Breeding Science 63: 292–300 (2013)
doi:10.1270/jsbbs.63.292

Construction of integrated linkage map of a recombinant inbred line population of white lupin (*Lupinus albus* L.)

Cina Ann Vipin[1)], David J. Luckett[1)], John D.I. Harper[1,3)], Gavin J. Ash[1,3)], Andrzej Kilian[2)], Simon R. Ellwood[4)], Huyen T.T. Phan[4)] and Harsh Raman*[1)]

[1)] *Graham Centre for Agricultural Innovation (an alliance between NSW Department of Primary Industries and Charles Sturt University), Wagga Wagga Agricultural Institute*, Wagga Wagga, NSW 2650, Australia
[2)] *Diversity Arrays Technology P/L*, 1 Wilf Crane Crescent, Yarralumla, Canberra, ACT 2600, Australia
[3)] *School of Agricultural and Wine Sciences, Charles Sturt University*, Wagga Wagga, NSW 2678, Australia
[4)] *Department of Environment and Agriculture, Curtin University*, Kent Street, Bentley, Perth, Western Australia 6102, Australia

We report the development of a Diversity Arrays Technology (DArT) marker panel and its utilisation in the development of an integrated genetic linkage map of white lupin (*Lupinus albus* L.) using an F_8 recombinant inbred line population derived from Kiev Mutant/P27174. One hundred and thirty-six DArT markers were merged into the first genetic linkage map composed of 220 amplified fragment length polymorphisms (AFLPs) and 105 genic markers. The integrated map consists of 38 linkage groups of 441 markers and spans a total length of 2,169 cM, with an average interval size of 4.6 cM. The DArT markers exhibited good genome coverage and were associated with previously identified genic and AFLP markers linked with quantitative trait loci for anthracnose resistance, flowering time and alkaloid content. The improved genetic linkage map of white lupin will aid in the identification of markers for traits of interest and future syntenic studies.

Key Words: *Lupinus albus*, genetic linkage mapping, molecular markers, Diversity Array Technology.

Introduction

Of more than 300 species of *Lupinus*, only *L. albus* (white lupin), *L. angustifolius* (blue or narrow-leafed lupin), *L. luteus* (yellow lupin) and *L. mutabilis* (pearl lupin) are widely cultivated (Mulayim *et al.* 2002). White lupin (2n = 50), originated in the Balkan region of the north east Mediterranean is now distributed throughout the Mediterranean region, and from the Azores Islands across North Africa to Ethiopia and Kenya. White lupin is a legume crop of the smooth-seed group known as the Malacosperamae. It is principally exhibits self-pollination, although partial out-crossing is often observed. In Australia, white lupin is grown as a crop mainly in Western Australia and New South Wales.

Lupin seeds are valued for their nutritional components, and contain 30–40% protein, 30% carbohydrate, 6% oil, vitamins and fibre (Lagunes-espinoza *et al.* 2000). They are used as a source of protein for human food and animal feed (Lin *et al.* 2009). In Australia, lupins are grown as a break crop to reduce disease build-up in mainly cereals-based cropping systems and to maintain soil fertility through the fixation of atmospheric nitrogen (Jones *et al.* 2003). White lupin is well adapted to soils having low phosphorus due to its proteoid roots (Neumann and Martinoia 2002).

Communicated by M. Ishimoto
Received October 18, 2012. Accepted May 14, 2013.
*Corresponding author (e-mail: harsh.raman@dpi.nsw.gov.au)

Conventional breeding methods have played an important role in the development of improved varieties for yield, resistance to pests and diseases and desired quality characteristics such as ‘sweet’ seed with low alkaloid content. In order to harness increased genetic gains in lupin breeding programs, more efficient tools for the selection of desirable traits are required. Molecular markers offer useful tools to increase selection efficiency through marker-assisted selection. Genome-wide molecular markers and genetic linkage maps are also essential for characterisation of germplasm, assessment of genetic diversity, the study of inheritance of both qualitative and quantitative traits, map-based gene cloning and for comparative genomic studies.

In white lupin, the first genetic linkage map was published with more than 300 genic and Amplified Fragment Length Polymorphism (AFLP) markers and these were further utilised to identify quantitative trait loci (QTL) for anthracnose resistance, flowering time and alkaloid content (Phan *et al.* 2007). More recently, high resolution melt analysis based Sequence Tagged site (STS) markers were also developed, mapped and utilised for mapping QTL for flowering time, alkaloid synthesis and stem height (Croxford *et al.* 2008). However, further increased marker density is required to saturate the white lupin genome and provide candidate markers suitable for marker-assisted selection in the breeding programs. Markers based on known genomic sequences, or those from which the sequence can be easily obtained (including Diversity Array Technology—DArT

markers), allow comparative mapping approaches to be applied to identify additional markers in regions of interest based on colinearity with related species which have more sequence resources.

The current molecular marker technologies such as AFLP, Simple Sequence Repeats (SSR) and STS have limitations that affect the level of uniform genome coverage, reproducibility and technical and time demands. Most of these markers are gel based and have limited capability to rapidly assay large numbers of marker loci. Some of these limitations can be overcome by specialised technology such as high throughput capillary electrophoresis, which can improve allelic discrimination ability, reproducibility and speed. However, the dependence on pre-existing DNA sequence information and high assay costs of the majority of marker technologies remain major limitations for many genetic improvement programs (Wittenberg *et al.* 2005). DArT has overcome these limitations and has been applied as an alternative to the gel-based marker technologies in various breeding and germplasm development programs and elsewhere (Akbari *et al.* 2006, Jaccoud *et al.* 2001, Raman *et al.* 2012, Wenzl *et al.* 2004, Wittenberg *et al.* 2005, Xia *et al.* 2005, Yang *et al.* 2006).

