Theor Appl Genet (2014) 127:1237–1249
DOI 10.1007/s00122-014-2294-y

ORIGINAL PAPER

New evidence of ancestral polyploidy in the Genistoid legume *Lupinus angustifolius* L. (narrow-leafed lupin)

Magdalena Kroc · Grzegorz Koczyk · Wojciech Święcicki · Andrzej Kilian · Matthew N. Nelson

Received: 23 August 2013 / Accepted: 21 February 2014 / Published online: 15 March 2014

Abstract

Key message **This is the first clear evidence of duplication and/or triplication of large chromosomal regions in a genome of a Genistoid legume, the most basal clade of Papilionoid legumes.**

Abstract *Lupinus angustifolius* L. (narrow-leafed lupin) is the most widely cultivated species of Genistoid legume, grown for its high-protein grain. As a member of this most basal clade of Papilionoid legumes, *L. angustifolius* serves as a useful model for exploring legume genome evolution. Here, we report an improved reference genetic map of *L. angustifolius* comprising 1207 loci, including 299 newly developed Diversity Arrays Technology markers and 54 new gene-based PCR markers. A comparison between the *L. angustifolius* and *Medicago truncatula* genomes was performed using 394 sequence-tagged site markers acting as bridging points between the two genomes. The improved *L. angustifolius* genetic map, the updated *M. truncatula* genome assembly and the increased number of bridging points between the genomes together substantially enhanced the resolution of synteny and chromosomal colinearity between these genomes compared to previous reports. While a high degree of syntenic fragmentation was observed that was consistent with the large evolutionary distance between the *L. angustifolius* and *M. truncatula* genomes, there were striking examples of conserved colinearity of loci between these genomes. Compelling evidence was found of large-scale duplication and/or triplication in the *L. angustifolius* genome, consistent with one or more ancestral polyploidy events.

Communicated by A. H. Paterson.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-014-2294-y) contains supplementary material, which is available to authorized users.

M. Kroc · W. Święcicki
Department of Genomics, Institute of Plant Genetics, Polish Academy of Sciences, Strzeszynska 34, 60-479 Poznan, Poland

G. Koczyk
Department of Biometry and Bioinformatics, Institute of Plant Genetics, Polish Academy of Sciences, Strzeszynska 34, 60-479 Poznan, Poland

A. Kilian
Diversity Arrays Technology Pty Ltd, 1 Wilf Crane Crescent, Yarralumla, Canberra, ACT 2600, Australia

M. N. Nelson
School of Plant Biology, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

M. N. Nelson (✉)
The UWA Institute of Agriculture, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia
e-mail: matthew.nelson@uwa.edu.au

Introduction

Narrow-leafed lupin (*Lupinus angustifolius* L.) is one of four lupin grain crop species that together produce around 1.1 million tonnes of grain annually (FAO 2011). It was first cultivated as a green manure and forage crop in the late nineteenth century in Northern Europe and only became a significant grain crop as domestication traits including reduced seed alkaloids were incorporated through systematic breeding efforts in Europe and Australia (Berger et al.

2012; Brummund and Święcicki 2011; Wolko et al. 2011). Currently, narrow-leafed lupin grain is produced mainly in Australia and Eastern Europe for animal feed though there is increasing interest in its use in the human diet due to its high protein and fibre content, and low glycaemic index (Foley et al. 2011).

Narrow-leafed lupin is a diploid, self-pollinating legume species ($2n = 40$) with an estimated physical genome size of 924 Mb (1C, cv. Sonet) (Kasprzak et al. 2006). It belongs to the genus *Lupinus* L. which comprises around 267 species including both annual and perennial species and ranging from tiny herbaceous plants to medium-sized trees (Drummond et al. 2012). *Lupinus* has among the highest speciation rates known for any genus with particular rapid speciation observed in the Andes (Hughes and Eastwood 2006). *Lupinus* is part of the basal tribe Genisteae within the phylogenetic clade Genistoid in the Papilionoideae subfamily of legumes (Drummond et al. 2012; Lavin et al. 2005). Lupins are highly diverged from all the agriculturally important legumes and model species, most of which belong to two other Papilionoideae clades: Galegoid (cool season legumes) and Phaseoloid (warm season legumes) clades from which they separated about 56 million years ago (Lavin et al. 2005; Zhu et al. 2005).

Polyploidy has played a crucial role in angiosperm genome evolution including the genus *Lupinus*. Multiple rounds of polyploidy provide a rich source of gene redundancy that permits rapid diversification of function and/or expression of genes and contributes to the genome plasticity of angiosperms (Leitch and Leitch 2008). A whole genome duplication (WGD) event was associated with the origin of the angiosperms (De Bodt et al. 2005). An analysis of a large sample of plant transcriptomes revealed evidence of a whole genome triplication (WGT) event associated with the early diversification of the eudicots (Jiao et al. 2012). Recent studies confirm the occurrence of a further whole genome duplication (WGD) event early in Papilionoid legume evolution, which may also have included the basal Genistoid clade containing the *Lupinus* genus (Cannon et al. 2010; Young and Bharti 2012). An additional WGD event occurred in the lineage of *Glycine max* around 13 million years ago (Schmutz et al. 2010). The occurrence of additional Genistoid-specific polyploidy event(s) in the genus *Lupinus* is supported by variation in chromosome numbers, nuclear DNA content, duplication of isozyme markers and DNA markers, and duplicated genes in transcriptome and genome survey sequences (Naganowska et al. 2003; Nelson et al. 2006; Parra-Gonzalez et al. 2012; Wolko and Weeden 1989; Yang et al. 2013). However, Nelson et al. (2006) found little evidence of conserved genetic linkage between 58 pairs of duplicated markers which indicated that this additional polyploidy event(s) in the *Lupinus* lineage was likely to be ancient.

Comparative mapping and synteny analysis are powerful tools for evaluating evolutionary relationships among related taxa and for understanding the structural changes that differentiate the genomes of present-day species (Bertioli et al. 2009; Ellwood et al. 2008; Young and Bharti 2012). Genome conservation of related species has been demonstrated in many plant families with Poaceae, Brassicaceae and Solanaceae being well-known examples (International_Brachypodium_Initiative 2010; Lagercrantz and Lydiate 1996; Tomato_Genome_Consortium 2012). In recent years, a growing number of studies also demonstrated substantial conservation among legume genomes giving hope that genomic information gathered from model genomes can be successfully applied to crop legume improvement (Cannon et al. 2009; Choi et al. 2004; Zhu et al. 2005). However, the ability to transfer knowledge between species reduces as the extent of structural changes increases (Bertioli et al. 2009; Cannon et al. 2009). Since *Lupinus* diverged from the model legume species early in Papilionoid evolution, it is expected that the many structural changes will differentiate the *Lupinus* and model legume genomes. Although recent studies confirm the widespread synteny among legumes, the extent of conservation is still not well understood in the Genistoid clade including the economically and ecologically significant genus *Lupinus* (Lambers et al. 2013). Previous studies revealed short regions of conserved synteny between the *L. angustifolius* and two model species: *Medicago truncatula* and *Lotus japonicus* (Nelson et al. 2006, 2010). Similarly, the structure of the *Lupinus albus* genome appears to be highly rearranged relative to the *M. truncatula* genome (Phan et al. 2007). A low-density survey sequence of the *L. angustifolius* genome was recently reported with a small proportion of scaffolds assigned to linkage groups, but a synteny analysis to other legume genomes was not attempted (Yang et al. 2013). Previous synteny studies suffered from low resolution due to the low number of bridging points, which in case of Nelson et al. (2006) and Phan et al. (2007) was exacerbated by the use of a rudimentary *M. truncatula* reference genome sequence.

