CSIRO PUBLISHING

Crop & Pasture Science, 2011, **62**, 765–775
http://dx.doi.org/10.1071/CP11046

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

Shahidul Islam[A], *Wujun Ma*[B,C], *Junhong Ma*[B], *Bevan J. Buirchell*[B], *Rudi Appels*[C], *and Guijun Yan*[A,D]

[A]School of Plant Biology, Faculty of Natural and Agricultural Sciences and The UWA Institute of Agriculture, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia.
[B]Department of Agriculture and Food Western Australia, 3 Baron-Hay Court, South Perth, WA 6151, Australia.
[C]Centre for Comparative Genomics, Murdoch University, 90 South Street, Murdoch, WA 6150, Australia.
[D]Corresponding author. Email: guijun.yan@uwa.edu.au

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.

Additional keywords: breeding, diversity, MALDI-TOF, phylogeny, protein profiling.

Received 24 February 2011, accepted 24 September 2011, published online 10 November 2011

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; Howieson *et al*. 2000; Guemes-Vera *et al*. 2008). 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 (Gladstones 1994; Cowling *et al*. 1998) 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 (Stejskal and Griga 1995; Hsieh *et al*. 2001; Fahmy and Salama 2002). 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 (Cunsolo *et al*. 2004; Alberghina *et al*. 2005; Muccilli *et al*. 2005; Chen *et al*. 2007; Liu *et al*. 2009) with high throughput capability. The MALDI-TOF-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; Yahata *et al*. 2005; Magni *et al*. 2007) 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 µM. 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 (0.5 M NaCl) was added at the ratio of 15 ml/g. Protein was extracted by stirring at 4°C 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 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 of reproducibility.

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.

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	–	–	

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 (Fig. 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. A few (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.