This paper reports the (i) development of DArT markers in white lupins and (ii) the construction of an integrated genetic linkage map based upon newly developed DArT markers together with the existing AFLP and genic markers that were used to generate the first map of white lupins utilising a recombinant inbred line (RIL) population from Kiev Mutant/P27174 (Phan *et al.* 2007).

Materials and Methods

Mapping population

A subset of 96 F_8 RILs was chosen for DArT analysis. These included the parents Kiev Mutant (maternal parent, introduced into Australia by CSIRO in 1982) and P27174 (paternal parent, an Ethiopian landrace) that were used for the construction of the first linkage map and to locate loci linked to anthracnose resistance, flowering time and alkaloid content (Phan *et al.* 2007).

Development of complexity reduction methods, library creation and array development

The selection of a suitable complexity reduction method is critical in DArT technology. Previously, various combinations of the restriction enzyme (RE) *PstI* with frequently cutting RE's such as *TaqI*, *BstNI*, *AluI*, *BanII*, *MseI*, *HaeIII*, *MspI*, were efficiently used to prepare genomic representation in different plant species including barley, wheat, *Arabidopsis* and sorghum (Akbari *et al.* 2006, Mace *et al.* 2008, Wenzl *et al.* 2006, Wittenberg *et al.* 2005). Initially *PstI*, the primary rare cutting restriction enzyme in conjunction with eight frequently cutting restriction enzymes, *TaqI*, *BstNI*, *AluI*, *BanII*, *MseI*, *HaeIII*, *MspI* and *TaqI* + *MfeI* (New England Biolabs; NEB, USA) were tested to find suitable enzyme combinations for the complexity reduction method. Of the different enzyme combinations tested, *PstI/AluI* and *PstI/MspI* performed best, as they produced homogenous DNA smears on gels instead of distinct bands. Genomic representations from the parental lines, Kiev Mutant and P27174 were prepared to make the library.

DNA digestion and adapter ligation was performed using 100 ng of genomic DNA in buffer containing 10 mM Tris-OAc, 50 mM $KH(O_2CCH_3)_2$, 10 mM $Mg(CH_3COO)_2$ and 5 mM DTT. A *PstI* adapter (5′-CAC GAT GGA TCC AGT GCA-3′ annealed with 5′-CTG GAT CCA TCG TGCA-3′) was simultaneously ligated to the complimentary overhangs with 3 units of T4 DNA ligase (New England Biolabs, Australia) in a total volume of 7.7 µl. Reactions were incubated for 37°C for 2 hours, followed by 2 hours at 60°C as required by the enzyme combinations. One µl of the digestion ligation reaction product was used as a template for PCR amplification in a 50 µl reaction using DArT *PstI* + 0 primer (5′-CAG TCA AGT TAG ATG GTG CAG-3′) using PCR cycles as 94°C for 1 min, 94°C for 20 sec, 58°C for 40 sec, 72°C for 1 min repeat step 2 for another 29 times, 72°C for 7 min and hold at 10°C. PCR products were analysed on 1.2% agarose gel and all targets with an aberrant migration (downshifted targets) or reduced product yield were removed from analysis.

Cloning and library construction for the DArT array

Amplification products (representations) from both the *PstI/MspI* and *PstI/AluI* of the parental lines (P27174 and Kiev Mutant) were ligated into the pCR2.1-TOPO vector using the TOPO cloning kit (Invitrogen Life Technologies Corporation, Carlsbad, CA, USA) and transformed into the heat-shock competent TOP10 *Escherichia coli* strain (Invitrogen) according to manufacturer's instructions. Transformants were selected on LB (Luria-Bertani) medium containing ampicillin (100 µg/ml) and X-gal (40 µg/ml), incubated at 37°C for 16 hour. Transformed bacterial colonies were tooth picked into 384 well plates containing LB medium supplemented with 4.4% glycerol, 8.21 g/L K_2HPO_4, 1.80 g/L KH_2PO_4, 0.50 g/L Na_3-citrate, 0.10 g/L $MgSO_4 \cdot 7H_20$, 0.90 g/L $(NH_4)_2SO_4$ and 100 µg/ml ampicillin and were incubated for approximately 22 hours at 37°C. A set of 3,072 *L. albus* clones was created from each method of *PstI/AluI* and *PstI/MspI* combinations and were stored at –80°C in eight 384 well plates. This library was extended with six more plates of *L. albus* (2,304 clones) and three plates each of *L. angustifolius* (1,152 clones) and *L. luteus* (1,152 clones) using *PstI/BstNI* developed previously by DArT P/L (http://www.diversityarrays.com/) Yarralumla, Australia), making a 20 plate array with a total of 7,680 clones (Table 1).

An aliquot of 0.5 µl of the culture was taken to amplify with 0. 2 µM each of the M13 forward and reverse primers using the PCR program 95°C for 4 min, 57°C for 35 sec, 72°C for 1 min followed by 35 cycles of 94°C for 35 sec, 52°C for 35 sec and 72°C for 1 min and the final extension of 72°C for 7 min. The PCR products were then precipitated

Table 1. Library development and polymorphism revealed using different restriction enzymes for DArT array construction in *Lupinus* species. Number of DArT clones, polymorphic clones and polymorphism rate represent to individual *Lupinus* species/restriction enzyme combination