In the present comparative analysis, we employed an improved genetic map of the *L. angustifolius* genome with an almost threefold increase in the number of sequence-based genetic markers. In addition, an improved assembly of the *M. truncatula* genome (Young et al. 2011) was enlisted in this updated synteny analysis. The study not only improved the resolution of synteny between these two genomes, but also revealed compelling evidence of one or more ancient polyploidy events shaping the *L. angustifolius* genome.

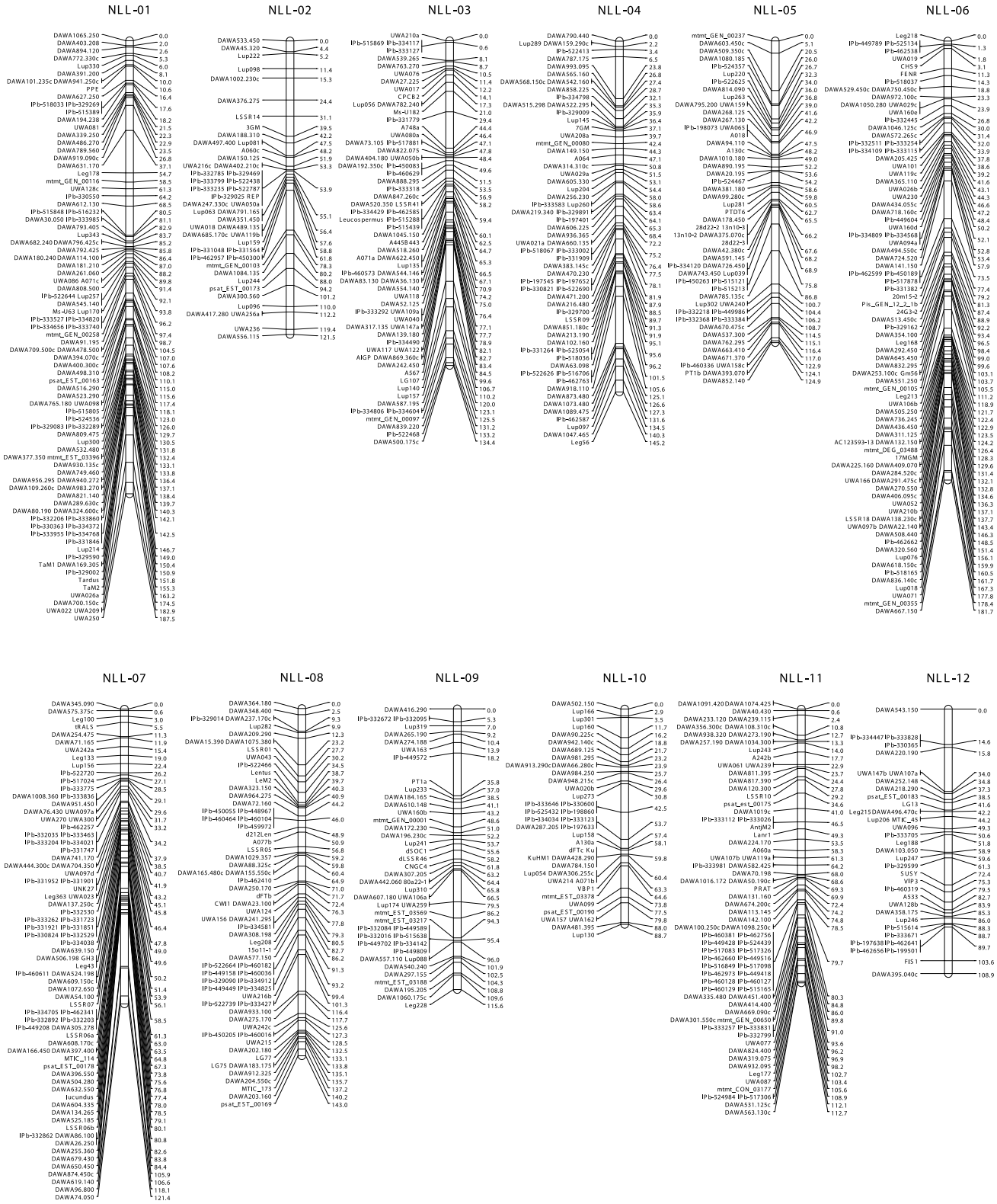

Fig. 1 Linkage map of the *Lupinus angustifolius* genome comprising 1,200 markers and seven trait loci distributed over 20 linkage groups (NLL-01 to NLL-20) and three small clusters (Cluster-1 to Cluster-3). Genetic distances are in Kosambi centiMorgans. Framework and redundant markers are presented here. A more detailed genetic map including attached markers is provided in Figure S1

