operated at linear mode. Each mass spectrum was the result of 500 laser shots on 10 random positions of the spot. The mass accuracy for linear mode was 1.01 ppm. The mass range was set to 5,000 to 80,000 Da. The machine was calibrated by using a Sequazyme Peptide Mass Standards Kit (Applied Biosystem, USA) following sinapic acid matrix-calibration mixture 3. The results from MALDI-TOF were analysed using the Voyage machine companion software, Data Explorer, to produce the protein spectrum profiles.

Two-dimensional gel electrophoresis: Iso-electric focusing (IEF) was conducted on 17 cm IPG strips with pH 3-10. The strips were rehydrated with the buffer (7 M urea, 2 M thiourea, 2% CHAPS, 65 mM DTT and 2% IPG buffer) containing 1100 µg of protein for 12 h. Strips were focussed at 60,000 Vh, with a maximum of 10,000 V, at 20 ^{o}C using Protein IEF cell (BioRad). Before running SDS-PAGE, the strips were equilibrated with 50 mM Tris-HCl (pH 8.8), 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS and 0.002% bromophenol blue, containing 65 mM DTT for 15 min and another 10 min by substituting DTT with 135 mM iodoacetamide in the same buffer.

Protein separation was carried out on 12% acrylamide/bis (31.5:1) gels, using Protean II Xi cell (Bio-Rad). The running buffer consists of 2.5 mM Tris-Base, 19.2 mM glycine and 0.01% SDS. The gels were stained by Coomassie Brilliant Blue (CBB). Protein standards (Bio-Rad) were used to estimate the molecular size of the proteins. To minimize experimental variability all samples were run three times with individual extraction and IEF.

The gels were analysed by a 2-D Proteomic Imaging System (PerkinElmer) using software ProScan 4.0. The digital gel maps of different samples were analysed and compared by using software Progenesis Same Spots (Nonlinear Dynamics). Master gels were generated for each sample by matching all the available gels. Normalisation was carried out by determining the gain factor for each sample which can be modelled as

$y_{i,}/y'_i = 1/\alpha_k$ where y_i is the measured abundance of feature i on sample k, $1/\alpha_k$ is the gain factor for sample k and y'_i is the normalised abundance of feature i on sample k (Nonlinear Dynamics n.d.).

Protein identification by MS/MS: Protein spots of interest were excised from CBB stained 2-D gels and were identified by mass spectrometric peptide sequencing. To avoid the overlapping parts of closely related spots, the centre portions of each spot was sampled. The spots were analysed by Proteomics International Ltd Pty, UWA, Perth, Australia. Protein samples were trypsin digested and the resulting peptides were extracted according to standard techniques (Bringans et al. 2008). Peptides were analysed by electrospray ionisation mass spectrometry using the Ultimate 3000 nano HPLC system (Dionex) coupled to a 4,000 Q TRAP mass spectrometer (Applied Biosystems). Tryptic peptides were loaded onto a C18 PepMap100, 3 μm (LC Packings) and separated with a linear gradient of water/acetonitrile/0.1% formic acid (v/v).

Spectra were analysed to identify proteins of interest using Mascot sequence matching software (Matrix Science) with taxonomy set to Viridiplantae. All searches used the Ludwig NR. The software was set to allow 1 missed cleavage, a mass tolerance of ± 1.2 Da for peptides and ± 0.6 for fragment ions. The peptide charges were set at 1+, 2+ and 3+, and the significance threshold at $P<0.05$. Generally a match was accepted where two or more peptides from the same protein were present in a protein entry in the Viridiplantae database.

Results

Two-dimensional gel electrophoresis of protein extracted after reducing and denaturing

The solubilisation of wheat proteins in the reducing/denaturing conditions indicated that proteins in 70-100 kDa (pI 6-8) ; 32-37 kDa (pI 6.5- 8.0) ; 20-24 kDa (pI 6.5 - 7.0) and 10-15 kDa (pI 6.5-7.0) ranges were relatively unaffected by the baking process (**Figure 1 A**, **B**). This allowed the identity of proteins in the baked wheat bread to be assigned based on their electrophoretic mobility and compared to the flour (**Figure 1A**,**D**). The comparison of normal wheat bread and lupin-wheat bread indicated that specific high and low molecular weight wheat proteins that characterise important features of bread could be extracted from the lupin-wheat bread under denaturing plus reducing condition (**Figure 1 B**, **E**).

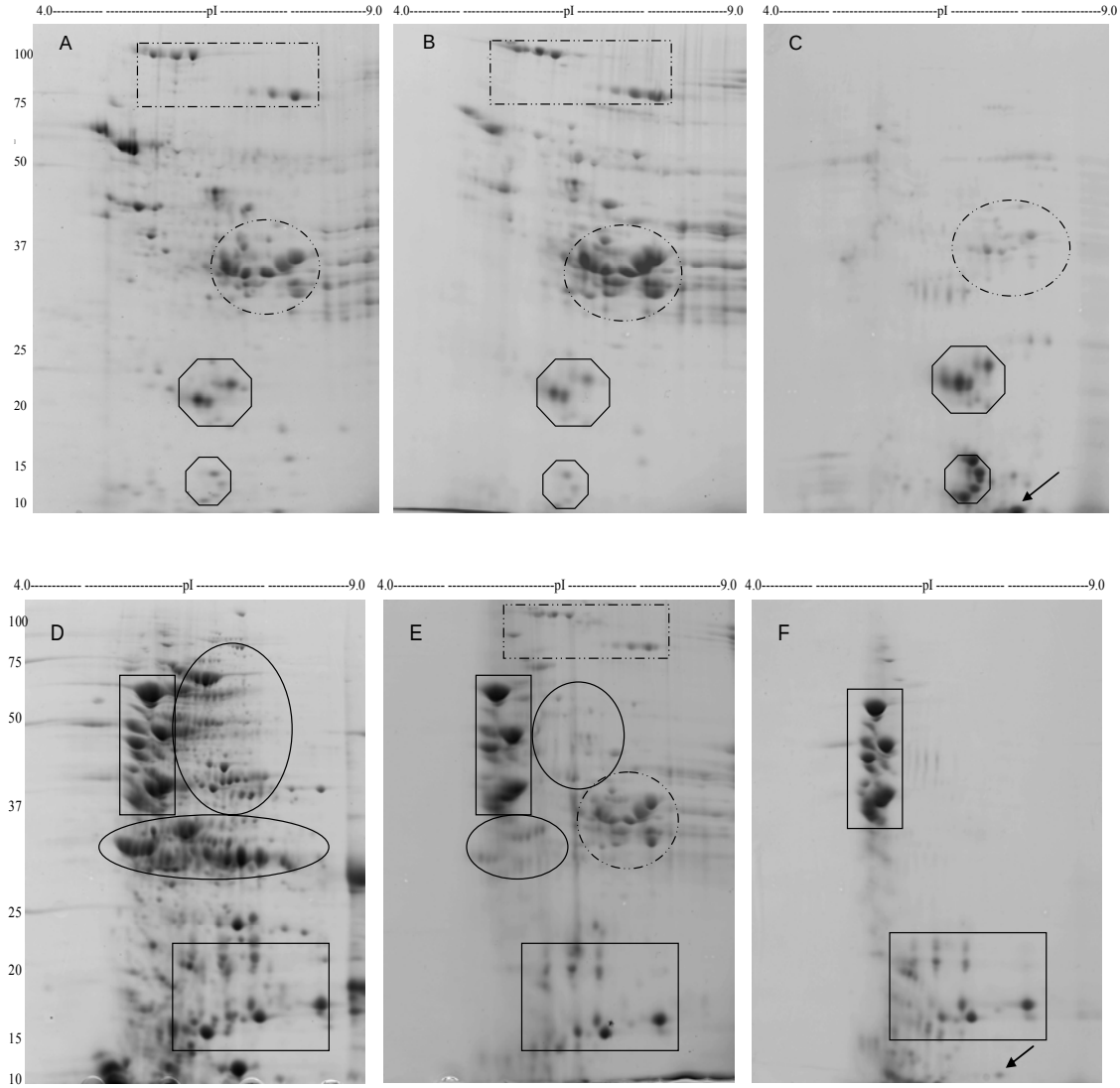


Figure 1: Protein profiles of lupin and wheat flours, lupin-wheat bread and normal wheat bread by two-dimensional gel electrophoresis. Proteins enclosed by rectangle with continuous line indicate the α-conglutins. Proteins enclosed by circles with continuous line indicate the β-conglutins. Proteins enclosed by rectangle with dotted line indicate the HMW glutens of wheat. Proteins enclosed by circles with dotted line indicate the wheat gliadins. The proteins enclosed by octagons are LMW wheat proteins. The lines and numbers on the top of each gel indicate the pH values and the numbers on the left of gels indicate molecular weight of the proteins in kDa. Panel A: Wheat flour with reduction and denaturing; B: Wheat bread with reduction and denaturing; C: Wheat bread without reduction and denaturing; D: Lupin flour with reduction and denaturing; E: Lupin-wheat bread with reduction and denaturing; F: Lupin-wheat bread without reduction and denaturing. The arrow marked is the only wheat protein extracted from lupin wheat bread with non-reducing and non-denaturing buffer.

In the case of lupin proteins, a wide range of proteins (10 – 100 kDa, pI 5-8.5) were found not to be extracted from the lupin-wheat bread. The largest group of proteins in this class was in the molecular mass range 27 – 33 kDa (pI 5-8) (**Figures 1 D**, **E**). In addition, some lupin proteins in the 37 - 85 kDa range (pI 6-7.5) were not extracted from lupin-wheat bread. Most of these proteins were β-conglutins as identified by mass spectrometry (**Table 1**; **Figure 2**) and comparisons to previous studies in *Lupinus angustifolius* (Goggin et al. 2008; Sirtori et al. 2010).

A group of proteins from lupin (molecular weight ranges 35 - 70 kDa, pI 5-6) was clearly extracted from lupin-wheat bread (**Figures 1 D**, **E**). This group of lupin proteins plus the entities in the molecular weight range 15 - 22 kDa (pI 6 - 8) were identified as α-conglutin (**Table 1**; **Figure 2**). In addition, a few proteins corresponding to the β-conglutin with a molecular mass range 32-37 kDa (pI 4.5-5.5) and 45-60 kDa (pI 5.5-7) could also be extracted from lupin-wheat bread (**Figures 1 D**, **E**). At least three γ-conglutins (**Table 1**; **Figure 2**) were extracted from lupin-wheat bread.

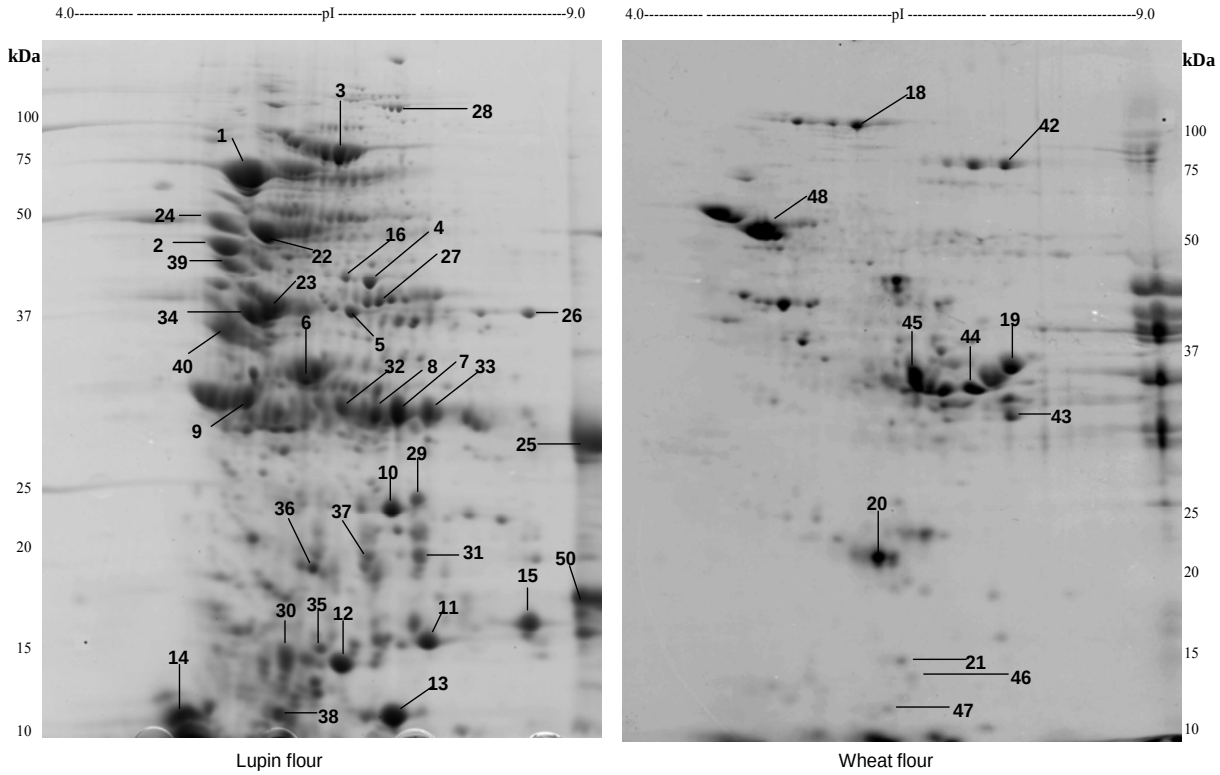


Figure 2: Two-dimensional gel map of lupin and wheat flour showing the protein spots having significance in lupin-wheat bread which were identified by mass spectrometry.

Table 1: Identification of the Lupin and Wheat Proteins from 2-D Gels by Mass Spectrometry.

Origin of proteins	Extractability from lupin-wheat bread	Spot Numbers[1]	Protein Identification by MS/MS	Matching NCBInr accession (GI)
Lupin proteins	Extracted at both reducing /denaturing and non-reducing/non-denaturing condition	1	Not clearly identified, hold the peptide TLTSLDFPILR which is part of alpha conglutin	
		23^{a+},31^{a+}, 36^{a+},37^{a+}	α-conglutin	2313076
		11^{a}, 12^{c},15^{c},22^{a},30^{c},34^{b},35^{b}	α-conglutin (legumin like seed storage protein)	224184735
		24^{c}, 40^{c}	α-conglutin (legumin like protein)	85361412
		2^{b}, 39^{a+}	β- conglutin	149208401
		13^{b},38^{b}	γ-Conglutin	662366
	Extracted at reducing/denaturing condition	29^{b}	α-conglutin	2313076
		3^{b}, 9^{b},16^{c}	β- conglutin	149208401
		6^{a+},7^{b},8^{b}, 32^{b}	β- conglutin	169950562
		27^{a+}	β- conglutin	149208403
		25^{b}	γ-Conglutin	662366
	Not extracted	4^{b}, 5^{b}, 50^{a}	β- conglutin	149208403
		10^{b}	β- conglutin	46451223
		33^{b}	β- conglutin	169950562
		26^{a+}	Glyceraldehyde-3 phospahte dehydrogenase	62816190
		14^{b}	δ-Conglutin	116181
		28^{c}	γ-Conglutin	662366
Wheat Protein	Extracted at reducing/denaturing condition	18^{b}	HMW glutenin	269849175
		42^{a}	HMW glutenin	162415983
		19^{c}	Gamma gliadin	209971935
		43^{c}	Alpha gliadin	282721196
		45^{a}	Glyceraldehyde-3-phosphate dehydrogenase	148508784
		48^{a+}	Beta amylase	32400764
		21, 44,46,47	Not Identified	N/A
	Not extracted	20^{b}	Superoxide dismutase	1654387

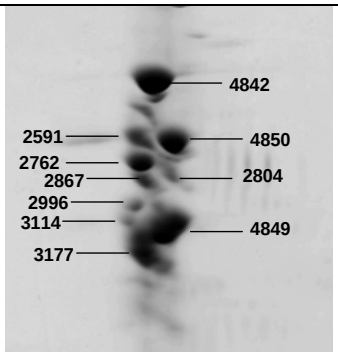
[1]Proteins were obtained from 2-D gels as shown in the **Figure 2**. Spots from all protein groups significant to lupin-wheat bread were selected Sequence coverage, a+: >30%; a: 20-29%, b: 10-19%; c: 4-9%

* β- conglutin and α-conglutin are sometimes described as vicilin-like proteins and legumin-like proteins respectively, but we have used β-conglutin and α-conglutin consistently to avoid the confusion.

The qualitative difference between the 2-D gels (**Figures 1 D**, **E**) indicating that the α-conglutins group of proteins were relatively more prominent in the protein complement extracted from bread relative to flour could be quantitated using standard software 'Progenesis Samespot'. In **Table 2** the output confirmed that the α-conglutins group at 35 to 70 kDa were relatively more prominent in the protein from lupin-wheat bread compared to lupin flour.

Table 2: Differential protein intensity of the α-conglutins group (35-70kda) between lupin flour and lupin-wheat bread that quantitated the qualitative difference between the 2-d gels

Spot number	Anova (p)	Average Normalised volumes ($x10^6$) Lupin-wheat bread	Lupin flour
2804	0.003	13.49	2.08
2996	0.017	14.16	5.25
3114	0.024	25.19	10.25
4849	0.026	102.70	38.27
3177	0.031	86.36	29.51
4842	0.040	78.13	33.28
2867	0.045	17.66	9.39
2591	0.065	41.46	20.12
2762	0.089	40.98	22.62
4850	0.214	73.42	43.70



Two-dimensional gel electrophoresis of protein extracted without reducing and denaturing (0.5 M NaCl extraction)

The extent of binding within the bread matrix was examined using a milder extraction procedure (0.5 M NaCl). As expected under these conditions, the HMW glutenin subunit proteins were not extracted (**Figure 1 C**) from wheat bread because it is well known that these proteins are bound within the bread gluten matrix by disulfide bonding (via the cysteine amino acids). In contrast the wheat gliadins (32 to 37 kDa molecular weight range with pI 6.5 to 8.0) and low molecular weight protein groups (20 to 24 kDa molecular weight range with pI 6.5 to 7.0 and 10 to 15 kDa molecular weight range with pI 6.5 to 7.0) were extracted under this condition (**Figure 1 C**). In the lupin-wheat bread the above mentioned groups of wheat protein all became unextractable (**Figures 1 C, F**). The only indication of a wheat protein being extractable came from the present of the faint spot at 10kDa, pI 7.5 (**Figures 1 C**, **F**).

In the case of lupin proteins, all the β-conglutins became unextractable in lupin-wheat bread (**Figures 1 D**, **F**) using 0.5 M NaCl. In contrast, most α-conglutins (molecular weight range 35 - 70 kDa at pI 5-6; molecular weight range 15 - 22 kDa at pI 6 to 8) and

two γ-conglutin appeared as extractable as observed for the reducing and denaturing condition (**Figure 1 F**, **Table 1**).

Two-dimensional gel protein identification by mass spectrometry (MS/MS)

Mass spectrometric peptide sequence of the proteins of interest from 2-D gel provided a successful identification of the proteins (**Table 1**). In total, 38 lupin proteins and 11 wheat proteins representing the major protein groups of lupin-wheat bread were studied. Most of the proteins (45 out of 49) were matched positively with respective proteins in the database. Some protein spots such as the number 2 and 39 (**Figure 2**) were reported as α-conglutin previously but in our identification they were more closely matched with β-conglutins. Moreover, one very dominant lupin protein, shown as spot number 1 in **Figure 2** was not clearly identified even though it was claimed to be an α-conglutin (Sirtori et al. 2010). The peptide sequence of this protein from both the lupin flour and lupin-wheat bread gels provided 11 common peptides (**Table 3**) which could not be assigned to known conglutin protein. In contrast, another protein (spot number 22) (**Figure 1**) could be clearly identified as α-conglutin based on the common peptides (**Table 3**) of protein sequences from lupin flour and lupin-wheat bread gels. Most of the β and γ conglutins showed relatively higher sequence coverage of peptides to known corresponding proteins compared to α-conglutin due to lack of information regarding this group of protein in the database.