Species	Restriction enzymes	Genotype ID	Country of origin	Accessions type	Number of clones/*Lupinus* species[#]	Number of polymorphic clones[§]	Polymorphism (%)*
Lupinus albus	*PstI/MspI*	P27174	Ethiopia	Landrace	1536	422	27.5
		P25758	Greece	Landrace			
		Kiev-mutant	Ukraine	Cultivar			
Lupinus albus	*PstI/AluI*	P27174	Ethiopia	Landrace	1536	349	22.7
		P25758	Greece	Landrace			
		Kiev-mutant	Ukraine	Cultivar			
Lupinus albus	*PstI/BstNI*	24706	Ukraine	Landrace	2304	270	11.7
		27715	Unknown	Landrace			
		28530	Unknown	Landrace			
		28548	Unknown	Landrace			
		Amiga	Chile	Cultivar			
		Andromeda	Australia	Cultivar			
		Astra	France	Cultivar			
		ESTA-1	South Africa	Cultivar			
		Hamburg	Germany	Cultivar			
		Kiev-mutant	Ukraine	Cultivar			
		Madeira	Portugal	Cultivar			
		Mini bean	Australia	Cultivar			
		Minori	Germany	Cultivar			
		Multolupa-2	Germany	Cultivar			
Lupinus luteus	*PstI/BstNI*	2D12	Unknown	Unknown	1152	150	13.0
		2E 12	Unknown	Unknown			
Lupinus angustifolius	*PstI/BstNI*	P27255	Unknown	Breeding line	1152	170	14.8
		83A476	Unknown	Unknown			

[#] Number of clones represent to the total number of amplicons created for an individual library.

[§] Number of polymorphic clones represent to the number of polymorphic DArT markers: polymorphism was calculated in four genotypes P27174, Kiev Mutant (the parents of the population used to generate the map), P25758, Kiev Mutant (mapping population generated to map resistance loci to Pleiochaeta Root Rot).

* Polymorphism (%) is calculated as the number of polymorphic clones/total number of clones ×100.

with equal volume of isopropanol at room temperature and washed with 100 µl of 77% ethanol. The ethanol was removed and the products were air-dried and dissolved with DArT spotting buffer 50% DMSO, 1.5 M sorbitol. 0.1 M Triethanolamine. HCl, 0.5% (w/v) dextran, 0.02% (w/v) CHAPS (detergent).

Array development

Amplicons suspended in the spotting buffer were arrayed with two replicates onto poly-L-lysine coated micro-array slides (Erie Scientific) using a Microgrid II arrayer (Biorobotics). The arrays were spotted in 50% relative humidity and then dried in the incubator for 24 hours to allow the DNA adhere to the slides. Slides were processed by immersing in milli-Q water at 92°C for 2 min to denature the DNA and further in milli-Q water with 0.1 mM DTT and 0.1 mM EDTA and finally dried by centrifugation at 1,570 rpm for 7 minutes, followed by vacuum-drying for about 20 min.

Target preparation and labelling of genomic representations

A pilot test was carried out to test polymorphism among eight genotypes: P27174, Kiev Mutant-I, P25758, Kiev Mutant-II and two RILs from the P27174/Kiev Mutant-I and P25758, Kiev Mutant-II—the parental lines of mapping populations using DArT clones generated with different restriction enzymes (Table 1). The *PstI/MspI* DArT arrays revealed the highest polymorphism among arrays using different restriction enzymes and therefore selected for the present study. Genomic representations of 96 samples of the RIL population from Kiev Mutant/P27174 including the parents with two replications were prepared as above for library construction, generating 'targets' for hybridisation to the arrays. The amplification products were then precipitated using isopropanol, washed with 77% ethanol and dried for 12 hours at room temperature. All targets in the dried form were combined with 5 µl labelling mixture containing 1 × NEB buffer 2, 50 µM random decamers and labelling dNTPs (2 mM dATP, 2 mM dCTP, 2 mM dGTP, 0.2 mM dTTP), denatured to 95°C for 3 min and cooled to 25°C before labelling with 5 µl of Cy-dye mix containing 1 × NEB buffer 2, 25 µM cy3-dUTP or cy5-dUTP (Amersham Biosciences) and 2.5 units of Klenow exo-fragment of *E. coli* polymerase 1 (NEB). The labelling reactions were performed at 37°C for 3 hours.

Hybridisation to microarrays

Cy3 and Cy5 labelled targets were mixed and added with 5 µl of the Inactivation Mix (1 × NEB buffer 2, 60 mM

EDTA pH 8.0) and 50 µl of hybridisation buffer preheated to 65°C [FAM-labelled polylinker of the pCR2.1 vector, Express Hyb, (Clonetech, USA), 10 mg/ml herring sperm DNA (Promega, USA) and 2 mM EDTA pH 8.0]. The polylinker fragment was used as a reference to determine the amount of DNA spotted on the array. After denaturing the targets at 95°C for 3 min followed by 56°C for 5 min and maintained at 55°C for a maximum of 1 hour. Within this period of time the denatured targets ready to be hybridised were pipetted (60 µl) onto a 7,680-clone microarray slide, quickly covered with a glass coverslip, placed in a hybridisation chamber and incubated at 65°C overnight. Slides were then washed in four solutions sequentially, solution 1 (1 × SSC (Saline Sodium Citrate) with 0.1% SDS (Sodium Dodecyl Sulphate) for 5 min), solution 2 (1 × SSC for 5 min), solution 3 (0.2 × SSC for 1 min) and solution 4 (0.02 × SSC for 30 sec) and a final wash in milli Q water for 1.5 min. All the washing solutions are added with 0.5 M DTT (Dithiothreitol) and done at room temperature. The targets were quickly dried by centrifugation at 1570 × g for 7 min and vacuum-dried for about 30 min in a dark desiccator.