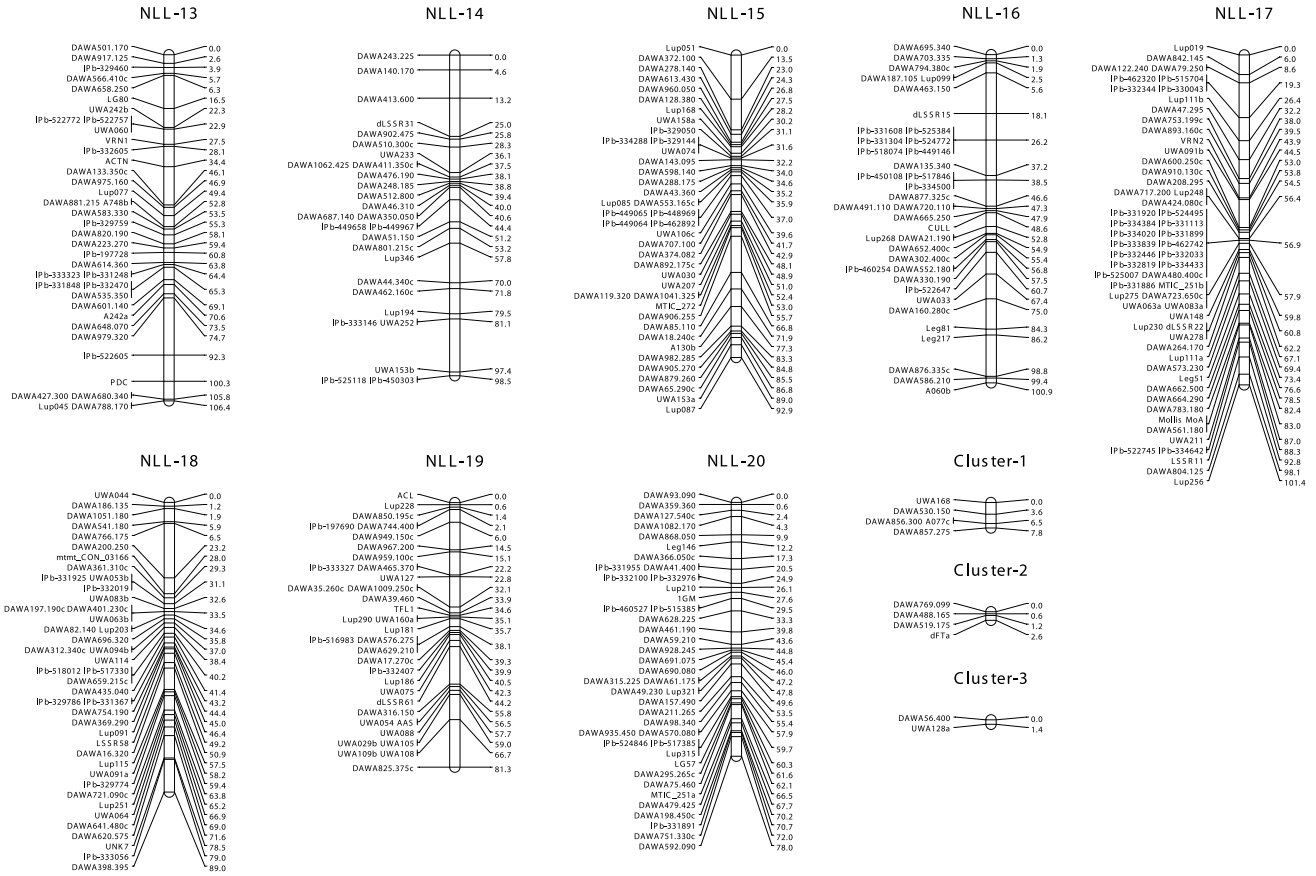

Fig. 1 continued

Materials and methods

Mapping population

The mapping population used in this study consisted of 112 recombinant inbred lines (RILs), derived from a cross between a domesticated line (83A:476) and a wild type (P27255), and was developed at the Department of Agriculture and Food Western Australia (Perth, Australia). The parental lines were selected on the basis of contrasting phenotypes for domestication loci: *Ku* (reduced vernalisation requirement for flowering), *iucundus* (low alkaloid content), *lentus* and *tardus* (both for reduced pod shattering), *mollis* (water-permeable seeds) and *leucospermus* (visual marker for domestication conferring white flowers and reduced pigmentation in many tissues). This mapping population and phenotyping of the domestication traits and anthracnose resistance locus *Lanr1* have been described in detail in previous mapping studies encompassing different subsets of RILs (Boersma et al. 2005, 2009; Li et al. 2011; Nelson et al. 2006, 2010; Yang et al. 2004).

DArT markers

Diversity array technology (DArT) marker assay was performed by Diversity Arrays Technology Pty Ltd (Canberra, Australia) according to Jaccoud et al. (2001) and the protocols described by Kilian et al. (2012). Briefly, a genomic representation of a mixture of DNA of both parental lines was generated after *Pst*I–*Taq*I digestion (as a complexity reduction method) and distributed on microarray slide. The resulting array was used to genotype the fluorescently labelled RIL individuals, prepared by using the same complexity reduction method. A total of 299 polymorphic loci were scored as present (1) or absent (2) with the aid of dedicated software DArTsoft. Using the parental control samples, the scoring phase was determined for each locus and data converted to 'A' (maternal parent allele) and 'B' (paternal parent allele). The DNA sequences of 149 DArT non-redundant clones were determined by conventional Sanger sequencing and can be accessed at the GSS database of GenBank (accession numbers KG701214 to KG701371).

Table 1 Summary of the updated reference genetic map of *Lupinus angustifolius* comprising 20 linkage groups (NLL-01 to NLL-20) and three small clusters

Linkage group	Total marker number	Framework markers[a]	Redundant markers[b]	Attached markers[c]	Genetic Length (cM)	Average spacing[d] (cM)	Trait locus
NLL-01	108	72	28	8	187.5	2.64	*Tardus*
NLL-02	51	28	19	4	121.5	4.50	
NLL-03	74	48	22	4	134.4	2.86	*Leucospermus*
NLL-04	67	50	15	2	145.2	2.96	
NLL-05	64	41	17	6	124.9	3.12	
NLL-06	93	72	17	4	181.7	2.56	
NLL-07	92	54	29	9	121.4	2.29	*Iucundus*
NLL-08	70	48	19	3	143	3.04	*Lentus*
NLL-09	51	34	11	6	115.6	3.50	
NLL-10	49	26	15	8	88.7	3.55	*Ku*
NLL-11	83	45	32	6	112.7	2.56	*Lanr1*
NLL-12	43	28	8	7	108.9	4.03	
NLL-13	43	31	8	4	106.4	3.55	
NLL-14	33	23	5	5	98.5	4.48	
NLL-15	43	34	7	2	92.9	2.82	
NLL-16	42	26	11	5	100.9	4.04	
NLL-17	63	31	29	3	101.4	3.38	*Mollis*
NLL-18	44	34	9	1	89	2.70	
NLL-19	37	26	9	2	81.3	3.25	
NLL-20	46	34	8	4	78	2.36	
Cluster-1	5	4	1	0	7.8	2.60	
Cluster-2	4	4	0	0	2.6	0.87	
Cluster-3	2	2	0	0	1.4	1.40	
TOTAL	1,207	795	319	93	2,345.7	2.95	7

[a] Framework markers are high-quality markers each with a unique position in the genetic map

[b] Redundant markers have identical positions as their respective framework markers

[c] Attached markers have ambiguous map positions and are placed in the most likely interval between framework markers

[d] Marker distances are calculated between framework markers only

New PCR-based markers

New PCR-based STS markers comprised length polymorphism, SNaPshot (Life Technologies Inc., Foster City, USA), cleaved amplified polymorphic sequence (CAPS) or derived-CAPS (Neff et al. 2002). Table S1 details the primer sequences, assay details and parental amplicon sizes for new PCR-based markers. Parental amplicon sequences can be accessed at the GSS database of GenBank (accession numbers KG701372 to KG701468).

Six new gene-based STS primer pairs (mtmt_EST_03396, mtmt_DEG_03488, tRALS, mtmt_GEN_00650, SHK75 and mtmt_EST_03280) were developed within the Sixth European Union Framework Programme's Grain Legumes Integrated Project (GLIP) and implemented in *L. angustifolius* using the method described by Nelson et al. (2010). Seven new intron-targeted STS primer pairs designed to amplify orthologous single or low copy genes LG1, LG11, LG13, LG16, LG75, LG96 and LG107) were provided by Prof. Richard Oliver and Dr. Simon Ellwood (Curtin University, Perth, Australia). Thirty gene-based legume anchor primer pairs (prefixed with "Leg") were developed by Fredslund et al. (2006) and provided by Prof. Jens Stougaard (University of Aarhus, Denmark) (Table S1). The marker assay Gm56 targeting a candidate gene for pod shattering was provided by Dr. Varma Penmetsa (UC Davis, USA).