Direct mass spectrometric (MALDI-TOF) survey of protein changes during bread baking

In the present research work a survey of protein modification was carried out using mass spectrometry (MALDI-TOF) for a higher resolution study of the low molecular weight proteins using non-reducing and non-denaturing (0.5 M NaCl) extraction. Many of the lupin and wheat proteins extracted from flour could be identified in the extracts from lupin-wheat bread (**Figure 3 A**). A group of proteins in the molecular weight range 13-16 kDa were very similar in the lupin flour and lupin-wheat bread extractions (**Figure 3 A**) and may correspond to the small sub-unit of α-conglutin based on the comparison of 2-D gels and MALDI-TOF profiles of the samples but further investigation of this possibility is required. Another group of proteins in the molecular weight range 20-22 kDa was extracted from the lupin-wheat bread and were the same as those found in lupin flour, perhaps corresponding to the basic sub-unit of α-conglutin (Duranti et al. 2008; Sirtori et al. 2010).

Table 3: List of the Common Peptides as Identified by MS/MS Peptide Sequencing Among the Same Proteins from Lupin Flour and Lupin-Wheat Bread

Common peptides* of protein spot number 1 (see Figure 2) from lupin flour and the corresponding spot from lupin-wheat bread	**Common peptides of protein spot number 22 (see Figure 2) from lupin flour and the corresponding spot from lupin-wheat bread**
AGPVR	AGMPK
ASLKVGEEEEEEEAGDGR	FYLAGNPEEEYPETQQQR
CAGQGR	GDEGQEEEETTTTTEER
CGAKVEFK	GGHEEEEVEEER
EQLATFR	GGHEEEEVEEERGR
GISILRR	GGKDH
IRNQEEFEQEIGR	GKPSESGPFNLR
KPSSPK	GSVVLSERGDGAAVPR
KYETTEQGR	HTRGDEGQEEEETTTTTEER
NKMSVIPYASAIGSIMYAMLCTR	IVEFQSNPNTLILPK
XEEXR	KGKPSESGPFNLR
	KITMPSSTQGFTY
	LLGFGINANENQR
	NFLAGSEDNVIR
	NNILSGFDPQFLSQALNIDEDTVHK
	NTLEATFNTR
	NTLEATFNTRYEEIQR
	QIIRVEEGLGVISPK
	QRVDTYWDLLSPK
	RFYLAGNPEEEYPETQQQR
	RGQEQSYQDEGVIVR
	TNRLENLQNYR
	VEEGLGVISPK
	YQAMKAGPDGEVVSLR

*These peptides could not be assigned to known conglutin protein.

At least five lupin proteins (at the molecular mass range 7 kDa to 9 kDa) corresponding to the large sub-unit of δ-conglutin (Duranti et al. 2008) were not extracted from the lupin-wheat bread (**Figure 3 B**). Similarly, two lupin proteins (24300 and 25860 Da) corresponding to the intermediate molecular weight sub-unit of β-conglutin (Duranti et al. 2008) were not extracted from the lupin-wheat bread. In total, 14 lupin proteins out of 24 became extractable from lupin-wheat bread (**Table 4**). In the case of wheat protein 25 out of 38 (**Table 4**, **Figure 3 A**) were found as extractable from the lupin-wheat bread.

In addition some new protein entities were identified (**Figure 3 C Table 4**). Proteins in the molecular weights of 39030, 39230 (**Figure 3 C**) and 19760 Da (**Table 4**) were only present in the lupin-wheat bread and could not be identified in either lupin or wheat flour or wheat bread.

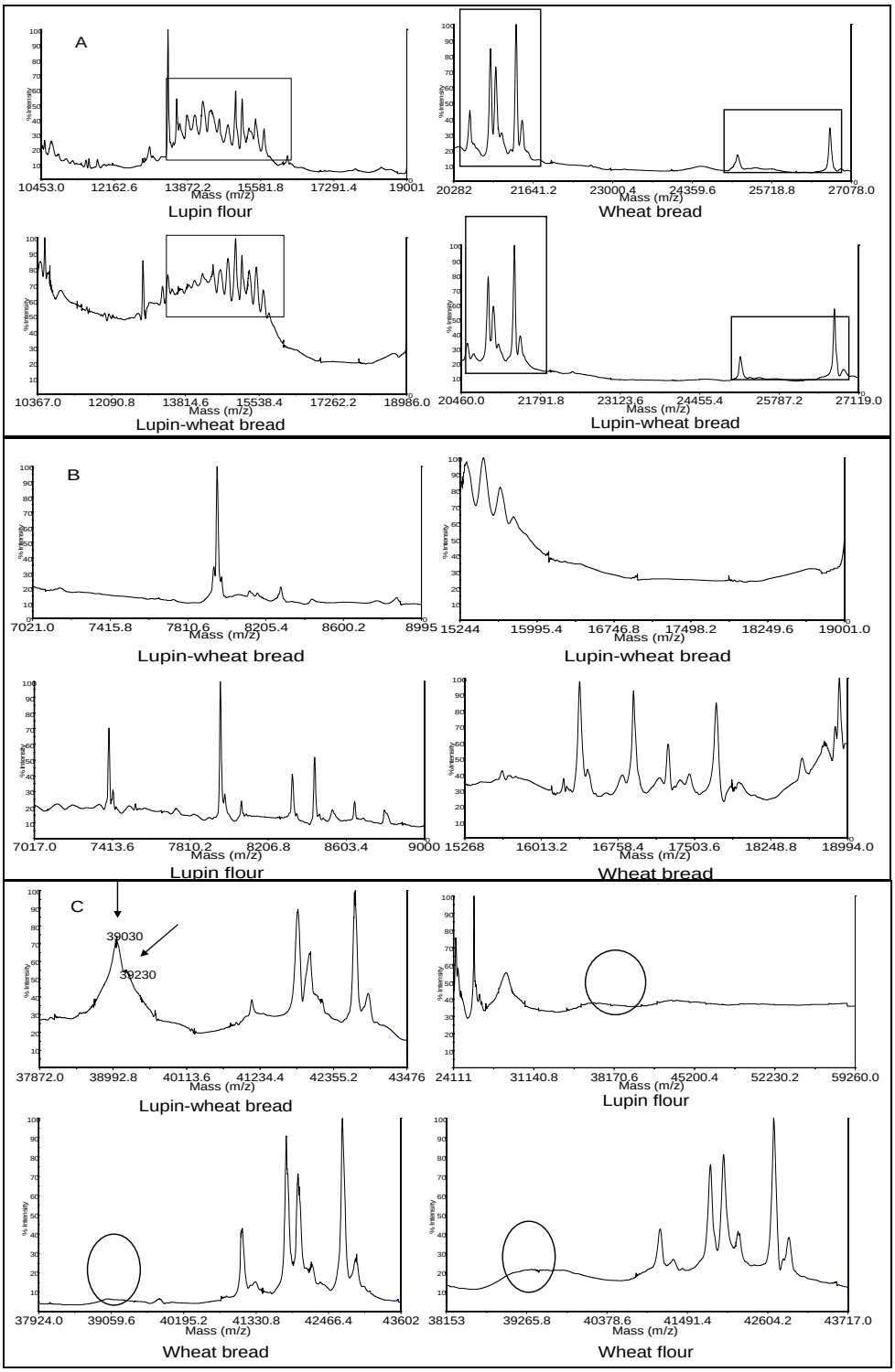


Figure 3: MALDI-TOF protein profiles of non-reduced and non-denatured (0.5 M NaCl) extracts showed different extractability of proteins from lupin-wheat bread.
A: The boxes are showing some readily extractable lupin and wheat proteins from the lupin-wheat bread.
B: Showing part of lupin and wheat proteins which could not be extracted from lupin-wheat bread.
C: Proteins indicated by arrows are unique to lupin-wheat bread and could not be identified in lupin or wheat flour or wheat bread (as the corresponding molecular weight range showed by the circles).

Table 4: Comparative List of the Proteins as Identified by Direct Mass Spectrometry (MALDI-TOF) from Lupin-Wheat Bread, Normal Wheat Bread and Lupin Flour Following Non-reducing and Non-denaturing (0.5 M NaCl) Extraction.

Lupin-wheat bread	Wheat bread	Lupin flour	Lupin-wheat bread	Wheat bread	Lupin flour
-	51430	-	19550	19550	-
-	50180	-	19400	19408	19400
-	49000	-	17240	17240	-
42880	42880	-	-	17716	-
42690	42690	-	-	17247	-
42203	42203	-	-	16913	-
41990	41990	-	-	16387	-
41810	41810	-	15600	-	15600
41300	41300	-	15400	-	15400
41110	41110	-	15300	-	15300
39030	-	-	15100	-	15100
39230	-	-	14900	-	14900
33940	33940	-	14800	-	14800
31270	31270		14600	-	14600
-	-	28000	13417	-	13417
26900	26900	-	-	11596	-
26710	26710	-	-	-	8645
-	-	25860	-	-	8535
25130	25130	-	-	-	8442
-	-	24300	-	-	8329
21450	21450	-	-	-	8070
21349	21349	21300	8290	8290	-
-	21100	-	7960	-	7960
21000	21000	21000	-	-	7396
20900	20900	20900	-	7367	-
-	-	20845	-	7174	-
20560	20560	-	-	7032	-
20020	20020	-	-	6992	-
19880	19880	-	6600	-	6600
19760	-	-	5363	5363	-
19710	19710	-			

Proteins are presented by their molecular weight in Da and '-' indicates the absence of protein to the corresponding sample.

Discussion

Recent reports have demonstrated that specific lupin protein bioactivities provide some health attributes. For example, γ-conglutin is capable of interacting with the mammalian protein hormone insulin and has effects in lowering blood glucose (Magni et al. 2004). The other reported bioactivities of lupin protein include plasma cholesterol and triglyceride lowering effects (Sirtori et al. 2004), anti-hypertensive properties (Pilvi et al. 2006), and angiotensin converting enzyme (ACE) inhibitory activity (Yoshie-Stark et al. 2004). Diversity in extractability of different conglutins (lupin proteins) from lupin-wheat bread as reported in the present study could account for some of the health attributes of these novel breads.

Two-dimensional gel electrophoresis was used as the primary basis for the protein analysis in this study. This allowed the broad range of intact proteins to be assayed and their extractability, as affected by baking, to be examined (Goggin et al. 2008; Magni et al. 2007; Sirtori et al. 2010). The MALDI-TOF based analyses provided a high resolution analysis of the lower molecular weight proteins. The results showed that many of the proteins from both lupin and wheat remained qualitatively unchanged (judged by their electrophoretic behaviour) in baked lupin-wheat bread while others, such as the β-conglutins, were incorporated into the bread matrix and could not be extracted.

Decrease of protein extractability due to baking

The demonstrated decrease of both the lupin and wheat protein extractability caused by baking relates to the high temperatures during baking and the presence of a range of ingredients that would enhance the interaction and formation of protein linkages into a matrix (Xu et al. 1999). Up to 80% loss of protein extractability due to baking in wheat bread has been reported (Singh 2005). Disulfide bonds were considered to be the major basis of the protein-protein interaction in baking and associated loss of extractability (Singh 2005) The lack of extraction of β-conglutins from bread even under reducing and denaturing condition indicates that covalent bonds other than disulfide bonds are involved in linking these proteins in to the very high molecular weight protein matrix. For example the reduction of protein extractability could result from tyrosine cross-linking among the proteins (Tilley et al. 2001). It has also been noted that lupin-wheat heteropolymer formation could be catalysed by transglutaminase (Bonet et al. 2006). Protein binding to starch, as reported for soy protein interaction (Ryan and Brewer 2007) could also alter the extractability of the β-conglutins.

Lupin and wheat protein interaction

The current study revealed lupin and wheat protein interaction during the baking of lupin-wheat bread. The added proteins from lupin flours have been suggested to interact with the wheat gluten network by cross-linking (Doxastakis et al. 2002; Roccia et al. 2009). In studies of the gluten network in soy-wheat breads, the soy protein retained in this network was held by strong association among proteins (Ribotta et al. 2005b).

The wheat proteins extracted from wheat bread indicate that they were unaffected by baking process and were expected to be extractable from lupin-wheat bread. However, a

number of wheat proteins observed in wheat bread became unextractable from lupin-wheat bread. This observation indicates the occurrence of interactions of lupin protein with certain wheat proteins to produce a new polymer of high molecular weight that have a low extractability. In addition, some unchanged (judged by their electrophoretic behaviour) proteins from lupin flour, wheat bread and lupin-wheat bread indicated that not all proteins are involved in such interactions. Some soy proteins have been reported not to interact with wheat protein in bread baking (Ribotta et al. 2005b) which supports our observation.

The analysis of the low molecular weight proteins using MALDI-TOF showed 10 lupin proteins those were extracted from lupin flour could not be extracted from lupin-wheat bread and hence were deduced to be embedded in the bread matrix. In contrast 3 new peaks were appeared in lupin-wheat bread and are speculated to be new addition of either lupin-lupin or lupin-wheat protein interaction and are currently under further investigation.

Two subgroups of conglutins

In this study the two major sub groups of conglutins showed very different solubilities after the baking process and were identified as α-conglutin and β-conglutin by mass spectrometry. Under both reducing/denaturing condition and at non-reducing/non-denaturing condition most of the β-conglutins were not extracted from the lupin-wheat bread. These findings indicate that this sub group of proteins is very sensitive to the baking process and are incorporated into a very high molecular weight matrix. The loss of extractability of β-conglutins from lupin-wheat bread suggested that they may not be available for generating an allergic response. However the fact that most of the β-conglutins have been reported as allergenic proteins (Goggin et al. 2008) indicates that the degree of binding into the bread matrix may not necessarily relate to allergenicity. It is however possible that the level of extraction of protein into a range of different buffers may provide a guide to their biological availability with respect to food attributes. For example the easily extracted α-conglutins and one γ-conglutin are predicted to be more accessible for biological effects associated with the consumption of lupin-wheat products.

On the basis of thermal denaturation experiments, it has been suggested that α-conglutins (legumin like) proteins were significantly more stable to denaturation than β-

conglutins (vicillin like) protein (Sirtori et al. 2010). The structural features of these protein classes which may therefore account for differences in extractability could relate to difference in thermal stability. The high temperature during bread-making could denature the vicillin like proteins β-conglutins more readily and provide greater opportunities for the incorporation into the complex bread matrix. In contrast the more thermally stable legumin like α-conglutins would lead to greater retention of a folded structure and reduced exposure of internal amino acids to cross-linking and a relative independence from the lupin wheat bread matrix. The relative independence of the legumin like α-conglutins from the bread matrix would be consistent with being able to extract this protein class under mild condition. Alternatively, some (currently unknown) feature of the amino acid sequences of α and β conglutins could result in large differences in proteolytic degradation during the breadmaking process.

Mass spectrometric identification of lupin-wheat bread proteins

Peptide sequencing of proteins by MS/MS showed positive identification for most of the β, γ and δ and a number of α-conglutins. Sequences of α-conglutins are still lacking in the database and a single low molecular weight (131 amino acids; NCBInr 2313076) entity is available. As a result, additional information based on EST sequences was used for the identification of the α-conglutins. Four protein spots located in the predominantly α-conglutin region (**Figure 1 D**) (molecular weight range 35-70 kDa) were matched to legumin-like (NCBInr 85361412) and seed storage (NCBInr 224184735) proteins of lupin.

As expected, mass spectrometric peptide sequence of the same proteins from lupin flour and lupin-wheat bread provided a number of common peptides (**Table 3**) and confirmed the identification of the proteins. One of the most striking α-conglutins (spot number 1) (**Figure 2**) showed 11 common peptides with its corresponding protein spot from lupin-wheat bread. However there was no match of these peptides in the data bases examined which suggests the necessity of further investigation.

Most of the wheat proteins (including HMW glutenins and gliadins) extractable from lupin-wheat bread were identified, although one group of wheat proteins (at 10 to 15 kDa molecular weight range with 6.5 to 7.0 pI) could not be identified due to the lack of sequence information in the database.

Among the lupin proteins, most of the α-conglutins were not affected by the baking process and could be readily extracted from the lupin-wheat bread even at non-reducing condition. In contrast, most of the β-conglutins appeared to undergo interactions in baking that resulted in loss of extractability which are speculated to be due to cross-linking and/or aggregation in the formation of large protein polymer. The higher thermal stability of α-conglutins relative to β-conglutins was suggested as one possible contribution to the distinct properties of these proteins because the unfolding of β-conglutins at lower temperature would allow more of the molecules to participate in network reactions.

Acknowledgements

The authors are grateful to Junhong Ma for her technical support to the laboratory works and to Rhonda Foley and Karam Singh, from CSIRO to supply EST sequences for lupin protein identification. Islam Shahidul thanks the University of Western Australia for a SIRF scholarship. The project was partially supported by the School of Plant Biology and the Faculty of Natural and Agricultural Sciences of UWA; Centre for Food and Genomic Medicine, WA and the Department of Agriculture and Food Western Australia.

Abbreviations

CHAPS: 3-[(3-cholamidopropyl)dimethylammonio]-1- propanesulfonate; DTT : dithiothreito; EST: expressed sequence tag; IEF: isoelectric focusing; IPG: immobilized pH gradient; MALDI-TOF: matrix-assisted laser desorption/ionisation mass spectrometry-time of flight; NCBI: national center for biotechnology information; SDS-PAGE: sodium dodecyl sulfatepolyacrylamide gel electrophoresis; Tris: tris(hydroxymethyl)- aminomethane

Supporting information

Details of the protein identification using mass spectrometric peptide sequence by matching with database is presented in **Appendix III**.

Chapter 8

General Discussion

8.1 Introduction

Since seed storage protein is the major component of nutritional attributes of lupin grain, the overall aim of this thesis was to deduce seed protein diversity and to characterise proteins related to health attributes. This aim has been achieved by a series of experiments using a set of proteomic technologies including high throughput MALDI-TOF mass spectrometry and two-dimensional gel electrophoresis followed by peptide sequencing for protein identification. The studies successfully produced and analysed protein profiles among species, cultivars and baked products that characterised lupin seed proteomes with potential application in breeding and the food industry.

8.2 Summary of major outcomes

The major outcome of this thesis is the successful application of MALDI-TOF mass spectrometry, two dimensional gel electrophoresis and MS/MS peptide sequencing to:

1. Reveal proteomic diversity across the lupin species and cultivars.
2. Recognise species and cultivar-specific protein mass peaks and rare protein mass peaks as protein markers.
3. Establish phylogenetic relationships across lupin species and cultivars based on seed protein profile.
4. Quantify influence of environment × genetic interactions on seed protein profile of lupin cultivars.
5. Characterise expression of allergenic proteins among potential lupin cultivars. and
6. Elucidate differential extraction of lupin proteins from baked product.

The outcomes of this thesis are discussed below in respect to their implication.