Image analysis and data extraction

Images of the arrays were produced using a TECAN LS300 confocal microarray scanner (Tecan, Austria) at a resolution of 20 µm per pixel with three images on each slide. This included one reference image using a 488 nm laser with a 520 nm filter combination for the measurement of the signal from the polylinker fragment of the cloning vector common to all fragments on the array and one of the two target images displaying hybridised target labelled with Cy3 using 543 nm laser with 590 nm filter and the other target labelled with Cy5 using 633 nm/670 nm laser filter. The resulting TIF images were analysed using DArTSoft version 7.3, a software program developed by DArT P/L. Polymorphic markers were scored as 1's and 0's based on the membership probability estimates computed by the clustering algorithm. The clustering algorithm also provided a probability estimate for each individual genotype call. Markers that showed conflicting scores between the replicates or could not be scored in either replicates were recorded as 'X' (missing). The segregation ratio at each marker locus was tested for deviation from the expected Mendelian segregation ratio (1 : 1) by chi-squared tests. DArT analysis includes replication as a part of quality assurance. Every sample is labelled with Cy3 and Cy5 dyes. Reproducibility is calculated automatically by the software and the average reproducibility is taken as the average of all the selected markers of high reproducibility within the threshold. Thresholds or criteria selected for the map construction were 95% of reproducibility, 80% of call rate and 74% of *P* value. The reproducibility is obtained from the replicate individuals that were supposed to give identical results. Call rate is the percentage of DNA samples with binary ('0' or '1') allele calls and *P* value is the variance of the relative target hybridisation intensity between allelic states as a percentage of the total variance. The DArT toolbox was used for score merge analysis to check the quality of data from replicates.

Genetic mapping

Genetic linkage map was constructed using DArT linkage group and marker ordering software (Hudson *et al.* 2012). The DArT marker ordering system follows a three step process similar to the one described previously (Cheema and Dicks 2009, Raman *et al.* 2013). The optimum marker order is posed as the travelling salesman path within the group as described by Wu *et al.* (2008). Finally, the Kosambi function (Kosambi 1944) was used to convert recombination frequencies into genetic distance in centiMorgan (cM). Mapchart software was used for displaying the linkage groups graphically. All the DArT markers in the map were prefixed as 'lpms' where 'l' stands for lupin, 'p' stands for primary restriction enzyme used (*PstI*), 'ms' stands for secondary enzyme *MspI* and the numbers correspond to the clone ID.

Results

Mapping of DArT markers

High quality marker data are critical to the construction of good genetic maps. A combination of marker quality parameters automatically generated from DArTsoft, the software developed in-house was used as quality thresholds for the identification of polymorphic clones such as reproducibility, call rate and *P* values. Initially, we incorporated 92 DArT markers that had the parameters: call rate > 80%, reproducibility > 95% and *P* value > 74%. Subsequently, we included 44 more markers as 'attached markers' with less stringent thresholds of reproducibility 92%, call rate 73% and *P* value 64% (Supplemental Table 1). All the markers exhibited a normal segregation ratio of 1 (paternal alleles):1 (maternal alleles) at each DArT locus with the exception of 24 distorted DArT markers (Supplemental Table 1). A total of 136 marker loci with the call rate > 70% were merged with the segregation data of 220 AFLP and 105 gene based markers that were scored in a previous study (Phan *et al.* 2007). The markers that did not show linkage with other markers below threshold were excluded from the dataset and not further used for linkage map construction. The integrated linkage map based upon 441 DArT, AFLP and gene based markers covered 2,169 cM, with an average interval distance of 4.6 cM where as the average spacing between markers was 12.6 cM, with the recombination fraction of 0.27 for the previous map (Phan *et al.* 2007).

Distribution of DArT loci in the white lupin genome

In total, 28 linkage groups were identified in the first map of white lupin (Phan *et al.* 2007). With the addition of DArT markers along with the AFLPs and gene-based markers, the new integrated map comprised 38 linkage groups which varied in length from 2 cM to 153 cM (Table 2). Distribution of DArT markers across the genome in relation to AFLP and

Table 2. Characteristics of integrated map of an RIL population of Kiev Mutant/P27174 with DArT, genic and AFLP markers

Linkage groups[a]	Length (cM)	Total markers	No. of AFLP markers	No. of genic markers	No. of DArT markers	Frame work markers[b]	Attached markers	Delegate markers
*LG1	55.9	16	10	0	6	13	2	3
LG1-II	130.7	20	8	7	5	18	5	2
*LG2-I	42.5	8	5	2	1	7	2	1
LG2-II	92.7	24	11	4	9	18	11	6
*LG3-I	33.1	15	4	0	11	9	8	6
LG3-II	53	16	13	3	0	12	0	4
LG4	151.3	20	13	5	2	20	1	0
*LG5-I	54.2	13	6	3	4	11	4	2
LG5-II	99.4	9	5	3	1	9	0	0
LG6	156	18	11	6	1	14	1	4
LG7	88	18	11	4	3	14	5	4
LG8-I	49.7	14	5	3	6	10	6	4
LG8-II	24.5	5	5	0	0	5	0	0
LG9	79.8	12	12	0	0	9	3	3
LG10	122.6	34	11	2	21	17	7	17
*LG11-I	64.9	6	5	1	0	6	1	0
LG11-II	18.4	5	1	4	0	4	0	1
LG12	50.8	13	7	4	2	13	4	0
LG13-I	45.2	10	3	5	2	9	2	1
LG13-II	5.3	2	1	1	0	2	0	0
LG14	36	7	7	0	0	7	2	0
LG15	54.4	8	4	3	1	8	2	0
LG16	66.1	16	5	3	8	10	0	6
LG17	69.0	20	3	6	11	11	3	9
LG18	67.9	10	5	3	2	10	3	0
LG19	44	12	4	3	5	8	9	4
LG20-I	39.8	8	4	0	4	6	3	2
*LG20-II	7.5	2	0	2	0	2	0	0
LG21	72.2	13	3	4	6	11	1	2
LG22	30.3	6	5	1	0	6	2	0
LG23	33.3	6	3	3	0	5	3	1
LG24	46.2	7	2	3	2	5	2	2
LG25	2.1	2	1	1	0	2	0	0
LG26	14.1	4	3	0	1	4	1	0
LG27	22.9	3	0	3	0	4	1	0
LG28	13.8	3	3	0	0	3	0	0
LG29	11.2	3	0	0	3	2	0	1
LG30	16.9	5	0	1	4	3	2	2
LG31	8.8	5	0	0	5	3	0	2
LG32	6.5	4	1	0	3	2	0	2
LG33	19.4	4	2	0	2	3	0	2
LG34	25.6	7	1	1	5	4	0	1
LG35	8.6	2	0	2	0	2	0	0
LG36	7.4	2	2	0	0	2	0	0
LG37	13.8	2	2	0	0	2	0	0
LG38	12.8	2	2	0	0	2	0	0
Total	2169	441	209	96	136	347	96	94

[a] Linkage groups (LG) of integrated map were co-mapped/integrated with previous map (Phan *et al.* 2007) on the basis of a common set of markers.