Eight flowering time gene homologue markers (dFTa, dFTb, dFTc, dSOC1, TFL1, VIN3, VIP3 and VRN1) were developed by designing primers in conserved regions identified by aligning legume ESTs with *Arabidopsis thaliana* flowering time genes with the aid of Vector NTI (Invitrogen, Carlsbad, California). The primer pair AC123593-13 was designed in the same way except using the *M. truncatula* BAC clone sequence AC123593 as the template. The resulting PCR amplicons from parents 83A:476 and P27255 were sequenced to confirm the successful

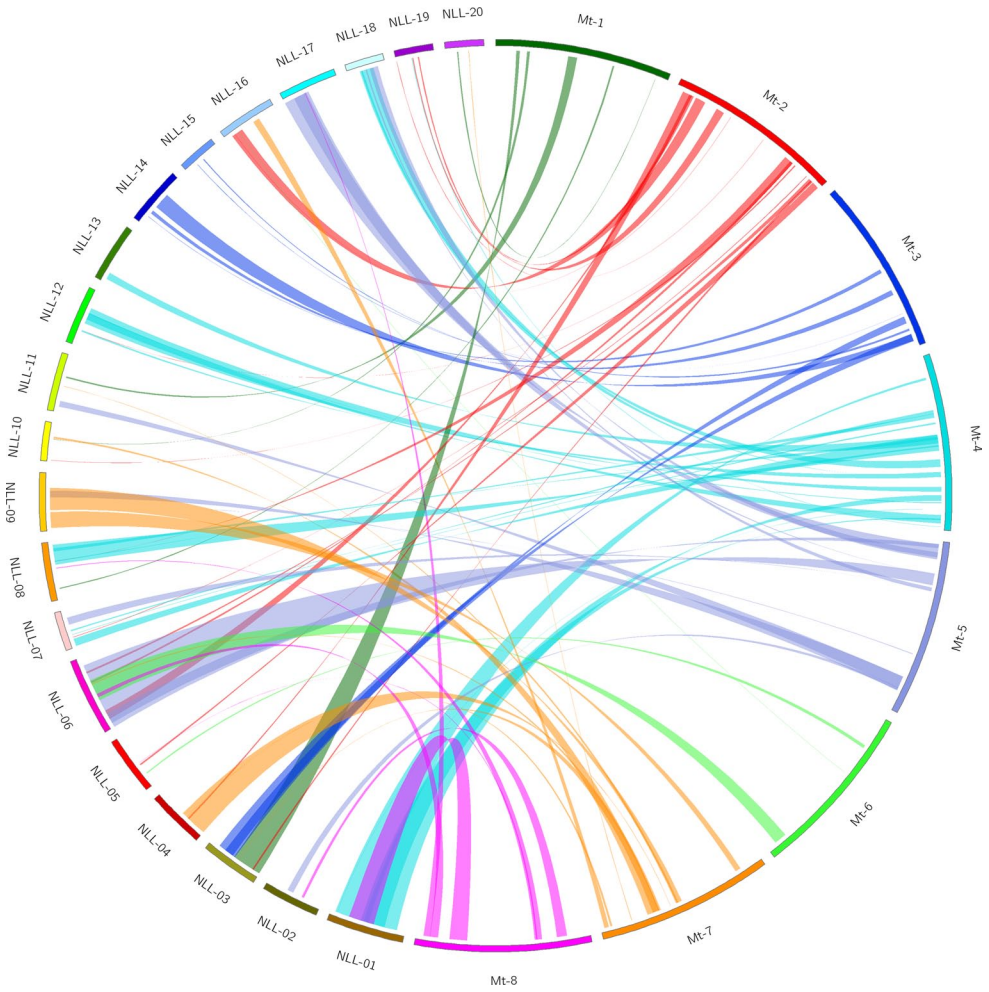

Fig. 2 Circos plot of synteny between the *Medicago truncatula* reference genome and the *Lupinus angustifolius* genetic map. Lines linking syntenic regions are coloured according to *M. truncatula* chromosome. The chromosomes and linkage groups are not drawn to scale

amplification of the targeted genes and to identify SNP polymorphisms, which were then assayed in the RIL population using length, CAPS or dCAPS assays (Table S1). An additional flowering time gene homologue marker developed for pea was also used (VRN2; Hecht et al. 2005).

Genetic mapping

The previous version of the *L. angustifolius* reference map comprised 1,090 marker and trait loci (Nelson et al. 2010). Seven trait loci and the highest quality markers from that map were selected for this current study: 208 PCR-based sequence-tagged site (STS), 157 restriction fragment length polymorphism (RFLP) and 492 microsatellite-anchored fragment length polymorphism (MFLP) markers (864 loci in total; Table S1). When combined with the 353 new DArT and PCR-based STS markers, there were 1,217 loci in total, which were then subjected to linkage mapping.

Linkage mapping was performed with the aid of MultiPoint 2.1 (MultiQTL Ltd, Haifa, Israel), which uses the 'evolutionary optimisation strategy' (Mester et al. 2003) to perform multi-locus ordering of linkage groups. We used the approach described in detail by Nelson et al. (2010) and Raman et al. (2012) with some modifications. Briefly, redundant markers were set aside before commencement of clustering analysis. Iterative clustering analysis was conducted at a recombination frequency: rf = 0.12, 0.15 and 0.18 and then increased at 0.02 increments until a maximum of rf = 0.28. At each stage, multi-point analysis was conducted and resulting groups merged as rf was incrementally increased. Jack-knife analysis was performed on the rf = 0.28 linkage groups to identify markers that had a destabilising effect on locus order, which were then temporarily set aside. The remaining markers were used to construct the framework map with genetic intervals size transformed to account for multiple meioses involved in the development of the RIL population and expressed in Kosambi centiMorgans (cM). Redundant markers were then assigned to their representative framework markers and destabilising markers were assigned (or 'attached') to the most likely intervals between framework markers.