8.2.1 Pioneering MALDI-TOF for quantifying proteomic diversity in lupin

Matrix-Assisted Laser Desorption Ionization Time-Of-Flight (MALDI-TOF) mass spectrometric technology was used for seed protein profiling for the first time in lupin. Traditional PAGE-based proteomic technology, particularly one- and two-dimensional gel electrophoresis is a powerful tool for high resolution comparative proteomic analysis. However, the limitations of time and labour requirements, poor efficiency to study exact molecular weight, and difficulties in identifying very low molecular weight proteins make these techniques unfeasible for analysing large numbers of samples

common in breeding programs. Thus this thesis aimed to develop a fast, high throughput, proteomic approach for lupin seed protein profiling to assist breeding programs in improving proteomic attributes. MALDI-TOF mass spectrometry has high throughput capability that provides accurate molecular weight data which is relatively fast compared to other common separation methods (Chen et al. 2007).

Unlike previous investigations on lupin species which used traditional HPLC or one-dimensional gel electrophoresis (Przybylska and Zimniak-Przybylska 1995; Salmanowicz 1999; Zimniak-Przybylska and Przybylska 1997; Salmanowicz and Przybylska 1994) that does not recognise much polymorphism, this study used MALDI-TOF mass spectrometry which recognised considerable proteomic diversity and species-specific proteomic characters (Chapter 3). Previous seed proteomic studies in lupins included limited numbers of species form either Old World or New World (Przybylska and Zimniak-Przybylska 1995; Salmanowicz 1999; Zimniak-Przybylska and Przybylska 1997; Salmanowicz and Przybylska 1994). This study included species from both groups for wider comparison. The results indicate a higher seed proteomic diversity in the Old World smooth-seeded species than Old World rough-seeded species that is comparable to most of morphological (Gladstones 1974; 1984) and chemotaxonic (Williams *et al.* 1983) studies in lupin.

The phylogenetic relationship among species based on mass spectrometric fingerprinting that clustered Old World smooth- and rough-seeded and New World lupin species separately, broadly agrees with the clustering based on morphological characters, life history, cytogenetics, DNA polymorphism and electrophoreses of seed proteins (Ainouche and Bayer 1999; Christopher 2008; Dunn 1984; Naganowska et al. 2003). However, this study demonstrated unique clustering of two New World species *L. mutabilis* and *L. succulentus* with Old World species indicating seed proteomic diversity is not always compatible with other characteristics.

The other experiment revealed substantial seed proteomic diversity across narrow-leafed lupin cultivars (Chapter 4) released in Australia since 1968. Expression pattern of the rare mass peaks across the cultivars generally agree with their pedigree history (Cowling 1999), indicating that seed protein variation is predominantly genetically controlled. Phylogenetic relationships across narrow-leafed lupin cultivars—based on seed protein mass peak profiles—demonstrated two major clusters, generally

compatible with pedigree: (1) cultivars generated from early crosses comprising pre-wild and primary wild crosses, and (2) cultivars developed through more complex crosses with different wild types.

The results demonstrate that MALDI-TOF protein mass peak profile analysis is very efficient at characterising lupin seed proteomic diversity. The analysis also identified cultivar- and species-specific proteins and rare proteins which are useful markers for cultivar or variety identification. The phylogenetic relationships based on MALDI-TOF protein profiles among species and cultivars have potential application in lupin breeding for the slection of parents to improve protein attributes.

Although seed protein of legumes is largely genetically controlled (Bolon et al. 2010; Timmerman-Vaughan et al. 2005) different quality attributes including protein contents were influenced by environmental variation at different growing locations (Cowling and Tarr 2004). Thus knowledge of environmental effects on lupin seed protein expression will help breeders to use the proteomic diversity for genetic improvement of protein attributes (Cowling and Tarr 2004). Based on this understanding, environmental influence on seed protein mass peak profiles of five narrow-leafed lupin cultivars grown in three environmental conditions were investigated (Chapter 5) using MALDI-TOF mass spectrometry. The result demonstrate that lupin seed protein mass peak expression was significantly influenced by cultivar but not by environment, indicating that seed protein mass peak expression is mostly genetically controlled. This study suggests that protein variation across genotypes (Chapter 3 and 4) might be useful for breeding new lupin cultivars targeting specific proteins.

8.2.2 High resolution characterisation of lupin proteome for healthy food

Substantial seed proteomic diversity across narrow-leafed lupin cultivars revealed by MALD-TOF mass spectrometry (Chapter 4) suggests that a high resolution study of lupin cultivars would characterise expression of proteins associated with health-related issues. Two-dimensional gel electrophoresis followed by MS/MS peptide sequencing was used to characterise proteins (Chapter 6). Considering the high labour and time intensity requirement, only four narrow-leafed lupin cultivars were selected—Uniharvest, Tanjil, Yorrel and Coromup—based on mass spectrometric relationships (Chapter 4). The cultivars covered the groups of pre-wild crosses, primary-wild crosses and complex-wild crosses, identified by Cowling (1999).

This study revealed that previously-reported lupin allergenic proteins had distinct differential expression across cultivars, which may provide a new insight into the protein attributes of lupin cultivars with respect to allergenicity and an important target for further breeding. For instance, eight differentially-expressed proteins were previously identified (Goggin et al. 2008) as allergenic. Cultivar Tanjil has all eight allergenic proteins highly expressed while cultivars Yorrel and Coromup have only two with the lowest expression level. On the other hand, 19 of 24 differentially-expressed proteins were identified as β-conglutin, the largest seed storage protein group of lupin. The results indicate that β-conglutins are valuable breeding targets with more variation while α, δ and γ conglutins had lower levels of variation. Bioactivities of certain cultivars might be attributed to differentially-expressed β-conglutins that have been assumed to have biological activities (Duranti et al. 2008). Overall, differentially-expressed proteins such as cultivar-specific proteins will be valuable markers for cultivar identification and for screening parental lines of low allergenicity in breeding.

The experiment reported in Chapter 7 of the thesis focused on a proteomic investigation of a lupin baked product (lupin-wheat bread) having different health benefits where lupin protein had been identified as the main contributor (Lee et al. 2009; Lee et al. 2006). Protein interactions and formation of linkages into the bread-gluten matrix during dough preparation and baking could affect protein characteristics and extractability (Xu et al. 1999). Thus this thesis characterises lupin protein interactions in baking and reveales their differential extraction from lupin-wheat bread that could signify the relationships of these properties to possible health attributes.

The results demonstrate that qualitatively many proteins from both lupin and wheat flour remained unchanged after baking, as per electrophoretic behaviour, whereas some were incorporated into the bread gluten matrix and became unextractable. Disulfide bonds are considered the major basis of the protein–protein interaction in baking and associated loss of extractability (Singh 2005) . Two major sub-groups of lupin proteins, namely α-conglutin and β-conglutin, had very different solubilities after baking. Particularly, most α-conglutins (from a wider range of molecular weights) could be readily extracted from the lupin–wheat bread even at low salt (0.05 M Nacl) and non-reducing/non-denaturing extraction conditions. On the other hand, most β-conglutins could not be extracted from lupin–wheat bread indicating that they were trapped in the bread gluten matrix. A comparatively higher stability of α-conglutins than β-conglutins

in response to thermal denaturation (Sirtori et al. 2010) is speculated to account for this difference. The demonstrated loss of extractability of β-conglutins from lupin–wheat bread suggests that they may not be rapidly available for generating bio-activity of relevance for human health. In contrast, easily extractable α-conglutins and γ-conglutins are predicted to be more accessible for biological effects immediately following consumption of lupin–wheat products (before extensive digestion).

Detail proteome analyses of mature seed of *L. angustifolius* using high resolution PAGE based studies in this thesis were discussed (Chapters six and seven) in terms of compatibility of observations with some early researches on this species (Goggin et. al. 2008; Sirtori et. al. 2010). On the other hand, comparison of these results with several PAGE based studies on seed proteins of a closely related species, *L. albus* (Brambilla et. al. 2009; Wait et. al. 2005; Magni et. al. 2007; Goggin et. al. 2008) indicates that there is a difference in seed protein expression between these two species. However, the peptide sequence analysis indicates a reasonably close matching of proteins between the species as attributed by several identical tryptic peptides. For example, the amino acid sequence NTLEATFTNR was found in several proteins of both species. The presence of very close peptides such as ILLGNEDEQEDEEQR vs ILLGNEDEQEYEEQR in the two species also indicates a close relationship.

8.2.3 New diagnostic tools for proteins from baked products and for breeding

This study developed two protein-based diagnostic approaches: (1) to identify breeding material through MALDI-TOF mass spectrometric fingerprinting where specific protein mass peaks work as molecular markers, and (2) to identify the availability of specific proteins after baking as per their electrophoretic behaviour that could be used as a new quality assurance tool for baked products.

The results presented in Chapters 3 and 4 represent how MALDI-TOF mass spectrometric analysis can be used to produce lupin genotype (cultivar/species) specific and rare mass peaks. These mass peaks will be useful to identify appropriate plant material for breeding. For example, in the first step of a breeding program, fingerprinting probable parent material using this approach will allow breeders to understand the proteomic diversity level among crossing material and to fix target protein mass peaks for crossing. In the second step, offspring can be tested for target mass peaks to confirm their segregation pattern. Fast and high throughput capability of

MALDI-TOF mass spectrometry suggests this diagnosis process would be efficient at analysing large populations in a breeding program, something that is not feasible with traditional PAGE-based approaches.

This thesis also demonstrates a two-dimensional gel electrophoresis based diagnostic approach for identifying grain proteins from baked products (Chapter 7). Grain protein extractability and their ultimate accessibility to overall digestion might be affected during baking due to cross linking and protein degradation (Gerrard et al. 2001). Thus identifying proteins from the baked product is challenging, particularly when protein from different sources such as lupin and wheat are mixed in bread preparation. This research found that grain proteins—for example, lupin α-conglutins and wheat gliadins—can be easily recognised in 2-DGE gels of baked products and identified using MS/MS peptide sequencing. Clear visibility of lupin and wheat proteins in the gel image of bread and unique peptide sequences proved the potential of this approach as a protein diagnostic tool for baked products, particularly to assure protein quality.

8.3 Limitations of thesis and future directions

Given the lack of research into lupin proteomic diversity and protein analysis of baked products it is obvious that this thesis has raised questions requiring further research.

8.3.1 Technical challenges for high molecular weight proteins

This research demonstrated MALDI-TOF as a high throughput approach to characterise lupin proteomic diversity. However, the method is limited for high resolution characterisation of high molecular weight proteins, particularly in comparison with protein spots in two-dimensional gels following the same protein extraction method. The approach would be more efficient if it matched MALDI-TOF mass peaks and protein spots of two-dimensional gel. This may be possible if a new methodology compatible with MALDI-TOF is developed for high resolution study of high molecular weight lupin proteins. Although this research optimised methodology for lupin protein, based on a number of available technologies, there is still scope for an extensive effort.

It may also be possible to identify a peptide(s) that defines specific high molecular weight proteins using other technology, for instance, the ‘triple quadrupole mass spectrometer’. Eventually MALDI-TOF analysis may be used to characterise specific peptides which provide highly accurate identification and quantification of specific

proteins in the sample. The technique has been used successfully for human protein analysis (reviewed in Appels et al. 2012).

8.3.2 Inclusion of more lupin genotypes to characterise protein diversity

This study investigated a limited number of New World lupin species. There is scope for considering more lupin genotypes further confirm the proteomic diversity. This would allow for a mass spectrometry based fingerprint collection of a wide range of lupin genotypes that can be used as a source of genetic data for further improvement of proteomic attributes. Meanwhile, the considerable differential expression of allergenic proteins across four narrow-leafed lupin cultivars suggests further testing of more cultivars to find even lower allergenic lupins.

8.3.3 Approaches towards whole proteome association of lupin

A detailed proteomic fingerprinting study using MALDI-TOF technology can associate between variation in a specific protein mass peak and variation in a phenotype such as allergenicity. In a breeding program, this could gradually build an association as carried for DNA level genome wide association (GWA). On the other hand, one target of the modern genomic era is to link proteomics to genomics to make it more functional. In lupin, interpretation of MALDI-TOF-based proteomic data with genomic data could facilitate research towards this target. For example, MALDI-TOF-based protein markers can complement generally used DNA-sequenced-based markers mapped onto existing linkage maps. Sequentially this approach can identify specific chromosomal locations of some seed storage protein controlling genes of lupin that have been domesticated in wheat (Chen et al. 2007).

8.3.4 Sequencing lupin conglutins

This study identified a practical challenge of the lupin proteomic study which is the lack of amino acid sequences in the databases, particularly, of α-conglutins. Similar observations have been reported by Sirtori et al. (2010). More effort is needed to advance lupin protein sequencing to help accurately characterise proteins. To achieve this, amino acid sequences of narrow-leafed lupin conglutins need to be deduced by sequencing cDNA clones. Moreover, since the whole genome assembling of lupin has progressed, identification of the lupin genome for all protein coding genes would advance proteomic studies (Appels et al. 2012).

8.3.5 Characterising biological function of lupin seed protein

Like many other grain products, most lupin seed proteins are storage proteins with a central role in providing energy during germination and early stages of seedling establishment (Duranti et al. 2008). However, different biological functions of lupin proteina have been reported including plasma cholesterol and triglyceride lowering effects (Sirtori et al. 2004), anti-hypertensive properties (Pilvi et al. 2006), angiotensin converting enzyme (ACE) inhibitory activity (Yoshie-Stark et al. 2004). There are different classes of proteins (reviewed in the Chapter 2) which are assumed to have different biological roles (Magni et al. 2004). Although characterising the functional role specific to different protein classes is crucial to discuss their diversity in regard to bioactivity, the biological function of lupin-specific proteins has not been adequately investigated. Thus more information on biological function of lupin conglutins will be helpful for future diversity-based studies.

8.4 Conclusions

The outcomes of this thesis have advanced our understanding of the lupin seed proteome, particularly with respect to protein diversity, expressional behaviour of proteins and their response to baking. Considerable proteomic diversity across lupin species' and cultivars have been revealed by characterising genotype-specific protein mass peaks and by deducing phylogenetic relationships. This thesis also demonstrates the greater influence of genetic diversity, relative to environmental variation, on lupin seed protein expression. The information of proteomic diversity and the methodology will be useful in lupin breeding programs. The high resolution study using two-dimensional gel electrophoresis demonstrated differential expression of allergenic proteins among cultivars and differential recovery of lupin proteins from the baked product and their overall accessibility to digestion. These findings have potential in the food industry and will help in the understanding of the health benefits of lupin-enriched food.

ABARE (2011) Australian Crop report. Australian Bureau of Agricultural and Resource Economics. www.abare.gov.au. 2011

Agro products (2011) Lupins. http://www.agriculturalproductsindia.com/cereals-pulses/cereals-lupins.html.

Ainouche A-K, Bayer RJ (1999) Phylogenetic relationships in *Lupinus* (Fabaceae: Papilionoideae) based on internal transcribed spacer sequences (ITS) of nuclear ribosomal DNA. American Journal of Botany 86 (4):590-607

Ainouche A, Bayer RJ (2000) Genetic evidence supports the new Anatolian lupine accession, *Lupinus anatolicus*, as an Old World 'rough-seeded' lupine (section Scabrispermae) related to *L. pilosus*. Folia Geobotanica 35 (1):83-95

Ainouche A, Bayer RJ, Misset MT (2004) Molecular phylogeny, diversification and character evolution in *Lupinus* (Fabaceae) with special attention to Mediterranean and African lupines. Plant Systematics and Evolution 246 (3-4):211-222

Ainouche A, Greinwald R, Witte L, Huon A (1996) Seed alkaloid composition of *Lupinus tassilicus* Maire (Fabaceae: Genisteae) and comparison with its related rough seeded lupin species. Biochemical Systematics and Ecology 24 (5):405-414

Alberghina G, Cozzolino R, Fisichella S, Garozzo D, Savarino A (2005) Proteomics of gluten: mapping of the 1Bx7 glutenin subunit in Chinese Spring cultivar by matrix-assisted laser desorption/ionization. Rapid Communications in Mass Spectrometry 19 (14):2069-2074. doi:10.1002/rcm.2032

Altenbach SB, Kothari KM, Lieu D (2002) Environmental conditions during wheat grain development alter temporal regulation of major gluten protein genes. Cereal Chemistry 79 (2):279-285

Aly MA, Rashed NAK, Hassan HZ (2000) Genetic variability and assessment of relationships among some Egyptian rice (*Oriza sativa* L.) cultivars as revealed by SDS– protein and PCR–RAPD analyses. . Egyptian journal of genetics and cytology 29:71-87

Amaral Franco JD, Pinto da Silva AR (1978) *Lupinus L.* . In: Heywood VH (ed) Flora Europaea, vol 2. Cambridge University Press, London, UK,

Anderson GH, Tecimer SN, Shah D, Zafar TA (2004) Protein source, quantity, and time of consumption determine the effect of proteins on short-term food intake in young men. The Journal of Nutrition 134 (11):3011-3015

Appels R, Barrero R, Bellgard M (2012) Advances in biotechnology and linking outputs to variation in complex traits: Plant and Animal Genome meeting January 2012. Functional & Integrative Genomics 12 (1):1-9. doi:10.1007/s10142-012-0270-7

Archer BJ, Johnson SK, Devereux HM, Bax AL (2004) Effect of fat replacement by inulin or lupin-kernel fibre on sausage patty acceptability, post-meal perceptions of satiety and food intake in men. British Journal of Nutrition 91:591–599

Badr A, El-Shazly HH, El-Enim MA (2000) Seed protein diversity and its implications on the relationships in the genus *Lathyrus* L. (fabaceae). In: Proc 1st Conf Biol (ICBS), Faculty of Sci, Tanta Univ 1:284-300

Baldwin MA (2004) Protein identification by mass spectrometry. Molecular & Cellular Proteomics 3 (1):1-9. doi:10.1074/mcp.R300012-MCP200

Barakat H (2004) Genetic fingerprinting and relationships of six soybeans (*Glycine max* L.) cultivars based on protein and DNA polymorphism. International Journal of Agriculture and Biology 6 (5):877-883

Baranyi M, Greilhuber J, Swiecicki W (1996) Genome size in wild *Pisum* species. TAG Theoretical and Applied Genetics 93 (5):717-721. doi:10.1007/bf00224067

Barrett J, Brophy PM, Hamilton JV (2005) Analysing proteomic data. International Journal for Parasitology 35 (5):543-553

Bertoglio JC, Calvo MA, Hancke JL, Burgos RA, Riva A, Morazzoni P, Ponzone C, Magni C, Duranti M (2011) Hypoglycemic effect of lupin seed β-conglutin in experimental animals and healthy human subjects. Fitoterapia 82 (7):933-938

Bettzieche A, Brandsch C, Schmidt M, Weibe K, Eder K, Stangl IG (2008) Differing effect of protein isolates from different cultivars of blue lupin on plasma lipoproteins of hypercholesterolemic rats. Biosci, Biotechnol, Biochem 72 (12):3114-3121

Bhardwaj HL, Hamama AA, Merrick LC (1998) Genotypic and environmental effects on lupin seed composition. Plant Foods for Human Nutrition (Formerly Qualitas Plantarum) 53 (1):1-13. doi:10.1023/a:1008060813257

Blagrove RJ, Gillespie JM (1975) Isolation, purification and characterisation of the seed globulins of *Lupinus angustifolius*. Functional Plant Biology 2 (1):13-27. doi:http://dx.doi.org/10.1071/PP9750013