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

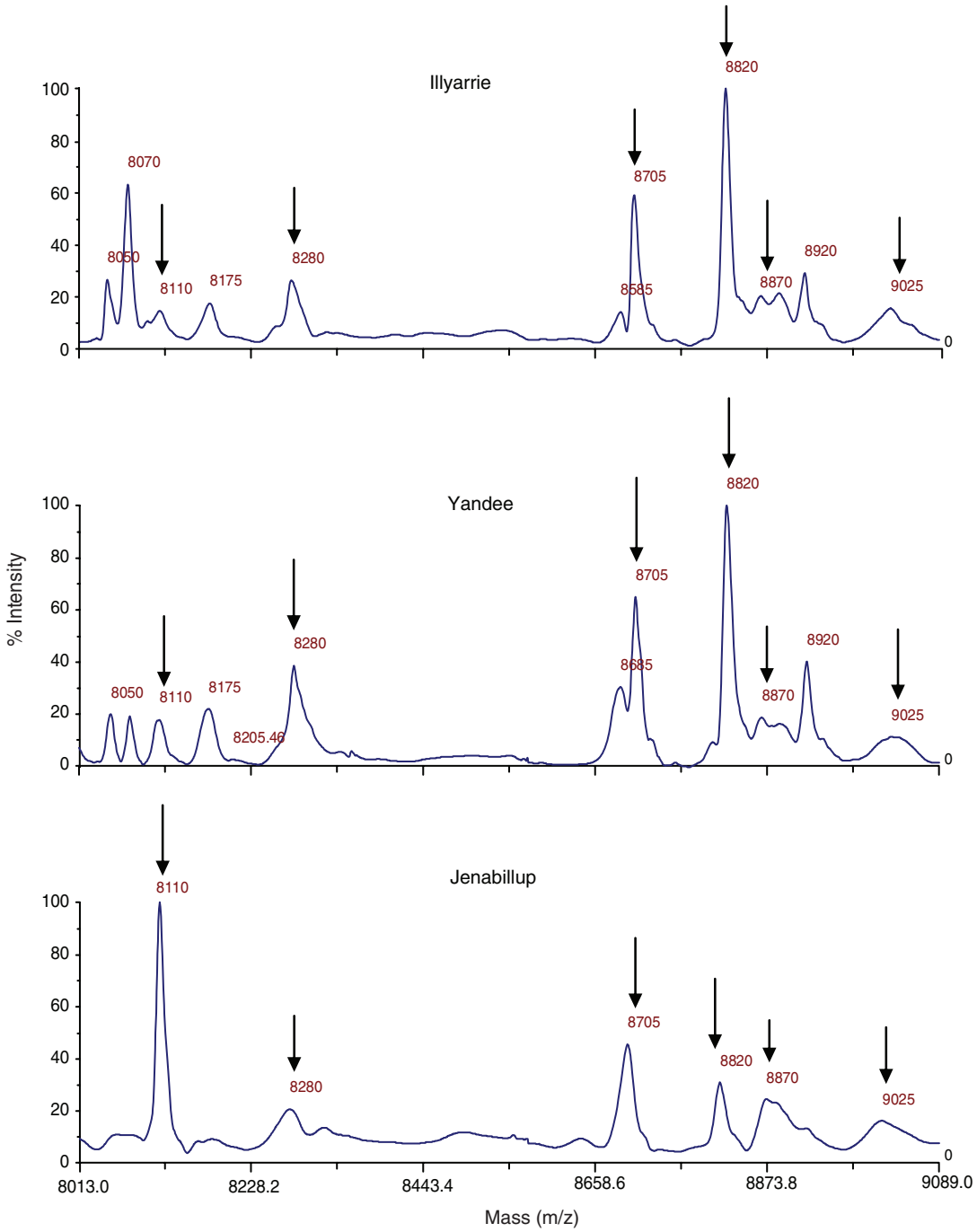

Fig. 1. MALDI-TOF outputs of narrow-leafed lupin protein profiles demonstrating easily visible and identifiable polymorphism of protein mass peaks among different cultivars. The numbers on the protein peaks indicate the molecular weight of the corresponding protein in Daltons. Two sister cultivars Illyarrie 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.

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 (Fig. 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

Table 2. List of the very common mass peaks[A] 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
13 127	23	Chittick, Coromup
13 926	21	Uniharvest, Illyarrie, Chittick, Geebung
14 092	23	Yandee, Coromup,
14 459	21	Chittick, Gungurru, Kalya, Belara
14 642	25	–
15 024	24	Chittick
15 182	24	Chittick
15 332	21	Unicrop, Marri, Illyarrie, Tallerack
15 393	24	Wonga
15 495	24	Illyarrie
15 694	23	Uniwhite, Warrah
16 242	23	Kalya, Moonah
21 350	20	Gungurru, Yorrel, Warrah, Merrit, Mandelup

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

Table 3. List of cultivar-specific mass peaks[A] 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, 11 962, 12 132, 12 965, 13 652, 16 320, 18 841, 19 813
Geebung	8	2258, 2603, 3402, 3536, 3740, 6562, 7640, 13 593
Wonga	5	5488, 10 728, 16 450, 16 904, 28 603
Uniharvest	4	2462, 40 000, 53 000, 60 000
Chittick	3	2420, 9498, 14 878
Gungurru	3	3665, 21 121, 22 008
Mandelup	3	7524, 13 322, 14 690
Tallerack	2	4993, 10 682
Danja	2	4611, 28 436
Moonah	2	17 774, 17 922
Marri	2	5932, 20 002
Uniwhite	2	10 632, 28 639
Illyarrie	1	3200
Merrit	1	3284
Unicrop	1	24 521
Tanjil	1	14 055
Warrah	1	17 245
Yorrel	1	10 423
Yandee	0	–
Myallie	0	–
Kalya	0	–
Belara	0	–
Quilinock	0	–
Jindalee	0	–
Jenabillup	0	–

[A]Cultivar-specific mass peaks are those specific to a single cultivar.

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.

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 (Fig. 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*. 2011). 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.

The 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

Table 4. List of rare mass peaks[A] 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
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
10 061	2	Yorrel, Belara

(*continued next page*)

Table 4. (*continued*)

Protein mass peaks (molecular weight in Dalton)	Number of cultivars with the mass peak	Name of the specific lupin cultivars with the mass peak
10 521	2	Myallie, Tanjil
10 533	3	Mandelup, Yorrel, Merrit
10 666	4	Warrah, Belara, Moonah, Coromup
11 214	2	Tanjil, Coromup
11 562	4	Mandelup, Coromup, Marri, Geebung
11 797	4	Uniharvest, Marri, Illyarrie, Chittick
11 898	4	Illyarrie, Yorrel, Warrah, Coromup
12 468	4	Danja, Moonah, Quilinock, Jenabillup
12 561	5	Myallie, Tanjil, Gungurru, Jindalee, Mandelup
12 834	4	Uniharvest, Tallerack, Myallie, Wonga
12 858	2	Tanjil, Coromup
13 106	2	Unicrop, Chittick
13 794	3	Kalya, Quilinock, Coromup
13 975	5	Chittick, Danja, Geebung, Gungurru, Warrah
13 997	2	Jindalee, Tanjil
15 068	3	Tanjil, Uniharvest, Chittick
15 349	4	Illyarrie, Unicrop, Marri, Tallerack
15 542	3	Unicrop, Tallerack, Uniharvest
16 181	2	Jindalee, Moonah
17 082	4	Geebung, Quilinock, Mandelup, Jenabillup
17 213	3	Geebung, Quilinock, Mandelup
18 396	4	Uniharvest, Marri, Yandee, Mandelup
18 496	3	Belara, Uniharvest, Chittick
19 177	4	Uniwhite, Uniharvest, Danja, Geebung
19 471	2	Geebung, Mandelup
19 750	4	Chittick, Wonga, Illyarrie, Belara
21 010	3	Marri, Tallerack, Jenabillup
21 330	4	Yorrel, Warrah, Gungurru, Mandelup
21 565	4	Unicrop, Illyarrie, Yandee, Mandelup
23 000	2	Uniwhite, Uniharvest
23 777	2	Yorrel, Uniharvest
24 289	5	Uniharvest, Unicrop, Yandee, Geebung, Kalya
28 378	2	Warrah, Uniharvest
28 523	3	Illyarrie, Chittick, Geebung
28 698	2	Kalya, Myallie
28 803	2	Kalya, Wonga
31 000	3	Yandee, Chittick, Merrit
35 000	2	Merrit, Yorrel