[b] Framework markers are those that have consistent map order. Delegate markers map at the same location as the representative frame work marker for a specific locus. Attached markers are those that are excluded at the initial stage and later included in the map assigning them to best intervals on the framework map as described previously (Phan *et al.* 2007).

*Linkage groups identified by Phan *et al.* (2007) that formed additional groups in this study.

gene based markers is shown in Fig. 1. DArT markers were well distributed throughout the linkage groups and showed close linkage with the AFLP and genic markers. The positions of AFLP and genic markers were consistent with those mapped previously. However, 10 AFLP markers that were not mapped previously due to loose or no linkage could be grouped in the present study (LG30 to LG38). This demonstrates the value of combining different markers systems for the construction of linkage map with good coverage.

DArT markers were mapped in the majority (65.7%) of

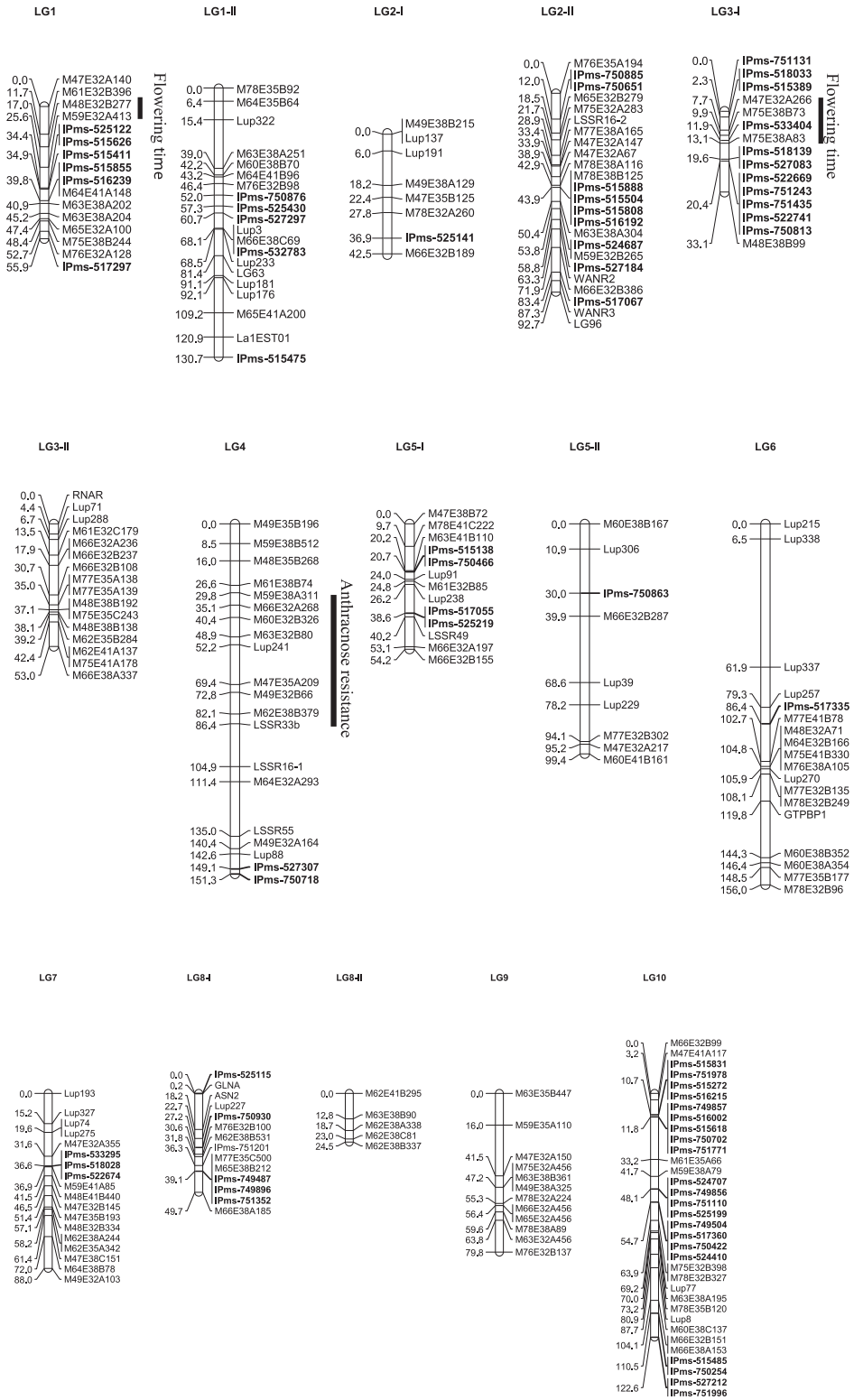

Fig. 1. The integrated map based on DArT, AFLP and genic markers of the RIL population derived from Kiev Mutant/P27174. Map distances (on the left) are given in cM (Kosambi 1944). DArT markers are highlighted in bold. Genomic locations and marker loci delimiting QTLs for flowering time (LG1 and LG3) and anthracnose resistance (LG4 and LG17) are marked with solid lines and marked in bold/italics from the previous map (Phan *et al.* 2007).