Blumenthal C, Bekes F, Gras PW, Barlow EWR, Wrigley CW (1995) Identification of wheat genotypes tolerant to the effects of heat stress on grain quality. Cereal Chemistry 72 (6):539-544

Boersma JG (2007) Contributions to the molecular genetics of the narrow-leaf lupin (*Lupinus angustifolius* L.)-mapping, marker development and QTL analysis. PhD Thesis, The University of Western Australia, Perth, Australia

Bolon Y-T, Joseph B, Cannon S, Graham M, Diers B, Farmer A, May G, Muehlbauer G, Specht J, Tu Z, Weeks N, Xu W, Shoemaker R, Vance C (2010) Complementary genetic and genomic approaches help characterize the linkage group I seed protein QTL in soybean. BMC Plant Biology 10 (1):41

Bonet A, Blaszczak W, Rosell CM (2006) Formation of homopolymers and heteropolymers between wheat flour and several protein sources by transglutaminase-catalyzed cross-linking. Cereal Chemistry Journal 83 (6):655-662. doi:10.1094/cc-83-0655

Bourgeois M, Jacquin F, Savois V, Sommerer N, Labas V, Henry C, Burstin J (2009) Dissecting the proteome of pea mature seeds reveals the phenotypic plasticity of seed protein composition. PROTEOMICS 9 (2):254-271. doi:10.1002/pmic.200700903

Boutraa T, Sanders FE (2001) Influence of water stress on grain yield and vegetative growth of two cultivars of bean (*Phaseolus vulgaris* L.). Journal of Agronomy and Crop Science 187 (4):251-257. doi:10.1046/j.1439-037X.2001.00525.x

Braly J, Hogganm R (2002) Dangerous grains. Why gluten cereal grains may be hazardous to your health. Penguin Group, New York, N.Y., USA

Brambilla F, Resta D, Isak I, Zanotti M, Arnoldi A (2009) A label-free internal standard method for the differential analysis of bioactive lupin proteins using nano HPLC-Chip coupled with Ion Trap mass spectrometry. Proteomics 9 (2):272-286. doi:10.1002/pmic.200800317

Brand TS, Brandt DA, Cruywagen CW (2004) Chemical composition, true metabolisable energy content and amino acid availability of grain legumes for poultry. South African Journal of Animal Science 34 (2):116-122

Breiteneder H, Radauer C (2004) A classification of plant food allergens. The Journal of allergy and clinical immunology 113 (5):821-830

Bressani R (1973) Nutritional improvement of food legumes by breeding: based on proceedings of a symposium sponsored by PAG, held at the Food and Agriculture Organization, Rome, Italy 3-5 July 1972, and PAG statement 22, "Upgrading human nutrition through the improvement of food legumes". In: Milner M (ed). Wiley,

Bringans S, Eriksen S, Kendrick T, Gopalakrishnakone P, Livk A, Lock R, Lipscombe R (2008) Proteomic analysis of the venom of *Heterometrus longimanus* (Asian black scorpion). Proteomics 8 (5):1081-1096. doi:10.1002/pmic.200700948

Buirchell BJ (2008) Narrow-leafed lupin breeding in Australia-where to from here? Paper presented at the Lupin for Health and Wealth:12th International Lupin Confrence, Frementle, Western Australia,

Bushuk W (1998) Wheat breeding for end-product use. Euphytica 100 (1):137-145. doi:10.1023/a:1018368316547

Bushuk W, Bekes F (2002) Contribution of protein to flour quality. In: Salgo A, Tomoskozi S, Lasztity R (eds) Novel raw materials, technologies and products– new challenge for the quality control. International Association for Cereal Science and Technology (ICC), Budapest, pp 14-19

Butow BJ, Ma W, Gale KR, Cornish GB, Rampling L, Larroque O, Morell MK, Békés F (2003) Molecular discrimination of Bx7 alleles demonstrates that a highly expressed high-molecular-weight glutenin allele has a major impact on wheat flour dough strength. TAG Theoretical and Applied Genetics 107 (8):1524-1532. doi:10.1007/s00122-003-1396-8

Capraro J, Magni C, Fontanesi M, Budelli A, Duranti M (2008) Application of two-dimensional electrophoresis to industrial process analysis of proteins in lupin-based pasta. LWT - Food Science and Technology 41 (6):1011-1017

Carrao-Panizz MC, Kwanyuen P, Erhan SZ, Lopes IdON (2008) Genetic variation and environmental effects on beta-conglycinin and glycinin content in Brazilian soybean cultivars. Pesq agropec bras, Brasilia 43 (9):1105-1114

Carvalho IS, Chaves M, Pinto Ricardo C (2005) Influence of water stress on the chemical composition of seeds of two lupins (*Lupinus albus* and *Lupinus mutabilis*). Journal of Agronomy and Crop Science 191 (2):95-98. doi:10.1111/j.1439-037X.2004.00128.x

Casero P, Duranti M, Cerletti P (1983) Heterogeneity of subunit composition in lupin globulins. Comparison of proteins from a single seed and from a number of seeds. Journal of the Science of Food and Agriculture 34 (10):1113-1116. doi:10.1002/jsfa.2740341011

Cerletti P, Fumagalli A, Venturin D (1978) Protein composition of seeds of *Lupinus albus*. Journal of Food Science 43 (5):1409-1411. doi:10.1111/j.1365-2621.1978.tb02504.x

Chen G, Pramanik BN (2009) Application of LC/MS to proteomics studies: current status and future prospects. Drug Discovery Today 14 (9-10):465-471

Chen J, Lan P, Tarr A, Yan YM, Francki M, Appels R, Ma W (2007) Matrix-assisted laser desorption/ionization time-of-flight based wheat gliadin protein peaks are useful molecular markers for wheat genetic study. Rapid Communications in Mass Spectrometry 21 (17):2913-2917. doi:10.1002/rcm.3151

Christopher SD (2008) Diversification of *Lupinus* (Leguminosae) in the western New World: Derived evolution of perennial life history and colonization of montane habitats. Molecular Phylogenetics and Evolution 48 (2):408-421

Clark R, Johnson S (2002) Sensory acceptability of foods with added lupin (*Lupinus angustifolius*) kernel fiber using pre-set criteria. Journal of Food Science 67 (1):356-362. doi:10.1111/j.1365-2621.2002.tb11410.x

Clarke K, Gorley R (2005) PRIMER v6. User manual/tutorial. Primer-E Ltd, Plymouth,UK

Conterato IF, Schifino-Wittmann MT (2006) New chromosome numbers, meiotic behaviour and pollen fertility in American taxa of *Lupinus* (Leguminosae): contributions to taxonomic and evolutionary studies. Botanical Journal of the Linnean Society 150 (2):229-240. doi:10.1111/j.1095-8339.2006.00443.x

Corbellini M, Canevar MG, Mazza L, Ciaffi M, Lafiandra D, Borghi B (1997) Effect of the duration and intensity of heat shock during grain filling on dry matter and protein accumulation, technological quality and protein composition in bread and durum wheat. Functional Plant Biology 24 (2):245-260. doi:http://dx.doi.org/10.1071/PP96067

Coruzzi GM, Zhou L (2001) Carbon and nitrogen sensing and signaling in plants: emerging "matrix effects". Current Opinion in Plant Biology 4 (3):247-253

Cowling WA (1999) Pedigrees and characteristics of narrow-leafed lupin cultivars released in Australia from 1967 to 1998. Lupin breeding Bulletin vol 4365. Agriculture Western Australia, Perth

Cowling WA, Huyghe C, Swiecicki W (1998) Lupin breeding. In 'Lupins as crop plants: biology, production and utilization'. JS Gladstones, C A Atkins, J Hamblin edn. CAB International, London

Cowling WA, Tarr A (2004) Effect of genotype and environment on seed quality in sweet narrow-leafed lupin (*Lupinus angustifolius* L.). Australian Journal of Agricultural Research 55 (7):745-751. doi:http://dx.doi.org/10.1071/AR03223

Cozzolino R, Giorgi SD, Fisichella S, Garozzo D, Lafiandra D, Palermo A (2001) Matrix-assisted laser desorption/ionization mass spectrometric peptide mapping of high molecular weight glutenin subunits 1Bx7 and 1Dy10 in Cheyenne cultivar. Rapid Communications in Mass Spectrometry 15 (10):778-787. doi:10.1002/rcm.299

Crespo JF, Rodriguez J, Vives R, James JM, Reano M, Daroca P, Burbano C, Muzquiz M (2001) Occupational IgE-mediated allergy after exposure to lupine seed flour. The Journal of allergy and clinical immunology 108 (2):295-297

Cristofolini G (1989) A serological contribution to the systematics of the genus *Lupinus* (*Fabaceae*). Plant Systematics and Evolution 166 (3):265-278. doi:10.1007/bf00935955

Cunsolo V, Foti S, Saletti R, Gilbert S, Tatham AS, Shewry PR (2004) Structural studies of the allelic wheat glutenin subunits 1Bx7 and 1Bx20 by matrix-assisted laser desorption/ionization mass spectrometry and high-performance liquid chromatography/electrospray ionization mass spectrometry. Journal of Mass Spectrometry 39 (1):66-78. doi:10.1002/jms.558

Czubaszek A, Karolini-Skaradzin´ska Z (2005) Effects of wheat flour supplementation with oat products on dough and bread quality. Polish Journal of Food and Nutrition Sciences 14 (3):281-286

Damodaran S, Wood TD, Nagarajan P, Rabin RA (2007) Evaluating peptide mass fingerprinting-based protein identification. Genomics, Proteomics & Bioinformatics 5 (3&4):152-157

Daniel C, Triboi E (2000) Effects of temperature and nitrogen nutrition on the grain composition of winter wheat: effects on gliadin content and composition. Journal of Cereal Science 32 (1):45-56

Dekker LJ, Dalebout JC, Siccama I, Jenster G, Sillevis Smitt PA, Luider TM (2005) A new method to analyze matrix-assisted laser desorption/ionization time-of-flight peptide profiling mass spectra. Rapid Communications in Mass Spectrometry 19 (7):865-870. doi:10.1002/rcm.1864

Derbyshire E, Wright DJ, Boulter D (1976) Legumin and vicilin, storage proteins of legume seeds. Phytochemistry 15 (1):3-24

Dervas G, Doxastakis G, Hadjisavva-Zinoviadi S, Triantafillakos N (1999) Lupin flour addition to wheat flour doughs and effect on rheological properties. Food Chemistry 66 (1):67-73

Dijkstra DS, Linnemann AR, van Boekel TAJS (2003) Towards sustainable production of protein-rich foods: appraisal of eight crops for Western Europe. part II:

Analysis of the technological aspects of the production chain. Critical Reviews in Food Science and Nutrition 43 (5):481-506. doi:10.1080/10408690390246332

Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. Science 312 (5771):212-217. doi:10.1126/science.1124619

Don C, Lichtendonk W, Plijter JJ, Hamer RJ (2003a) Glutenin macropolymer: a gel formed by glutenin particles. Journal of Cereal Science 37 (1):1-7

Don C, Lichtendonk WJ, Plijter JJ, Hamer RJ (2003b) Understanding the link between GMP and dough: from glutenin particles in flour towards developed dough. Journal of Cereal Science 38 (2):157-165

Dooper MMBW, Holden L, Faeste CK, Thompson KM, Egaas E (2007) Monoclonal antibodies against the candidate lupin allergens α-conglutin and β-conglutin. International Archives of Allergy and Immunology 143 (1):49-58

Doxastakis G, Zafiriadis I, Irakli M, Marlani H, Tananaki C (2002) Lupin, soya and triticale addition to wheat flour doughs and their effect on rheological properties. Food Chemistry 77 (2):219-227

Dunn DB, (1984) Cytotaxonomy and distribution of new world lupin species. In: 3rd International Lupin Conference, La Rochelle, France, International Lupin Association, pp 67-85

Dunn DB, Gillett J (1966) Lupines of Canada and Alaska. Queen's Press, Ottawa, Canada

DuPont FM, Altenbach SB (2003) Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. Journal of Cereal Science 38:133-146

DuPont FM, Altenbach SB, Chan R, Cronin K, Lieu D (2000). In: Shewry PR, Tatham AS (eds) Interactions between fertilizer, temperature and drought in determining flour composition and quality for bread wheat. Gluten Royal Society of Chemistry, Cambridge, UK, pp 488-491

DuPont FM, Chan R, Altenbach S, Hurkman WJ, Tanaka CK (1998). Effect of heat stress on flour composition and quality for several American bread wheats. In: Slinkard AE (ed) 9th International Wheat Symposium, Saskatoon, Saskatchewan, University Extension Press, University of Saskatchewan, pp 16–17

Duranti M (2006) Grain legume proteins and nutraceutical properties. Fitoterapia 77:67–82

Duranti M, Consonni A, Magni C, Sessa F, Scarafoni A (2008) The major proteins of lupin seed: characterisation and molecular properties for use as functional and nutraceutical ingredients. Trends in Food Science & Technology *19*:624-633

Duranti M, Horstmann C, Gilroy J, Croy RRD (1995) The molecular basis for N-glycosylation in the 11S globulin (legumin) of lupin seed. Journal of Protein Chemistry 14 (2):107-110. doi:10.1007/bf01888368

Duranti M, Restani P, Poniatowska M, Cerletti P (1981). The seed globulins of *Lupinus albus*. Phytochemistry 20 (9):2071-2075

Duranti M, Sessa F, Carpen A (1992) Identification, purification and properties of the precursor of conglutin β, the 7s storage globulin of *Lupinus albus* L. seeds. Journal of Experimental Botany 43 (10):1373-1378. doi:10.1093/jxb/43.10.1373

Eastwood RJ, Drummond CS, Schifino-Wittmann MT, Hughes CE (2008) Diversity and evolutionary history of lupins – insights from new phylogenies. Paper presented at the 12th International Lupin Conference, Fremantle, Australia,

Egana JI, Uauy R, Cassorla X, Barrera G, Yanez E (1992) Sweet lupin protein quality in young men. The Journal of Nutrition 122 (12):2341-2347

Eggum BO, Tomes G, Beames RM, Datta FU (1993) Protein and energy evaluation with rats of seed from 11 lupin cultivars. Animal Feed Science and Technology 43 (1&2):109-119

Eliasson AC, Larsson K (1993) Cereals in breadmaking: a molecular colloidal approach. Marcel Dekker, New York

Fahmy KH, Salama SI (2002) Biochemical and genetic finger prints for some soybean [*Glycin max* (L.) Merr.] cultivars resistant to *Etiella zinekenlla* (Triet.). Egyptian Journal of Genetics 31:309-329

FAOSTAT database (2011) FAO in http://faostat.fao.org/site/567/DesktopDefault.aspx.

Figuerola FE, Estevez AM, Castillo E (1987) Supplementation of wheat flour with chickpea (*Cicer arietinum*) flour. I. Preparation of flours and their properties for bread making. Arch Latinoam Nutr 37:378-387

Finnie C, Maeda K, Ostergaard O, Bak-Jensen KS, Larsen J, Svensson B (2004) Aspects of the barley seed proteome during development and germination. Biochemical Society Transactions 32:517–519

Foley R, Gao L-L, Spriggs A, Soo L, Goggin D, Smith P, Atkins C, Singh K (2011) Identification and characterisation of seed storage protein transcripts from *Lupinus angustifolius*. BMC Plant Biology 11 (1):59

Ford R (2008) The Gluten Syndrome. Is wheat causing you harm? RRS Global LT, Christchurch, New Zealand

French B, Shea G, Buirchell BJ (2008) Introduction and history. Producing lupins. Department of Agriculture and Food Western Australia, South Perth, W.A.

Fudiyansyah N, Petterson DS, Bell RR, Fairbrother AH (1995) A nutritional, chemical and sensory evaluation of lupin (*L. angustifolius*) tempe. International Journal of Food Science & Technology 30 (3):297-305. doi:10.1111/j.1365-2621.1995.tb01378.x

Galili G, Tang G, Zhu X, Gakiere B (2001) Lysine catabolism: a stress and development super-regulated metabolic pathway. Current Opinion in Plant Biology 4 (3):261-266

Gallagher E, Gormley TR, Arendt EK (2003) Crust and crumb characteristics of gluten free breads. Journal of Food Engineering 56 (2&3):153-161

Gallardo K, Job C, Groot SPC, Puype M, Demol H, Vandekerckhove Jl, Job D (2002) Proteomics of Arabidopsis seed germination. A comparative study of wild-type and gibberellin-deficient seeds. Plant Physiology 129 (2):823-837. doi:10.1104/pp.002816

Gallardo K, Le Signor C, Vandekerckhove Jl, Thompson RD, Burstin J (2003) Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. Plant Physiology 133 (2):664-682. doi:10.1104/pp.103.025254

Gdala J, Jansman AJM, van Leeuwen P, Huisman J, Verstegen MWA (1996) Lupins (*L. luteus, L. albus*, *L. angustifolius*) as a protein source for young pigs. Animal Feed Science and Technology 62 (2-4):239-249

Gerrard JA, Fayle SE, Brown PA, Sutton KH, Simmons L, Rasiah I (2001) Effects of microbial transglutaminase on the wheat proteins of bread and croissant dough. Journal of Food Science 66 (6):782-786. doi:10.1111/j.1365-2621.2001.tb15172.x

Gladstones JS (1970) Lupins as crop plants. Field Crop Abstract 2 (23):123-148

Gladstones JS (1974) Lupins of the Mediterranean region and Africa (trans: Agriculture ADo). vol technical bulletin26. Australian Department of Agriculture, Perth, Australia

Gladstones JS (1984) Present situation and potential of Mediterranean/African *Lupinus* for crop production. Paper presented at the 3rd International Lupine Conference, La Rochelle, France,

Gladstones JS (1994) A historical view of lupins in Australia. Proceedings of the First Australian Lupin Technical Symposium. Department of Agriculture Western Australia, Perth, Western Australia

Gladstones JS (1998) Distribution, origin, taxonomy, history and importance of lupins. Lupins as crop plants: Biology, Production and Utilization. CAB International, London

Gladstones JS, Crosbie GB (1979) Lupin wild types introduced into Western Australia to 1973: collection site data, preliminary ratings of field characteristics and disease ractions, and measurement of seed protein and oil contents. . Technical Bulletin, vol 43. Department of Agriculture of Western Australia, Perth, Australia

Goggin DE, Mir G, Smith WB, Stuckey M, Smith PMC (2008) Proteomic analysis of lupin seed proteins to identify conglutin β as an allergen, Lup an 1. Journal of Agricultural and Food Chemistry 56:6370-6377

Gorecka D, Lampart-Szczapa E, Janitz W, Sokolowska B (2000) Composition of fractional and functional properties of dietary fiber of lupines (*L. luteus* and *L. albus*). Food / Nahrung 44 (4):229-232. doi:10.1002/1521-3803(20000701)44:4<229::aid-food229>3.0.co;2-i

Guarneri F, Guarneri C, Benvenga S (2005) Identification of potentially cross-reactive peanut-lupine proteins by computer-assisted search for amino acid sequence homology. International Archives of Allergy and Immunology 138 (4):273-277

Guemes-Vera N, Arciniega RO, Davila OG (2004) Structural analysis of seed, flour, protein concentrates and isolates of *Lupinus mutabilis*, and their mixtures behavior with wheat flour. Lebensm Wiss U Technol 37:283–290

Guemes-Vera N, Pena BRJ, Jimenez MC, D´avila OG, Calderon DG (2008) Effective detoxification and decoloration of *Lupinus mutabilis* seed derivatives, and effect of these derivatives on bread quality and acceptance. Journal of the Science of Food and Agriculture 88:1135–1143