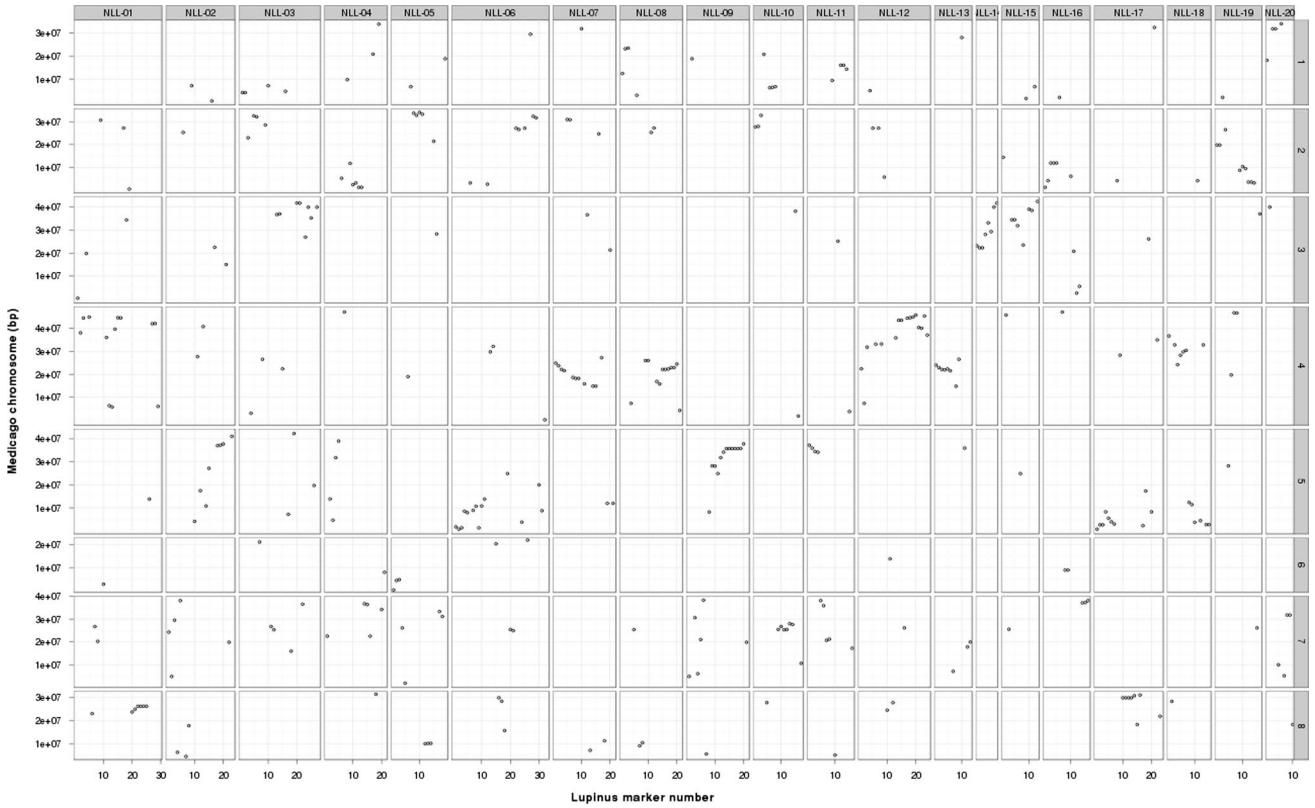

Fig. 3 Dot plot of *Lupinus angustifolius* linkage groups vs. *Medicago truncatula* chromosomes. Markers within *L. angustifolius* linkage groups (*x*-axis) are presented in sequential order without scale. Positions of loci within *M. truncatula* chromosomes are drawn in base pair scale (*y*-axis)

Synteny analyses

DNA sequences corresponding from the STS markers (PCR-based, RFLP and DArT) used to generate the genetic map of *L. angustifolius* were used to identify orthologous loci in the *M. truncatula* reference genome version Mt3.5.1 (Young et al. 2011) using BLASTn analysis with a minimum *e*-value of 1 e^{-5} and minimum bit score of 50. The two most significant matches were retained for synteny analysis.

Patterns of conserved locus order between the *L. angustifolius* and *M. truncatula* genomes were visualised using a custom ggplot2 script (Wickham 2009), Strudel (Bayer et al. 2011) and Circos v0.63 (Krzywinski et al. 2009). For Circos visualisation, the links were pre-processed with the "bundlelinks" utility script, grouping together markers mapped to the same *L. angustifolius* linkage group and occupying positions within *M. truncatula* genome spaced less than 100 kbp apart.

Results

An improved genetic map for *Lupinus angustifolius*

A new linkage map of narrow-leafed lupin genome incorporating new PCR-based and DArT markers was constructed with the aid of MultiPoint software. The map comprised 1,200 markers and 7 trait loci distributed over 20 linkage groups (NLL-01 to NLL-20) and three small clusters (Fig. 1; Fig. S1; Table S1). A further ten markers remained unlinked (Table S1). The new map was 2,345.7 cM in length with linkage groups ranging from 78 to 187.5 cM and the average spacing between unique framework markers was 2.95 cM (Table 1). The markers are well distributed with just 13 intervals exceeding 15 cM and 1 interval exceeding 20 cM (Fig. 1; Fig. S1). All linkage groups contained at least one marker from the major marker types (PCR-based STS, RFLP, DArT and MFLP). The map included 566 markers with associated DNA sequences that could potentially be used to align the genetic map of *L. angustifolius* with the reference genome sequence of the model legume *M. truncatula*.

Synteny between *Lupinus angustifolius* and *Medicago truncatula* genomes

Synteny between the genomes of *L. angustifolius* and *M. truncatula* was assessed by comparing the positions of STS markers in the *L. angustifolius* genetic map with the positions of putatively orthologous sequences in the reference genome of *M. truncatula*. Of the 566 STS markers