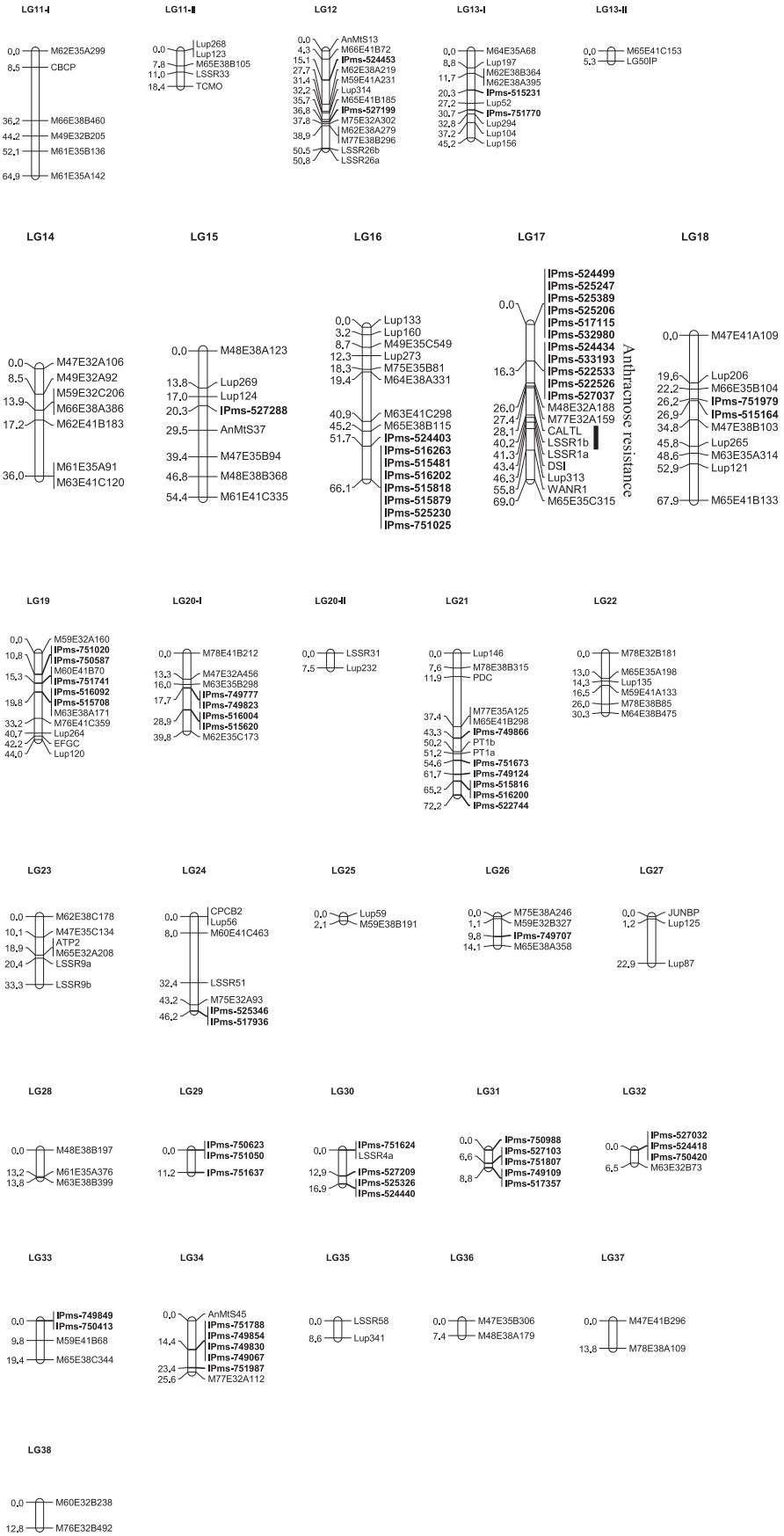

Fig. 1. (continued)

the previously-published linkage groups (LGs) (Table 2). However, there were 13 LGs without DArT markers, and eight of these LGs comprised no more than 3 genic/AFLP markers. Two LGs: LG29 and LG31 contained only DArT markers. Although DArT markers were distributed across the genome in a similar way to the non-DArT markers, there are significant differences in the number of markers among LGs. For example, LG3 and LG10 have quite high marker density (31 and 34 markers), whereas LG6 has low marker density (18 markers) with the largest gap spanning more than 55.4 cM (Fig. 1). In total, 8 linkage groups (LG1, LG2, LG3, LG5, LG8, LG11, LG13 and LG20) from the previous map (Phan *et al.* 2007) were separated to form additional linkage groups due to the poor linkage with DArT markers.

Discussion

We have mapped DArT markers for the first time in white lupin utilising a mapping population derived from Kiev Mutant/P27174 (Phan *et al.* 2007). These markers were selected based on their high call rate, scoring reproducibility and *P* values. When DArT markers were compared with microsatellites and other single locus PCR based markers, the DArT technology simultaneously typed thousands of loci in a single assay and provides a cost-effective and sequence independent tool for whole genome mapping (Wenzl *et al.* 2004). However, this technology depends on a complex process and requires specialised equipment and therefore can not be used routinely in most laboratories. DArT markers are also dominant and can not detect heterozygous/heterogeneous genotypes within an individual plant. Dominant markers are less informative for constructing linkage maps especially in populations derived from intercrosses (e.g. F_2). Furthermore, for a species with low level of sequence diversity, DArT markers may not be the first method of choice for genetic analysis. Nevertheless, our results confirmed that high quality DArT markers can be used for construction of genetic linkage maps in white lupin.