Guillamon E, Rodríguez J, Burbano C, Muzquiz M, Pedrosa MM, Cabanillas B, Crespo JF, Sancho AI, Mills ENC, Cuadrado C (2010) Characterization of lupin major allergens (*Lupinus albus* L.). Molecular Nutrition & Food Research 54 (11):1668-1676. doi:10.1002/mnfr.200900452

Gupta S, Buirchell BJ, Cowling WA (1996) Interspecific reproductive barriers and genomic similarity among the rough-seeded *Lupinus* species. Plant Breeding 115 (2):123-127. doi:10.1111/j.1439-0523.1996.tb00886.x

Hager JW (2002) A new linear ion trap mass spectrometer. Rapid Communications in Mass Spectrometry 16 (6):512-526. doi:10.1002/rcm.607

Hajduch M, Casteel JE, Hurrelmeyer KE, Song Z, Agrawal GK, Thelen JJ (2006) Proteomic analysis of seed filling in *Brassica napus*. developmental characterization of metabolic isozymes using high-resolution two-dimensional gel electrophoresis. Plant Physiology 141 (1):32-46. doi:10.1104/pp.105.075390

Hajduch M, Ganapathy A, Stein JW, Thelen JJ (2005) A systematic proteomic study of seed filling in soybean. Establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. Plant Physiology 137 (4):1397-1419. doi:10.1104/pp.104.056614

Hall RS, Thomas SJ, Johnson SK (2005) Australian sweet lupin flour addition reduces the glycaemic index of a white bread breakfast without affecting palatability in healthy human volunteers. Asia Pacific Journal of Clinical Nutrition 14:91-97

Hanelt P (1960) The Lupins. A Ziemsen, Wittenberg

He J, Gu D, Wu X, Chen J, Duan X, Chen J, Whelton PK (2005) Effect of soybean protein on blood pressure: a randomized, controlled trial. Annals of Internal Medicine 143 (1):1-9

He J, Whelton PK (1999) Effect of dietary fiber and protein intake on blood pressure: a review of epidemiologic evidence. Clinical and Experimental Hypertension 21 (5-6):785-796. doi:10.3109/10641969909061008

Hefle SL, Lemanske RF, Bush RK (1994) Adverse reaction to lupine-fortified pasta. The Journal of allergy and clinical immunology 94 (2):167-172

Heyn CC, Herrnstadt I (1977) Seed coat structure of old-world *Lupinus* species. Botaniska Notiser 130 (4):427-435

Hill GD (1986) Recent developments in the use of lupins in animal and human nutrition. . Paper presented at the 4th International Lupin Conference, Geraldton, Australia. ,

Hill GD (1998). In: Gladstones JS, Atkins C, Hamblin J (eds) Lupins as Crop Plant.s. CAB International, Wallingford Oxon, UK,

Hodgson JM, Burke V, Beilin LJ, Puddey IB (2006) Partial substitution of carbohydrate intake with protein intake from lean red meat lowers blood pressure in hypertensive persons. The American Journal of Clinical Nutrition 83 (4):780-787

Holden L, Sletten GBG, Lindvik H, Faeste CK, Dooper MMBW (2008) Characterization of IgE binding to lupin, peanut and almond with sera from lupin-allergic patients. International Archives of Allergy and Immunology 146 (4):267-276

Holt SHA, Delargy HJ, Lawton CL, Blundell JE (1999) The effects of high-carbohydrate vs high-fat breakfasts on feelings of fullness and alertness, and subsequent food intake. International Journal of Food Sciences and Nutrition 50 (1):13-28

Horneffer V, Foster TJ, Velikov KP (2007) Fast characterization of industrial soy protein isolates by direct analysis with matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Journal of Agricultural and Food Chemistry 55 (26):10505-10508. doi:10.1021/jf071321y

Hoseney RC (1986) Principles of Cereal Science and Technology. Am. Assoc.Cereal Chem., St Paul, MN

Hoveland CS, Townsend CE (1985) Other legumes. In: Heath ME, Barnes RF, Metcalfe DS (eds) Forages. Iowa State University Press, pp 146–153

Howieson JG, O'Hara GW, Carr SJ (2000) Changing roles for legumes in Mediterranean agriculture: developments from an Australian perspective. Field Crops Research 65 (2-3):107-122. doi:Doi: 10.1016/s0378-4290(99)00081-7

Hsieh JS, Hsieh KL, Tsai YC, Hsing YI (2001) Each species of *Glycine* collected in Taiwan has a unique seed protein pattern. Euphytica 118 (1):67-73

Hu S, Michels DA, Fazal MA, Ratisoontorn C, Cunningham ML, Dovichi NJ (2004) Capillary sieving electrophoresis/micellar electrokinetic capillary chromatography for two-dimensional protein fingerprinting of single mammalian cells. Analytical Chemistry 76 (14):4044-4049. doi:10.1021/ac0498314

Huebner FR, Nelsen TC, Chung OK, Bietz JA (1997) Protein distributions among hard red winter wheat varieties as related to environment and baking quality1. Cereal Chemistry Journal 74 (2):123-128. doi:10.1094/cchem.1997.74.2.123

Hughes C, Eastwood R (2006) Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. Proceedings of the National Academy of Sciences 103 (27):10334-10339. doi:10.1073/pnas.0601928103

Hung TV, Papolais M, Nithianandan V, Jiang HH, Versteeg K (1990) Utilisation of lupin seeds for human consumption. Paper presented at the 'Food Pacific' Convention, Gold Coast, Australia,

Islam N, Woo S-H, Tsujimoto H, Kawasaki H, Hirano H (2002) Proteome approaches to characterize seed storage proteins related to ditelocentric chromosomes in common wheat (*Triticum aestivum* L.). PROTEOMICS 2 (9):1146-1155. doi:10.1002/1615-9861(200209)2:9<1146::aid-prot1146>3.0.co;2-6

Islam S, Ma W, Ma J, Buirchell BJ, Appels R, Yan G (2011a) Diversity of seed protein among the Australian narrow-leafed lupin (*Lupinus angustifolius* L.) cultivars. Crop and Pasture Science 62 (9):765-775. doi:http://dx.doi.org/10.1071/CP11046

Islam S, Ma W, Yan G, Gao L, Appels R (2011b) Differential recovery of lupin proteins from the gluten matrix in lupin-wheat bread as revealed by mass spectrometry and two-dimensional electrophoresis. Journal of Agricultural and Food Chemistry 59 (12):6696-6704. doi:10.1021/jf201293p

Islam S, Yan G, Appels R, Ma W (2012) Comparative proteome analysis of seed storage and allergenic proteins among four narrow-leafed lupin cultivars. Food Chemistry 135 (3):1230-1238

Jenkins DJA, Leeds AR, Gassull MA, Cochet B, Alberti KGMM (1977) Decrease in postprandial insulin and glucose concentrations by guar and pectin. Annals of Internal Medicine 86 (1):20-23. doi:10.1059/0003-4819-86-1-20

Johansson E, Prieto-Linde ML, Jonsson JO (2001) Effects of wheat cultivar and nitrogen application on storage protein composition and breadmaking quality. Cereal Chemistry 78 (1):19-25

Johnson SK, Gray DM (1993) Ingredients derived from lupin - strong potential for a range of dietary fibre applications, vol 5. International Food Ingredients. Cmp Information Ltd London, UK

Johnson SK, McQuillan PL, Sin JH, Ball MJ (2003) Sensory acceptability of white bread with added Australian sweet lupin (*Lupinus angustifolius*) kernel fibre and its glycaemic and insulinaemic responses when eaten as a breakfast. Journal of the Science of Food and Agriculture 83 (13):1366-1372. doi:10.1002/jsfa.1552

Kass E, Wink M (1997) Molecular phylogeny and phylogeography of *Lupinus* (Leguminosae) inferred from nucleotide sequences of the rbcL gene and ITS 1 + 2 regions of rDNA Plant Systematics and Evolution 208 (3):139-167. doi:10.1007/bf00985439

Kenny S, Wehrle K, Stanton C, Arendt EK (2000) Incorporation of dairy ingredients into wheat bread: effects on dough rheology and bread quality. European Food Research and Technology 210 (6):391-396. doi:10.1007/s002170050569

KeShun L (1997) Soybeans: chemistry, technology and utilisation. Chapman and Hall, London, UK

Klos P, Poreba E, Springer E, Lampart-Szczapa E, zefiak AG (2010) Identification of a specific IgE-binding protein from narrow-leafed lupin (*L. angustifolius*) seeds. Journal of Food Science 75 (1):H39-H43

Kottapalli KR, Payton P, Rakwal R, Agrawal GK, Shibato J, Burow M, Puppala N (2008) Proteomics analysis of mature seed of four peanut cultivars using two-dimensional gel electrophoresis reveals distinct differential expression of storage, anti-nutritional, and allergenic proteins. Plant Science 175 (3):321-329

Kussmann M, Nordhoff E, Rahbek-Nielsen H, Haebel S, Rossel-Larsen M, Jakobsen L, Gobom J, Mirgorodskaya E, Kroll-Kristensen A, Palm‖ L, Roepstorff P (1997) Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes.

Journal of Mass Spectrometry 32 (6):593-601. doi:10.1002/(sici)1096-9888(199706)32:6<593::aid-jms511>3.0.co;2-d

Lampart-Szczapa E (1996) Preparation of protein from lupin seeds. Nahrung 40 (2):71 - 74

Lampart-Szczapa E, Korczak J, Nogala-Kalucka M, Zawirska-Wojtasiak R (2003) Antioxidant properties of lupin seed products. Food Chemistry 83 (2):279-285

Lasonder E, Ishihama Y, Andersen JS, Vermunt AMW, Pain A, Sauerwein RW, Eling WMC, Hall N, Waters AP, Stunnenberg HG, Mann M (2002) Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry. Nature 419 (6906):537-542

Leduc V, Moneret-Vautrin DA, Guerin L (2002) Allergénicité de la farine de lupin. Allerg Immunol 34:213-217

Lee YP, Mori TA, Puddey IB, Sipsas S, Ackland TR, Beilin LJ, Hodgson JM, (2009) Effects of lupin kernel flour–enriched bread on blood pressure: a controlled intervention study1–3 American Journal of Clinical Nutrition 89:1-7

Lee YP, Mori TA, Sipsas S, Barden A, Puddey IB, Burke V, Hall RS, Hodgson JM (2006) Lupin-enriched bread increases satiety and reduces energy intake Acutely. American Journal of Clinical Nutrition 84:975– 980

Lefebvre J, Mahmoudi N, (2007) The pattern of the linear viscoelastic behaviour of wheat flour dough as delineated from the effects of water content and high molecular weight glutenin subunits composition. Journal of Cereal Science 45:49-58

Liang XQ, Luo M, Holbrook CC, Guo BZ (2006) Storage protein profiles in Spanish and runner market type peanuts and potential markers. BMC Plant Biology 6 (1):24

Lindsay MPMP, Skerritt JHJH (1999) The glutenin macropolymer of wheat flour doughs - structure-function perspectives. Trends in Food Science and Technology 10 (8):247-253

Lindvik H, Holden L, Lovik M, Cvancarova M, Halvorsen R (2008) Lupin sensitization and clinical allergy in food allergic children in Norway. Acta Pædiatrica 97 (1):91-95. doi:10.1111/j.1651-2227.2007.00568.x

Liu L, Wang A, Appels R, Ma J, Xia X, Lan P, He Z, Bekes F, Yan Y, Ma W (2009) A MALDI-TOF based analysis of high molecular weight glutenin subunits for wheat breeding. Journal of Cereal Science 50 (2):295-301

Liu Z, Schey KL (2005) Optimization of a MALDI TOF-TOF mass spectrometer for intact protein analysis. Journal of The American Society for Mass Spectrometry 16 (4):482-490

Lou X, van Dongen JLJ, Meijer EW (2010) Generation of CsI cluster ions for mass calibration in matrix-assisted laser desorption/ionization mass spectrometry. Journal of the American Society for Mass Spectrometry 21 (7):1223-1226

Lqari H, Vioque J, Pedroche J, Millan F (2002) *Lupinus angustifolius* protein isolates: chemical composition, functional properties and protein characterization. Food Chemistry 76 (3):349-356

Lupin.org (2011) About Lupins. http://www.lupins.org/lupins/.

Ma W, Anderson O, Kuchel H, Bonnardeaux Y, Collins H, Morell MK, Langridge P, Appels R (2009) Genomics of quality traits. genetics and genomics of the triticeae. Springer, New York

Ma W, Appels R, Bekes F, Larroque O, Morell MK, Gale KR (2005) Genetic characterisation of dough rheological properties in a wheat doubled haploid population: additive genetic effects and epistatic interactions. TAG Theoretical and Applied Genetics 111 (3):410-422. doi:10.1007/s00122-005-2001-0

MacRitchie F (1980) Studies of gluten protein from wheat flours. Cereal Foods World 25:382-385

MacRitchie F (1987) Evaluation of contributions from wheat protein fractions to dough mixing and breadmaking. Journal of Cereal Science 6 (3):259-268

Magni C, Herndl A, Sironi E, Scarafoni A, Ballabio C, Restani P, Bernardini R, Novembre E, Vierucci A, Duranti M (2005) One- and two-dimensional electrophoretic identification of IgE-binding polypeptides of *Lupinus albus* and other legume seeds. Journal of Agricultural and Food Chemistry 53 (11):4567-4571. doi:10.1021/jf0500785

Magni C, Scarafoni A, Herndl A, Sessa F, Prinsi B, Espen L, Duranti M (2007) Combined 2D electrophoretic approaches for the study of white lupin mature seed storage proteome. Phytochemistry 68 (7):997-1007

Magni C, Sessa F, Accardo E, Vanoni M, Morazzoni P, Scarafoni A, Duranti M (2004) Conglutin γ, a lupin seed protein, binds insulin in vitro and reduces plasma glucose levels of hyperglycemic rats. The Journal of nutritional biochemistry 15 (11):646-650

Majoul T, Bancel E, Triboï E, Ben Hamida J, Branlard G (2003) Proteomic analysis of the effect of heat stress on hexaploid wheat grain: Characterization of heat-responsive proteins from total endosperm. PROTEOMICS 3 (2):175-183. doi:10.1002/pmic.200390026

Margiotta B, Colaprico G, D'Ovidio R, Lafiandra D (1993) Characterization of high mr subunits of glutenin by combined chromatographic (RP-HPLC) and electrophoretic separations and restriction fragment length polymorphism (RFLP) analyses of their encoding genes. Journal of Cereal Science 17 (3):221-236

Mariotti F, Pueyo ME, Tome D, Mahe S (2002) The bioavailability and postprandial utilisation of sweet lupin (*Lupinus albus)*-flour protein is similar to that of purified soyabean protein in human subjects: a study using intrinsically 15N-labelled proteins. 87:315-323

Martins JM, Riottot M, de Abreu MC, Viegas-Crespo AM, Lanca MJ, Almeida JA, Freire JB, Bento OP (2005) Cholesterol-lowering effects of dietary blue lupin (*Lupinus angustifolius* L.) in intact and ileorectal anastomosed pigs. Journal of Lipid Research 46 (7):1539-1547. doi:10.1194/jlr.M500129-JLR200

Maruyama N, Maruyama Y, Tsuruki T, Okuda E, Yoshikawa M, Utsumi S (2003) Creation of soybean β-conglycinin β with strong phagocytosis-stimulating activity. Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics 1648 (1-2):99-104

Matheu V, Manuel de Barrio V, Sierra Z, Gracia-Bara MT, Tornero P, Baeza ML (1999) Lupine-induced anaphylaxis. Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology 83 (5):406-408

Millan F, Alaiz M, Hernandez-Pinzon I, Sanchez-Vioque R, Bautista J (1995) Study of neutral lipids of *Lupinus mutabilis* meal and isolates. Journal of the American Oil Chemists' Society 72 (4):467-471. doi:10.1007/bf02636091

Monteiro S, Freitas R, Rajasekhar BT, Teixeira AR, Ferreira RB (2010) The unique biosynthetic route from *Lupinus* beta-Conglutin gene to blad. PLoS One 5 (1):11

Moore JC, DeVries JW, Lipp M, Griffiths JC, Abernethy DR (2010) Total protein methods and their potential utility to reduce the risk of food protein adulteration. Comprehensive Reviews in Food Science and Food Safety 9 (4):330-357. doi:10.1111/j.1541-4337.2010.00114.x

Moore MM, Schober TJ, Dockery P, Arendt EK (2004) Textural comparison of gluten-free and wheat based doughs, batters and breads. Cereal Chemistry 81:567-575

Moreno-Ancillo A, Gil-Adrados AC, Dominguez-Noche C, Cosmes PM (2005) Lupine inhalation induced asthma in a child. Pediatric Allergy and Immunology 16 (6):542-544. doi:10.1111/j.1399-3038.2005.00299.x

Morr C (1979) Technical problems and opportunities in using vegetable proteins in dairy products. Journal of the American Oil Chemists' Society 56 (3):383-385. doi:10.1007/bf02671510

Mubarak AE (2001) Chemical, nutritional and sensory properties of bread supplemented with lupin seed (*Lupinus albus*) products. Nahrung Food 45 (4):241-245

Muccilli V, Cunsolo V, Saletti R, Foti S, Masci S, Lafiandra D (2005) Characterization of B- and C-type low molecular weight glutenin subunits by electrospray ionization mass spectrometry and matrix-assisted laser desorption/ ionization mass spectrometry. PROTEOMICS 5 (3):719-728. doi:10.1002/pmic.200401029

Muller M, Dues G, Balconi C, Salamini F, Thompson RD (1997) Nitrogen and hormonal responsiveness of the 22 kDa α-zein and b-32 genes in maize endosperm is displayed in the absence of the transcriptional regulator Opaque-2. The Plant Journal 12 (2):281-291. doi:10.1046/j.1365-313X.1997.12020281.x

Muller M, Knudsen S (1993) The nitrogen response of a barley C-hordein promoter is controlled by positive and negative regulation of the GCN4 and endosperm box. The Plant Journal 4 (2):343-355. doi:10.1046/j.1365-313X.1993.04020343.x

Muntz K, Blattner FR, Shutov AD (2002) Legumains, a family of asparagine-specific cysteine endopeptidases involved in propolypeptide processing and protein breakdown in plants. journal of plant physiology 159:1287-1293

Murphy PA, Resurreccion AP (1984) Varietal and environmental differences in soybean glycinin and .beta.-conglycinin content. Journal of Agricultural and Food Chemistry 32 (4):911-915. doi:10.1021/jf00124a052

Naganowska B, Wolko B, Sliwinska E, Kaczmarek Z (2003) Nuclear DNA content variation and species relationships in the genus *Lupinus* (Fabaceae). Annals of Botany 92 (3):349-355. doi:10.1093/aob/mcg145

Ng PKW, Bushuk W (1987) Glutenin of Marquis wheat as a reference for estimating molecular weights of glutenin subunits by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. Cereal Chemistry 64:324-327

Nickols-Richardson SM, Coleman MD, Volpe JJ, Hosig KW (2005) Perceived hunger is lower and weight loss is greater in overweight premenopausal women consuming a low-carbohydrate/high-protein vs high-carbohydrate/low-fat diet. Journal of the American Dietetic Association 105 (9):1433-1437

Nonlinear Dynamics (n.d.) Progenesis samespots normalisation. http://www.nonlinear.com/support/progenesis/samespots/faq/normalisation.aspx . Accessed 15 March 2011

Nowacki E (1963) Inheritance and biosynthesis of alkaloids in lupin. Genetica Polonica 4:161–202

Olguin MC, Hisano N, D'Ottavio AE, Zingale MI, Revelant GC, Calderari SA (2003) Nutritional and antinutritional aspects of an Argentinian soy flour assessed on weanling rats. Journal of Food Composition and Analysis 16 (4):441-449

Oliver JR, Allen HM (1992) The prediction of bread baking performance using the farinograph and extensograph. Journal of Cereal Science 15 (1):79-89

Olsen JV, Mann M (2004) Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. Proceedings of the National Academy of Sciences of the United States of America 101 (37):13417-13422. doi:10.1073/pnas.0405549101

Panozzo JF, Eagles HA (2000) Cultivar and environmental effects on quality characters in wheat. II. Protein. Australian Journal of Agricultural Research 51 (5):629-636. doi:http://dx.doi.org/10.1071/AR99137

Parisot L, Aparicio C, Moneret-Vautrin DA, Guerin L (2001) Allergy to lupine flour. Allergy 56 (9):918-919. doi:10.1034/j.1398-9995.2001.00254.x

Pazy P, Heyn CC, Herrnstadt I, Plitmann U (1977) Studies in populations of the Old World *Lupinus* species. I. Chromosomes of the east-mediterranean lupines. Israel Journal of Botany 26:115–127

Peeters KABM, Nordlee JA, Penninks AH, Chen L, Goodman RE, Bruijnzeel-Koomen CAFM, Hefle SL, Taylor SL, Knulst AC (2007) Lupine allergy: Not simply cross-reactivity with peanut or soy. Journal of Allergy and Clinical Immunology 120 (3):647-653

Pereira MA, Ludwig DS (2001) Dietary fiber and body-weight regulation: observations and mechanisms. Pediatric Clinics of North America 48 (4):969-980

Petterson DS (1998) Composition and food uses of lupins.In:Lupins as Crop Plants: Biology, Production, and Utilization. Gladstones J.S., Atkins C. and Hamblin J. edn. CABI, Oxon, U.K.