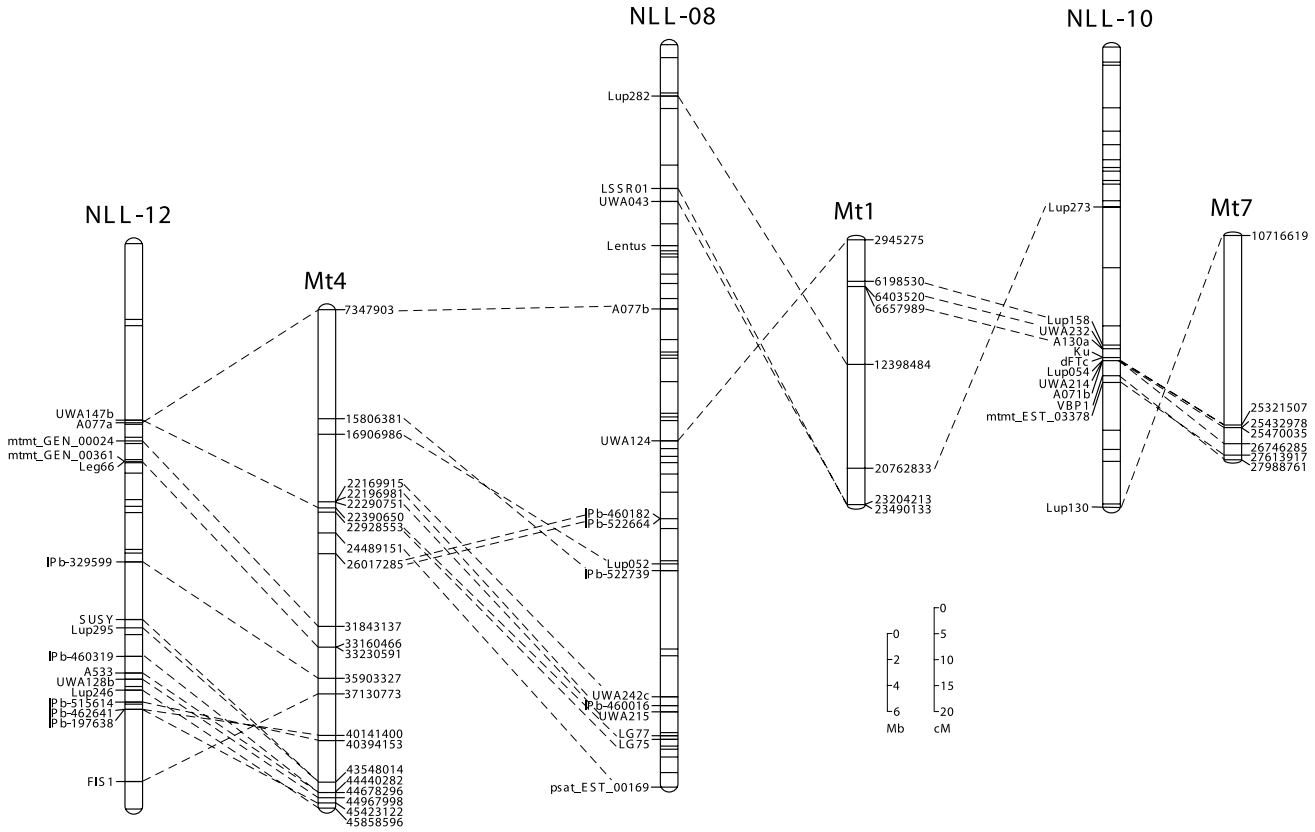

Fig. 4 Examples of synteny between three linkage groups of *Lupinus angustifolius* (NLL-08, NLL-10 and NLL-12) and three chromosomes of *Medicago truncatula* (Mt1, Mt4 and Mt7). The pod shatter resistance gene, *Lentus*, is located on NLL-08, while the early flowering gene, *Ku*, is located on NLL-10. Names of loci showing syntenic regions are shown next to the *L. angustifolius* linkage groups (scaled in Kosambi centiMorgans, cM) and the start position of the homologous sequence on *M. truncatula* is presented next to each chromosome (scaled in megabases, Mb)

mapped in *L. angustifolius*, 410 (72.4 %) found one or more significant matches in the *M. truncatula* genome by BLASTn analysis. Of these, 16 markers were excluded from further analyses as they matched repetitive sequences within the *M. truncatula* genome leaving 394 markers for synteny analysis. Table S2 presents the BLASTn results of 394 markers against the two most significant matches in the *M. truncatula* genome. The primary match (i.e. the most significant) was considered the 'best match' (i.e. the most likely orthologous locus in the *M. truncatula* genome) for 367 (93.1 %) markers. For 27 (6.9 %) markers the second most significant match appeared to be the best match in the *M. truncatula* genome on the basis of conserved synteny and colinearity relative to neighbouring markers (Table S2).

Circos visualisation of genome-wide synteny revealed the high level of fragmentation of genome structure between *L. angustifolius* and *M. truncatula* with each *M. truncatula* chromosome sharing syntenic regions with two or more *L. angustifolius* linkage groups (Fig. 2). Dot-plot analysis was then used to examine these relationships in more detail (Fig. 3). At the segmental level, syntenic regions could be detected in all 20 linkage groups of *L. angustifolius* and all eight chromosomes of *M. truncatula* (Mt1–Mt8). The longest conserved block in the two genomes comprised 13 markers (NLL-09 and Mt5; Fig. 3). When evaluating 160 pairwise chromosome comparisons defined as 20 *L. angustifolius* × 8 *M. truncatula* chromosomes, there were 53 with at least 3 marker correspondences, indicating a high degree of syntenic fragmentation. Mt4 was the chromosome with the most correspondences (90) with *L. angustifolius* linkage groups, whereas Mt6 had the fewest (12). Detailed examples of selected regions of the *L. angustifolius* genome that showed extensive marker colinearity with Mt1, Mt4 and Mt7 are presented in Fig. 4. Four trait loci (*tardus*, *lentus*, *Lanr1* and *Ku*) fell within, or adjacent to, conserved syntenic blocks (Fig. 4; Fig. S1). Interestingly, the flowering time gene homologue marker dFTc mapped to the same genetic location as the flowering time locus, *Ku*, with no recombination detected between the two loci (Fig. 4).

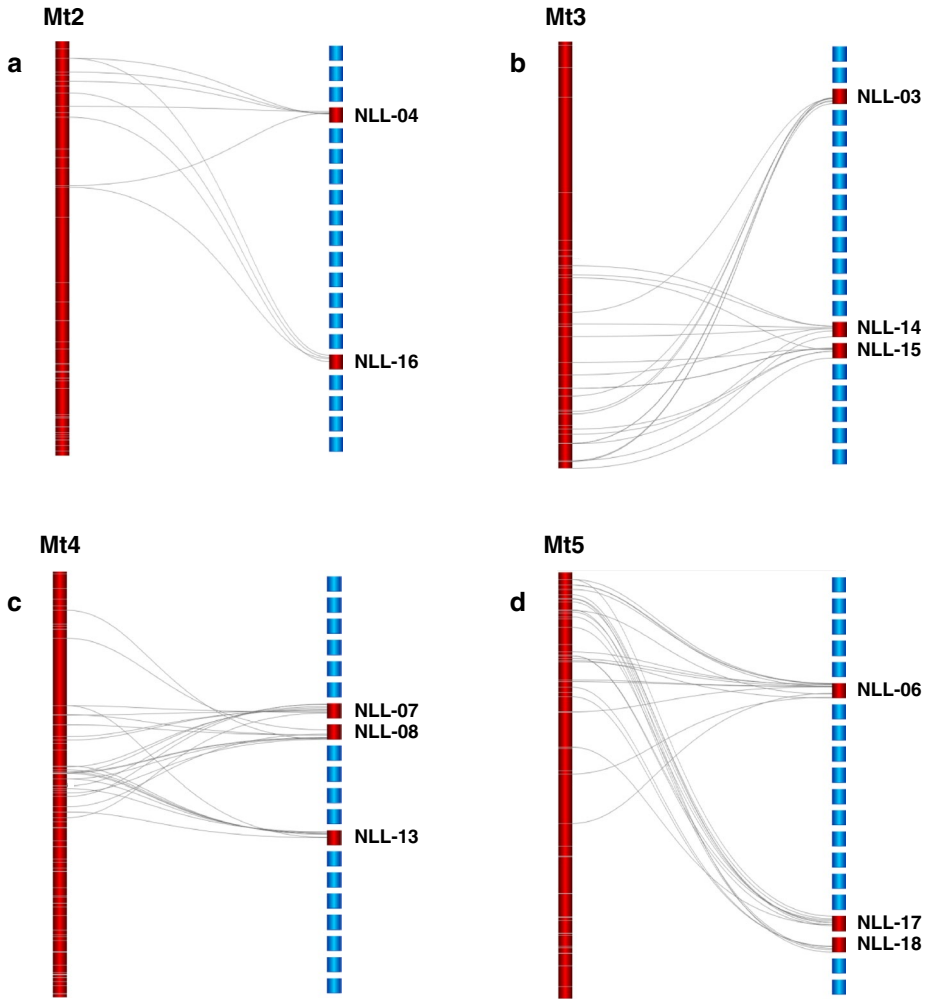

Fig. 5 Examples of duplication and triplication in the *Lupinus angustifolius* genome detected by comparison with the *Medicago truncatula* genome. *M. truncatula* chromosomes (Mt) are shown on the *left* of each panel and *L. angustifolius* linkage groups (NLL) on the *right*, with homologous loci in both genomes connected by *lines*. Chromosomes and linkage groups have normalised total lengths and drawn to scale within each chromosome (Mb scale) and linkage group (cM scale). Precise positions are presented in Table S2. **a** The top of Mt2 shares synteny with NLL-04 and NLL-16. **b** The bottom of Mt3 shares synteny with NLL-03, NLL-14 and NLL-15. **c** The middle of Mt4 shares synteny with NLL-07, NLL-08 and NLL-13. **d** The top of Mt5 shares synteny with NLL-06, NLL-17 and NLL-18

Evidence of triplication in the lupin genome

Strikingly, there were several instances of large *M. truncatula* chromosome segments matching more than one *L. angustifolius* genomic region (Fig. 3), which were investigated further using Strudel visualisation. Figure 5 shows four of the clearest examples of *M. truncatula* chromosome segments that each matched two or three regions of the *L. angustifolius* genome. Figure 6 shows in detail one of the triplicated regions on *L. angustifolius* NLL-07, NLL-08 and NLL-13 which corresponded to *M. truncatula* chromosome Mt4.