The average reproducibility of all the DArT markers was 98.6% which was similar to that obtained in barley (99.8%) (Wenzl *et al.* 2004), hop (100%) (Howard *et al.* 2011) and pigeonpea (99.7%). Call rate being 90.9% and which expresses the reliability of the final scores versus the maximum number of potential scores, was similar to sugarcane (92.5%) (Heller-Uszynska *et al.* 2010) and banana (91.6%) (Risterucci *et al.* 2009) was slightly lower than barley (95.0%) (Wenzl *et al.* 2004), wheat (99.2%) (Akbari *et al.* 2006), pigeonpea (96.0%) (Yang *et al.* 2006) and hop (97.6%) (Howard *et al.* 2011). The *P* value here was 86.8%, which agreed with the values obtained in hop (89.9%) (Howard *et al.* 2011) and was slightly higher compared to banana (81.4%) and sugarcane (80.7%) (Heller-Uszynska *et al.* 2010).

The new integrated genetic map was shorter (2,169 cM) than the map reported previously which spanned a total length of 2,951 cM (Phan *et al.* 2007). This discrepancy could be due to usage of different algorithms implemented in different mapping software packages, scoring discrepancies of multilocus AFLP markers, and/or the relatively small size of the population used in this study. Interestingly, Phan *et al.* (2007) utilised MultiPoint software version 1.2 (MultiQTL Ltd., Institute of Evolution, Haifa University, Israel), which utilises a TSP algorithm, similar to the DArT software implemented in this paper. However, there are different implementations of TSP embedded in the two programs, which may account for the differences in map length. The most likely explanation for the difference in the map distances is the different marker datasets, as we used a subset of the AFLP and genic markers from earlier study. Most importantly, a shorter map using larger number of markers strongly suggests that the quality of marker scores in the new dataset is higher compared to the initial map. Marker density varied across linkage group. This is consistent with other studies where some LGs or chromosomes of a species possess higher levels of polymorphism than others irrespective of marker type (Akbari *et al.* 2006, Nelson *et al.* 2006, Wenzl *et al.* 2006).

It is also interesting that DArT markers showed a distinct distribution across the genome as compared to AFLP markers. This could be attributed due to the use of the *EcoRI/MseI* enzyme combination used for AFLP analysis. In the present study, we used *PstI* enzyme, which is a methylation-sensitive restriction enzyme. It is apparent that the 'methylation filtration' effect arising from using *PstI* partly enriches genomic representations for hypo-methylated 'gene space' regions in barley (Wenzl *et al.* 2006). A similar picture of DArT marker distribution was observed in several other genomes for which high density DArT maps have been already been developed, for example wheat (Akbari *et al.* 2006) sorghum (Mace *et al.* 2008) and rye (Brągoszewska *et al.* 2009).

Some of DArT markers were mapped within the genomic regions that were associated with anthracnose resistance, flowering time and alkaloid content (Fig. 1) (Phan *et al.* 2007). For example, a suite of DArT markers were mapped within 15 cM from flanking markers associated with flowering time QTLs on LG1-1 (delimited with AFLP marker M61E32B396) and LG3-1 (delimited with AFLP markers M47E32A266 and M75E38B73) (Phan *et al.* 2007). Besides segregating for anthracnose, alkaloid content and flowering time, the population in this study also shows segregation for resistance to phomopsis blight (Cowley *et al.* 2011). Therefore, this integrated map will provide a platform to localise QTLs for resistance to phomopsis blight and the identification of linked molecular markers.

Our results reconfirmed that DArT provides high quality markers that can be used to construct high density genetic linkage maps for plants with no sequence information available. An additional advantage is that useful DArT clones can be readily sequenced to provide information for their conversion into PCR based markers and for linkage group alignment with genomes of other species for which aligned DNA sequence information is available. This can be advantageous in cases when there are not yet any cheap-to-assay markers

closely flanking a potential target QTL that could be used in foreground selection for the favourable allele.

In conclusion, we have more intensively saturated the first map of white lupin which was based on genic and AFLP markers with DArT markers. This new map can be used (i) as a 'seed map' to compare other maps that are based on multilocus AFLP markers, (ii) for various genetic analyses such as genetic diversity, structure and gene flow estimations, and gene mapping, QTL analysis and whole genome assembly studies.

Acknowledgements

This research was supported by Charles Sturt University and NSW Department of Primary Industries through the Graham Centre for Agricultural Innovation. Authors are thankful to Drs. Eric Huttner, Frank Detering, Vanessa Caig, Ling Xia (DArT P/L) Rosy Raman and Ray Cowley (NSW DPI) for their support.

Literature Cited

Akbari, M., P. Wenzl, V. Caig, J. Carling, L. Xia, S. Yang, G. Uszynski, V. Mohler, A. Lehmensiek, H. Kuchel *et al.* (2006) Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. Theor. Appl. Genet. 113: 1409–1420.

Brągoszewska, H.B., K.H. Uzyńska, P. Wenzl, G. Uszyński, A. Kilian and M.R. Trojanowska (2009) DArT markers for the rye genome—genetic diversity and mapping. BMC Genomics 10: 578.

Cheema, J. and J. Dicks (2009) Computational approaches and software tools for genetic linkage map estimation in plants. Briefings in Bioinformatics 10: 595–608.

Cowley, R., G. Ash, J. Harper, A. Smith, B. Cullis and D.J. Luckett (2011) Application of multi-phase experiments in plant pathology to identify genetic resistance to *Diaporthe toxica* in *Lupinus albus*. Euphytica 186: 655–669.

Croxford, A.E., T. Rogers, P.D.S. Caligari and M.J. Wilkinson (2008) High-resolution melt analysis to identify and map sequence-tagged site anchor points onto linkage maps: a white lupin (*Lupinus albus*) map as an exemplar. New Phytol. 180: 594–607.