Petterson DS, Crosbie GB (1990) Potential for lupins as food for humans. Food Australia 42 (5):266–268

Petterson DS, Sipsas S, Mackintosh JB (1997) The chemical composition and nutritive value of Australian pulses. Grains Research and Development Corporation, Canberra, Australia

Pilvi TK, Jauhiainen T, Cheng ZJ, Mervaala EM, Vapaatalo H, Korpela R (2006) Lupin protein attenuates the development of hypertension and normalises the vascular function of NaCl-loaded GotoeKakizaki rats. Journal of Physiology and Pharmacology 57:167-176

Planchuelo-Ravello AM Taxonomic studies of *Lupinus* in South America. In: 3rd International Lupin Conference, La Rochelle, France, 1984. International Lupin Association, pp 39-54

Planchuelo-Ravelo AM Taxonomic studies of Lupinus in South America. In: III International Lupine Congress, La Rochelle, France, 1984. pp 40–54

Plitmann U, Heyn CC, Pazy B (1980) Biological flora of Israel. 7 *Lupinus palaestinus* Boiss. and *L. pilosus* Murr. Israel Journal of Botany 28:108–130

Pomeranz Y (1988) Composition and functionality of wheat flour components. In: Pomeranz Y (ed) Wheat Chemistry and Technology 3rd edn. American Association of Cereal Chemists, St. Paul, MN, pp 219-370

Prak K, Maruyama Y, Maruyama N, Utsumi S (2006) Design of genetically modified soybean proglycinin A1aB1b with multiple copies of bioactive peptide sequences. Peptides 27 (6):1179-1186

Prieto A, Razzak E, Lindo DP, Alvarez-Perea A, Rueda M, Baeza ML (2010) Recurrent anaphylaxis due to lupin flour: primary sensitization through inhalation. Journal of Investig Allergol Clin Immunol 20 (1):76-79

Przybylska J, Zimniak-Przybylska Za (1995) Electrophoretic patterns of seed globulins in the Old-World *Lupinus* species Genetic Resources and Crop Evolution 42 (1):69-75. doi:10.1007/bf02310686

Radcliffe M, Scadding G, Brown HM (2005) Lupin flour anaphylaxis. The Lancet 365 (9467):1360

Randall PJ, Moss HJ (1990) Some effects of temperature regime during grain filling on wheat quality. Australian Journal of Agricultural Research 41 (4):603-617. doi:http://dx.doi.org/10.1071/AR9900603

Rayburn AL, Biradar DP, Nelson RL, McCloskey R, Yeater KM (2004) Documenting Intraspecfic Genome Size Variation In Soybean. Crop Sci 44 (1):261-264. doi:10.2135/cropsci2004.2610

Restani P, Duranti M, Cerletti P, Simonetti P (1981) Subunit composition of the seed globulins of *Lupinus albus*. Phytochemistry 20 (9):2077-2083

Ribotta PD, Arnulphi SA, Leon AE, Anon MC (2005a) Effect of soybean addition on the rheological properties and breadmaking quality of wheat flour. Journal of the Science of Food and Agriculture 85 (11):1889-1896. doi:10.1002/jsfa.2191

Ribotta PD, Leon AE, Perez GT, Anon MC (2005b) Electrophoresis studies for determining wheat–soy protein interactions in dough and bread. European Food Research and Technology 221 (1):48-53. doi:10.1007/s00217-005-1135-2

Riggins R, Sholars T (1993) Lupinus. In: Hickman JC (ed) The Jepson manual. higher plants of California. University of Calfornia Press, Berkeley,CA, pp 622–636

Roccia P, Ribotta PD, Perez GT, Leon AE, (2009) Influence of soy protein on rheological properties and water retention capacity of wheat gluten. LWT - Food Science and Technology 42:358–362

Rojas-Hijazo B, Garces MM, Caballero ML, Alloza P, Moneo I (2006) Unsuspected lupin allergens hidden in food. International Archives of Allergy and Immunology 141 (1):47-50

Rosell CM, Collar C (2009) Effect of temperature and consistency on wheat dough performance. International Journal of Food Science & Technology 44 (3):493-502. doi:10.1111/j.1365-2621.2008.01758.x

Ruuska SA, Girke T, Benning C, Ohlrogge JB (2002) Contrapuntal networks of gene expression during *Arabidopsis* seed filling. The Plant Cell Online 14 (6):1191-1206. doi:10.1105/tpc.000877

Ryan KJ, Brewer MS (2007) In situ examination of starch granule-soy protein and wheat protein interactions. Food chemistry 104 (2):619-629. doi:DOI: 10.1016/j.foodchem.2006.12.037

Ryan KJ, Homco-Ryan CL, Jenson J, Robbins KL, Prestat C, Brewer MS (2002) Lipid extraction process on texturezed soy flour and wheat gluten protein–protein interactions in a dough matrix. Cereal Chemistry 79:434–438

Sabanis D, Tzia C (2009) Effect of rice, corn and soy flour addition on characteristics of bread produced from different wheat cultivars. Food and Bioprocess Technology 2 (1):68-79. doi:10.1007/s11947-007-0037-7

Salmanowicz BP (1999) Seed globulins in the Old World *Lupinus* species: Comparative study by HPLC. Genetic Resources and Crop Evolution 46 (4):409-417. doi:10.1023/a:1008608120931

Salmanowicz BP, Przybylska J (1994) Electrophoretic patterns of seed albumins in the Old-World *Lupinus* species (Fabaceae): Variation in the 2S albumin class. Plant Systematics and Evolution 192 (1):67-78. doi:10.1007/bf00985909

Salmanowicz BP, Weder JKP (1997) Primary structure of 2S albumin from seeds of Lupinus albus. Zeitschrift für Lebensmitteluntersuchung und -Forschung A 204 (2):129-135. doi:10.1007/s002170050049

Sanchez-Vioque R, Clemente A, Vioque J, Bautista J, Millan F (1999) Protein isolates from chickpea (*Cicer arietinum* L.): chemical composition, functional properties and protein characterization. Food Chemistry 64 (2):237-243

Santos CuN, Ferreira RB, Teixeira AR (1997) Seed Proteins of *Lupinus mutabilis*. Journal of Agricultural and Food Chemistry 45 (10):3821-3825. doi:10.1021/jf970075v

Sanz Penella JM, Collar C, Haros M (2008) Effect of wheat bran and enzyme addition on dough functional performance and phytic acid levels in bread. Journal of Cereal Science 48 (3):715-721

Sarkki ML (1980) Wheat gluten. In: Inglett GE, Munck L (eds) Cereals for food and beverages: recent progress in cereal chemistry and technology. Academic Press, New York, USA,

Scarafoni A, Di Cataldo A, Vassilevskaia TD, Bekman EP, Rodrigues-Pousada C, Ceciliani F, Duranti M (2001) Cloning, sequencing and expression in the seeds and radicles of two *Lupinus albus* conglutin γ genes. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression 1519 (1&2):147-151

Scarafoni A, Consonni A, Galbusera V, Negri A, Tedeschi G, Rasmussen P, Magni C, Duranti M (2008) Identification and characterization of a Bowman-Birk inhibitor active towards trypsin but not chymotrypsin in *Lupinus albus* seeds. Phytochemistry 69 (9):1820-1825

Senko M, Hendrickson C, Emmett M, Shi S, Marshall A (1997) External accumulation of ions for enhanced electrospray ionization fourier transform ion cyclotron resonance mass spectrometry. Journal of The American Society for Mass Spectrometry 8 (9):970-976. doi:10.1016/s1044-0305(97)00126-8

Sgarbieri VC, Galeazzi MAM (1978) Some physicochemical and nutritional properties of a sweet lupin (*Lupinus albus* var multolupa) protein. Journal of Agricultural and Food Chemistry 26 (6):1438-1442. doi:10.1021/jf60220a047

Shaw J, Roberts G, Grimshaw K, White S, Hourihane J (2008) Short communication: Lupin allergy in peanut-allergic children and teenagers. Allergy 63 (3):370-373. doi:10.1111/j.1398-9995.2007.01568.x

Shewry PR, Tatham AS, Halford NG (1999) The prolamins of the Triticeae. In: Shewry PR, Casey R (eds) Seed Proteins. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 35-78

Shewry PR, Tatham AS, Halford NG (2001) Nutritional control of storage protein synthesis in developing grain of wheat and barley. Plant Growth Regulation 34 (1):105-111. doi:10.1023/a:1013382803849

Siddique KHM, Sykes J (1997) Pulse production in Australia past, present and future. Australian Journal of Experimental Agriculture 37 (1):103-111. doi:http://dx.doi.org/10.1071/EA96068

Singh H (2005) A study of changes in wheat protein during bread baking using SE-HPLC. Food chemistry 90:247-250

Sipsas S, Seymour M (2008) End uses for lupin. Producing lupins. Department of Agriculture and Food Western Australia, South Perth, W.A.

Sirtori CR, Eberini I, Arnoldi A (2007) Hypocholesterolaemic effects of soya proteins: results of recent studies are predictable from the Anderson meta-analysis data. British Journal of Nutrition 97 (05):816-822. doi:doi:10.1017/S0007114507670810

Sirtori CR, Lovati MR, Manzoni C, Castiglioni S, Duranti M, Magni C, Morandi S, D'Agostina A, Arnoldi A (2004) Proteins of white lupin seed, a naturally isoflavone-poor legume, reduce cholesterolemia in rats and increase LDL receptor activity in HepG2 cells. Journal of Nutrition 134 (1):18-23

Sirtori CR, Resta D, Brambilla F, Zacherl C, Arnoldi A (2010) The effects of various processing conditions on a protein isolate from *Lupinus angustifolius*. Food Chemistry 120:496-504

Sirtori E, O'Kane F, Brambilla F, Arnoldi A *Lupinu angustifolius* vs *L. albus*: a combined chromatographic and electrophoretic analysis to highlight the differences in protein profile. In: Palta JA, Berger JB (eds) 12th International Lupin Conference, Fremantle, Western Australia, 2008. International Lupin Association, Canterbury, New Zealand,

Sirtori E, Resta D, Arnoldi A, Savelkoul HFJ, Wichers HJ (2011) Cross-reactivity between peanut and lupin proteins. Food Chemistry 126 (3):902-910

Skylas DJ, Copeland L, Rathmell WG, Wrigley CW (2001) The wheat-grain proteome as a basis for more efficient cultivar identification. PROTEOMICS 1 (12):1542-1546. doi:10.1002/1615-9861(200111)1:12<1542::aid-prot1542>3.0.co;2-k

Smith CP (1944) *Lupinus* L. In: Abrahams L (ed) Illustrated flora of the Pacific states, vol 2. Stanford University Press, Stanford, CA., pp 483–519

Smith PMC, Goggin DE, Mir GA, Cameron E, Colinet H, Stuckey M, Smith WB, Prescott SL, Soutter V, Loblay RH (2008) Characterisation of allergenic proteins in lupin seeds and the relationship between peanut and lupin allergens. Lupins for health and wealth Proceedings of the 12th International Lupin Conference, Fremantle, Western Australia, 14-18 September 2008:459 - 462

Soltis PS, Soltis DE (1995) Plant molecular systematics:inferences of phylogeny and evolutionary processes. In: al. MKHe (ed) Evolutionary biology, vol 28. Plenum Press, New York, NY, pp 139-194

Stejskal J, Griga M (1995) Comparative analysis of some isozymes and proteins in somatic and zygotic embryos of soybean (*Glycine max* [L.] Merr.). Journal of Plant Physiology 146 (4):497-502

Swofford DL (1998) PAUP: Phylogenetic analysis using parsimony. 4.0b2 edn. the Illinois Natural History Survey, Champaign, IL,

Szajli E, Feher T, Medzihradszky KF (2008) Investigating the quantitative nature of MALDI-TOF MS. Molecular & Cellular Proteomics 7 (12):2410-2418. doi:10.1074/mcp.M800108-MCP200

Tabe L, Hagan N, Higgins TJV (2002) Plasticity of seed protein composition in response to nitrogen and sulfur availability. Current Opinion in Plant Biology 5 (3):212-217

Talhinhas P, Leitao J, Neves-Martins J (2006) Collection of L*upinus angustifolius* L. germplasm and characterisation of morphological and molecular diversity. Genetic Resources and Crop Evolution 53 (3):563-578

Talhinhas P, Neves-Martins J, Leitao J (2003) AFLP, ISSR and RAPD markers reveal high levels of genetic diversity among *Lupinus* spp. Plant Breeding 122 (6):507-510. doi:10.1111/j.1439-0523.2003.00892.x

Tilley KA, Benjamin RE, Bagorogoza KE, Okot-Kotber BM, Prakash O, Kwen H (2001) Tyrosine cross-links: molecular basis of gluten structure and function. Journal of Agricultural and Food Chemistry *49*:2627-2632. doi:10.1021/jf010113h

Timmerman-Vaughan G, Mills A, Whitfield C, Frew T, Butler R, Murray S, Lakeman M, McCallum J, Russell A, Wilson D (2005) Linkage mapping of QTL for seed yield, yield components, and developmental traits in pea. Crop Sci 45:1336 - 1344

Torres A, Frias J, Granito M, Guerra M, Vidal-Valverde C (2007) Chemical, biological and sensory evaluation of pasta products supplemented with α-galactoside-free lupin flours. Journal of the Science of Food and Agriculture 87 (1):74-81. doi:10.1002/jsfa.2673

Tremblay A, Dumesnil JG, Depres JP (2002) Diet, satiety and obesity treatment. Britsh Journal of Nutrition 88 (213-214)

Triboi E, Abad A, Michelena A, Lloveras J, Ollier JL, Daniel C (2000) Environmental effects on the quality of two wheat genotypes: 1. quantitative and qualitative variation of storage proteins. European Journal of Agronomy 13 (1):47-64

Turner BL (1995) A new species of *Lupinus* (Fabaceae) from Oaxaca, Mexico: a shrub or tree mostly three to eight meters high. Phytologia 79:102-107

Van Barneveld RJ (1999) Understanding the nutritional chemistry of lupin (*Lupinus* spp.) seed to improve livestock production eficiency. Nutrition Research Reviews 12:203-230

Vandewater K, Vickers Z (1996) Higher-protein foods produce greater sensory-specific satiety. Physiology & Behavior 59 (3):579-583

Vaz AC, Pinheiro C, Martins JMN, Ricardo CPP (2004) Cultivar discrimination of Portuguese *Lupinus albus* by seed protein electrophoresis: the importance of considering "glutelins" and glycoproteins. Field Crops Research 87 (1):23-34

Wang H, Clouthier SG, Galchev V, Misek DE, Duffner U, Min C-K, Zhao R, Tra J, Omenn GS, Ferrara JLM, Hanash SM (2005) Intact-protein-based high-resolution three-dimensional quantitative analysis system for proteome profiling of biological fluids. Molecular & Cellular Proteomics 4 (5):618-625. doi:10.1074/mcp.M400126-MCP200

Wangen S (2009) Healthier without wheat. a new understanding of wheat allergies, Celiac disease and non-celiac gluten intolerance. Innate Health Publ. Sieattle, WS, USA

Weiße K, Brandsch C, Zernsdorf B, Nkengfack Nembongwe G, Hofmann K, Eder K, Stangl G (2010) Lupin protein compared to casein lowers the LDL cholesterol:HDL cholesterol-ratio of hypercholesterolemic adults. European Journal of Nutrition 49 (2):65-71. doi:10.1007/s00394-009-0049-3

Welsh SL, Atwood ND, Goodrish S, Higgins LC (1987) Great Basin Naturalist Memoirs, number 9, a Utah flora., vol 9. Brigham Young University Press, Provo, UT

Wieser H, Antes S, Seilmeier W (1998) Quantitative determination of gluten protein types in wheat flour by reversed-phase high-performance liquid chromatography. Cereal Chem 75:644-650

Williams CA, Demissie A, Harborne JB (1983) Flavonoids as taxonomic markers in old world *Lupinus* species. Biochemical Systematics and Ecology 11 (3):221-231

Wilson TA, Orthoefer F, Nicolosi RJ (2007) Soy protein concentrate lowers serum high-density lipoprotein cholesterol concentrations compared with casein in ovariectomized rats fed a low-fat, cholesterol-free diet. Nutr Res 27:417-422

Wink M, Meiβner C, Witte L (1995) Patterns of quinolizidine alkaloids in 56 species of the genus *Lupinus*. Phytochemistry 38 (1):139-153

Wolko B, Weeden NF (1990a) Isozyme number as an indicator of phylogeny in *Lupinus*. Genetica Polonica 31:179–187

Wolko B, Weeden NF (1990b) Relationships among lupin species as reflected by isozyme phenotype. Genetica Polonica 31:189–197

Wrigley CW, Bekes F, Bushuk W (2006) Chapter 1. Gluten: A balance of gliadin and glutenin. In: Wrigley CW, Bekes F, Bushuk W (eds) Gliadin and glutenin. The unique balance of wheat quality. AACCI Press, St Paul, Min., USA,

Xu F, Brown KM, Dybdal L, Forman TM, Fuglslang CC, Wagner P (1999) Controlled stepwise reduction of disulfide bonds and heat-induced modification of wheat dough proteins. Cereal Chemistry *76*:931-937. doi:doi:10.1094/CCHEM.1999.76.6.931

Xu J, Bietz JA, Carriere CJ (2007) Viscoelastic properties of wheat gliadin and glutenin suspensions. Food Chemistry 101 (3):1025-1030