Discussion

The key finding of this study was the first ever observation of duplicated and triplicated regions in the *L. angustifolius* genome that were present as single copies in the *M. truncatula* genome (Figs. 3, 5, 6). Such duplications and/or triplications were observed on parts of most *L. angustifolius* linkage groups, which points to WGD or WGT arising from polyploidy event(s) rather than multiple independent duplications of single chromosomes. Not all of the *L. angustifolius* genome showed clear duplication and/or triplication; therefore, it is likely that the polyploidy event(s) was ancient and subsequently chromosomes have undergone numerous rearrangements. The conclusion of ancient polyploidy event(s) is supported by previously presented evidence based on diversified genome size and chromosome numbers within the genus, isozyme and DNA marker duplication, and duplicated genes in transcriptome and genome survey sequences (Naganowska et al. 2003; Nelson et al. 2006; Parra-Gonzalez et al. 2012; Wolko and Weeden 1989; Yang et al. 2013).

The timing of the inferred polyploidy event(s) remains an open question. Based on the synteny-based analysis of this study and the above-mentioned published studies, it appears that at least one polyploidy event took place after the divergence of *Lupinus* from *M. truncatula* and other Papilionoid legumes around 56 million years ago (Lavin et al. 2005). This question is being addressed in ongoing

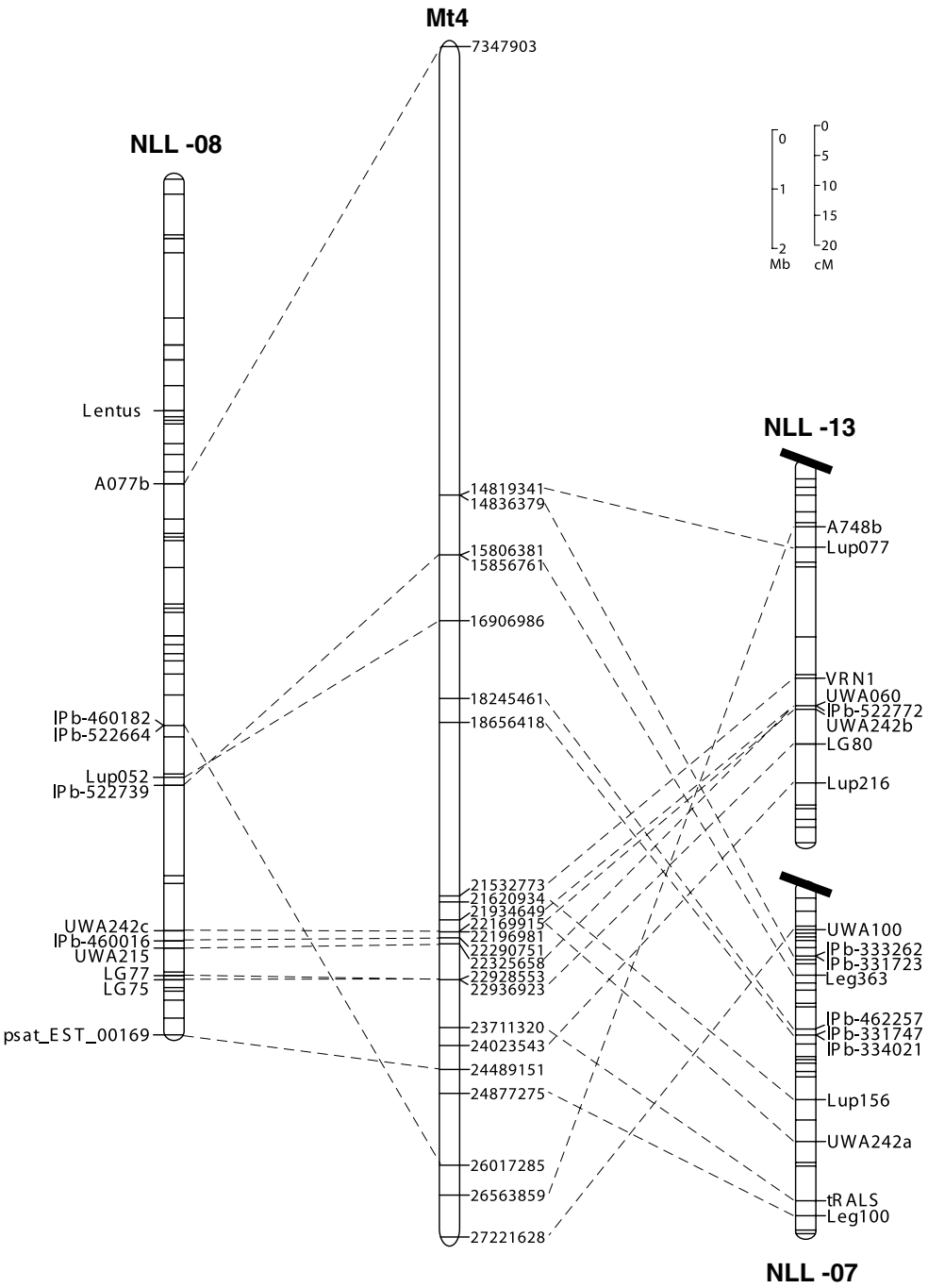

Fig. 6 An example of triplication in the *Lupinus angustifolius* genome (on linkage groups NLL-07, NLL-08 and NLL-13) relative to one *Medicago truncatula* chromosome (Mt4). Names of loci showing syntenic regions are shown next to the *L. angustifolius* linkage groups (scaled in Kosambi centiMorgans, cM) and the start position of the homologous sequence on *M. truncatula* is presented next to each chromosome (scaled in megabases, Mb). NLL-07 and NLL-08 are presented in reverse loci order and truncated (indicated by a *slanted*, *thick bar*)

collaborative projects that are estimating the divergence times of homoeologous gene sequences obtained from transcriptomes of members of the Genistoid clade (C. Hughes, G. Aitcheson, D. Filatov and M. Nelson, unpublished data) and across the legume family (S. Cannon, unpublished data). These ongoing analyses should provide more robust inferences about the timing of polyploidy event(s) in Genistoid genome evolution. The identification of homoeologous gene pairs would be aided by the availability of a high-quality reference genome. The genome survey sequence of *L. angustifolius* recently reported (Yang et al. 2013) could not resolve gene duplications, highlighting the need for the high-quality reference genome sequence currently being developed for *L. angustifolius* (Gao et al. 2011). The current analysis will guide the sequence assembly of the *L. angustifolius* reference genome as it navigates the complexities of a polyploid genome. Owing to the accumulating genomic and transcriptomic data, *L. angustifolius* may soon become the model genome for the other Genistoid legume species.