Heller-Uszynska, K., G. Uszynska, G. Huttner, M. Evers, J. Carling, V. Caig, K. Aitken, P. Jackson, G. Piperidis, M. Cox *et al.* (2010) Diversity arrays technology effectively reveals DNA polymorphism in a large and complex genome of sugarcane. Mol. Breed. 28: 37–55.

Howard, E.L., S.P. Whittock, J. Jakse, J. Carling, P.D. Matthews, G. Probasco, J.A. Henning, P. Darby, A. Cerenak, B. Javornik *et al.* (2011) High-throughput genotyping of hop (*Humulus lupulus* L.) utilising diversity arrays technology (DArT). Theor. Appl. Genet. 122: 1265–1280.

Hudson, C.J, J.S. Freeman, A.R.K. Kullan, C.D. Petroli, C.P. Sansaloni, A. Kilian, F. Detering, D. Grattapaglia, B.M. Potts, A.A. Myburg *et al.* (2012) A reference linkage map for Eucalyptus. BMC Genomics 13: 240.

Jaccoud, D., K. Peng, D. Feinstein and A. Kilian (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res. 29: e25.

Jones, R.A.C., B.A. Coutts and Y. Cheng (2003) Yield limiting potential of necrotic and non necrotic strains of *Bean yellow mosaic virus* in narrow-leafed lupin (*Lupinus angustifolius*). Aust. J. Agric. Res. 54: 849–859.

Kosambi, D.D. (1944) The estimation of map distance from recombination values. Annals of Eugenics 12: 172–175.

Lagunes-espinoza, L.C., C. Huyghe and J. Papineau (2000) Genetic variation for pod wall concentration in *Lupinus albus*. Plant Breed. 119: 421–425.

Lin, R., D. Renshaw, D. Luckett, J. Clements, G. Yan, K. Adhikari, B. Buirchell, M. Sweetingham and H. Yang (2009) Development of a sequence-specific PCR marker linked to the gene "*pauper*" conferring low-alkaloids in white lupin (*Lupinus albus* L.) for marker assisted selection. Mol. Breed. 23: 153–161.

Mace, E.S., L. Xia, D.R. Jordan, K. Halloran, D.K. Parh, E. Huttner, P. Wenzl and A. Kilian (2008) DArT markers: diversity analyses and mapping in *Sorghum bicolor*. BMC Genomics 9: 26.

Mulayim, M., A. Tamkoc and M. Babaoglu (2002) Sweet lupins versus local bitter genotype: agronomic characteristic as affected by different planting densities in the goller region of Turkey. Eur. J. Agr. 17: 181–189.

Nelson, M., H.T.T. Phan and S.R. Ellwood (2006) The first gene-based map of *Lupinus angustifolius* L-location of domestication genes and conserved synteny with *Medicago truncatula*. Theor. Appl. Genet. 113: 225–238.

Neumann, G. and E. Martinoia (2002) Cluster roots—an underground adaptation for survival in extreme environments. Trends Plant Sci. 7: 162–167.

Phan, H.T., S.R. Ellwood, K. Adhikari, M.N. Nelson and R.P. Oliver (2007) The first genetic and comparative map of white lupin (*Lupinus albus* L.): identification of QTLs for anthracnose resistance and flowering time and a locus for alkaloid content. DNA Res. 14: 59–70.

Raman, H., R. Raman, M.N. Nelson, M.N. Aslam, R. Rajasekaran, N. Wratten, W.A. Cowling, A. Kilian, A.G. Sharpe and J. Schondelmaier (2012) Diversity array technology markers: genetic diversity analyses and linkage map construction in rapeseed (*Brassica napus* L.). DNA Res. 19: 51–65.

Raman, H., R. Raman, A. Kilian, F. Detering, Y. Long, D. Edwards, I.A.P. Parkin, A.G. Sharpe, M.N. Nelson, N. Larkan *et al.* (2013) A consensus map of rapeseed (*Brassica napus* L.) based on diversity array technology markers: applications in genetic dissection of qualitative and quantitative traits. BMC Genomics 14: 277.

Risterucci, A.M., I. Hippolyte, X. Perrier, L. Xia, V. Caig, M. Evers, E. Huttner, A. Kilian and J.C. Glaszmann (2009) Development and assessment of diversity arrays technology for high-throughput DNA analyses in Musa. Theor. Appl. Genet. 119: 1093–1103.

Wenzl, P., J. Carling, D. Kudrna, D. Jaccoud, E. Huttner, A. Kleinhofs and A. Kilian (2004) Diversity Arrays Technology (DArT) for whole-genome profiling of barley. Proc. Natl. Acad. Sci. USA 101: 9915–9920.

Wenzl, P., H. Li, J. Carling, M. Zhou, H. Raman, E. Paul, P. Hearnden, C. Maier, L. Xia, V. Caig *et al.* (2006) A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. BMC Genomics 7: 206.

Wu, Y., P.R. Bhat, T.J. Close and S. Lonardi (2008) Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. PLoS Genet. 4: e1000212.

Wittenberg, A.H.J., T. Van der Lee, C. Cayla, A. Kilian, R.G. Visser and H.J. Schouten (2005) Validation of the high-throughput marker technology DArT using the model plant *Arabidopsis thaliana*. Mol. Genet. Genomics 274: 30–39.

Xia, L., K. Peng, S. Yang, P. Wenzl, M. Carmen de Vicente, M. Fregene and A. Kilian (2005) DArT for high-throughput genotyping of Cassava (*Manihot esculenta*) and its wild relatives. Theor. Appl. Genet. 110: 1092–1098.

Yang, S., W. Pang, G. Ash, J. Harper, J. Carling, P. Wenzl, E. Huttner, X. Zong and A. Kilian (2006) Low level of genetic diversity in cultivated Pigeonpea compared to its wild relatives is revealed by diversity arrays technology. Theor. Appl. Genet. 113: 585–595.