Yahata E, Maruyama-Funatsuki W, Nishio Z, Tabiki T, Takata K, Yamamoto Y, Tanida M, Saruyama H (2005) Wheat cultivar-specific proteins in grain revealed by 2-DE and their application to cultivar identification of flour. PROTEOMICS 5 (15):3942-3953. doi:10.1002/pmic.200402103

Yaklich RW (2000) β-conglycinin and glycinin in high-protein soybean seeds. Journal of Agricultural and Food Chemistry 49 (2):729-735. doi:10.1021/jf001110s

Yan Y, Hsam SLK, Yu JZ, Jiang Y, Ohtsuka I, Zeller FJ (2003) HMW and LMW glutenin alleles among putative tetraploid and hexaploid European spelt wheat (

Triticum spelta L.) progenitors. TAG Theoretical and Applied Genetics 107 (7):1321-1330. doi:10.1007/s00122-003-1315-z

Yang H, Boersma JG, You M, Buirchell BJ, Sweetingham MW (2004) Development and implementation of a sequence-specific PCR marker linked to a gene conferring resistance to anthracnose disease in narrow-leafed lupin (*Lupinus angustifolius* L.). Molecular Breeding 14 (2):145-151. doi:10.1023/b:molb.0000038003.49638.97

Yang H, Sweetingham MW, Cowling WA, Smith PMC (2001) DNA fingerprinting based on microsatellite-anchored fragment length polymorphisms, and isolation of sequence-specific PCR markers in lupin (*Lupinus angustifolius* L.). Molecular Breeding 7 (3):203-209. doi:10.1023/a:1011363205557

Yoshie-Stark Y, Bez J, Wada Y, Wasche A (2004) Functional Properties, Lipoxygenase Activity, and Health Aspects of *Lupinus albus* Protein Isolates. Journal of Agricultural and Food Chemistry 52 (25):7681-7689. doi:10.1021/jf049583c

Yuan H, Yan G, Siddique KHM, Yang H (2005) RAMP based fingerprinting and assessment of relationships among Australian narrow-leafed lupin (*Lupinus angustifolius* L.) cultivars. Australian Journal of Agricultural Research 56 (12):1339-1346. doi:doi:10.1071/AR05188

Zdunczyk Z, Juskiewicz J, Frejnagel S, Gulewicz K (1998) Influence of alkaloids and oligosaccharides from white lupin seeds on utilization of diets by rats and absorption of nutrients in the small intestine. Animal Feed Science and Technology 72 (1):143-154

Zhang L, Cheng L, Jiang L, Wang Y, Yang G, He G (2010) Effects of tannic acid on gluten protein structure, dough properties and bread quality of Chinese wheat. Journal of the Science of Food and Agriculture 90 (14):2462-2468. doi:10.1002/jsfa.4107

Zhang Q, Dong Y, An X, Wang A, Zhang Y, Li X, Gao L, Xia X, He Z, Yan Y (2008) Characterization of HMW glutenin subunits in common wheat and related species by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Journal of Cereal Science 47 (2):252-261

Zhu J, Khan K (2001) Effects of genotype and environment on glutenin polymers and breadmaking quality. Cereal chemistry 78 (2):125-130

Zimniak-Przybylska Z, Przybylska J (1997) Electrophoretic seed globulin patterns in some New World *Lupinus* species. Genetic Resources and Crop Evolution 44 (1):57-62. doi:10.1023/a:1008672005826

Appendix I

Figure 1 and **2** Showing weather Data of three lupin growing locations of Chapter 5

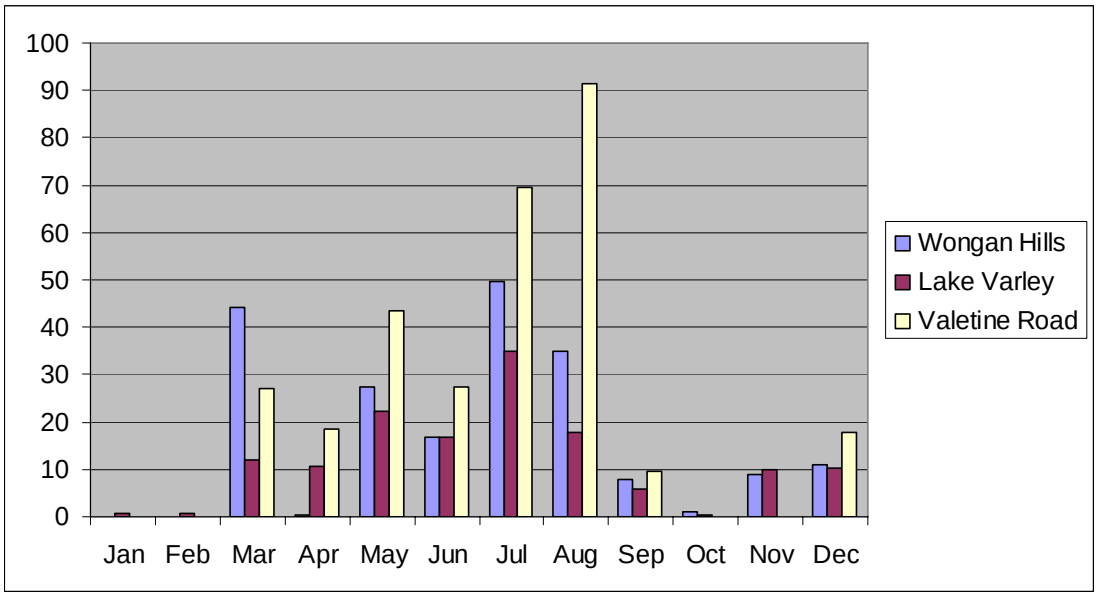


Figure 1: Total rain fall (mm) of the year 2010 of three locations

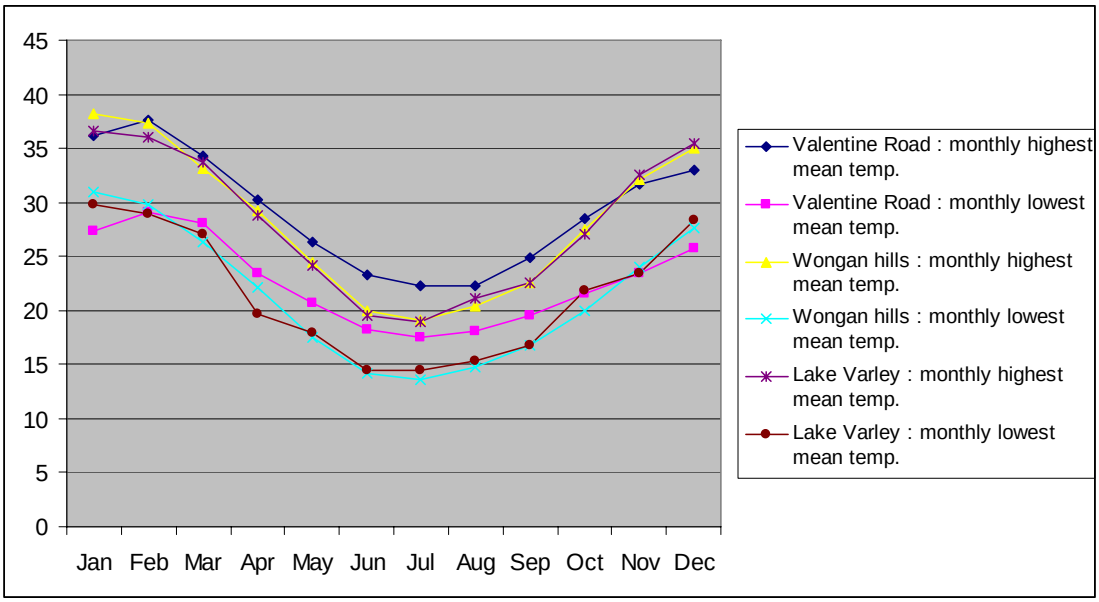


Figure 2: Average maximum and minimum temperatures (^{o}C) of the year 2010 of the three locations

Appendix II

Supporting table of Chapter 6: Matching of the mass spectrometric peptide sequences to identify the common proteins among the cultivars. Matching has been achieved using Mascot sequence matching software (Matrix Science) with the taxonomy set to Viridiplanate (Green Plants). Conglutin β and Conglutin α are sometimes described as vicilin-like proteins and legumin-like proteins respectively, but we have used the term Conglutin β and Conglutin α for all strong matching to avoid the confusion.

Spot no	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
31	Conglutin beta [*Lupinus angustifolius*]	gi|149208401	61433/5.34	70000/6.3	13%	480	EQIQELR LENLQNYR ATITIVNPDKR VSKEQIQELR DQQSYFSGFSK QAYNLEYGDALR KGKPSESGPFNLR TNRLENLQNYR
32	Conglutin gamma [*Lupinus angustifolius*]	gi|662366	48885/7.66	80000/6.8	4%	49	AGIALGAHHLEENLVVFDLER
33	Conglutin alpha [*Lupinus angustifolius*]	gi|2313076	15523/5.80	65000/5.4	7%	43	TLTSLDFPILR
34	Congultin alpha (Legumin-like protein) [*Lupinus albus*]	gi|85361412	58374/5.53	46000/5.3	6%	65	RPFYTNAPQEIYIQQGR RFYLSGNQEQEFLQYQEK
35	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	48000/5.8	22%	693	SQSQEQEDSHQKIR RFYLAGNPEEEYPETQQQR FYLAGNPEEEYPETQQQR FYLAGNPEEEYPETQQQRQQR NNILSGFDPQFLSQALNIDEDTVHK QIIRVEEGLGVISPK

Spot no	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
							GGHEEEEVEEER GGHEEEEVEEERGR HTRGDEGQEEEETTTTTEER GDEGQEEEETTTTTEER
36	Conglutin beta [*Lupinus angustifolius*]	gi|149208401	61433/5.34	48000/5.3	11%	327	ATITIVNPDKR, NTLEATFNTR, VSKEQIQELR, EQIQELRK, GKPSESGPFNLR, NQQQSYFANAQPQQKQQR
37	Conglutin beta (Fragment) [*Lupinus angustifolius*]	gi|149208401	61433/5.34	45000/5.3	33%	700	NTLEATFNTR ATITIVNPDKR VSKEQIQELR FGNFYEITPER KGKPSESGPFNLR ELTFPGSAQDVER HSDADYILVVLNGR RGQEQSYQDEGVIVR NFLAGSEDNVISQLDR NQQQSYFANAQPQQK ILLGNEDEQEDEEQR NTLEATFNTRYEEIQR NFLAGSEDNVISQLDREVK NQQQSYFANAQPQQKQQR LAIPINNPSNFYDFYPSSTK
38	Conglutin alpha (Fragment) [*Lupinus angustifolius*]	gi|2313076	15523/5.80	37000/5.7	35%	146	TLTSLDFPILR QGREEEEEEEEEER HRPHHHEEEEEEEEWSHQVR
39	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	37000/5.7	17%	378	SQSQEQEDSHQKIR RFYLAGNPEEEYPETQQQR FYLAGNPEEEYPETQQQR NNILSGFDPQFLSQALNIDEDTVHK QIIRVEEGLGVISPK VEEGLGVISPK WQEQEEEEEEKEEPR

Spot no	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
40	Conglutin alpha (Legumin-like protein) [*Lupinus albus*]	gi|85361412	58374/5.53	36000/5.3	6%	114	RPFYTNAPQEIYIQQGR RFYLSGNQEQEFLQYQEK
41	Conglutin beta (Fragment) [*Lupinus angustifolius*]	gi|149208401	61433/5.34	31000/5.6	13%	388	EQIQELR EQIQELRK VSKEQIQELR GKPSESGPFNLR KGKPSESGPFNLR LLGFGINANENQR ELTFPGSAQDVER NFLAGSEDNVISQLDREVK
42	Conglutin beta [*Lupinus angustifolius*]	gi|169950562	71853/5.82	32000/5.9	33%	692	YSDKLSK KQIQELR LENLQNYR SNKPIYSNK ATITIVNPDKR DQQSYFSGFSK NFLAGSEDNVIK TNRLENLQNYR LLGFGINANENQR ELTFPGSIEDVER GQEQSHQDEGVIVR RGQEQSHQDEGVIVR NTLEATFNTRYEEIER HAQSSSGEGKPSESGPFNLR LPAGTTSYILNPDDNQNLR NQQQSYFANAQPQQQQQR AIFIVVVDEGEGNYELVGIRDQQR
43	Conglutin beta (Fragment) [*Lupinus angustifolius*]	gi|149208401	61433/5.34	42000/6.4	7%	143	TNRLENLQNYR ATITIVNPDKR NTLEATFNTRYEEIQR
44	Conglutin beta [*Lupinus angustifolius*]	gi|149208403	54267/6.27	40000/6.6	17%	251	LAIPINNPGK LENLQNYR ATITIVNPDKR IIEFQSKPNTLILPK

Spot no	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
							NTLEATFNTRYEEIER LYDFYPSTTKDQQSYFSGFSK
45	Conglutin beta [*Lupinus angustifolius*]	gi|149208403	54267/6.27	38000/6.6	35%	906	RQEEEEEEEEWQPR QEEEEEEEEWQPR QRNPYHFSSNR NPYHFSSNR FQTYYR TNRLENLQNYR IIEFQSKPNTLILPK HSDADFILVVLNGR ATITIVNPDKR QVYNLEQGDALR QVYNLEQGDALRLPAGTTSYILNPDDNQNLR LPAGTTSYILNPDDNQNLR LAIPINNPGK LAIPINNPGKLYDFYPSTTK LYDFYPSTTK LYDFYPSTTKDQQSYFSGFSK DQQSYFSGFSK NTLEATFNTRYEEIER YEEIER
46	Glyceraldehyde-3-phosphate-dehydrogenase [*Lupinus albus*]	gi|62816190	32166/6.80	37000/8.0	51%	865	KVIISAPSK VINDRFGIVEGLMTTVHSITATQK FGIVEGLMTTVHSITATQK TVDGPSSKDWR AASFNIIPSSTGAAK VLPVLNGK LTGMAFR LTGMAFRVPTVDVSVVDLTVR GILGYTEDDVVSTDFIGDNR SSIFDAK AGIALNEK AGIALNEKYVK LVSWYDNEWGYSTR
47	Conglutin beta [*Lupinus*	gi|149208403	54267/6.27	37000/6.4	18%	390	FQTYYR

Spot no	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
	angustifolius]						LAIPINNPGK LENLQNYR NPYHFSSNR ATITIVNPDKR LYDFYPSTTK QVYNLEQGDALR TNRLENLQNYR NTLEATFNTRYEEIER
48	Conglutin beta [*Lupinus angustifolius*]	gi|169950562	71853/5.82	30000/6.5	14%	204	NTLEATFNTR TNRLENLQNYR LLGFGINANENQR IIEFQSKPNTLILPK LPAGTTSYILNPDDNQNLR NQQQSYFANAQPQQQQQR
49	Conglutin beta [*Lupinus angustifolius*]	gi|169950562	71853/5.82	30000/6.7	12%	196	SNKPIYSNK ATITIVNPDKR ELTFPGSIEDVER NQQQSYFANAQPQQQQQR LIKNQQQSYFANAQPQQQQQR AIFIVVVDEGEGNYELVGIRDQQR
50	Conglutin gamma [*Lupinus angustifolius*]	gi|662366	48885/7.66	27000/8.8	14%	268	RTPLMQVPLLLDLNGK TPLMQVPLLLDLNGK LGPMVK GLPNNVQGALGLGQAPISLQNQLFSHFGLK GLPNNVQGALGLGQAPISLQNQLFSHFGLKR QGEYFIQVNAIR
51	Conglutin alpha (Fragment) [*Lupinus angustifolius*]	gi|2313076	15523/5.80	24000/6.9	18%	87	TLTSLDFPILR WLGLAAEHGSIYK
52	Conglutin beta [*Lupinus albus*]	gi|46451223	62092/6.43	24000/6.6	10%	208	NTLEATFNTR TNRLENLQNYR LLGFGINADENQR IVEFQSKPNTLILPK NTLEATFNTRYEEIQR

Spot no	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
53	Conglutin alpha (Fragment) [*Lupinus angustifolius*]	gi|2313076	15523/5.80	21000/6.0	33%	183	HNIGQSTSPDAYNPQAGR LKTLTSLDFPILR TLTSLDFPILR WLGLAAEHGSIYK
54	Conglutin alpha (Fragment) [*Lupinus angustifolius*]	gi|2313076	15523/5.80	22000/6.4	35%	348	LRHNIGQSTSPDAYNPQAGR HNIGQSTSPDAYNPQAGR LKTLTSLDFPILR TLTSLDFPILR WLGLAAEHGSIYK
55	Conglutin alpha (Fragment) [*Lupinus angustifolius*]	gi|2313076	15523/5.80	22000/6.8	35%	373	LRHNIGQSTSPDAYNPQAGR HNIGQSTSPDAYNPQAGR HNIGQSTSPDAYNPQAGRLK LKTLTSLDFPILR TLTSLDFPILR WLGLAAEHGSIYK
56	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	19000/8.2	8%	148	RGQLLVVPQNFVVAHQAGDEGFEFIAFK GQLLVVPQNFVVAHQAGDEGFEFIAFK GIPAEVLANAFR
57	Conglutin beta [*Lupinus angustifolius*]	gi|149208403	54267/6.27	18000/8.8	29%	797	QRNPYHFSSNR NPYHFSSNR FQTYYR TNRLENLQNYR LENLQNYR IIEFQSKPNTLILPK ATITIVNPDK QVYNLEQGDALR QVYNLEQGDALRLPAGTTSYILNPDDNQNLR LPAGTTSYILNPDDNQNLR LAIPINNPGK LAIPINNPGKLYDFYPSTTK LYDFYPSTTKDQQSYFSGFSK DQQSYFSGFSK NTLEATFNTRYEEIER
58	Conglutin alpha (Seed	gi|224184735	57074/5.45	17000/7.0	20%	739	ENIADPTR

Spot no	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
	storage protein) [*Lupinus angustifolius*]						ENIADPTRADLYNPTAGR ADLYNPTAGR GRVQVVNSQGNSVFNDDLR VQVVNSQGNSVFNDDLR VQVVNSQGNSVFNDDLRR GQLLVVPQNFVVAHQAGDEGFEFIAFK TNDQATTSPLK TNDQATTSPLKQVFR GIPAEVLANAFR LSLNQVSELK
59	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	15000/6.2	9%	130	GQLLVVPQNFVVAHQAGDEGFEFIAFK GIPAEVLANAFR LSLNQVSELK
60	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	17000/6.3	19%	721	ENIADPTR ENIADPTRADLYNPTAGR ADLYNPTAGR GRVQVVNSQGNSVFNDDLR VQVVNSQGNSVFNDDLR VQVVNSQGNSVFNDDLRR RGQLLVVPQNFVVAHQAGDEGFEFIAFK GQLLVVPQNFVVAHQAGDEGFEFIAFK TNDQATTSPLK IPAEVLANAFR LSLNQVSELK
61	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	17000/5.8	6%	100	VEEGLGVISPK RFYLAGNPEEEYPETQQQR
62	Conglutin delta-2 large chain Tax_Id=3871 [*Lupinus angustifolius*]	gi|116181	9394/4.56	10000/4.6	18%	110	QQEQQLEGELEK QQEQQLEGELEKLPR
63	Conglutin gamma	gi|662366	48885/7.66	10000/6.8	11%	403	VGFNSNSLK SRVGFNSNSLK

Spot no	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
	[*Lupinus angustifolius*]						KISGGAPSVDLILDK ISGGAPSVDLILDKNDAVWR AGIALGAHHLEENLVVFDLER AGIALGAHHLEENLVVFDLERSR
64	Conglutin beta Tax_Id=3871 [Lupinus angustifolius]	gi|149208403	54267/6.27		30%	654	NPYHFSSNR LYDFYPSTTK DQQSYFSGFSK QQEEQEREHR TNRLENLQNYR HSDADFILVVLNGR EREQEQQPRPQR SQSEESQEEEHER QEEEEEEEWQPR RQEEEEEEEWQPR NTLEATFNTRYEEIER LPAGTTSYILNPDDNQNLR

Appendix III

Supporting table of Chapter 7: Matching of peptides sequenced by mass spectrometry to identify the proteins. Matching has been done using Mascot sequence matching software (Matrix Science) with the taxonomy set to Viridiplanate (Green Plants). Conglutin β and Conglutin α are sometimes described as vicilin-like proteins and legumin-like proteins respectively, but we have used the term Conglutin β and Conglutin α for all strong matching to avoid the confusion.