[A]Rare mass peaks are those specific to 2–5 cultivars.

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

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	–

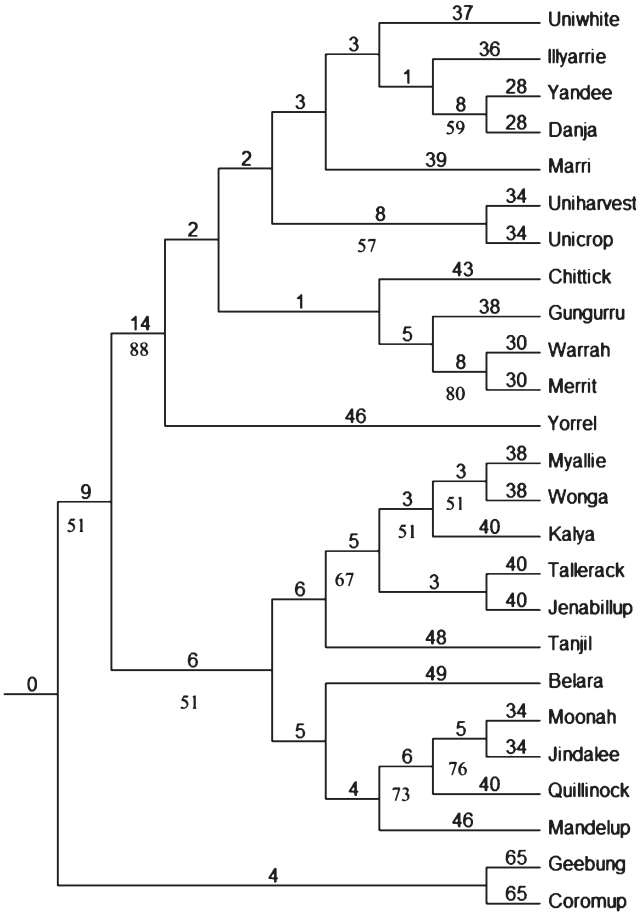

Fig. 2. Dendrogram showing relationship among 25 narrow-leafed lupin (*Lupinus angustifolius*) cultivars by unweighted pair-group method with arithmetic averages based on total character differences. Numbers above branches are branch length and those under the branches are bootstrap values based on 10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

marker-based studies (Yang *et al*. 2001, 2004; Yuan *et al*. 2005; Talhinhas *et al*. 2006).

The constructed phylogenetic tree (Fig. 2) based on the polymorphic seed protein mass peaks is in broad agreement with the pedigree information and DNA marker-based studies (Yuan *et al*. 2005; Talhinhas *et al*. 2006) of the NLL cultivars. Cultivars having common parental lines were grouped together as the seed proteins are generally genetically inherited (Timmerman-Vaughan *et al*. 2005; Bolon *et al*. 2010). 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, Belera, 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 (Fig. 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).

Acknowledgement

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

References

Alberghina G, Cozzolino R, Fisichell 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**, 2069–2074. doi:10.1002/rcm.2032

Aly MA, Rashed NAK, Hassan HZ (2000) Genetic variability and assessment of relationships among some Egyptian rice (*Oryza sativa* L.) cultivars as revealed by SDS–protein and PCR–RAPD analyses. *Egyptian Journal of Genetics and Cytology* **29**, 71–87.

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**, 877–883.

Bolon YT, 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**, 41. doi:10.1186/1471-2229-10-41

Buirchell BJ (2008) Narrow-leafed lupin breeding in Australia – where to from here? In 'Lupin for health and wealth: 12th International Lupin Conference'. (Eds JA Palta, JB Berger) pp. 226–230. (International Lupin Association: Fremantle)

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

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**, 2913–2917. doi:10.1002/rcm.3151

Cowling WA (1999) Pedigrees and characteristics of narrow-leafed lupin cultivars released in Australia from 1967 to 1998. In 'Lupin breeding bulletin'. 1–11. (Agriculture Western Australia: Perth)

Cowling WA, Huyghe C, Swiecicki W (1998) Lupin breeding. In 'Lupins as crop plants: biology, production and utilization'. (Eds JS Gladstones, CA Atkins, J Hamblin) pp. 93–120 (CAB International: London)

Cunsolo V, Foti S, Saletti R, Gilberts 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**, 66–78. doi:10.1002/jms.558

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**, 865–870. doi:10.1002/rcm.1864

Derbyshire E, Wright DJ, Boulter D (1976) Legumin and vicilin, storage proteins of legume seeds. *Phytochemistry* **15**, 3–24. doi:10.1016/S0031-9422(00)89046-9

Dervas G, Doxastakis G, Hadjisarva-Zinoviadi S, Triantafilakos N (1999) Lupin flour addition to wheat flour doughs and effect on rheological properties. *Food Chemistry* **66**, 67–73. doi:10.1016/S0308-8146(98)00234-9

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. doi:10.1016/j.tifs.2008.07.002

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.

Gladstones JS (1994) A historical view of lupins in Australia. In 'Proceedings of the First Australian Lupin Technical Symposium'. (Eds M Dracup, J Palta) pp. 1–38. (Department of Agriculture Western Australia: Perth)

Güémes-Vera N, Pena BRJ, Jimenez MC, Davila 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. doi:10.1002/jsfa.3152

Hill GD (1986) Recent developments in the use of lupins in animal and human nutrition. In '4th International Lupin Conference'. Geraldton, Australia. pp. 40–63. (Department of Agriculture for The International Lupin Association: Perth)

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**, 10505–10508. doi:10.1021/jf071321y

Howieson JG, O'Hara GW, Carr SJ (2000) Changing roles for legumes in Mediterranean agriculture: developments from an Australian perspective. *Field Crops Research* **65**, 107–122. doi:10.1016/S0378-4290(99)00081-7

Hsieh JS, Hsieh KL, Tsai YC, Hsingl YI (2001) Each species of *Glycin* collected in Taiwan has a unique seed protein pattern. *Euphytica* **118**, 67–73. doi:10.1023/A:1004075516984

Islam S, Ma W, Yan G, Gao L, 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**, 6696–6704. doi:10.1021/jf201293p

Kussmann M, Nordhoff E, Rahbek NH, Haebel S, Rossel LM, Jakobsen L, Gobom J, Mirgorodskaya E, Kroll KA, Palm L, Roepstorff P (1997) MALDI-MS sample preparation techniques designed for various peptide and protein analytes. *Journal of Mass Spectrometry* **32**, 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. *Die Nahrung* **40**, 71–74. doi:10.1002/food.19960400205

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**, 295–301. doi:10.1016/j.jcs.2009.05.006

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**, 1223–1226. doi:10.1016/j.jasms.2010.02.029

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. *Theoretical and Applied Genetics* **111**, 410–422. doi:10.1007/s00122-005-2001-0

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**, 997–1007. doi:10.1016/j.phytochem.2007.01.003

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**, 330–357. doi:10.1111/j.1541-4337.2010.00114.x

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**, 719–728. doi:10.1002/pmic.200401029

Müntz 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**, 1281–1293. doi:10.1078/0176-1617-00853

Petterson DS (1998) Composition and food uses of lupins. In 'Lupins as crop plants: biology, production, and utilization'. (Eds Gladstones JS, Atkins CA, Hamblin J) pp. 353–384(CAB International: London)

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

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

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

Swofford DL (1998) 'PAUP: phylogenetic analysis using parsimony.' (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**, 2410–2418. doi:10.1074/mcp.M800108-MCP200

Talhinhas P, Leitão J, Neves-Martins J (2006) Collection of *Lupinus angustifolius* L. germplasm and characterisation of morphological and molecular diversity. *Genetic Resources and Crop Evolution* **53**, 563–578. doi:10.1007/s10722-004-2684-0

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 Science* **45**, 1336–1344. doi:10.2135/cropsci2004.0436

Yahata E, Funatsuki WM, 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**, 3942–3953. doi:10.1002/pmic.200402103

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

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**, 145–151. doi:10.1023/B:MOLB.0000038003.49638.97

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**, 1339–1346. doi:10.1071/AR05188