Spot no	Spot origin	Extractability from lupin-wheat bread		Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
		with reduction/ denaturing	without reduction/ denaturing							
1	Lupin	Yes	Yes	Conglutin alpha [*Lupinus angustifolius*]	gi|2313076		65000/5.4	7%	43	TLTSLDFPILR
2	Lupin	Yes	Yes	Conglutin beta [*Lupinus angustifolius*]	gi|149208401	61433/5.34	48000/5.3	11%	327	ATITIVNPDKR, NTLEATFNTR, VSKEQIQELR, EQIQELRK, GKPSESGPFNLR, NQQQSYFANAQPQQKQQR
3	Lupin	Yes	No	Conglutin beta [*Lupinus angustifolius*]	gi|149208401	61433/5.34	70000/6.3	13%	480	EQIQELR LENLQNYR ATITIVNPDKR VSKEQIQELR DQQSYFSGFSK QAYNLEYGDALR KGKPSESGPFNLR TNRLENLQNYR
4	Lupin	No	No	Conglutin beta [*Lupinus angustifolius*]	gi|149208403	54267/6.27	40000/6.6	17%	251	LAIPINNPGK LENLQNYR ATITIVNPDKR IIEFQSKPNTLILPK NTLEATFNTRYEEIER LYDFYPSTTKDQQSYFSGFSK

Spot no	Spot origin	Extractability from lupin-wheat bread		Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
		with reduction/ denaturing	without reduction/ denaturing							
5	Lupin	No	No	Conglutin beta [*Lupinus angustifolius*]	gi|149208403	54267/6.27	37000/6.4	18%	390	FQTYYR LAIPINNPGK LENLQNYR NPYHFSSNR ATITIVNPDKR LYDFYPSTTK QVYNLEQGDALR TNRLENLQNYR NTLEATFNTRYEEIER
6	Lupin	Yes	No	Conglutin beta [*Lupinus angustifolius*]	gi|169950562	71853/5.82	32000/5.9	33%	692	YSDKLSK KQIQELR LENLQNYR SNKPIYSNK ATITIVNPDKR DQQSYFSGFSK NFLAGSEDNVIK TNRLENLQNYR LLGFGINANENQR ELTFPGSIEDVER GQEQSHQDEGVIVR RGQEQSHQDEGVIVR NTLEATFNTRYEEIER HAQSSSGEGKPSESGPFNLR LPAGTTSYILNPDDNQNLR NQQQSYFANAQPQQQQQR AIFIVVVDEGEGNYELVGIRDQQR
7	Lupin	Yes	No	Conglutin beta [*Lupinus angustifolius*]	gi|169950562	71853/5.82	30000/6.7	12%	196	SNKPIYSNK ATITIVNPDKR ELTFPGSIEDVER NQQQSYFANAQPQQQQQR LIKNQQQSYFANAQPQQQQQR AIFIVVVDEGEGNYELVGIRDQQR

Spot no	Spot origin	Extractability from lupin-wheat bread		Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
		with reduction/ denaturing	without reduction/ denaturing							
8	Lupin	Yes	No	Conglutin beta [*Lupinus angustifolius*]	gi|169950562	71853/5.82	30000/6.5	14%	204	NTLEATFNTR TNRLENLQNYR LLGFGINANENQR IIEFQSKPNTLILPK LPAGTTSYILNPDDNQNLR NQQQSYFANAQPQQQQQR
9	Lupin	Yes	No	Conglutin beta (Fragment) [*Lupinus angustifolius*]	gi|149208401	61433/5.34	31000/5.6	13%	388	EQIQELR EQIQELRK VSKEQIQELR GKPSESGPFNLR KGKPSESGPFNLR LLGFGINANENQR ELTFPGSAQDVER NFLAGSEDNVISQLDREVK
10	Lupin	No	No	Conglutin beta [*Lupinus albus*]	gi|46451223	62092/6.43	24000/6.6	10%	208	NTLEATFNTR TNRLENLQNYR LLGFGINADENQR IVEFQSKPNTLILPK NTLEATFNTRYEEIQR
11	Lupin	Yes	Yes	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	17000/7.0	20%	739	ENIADPTR ENIADPTRADLYNPTAGR ADLYNPTAGR GRVQVVNSQGNSVFNDDLR VQVVNSQGNSVFNDDLR VQVVNSQGNSVFNDDLRR GQLLVVPQNFVVAHQAGDEGFEFIAFK TNDQATTSPLK TNDQATTSPLKQVFR GIPAEVLANAFR LSLNQVSELK

Spot no	Spot origin	Extractability from lupin-wheat bread		Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
		with reduction/ denaturing	without reduction/ denaturing							
12	Lupin	Yes	Yes	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	15000/6.2	9%	130	GQLLVVPQNFVVAHQAGDEGFEFIAFK GIPAEVLANAFR LSLNQVSELK
13	Lupin	Yes	Yes	Conglutin gamma [*Lupinus angustifolius*]	gi|662366	48885/7.66	10000/6.8	11%	403	VGFNSNSLK SRVGFNSNSLK KISGGAPSVDLILDK ISGGAPSVDLILDKNDAVWR AGIALGAHHLEENLVVFDLER AGIALGAHHLEENLVVFDLERSR
14	Lupin	No	No	Conglutin delta-2 large chain Tax_Id=3871 [*Lupinus angustifolius*]	gi|116181	9394/4.56	10000/4.6	18%	110	QQEQQLEGELEK QQEQQLEGELEKLPR
15	Lupin	Yes	Yes	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	19000/8.2	8%	148	RGQLLVVPQNFVVAHQAGDEGFEFIAFK GQLLVVPQNFVVAHQAGDEGFEFIAFK GIPAEVLANAFR
16	Lupin	No	No	Conglutin beta (Fragment) [*Lupinus angustifolius*]	gi|149208401	61433/5.34	42000/6.4	7%	143	TNRLENLQNYR ATITIVNPDKR NTLEATFNTRYEEIQR

Spot no	Spot origin	Extractability from lupin-wheat bread		Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
		with reduction/ denaturing	without reduction/ denaturing							
17*	Lupin	yes	yes	Conglutin alpha (Legumin storage protein 3) [*Lotus japonicus*]	gi|206712284					INALEPDNR
18	Wheat	Yes	No	HMW glutenin subunit Ax2* [*Triticum aestivum*]	gi|269849175	88423/6.15	100000/6.5	15%	344	ELQEHSLK GGSFYPGETTPPQQLQQSILWGIPALLR QQPGQGQQLR GQQGQQSGQGQQLGQGQQGQQPGQK YYPTSPQQPGQEQQPR QLQQPEQGQQGQQPEQGQQGQQQR QWLQPR LEGGDALLASQ
19	Wheat	Yes	No	Gamma-gliadin [*Triticum urartu*]	gi|209971935		35000/7.5	4%	46	APFASIVADIGGQ
20	Wheat	No	No	Superoxide dismutase [*Triticum aestivum*]	gi|1654387	25283/7.9	23000/6.5	18%	95	NVRPDYLTNIWK LGWAIDEDFGSIEK KLSVETTPNQDPLVTK
21	Wheat	No	No	Not identified						
22	Lupin	Yes	Yes	Conglutin alpha (Seed storage protein)	gi|224184735	57074/5.45	48000/5.8	22%	693	SQSQEQEDSHQKIR RFYLAGNPEEEYPETQQQR FYLAGNPEEEYPETQQQR FYLAGNPEEEYPETQQQRQQR NNILSGFDPQFLSQALNIDEDTVHK QIIRVEEGLGVISPK

Spot no	Spot origin	Extractability from lupin-wheat bread with reduction/ denaturing	without reduction/ denaturing	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
				[*Lupinus angustifolius*]						GGHEEEEVEEER GGHEEEEVEEERGR HTRGDEGQEEEETTTTTEER GDEGQEEEETTTTTEER
23	Lupin	Yes	Yes	Conglutin alpha (Fragment) [*Lupinus angustifolius*]	gi|2313076	15523/5.80	37000/5.7	35%	146	TLTSLDFPILR QGREEEEEEEEEEER HRPHHHEEEEEEEEWSHQVR
24	Lupin	Yes	Yes	Congultin alpha (Legumin-like protein) [*Lupinus albus*]	gi|85361412	58374/5.53	46000/5.3	6%	65	RPFYTNAPQEIYIQQGR RFYLSGNQEQEFLQYQEK
25	Lupin	Yes	No	Conglutin gamma [*Lupinus angustifolius*]	gi|662366	48885/7.66	27000/8.8	14%	268	RTPLMQVPLLLDLNGK TPLMQVPLLLDLNGK LGPMVK GLPNNVQGALGLGQAPISLQNQLFSHFGLK GLPNNVQGALGLGQAPISLQNQLFSHFGLKR QGEYFIQVNAIR
26	Lupin	No	No	Glyceraldehyde-3-phosphate-dehydrogenase [*Lupinus albus*]	gi|62816190	32166/6.80	37000/8.0	51%	865	KVIISAPSK VINDRFGIVEGLMTTVHSITATQK FGIVEGLMTTVHSITATQK TVDGPSSKDWR AASFNIIPSSTGAAK VLPVLNGK LTGMAFR LTGMAFRVPTVDVSVVDLTVR GILGYTEDDVVSTDFIGDNR

Spot no	Spot origin	Extractability from lupin-wheat bread with reduction/ denaturing	without reduction/ denaturing	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
										SSIFDAK AGIALNEK AGIALNEKYVK LVSWYDNEWGYSTR
27	Lupin	Yes	No	Conglutin beta [*Lupinus angustifolius*]	gi|149208403	54267/6.27	38000/6.6	35%	906	RQEEEEEEEEWQPR QEEEEEEEEWQPR QRNPYHFSSNR NPYHFSSNR FQTYYR TNRLENLQNYR IIEFQSKPNTLILPK HSDADFILVVLNGR ATITIVNPDKR QVYNLEQGDALR QVYNLEQGDALRLPAGTTSYILNPDDNQNLR LPAGTTSYILNPDDNQNLR LAIPINNPGK LAIPINNPGKLYDFYPSTTK LYDFYPSTTK LYDFYPSTTKDQQSYFSGFSK DQQSYFSGFSK NTLEATFNTRYEEIER YEEIER
28	Lupin	No	No	Conglutin gamma [*Lupinus angustifolius*]	gi|662366	48885/7.66	80000/6.8	4%	49	AGIALGAHHLEENLVVFDLER
29	Lupin	Yes	No	Conglutin alpha (Fragment)	gi|2313076	15523/5.80	24000/6.9	18%	87	TLTSLDFPILR WLGLAAEHGSIYK

Spot no	Spot origin	Extractability from lupin-wheat bread with reduction/ denaturing	without reduction/ denaturing	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
30	Lupin	Yes	Yes	[*Lupinus angustifolius*] Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	17000/5.8	6%	100	VEEGLGVISPK RFYLAGNPEEEYPETQQQR
31	Lupin	Yes	Yes	Conglutin alpha (Fragment) [*Lupinus angustifolius*]	gi|2313076	15523/5.80	22000/6.8	35%	373	LRHNIGQSTSPDAYNPQAGR HNIGQSTSPDAYNPQAGR HNIGQSTSPDAYNPQAGRLK LKTLTSLDFPILR TLTSLDFPILR WLGLAAEHGSIYK
32	Lupin	Yes	No	Conglutin beta [*Lupinus angustifolius*]	gi|169950562	71853/5.82	30000/6.2	13%	325	TNRLENLQNYR ATITIVNPDKR LPAGTTSYILNPDDNQNLR SNKPIYSNK LLGFGINANENQR LIKNQQQSYFANAQPQQQQQR NQQQSYFANAQPQQQQQR
33	Lupin	No	No	Conglutin beta [*Lupinus angustifolius*]	gi|169950562	71853/5.82	30000/7.0	10%	229	TNRLENLQNYR IIEFQSKPNTLILPK ATITIVNPDKR LIKNQQQSYFANAQPQQQQQR NQQQSYFANAQPQQQQQR
34	Lupin	Yes	Yes	Conglutin	gi|224184735	57074/5.45	37000/5.7	17%	378	SQSQEQEDSHQKIR RFYLAGNPEEEYPETQQQR

Spot no	Spot origin	Extractability from lupin-wheat bread with reduction/ denaturing	without reduction/ denaturing	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
				alpha (Seed storage protein) [*Lupinus angustifolius*]						FYLAGNPEEEYPETQQQR NNILSGFDPQFLSQALNIDEDTVHK QIIRVEEGLGVISPK VEEGLGVISPK WQEQEEEEEEKEEPR
35	Lupin	Yes	Yes	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*]	gi|224184735	57074/5.45	17000/6.3	19%	721	ENIADPTR ENIADPTRADLYNPTAGR ADLYNPTAGR GRVQVVNSQGNSVFNDDLR VQVVNSQGNSVFNDDLR VQVVNSQGNSVFNDDLRR RGQLLVVPQNFVVAHQAGDEGFEFIAFK GQLLVVPQNFVVAHQAGDEGFEFIAFK TNDQATTSPLK IPAEVLANAFR LSLNQVSELK
36	Lupin	Yes	Yes	Conglutin alpha (Fragment) [*Lupinus angustifolius*]	gi|2313076	15523/5.80	21000/6.0	33%	183	HNIGQSTSPDAYNPQAGR LKTLTSLDFPILR TLTSLDFPILR WLGLAAEHGSIYK
37	Lupin	Yes	Yes	Conglutin alpha (Fragment) [*Lupinus angustifolius*]	gi|2313076	15523/5.80	22000/6.4	35%	348	LRHNIGQSTSPDAYNPQAGR HNIGQSTSPDAYNPQAGR LKTLTSLDFPILR TLTSLDFPILR WLGLAAEHGSIYK
38	Lupin	Yes	Yes	Conglutin gamma	gi|662366	48885/7.66	12000/5.7	11%	103	ISGGAPSVDLILDKNDAVWR AGIALGAHHLEENLVVFDLER VGFNSNSLK

Spot no	Spot origin	Extractability from lupin-wheat bread with reduction/ denaturing	without reduction/ denaturing	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
				[*Lupinus angustifolius*]						
39	Lupin	Yes	Yes	Conglutin beta (Fragment) [*Lupinus angustifolius*]	gi|149208401	61433/5.34	45000/5.3	33%	700	NTLEATFNTR ATITIVNPDKR VSKEQIQELR FGNFYEITPER KGKPSESGPFNLR ELTFPGSAQDVER HSDADYILVVLNGR RGQEQSYQDEGVIVR NFLAGSEDNVISQLDR NQQQSYFANAQPQQK ILLGNEDEQEDEEQR NTLEATFNTRYEEIQR NFLAGSEDNVISQLDREVK NQQQSYFANAQPQQKQQR LAIPINNPSNFYDFYPSSTK
40	Lupin	Yes	Yes	Conglutin alpha (Legumin-like protein) [*Lupinus albus*]	gi|85361412	58374/5.53	36000/5.3	6%	114	RPFYTNAPQEIYIQQGR RFYLSGNQEQEFLQYQEK
41*	Lupin	Yes	Yes	Conglutin alpha (Seed storage protein) [*Lupinus angustifolius*	gi|224184735	57074/5.45	50000/5.7	21%	522	VEEGLGVISPK GGHEEEEVEEER QIIRVEEGLGVISPK GGHEEEEVEEERGR GDEGQEEEETTTTTEER WQEQEEEEEEKEEPR FYLAGNPEEEYPETQQQR

Spot no	Spot origin	Extractability from lupin-wheat bread with reduction/ denaturing	without reduction/ denaturing	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
										HTRGDEGQEEEETTTTTEER RFYLAGNPEEEYPETQQQR NNILSGFDPQFLSQALNIDEDTVHK
42	wheat	Yes	No	High molecular weight glutenin subunit [*Triticum aestivum*]	gi|162415983	69714/8.33	85000/8.0	22%	428	QVVDQQLAGR QVVDQQLAGRLPWSTGLQMR LPWSTGLQMR SVAVSQVAR QYEQTVVPPK QGSYYPGQASPQQPGQGQQPGK WQEPGQGQQWYYPTSLQQPGQGQQIGK GQQGYYPTSLQQPGQGQQGYYPTSLQHTGQR GHYPASLQQPGQGQPGQR MEGGDALSASQ
43	wheat	Yes	No	Alpha-gliadin [*Triticum turgidum*]	gi|282721196		37000/7.6	5%	34	GSVQPQQLPQFEEIR
44	wheat	Yes	No	Not identified						
45	wheat	Yes	No	Glyceraldehyde-3-phosphate dehydrogenase [*Triticum aestivum*]	gi|148508784	36626/7.08	15000/6.6	27%	164	FGIVEGLMTTVHAMTATQK AASFNIIPSSTGAAK LAKPATYDQIK AAIKEESEGNLK GILGYVDEDLVSTDFQGDNR LVSWYDNEWGYSTR
46	wheat	No	No	Not identified						
47	wheat	No	No	Not idetified						
48	wheat	No	No	Beta amylase (Fragment) [*Triticum aestivum*]	gi|32400764	30872/8.6	50000/8.2	33%	461	YLQADFK AAAAMVGHPEWEFPR DAGQYNDAPQR TRFFVDNGTYLTEQGR FFVDNGTYLTEQGR

Spot no	Spot origin	Extractability from lupin-wheat bread with reduction/ denaturing	Extractability from lupin-wheat bread without reduction/ denaturing	Matched protein	Accession number	Theoretical MW/pI	Observed MW/pI	Sequence coverage	MOWSE score	Matched peptides
										FFLAWYSNNLIK ISGIHWWYK YDPTAYNTILR NARPHGINK
50	lupin	No	No	Conglutin beta [*Lupinus angustifolius*]	gi|149208403	54267/6.27	18000/8.8	29%	797	QRNPYHFSSNR NPYHFSSNR FQTYYR TNRLENLQNYR LENLQNYR IIEFQSKPNTLILPK ATITIVNPDK QVYNLEQGDALR QVYNLEQGDALRLPAGTTSYILNPDDNQNLR LPAGTTSYILNPDDNQNLR LAIPINNPGK LAIPINNPGKLYDFYPSTTK LYDFYPSTTKDQQSYFSGFSK DQQSYFSGFSK NTLEATFNTRYEEIER

* Spot number 17 and 41 are the corresponding spot of number 1 and 22 respectively extracted from the lupin-wheat bread.