The updated reference genetic map of *L. angustifolius* reported here comprised 1,207 loci, including 352 new, high-quality DArT and PCR-based STS markers. These new markers improved genome coverage with the number of unique framework loci increasing to 795 (Table 1) compared to 637 in the previous reference map (Nelson et al. 2010). The number of intervals >15 cM reduced from 18 (Nelson et al. 2010) to 13 (Fig. 1) in the current map. Intriguingly, three small clusters comprising markers generated by diverse PCR, RFLP and DArT technologies remained despite the increased marker density (Fig. 1). These clusters may represent ends of chromosomes that have low marker coverage, high recombination frequency, structural rearrangements, or a combination of the above. These chromosomal regions may present a challenge to incorporate in the more comprehensive *L. angustifolius* genome sequencing project that is currently underway (Gao et al. 2011) and may require additional cytogenetic analyses such as the BAC-FISH analysis developed by Lesniewska et al. (2011).

The most significant technical advance in the new map was the increased number of sequence-based markers (394) used as bridging points for comparing the *L. angustifolius* and *M. truncatula* genomes (Table S2). This was a substantial increase over the 147 bridging points used in the previous synteny analysis between *L. angustifolius* and *M. truncatula* (Nelson et al. 2006). This, along with the availability of an improved genome assembly of *M. truncatula*, permitted a higher-resolution analysis of synteny between the basal Papilionoid *L. angustifolius* with the model legume *M. truncatula*. While the overall impression of high differentiation between the two genomes remained unchanged (Fig. 2), there were much clearer examples of marker colinearity between the two genomes (Figs. 4, 6) compared to the previous study (Nelson et al. 2006). This improved delineation of conserved gene order in the basal Papilionoid genome of *L. angustifolius* will guide the reconstruction of the ancestral genomes of cool season and warm season currently underway (D. Cook, pers. comm.).

Four trait loci (*tardus*, *lentus*, *Lanr1* and *Ku*) in the *L. angustifolius* genetic map fell within, or closely adjacent to, regions of conserved synteny with *M. truncatula* (Fig. 4, Fig. S1). For example, the pod shattering gene *Lentus* on linkage group NLL-08 was located in the synteny block shared with Mt1 (Fig. 4) as well as in a conserved block of *Lotus japonicus* chromosome 5 (Nelson et al. 2010). It was elsewhere reported that Mt1 and Lj5 show synteny along their entire lengths (Cannon et al. 2006), which brings further strength to the possible synteny exploitation for identification of a candidate gene for *Lentus*. The region of *L. angustifolius* NLL-10 containing the flowering time locus *Ku* shared synteny with the region of *M. truncatula* Mt7 containing three *FT* homologues, the floral integrator gene encoding the florigen signal protein in plants (Turck et al. 2008). Intriguingly, the *FT*-derived marker dFTc showed no recombination with *Ku* (Fig. 1; Table S1). Further research is underway to determine if *FT* is indeed the gene underlying the *Ku* locus or is instead another closely linked flowering time gene such as *CONSTANS* (Pierre et al. 2011). If *FT* is demonstrated to be the gene underlying the *Ku* locus, this would be a strong validation of the synteny approach for transferring genomic knowledge from a model genome to a less well-resourced crop genome.

Acknowledgments Development of gene-based STS markers conducted by MK was supported by the Sixth European Union Framework Programme Grain Legumes Integrated Project and a grant from Ministry of Agriculture and Rural Development. We thank Richard Oliver and Simon Ellwood (Curtin University) and Jens Stougaard (University of Aarhus) for providing primers for STS markers, and for the technical assistance of Clare O'Lone in developing the flowering time gene homologue markers. The DArT genotyping was made possible by a small grant from the Centre for Legumes in Mediterranean Agriculture at the University of Western Australia (UWA). MNN especially thanks Doug Cook (UC Davis) for hosting a 3-month research visit where regular insightful conversations led to the recognition of ancient duplications in the narrow-leafed lupin genome. MNN also acknowledges the UWA Fay Gale Fellowship that supported that research visit.

Conflict of interest The authors declare no conflict of interest.

Ethical standards The authors declare that the experiments reported here comply with the current laws of Australia and Poland.

References

Bayer M, Milne I, Stephen G, Shaw P, Cardle L, Wright F, Marshall D (2011) Comparative visualization of genetic and physical maps with Strudel. Bioinformatics 27:1307–1308

Berger J, Buirchell B, Luckett D, Nelson M (2012) Domestication bottlenecks limit genetic diversity and constrain adaptation in narrow-leafed lupin (*Lupinus angustifolius* L.). Theor Appl Genet 124:637–652

Bertioli D, Moretzsohn M, Madsen L, Sandal N, Leal-Bertioli S, Guimaraes P, Hougaard B, Fredslund J, Schauser L, Nielsen A, Sato S, Tabata S, Cannon S, Stougaard J (2009) An analysis of synteny of *Arachis* with *Lotus* and *Medicago* sheds new light on the structure, stability and evolution of legume genomes. BMC Genom 10:45

Boersma JG, Pallotta M, Li CD, Buirchell BJ, Sivasithamparam K, Yang H (2005) Construction of a genetic linkage map using MFLP and identification of molecular markers linked to domestication genes in narrow-leafed lupin (*Lupinus angustifolius* L.). Cell Mol Biol Lett 10:331–344

Boersma J, Nelson M, Sivasithamparam K, Yang H (2009) Development of sequence-specific PCR markers linked to the *Tardus* gene that reduces pod shattering in narrow-leafed lupin (*Lupinus angustifolius* L.). Mol Breed 23:259–267

Brummund M, Święcicki W (2011) The recent history of lupin in agriculture. In: Naganowska B, Wolko B, Kachlicki P (eds) 13th International Lupin Conference, "Lupin crops—an opportunity for today, a promise for the future", Poznań, Poland, pp 15–26

Cannon SB, Sterck L, Rombauts S, Sato S, Cheung F, Gouzy J, Wang X, Mudge J, Vasdewani J, Schiex T, Spannagl M, Monaghan E, Nicholson C, Humphray SJ, Schoof H, Mayer KFX, Rogers J, Quetier F, Oldroyd GE, Debelle F, Cook DR, Retzel EF, Roe BA, Town CD, Tabata S, Van de Peer Y, Young ND (2006) Legume genome evolution viewed through the *Medicago truncatula* and *Lotus japonicus* genomes. Proc Natl Acad Sci USA 103:14959–14964

Cannon SB, May GD, Jackson SA (2009) Three sequenced legume genomes and many crop species: rich opportunities for translational genomics. Plant Physiol 151:970–977

Cannon SB, Ilut D, Farmer AD, Maki SL, May GD, Singer SR, Doyle JJ (2010) Polyploidy did not predate the evolution of nodulation in all legumes. PLoS One 5:e11630

Choi H-K, Mun J-H, Kim D-J, Zhu H, Baek J-M, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, Young ND, Cook DR (2004) Estimating genome conservation between crop and model legume species. Proc Natl Acad Sci USA 101:15289–15294

De Bodt S, Maere S, Van de Peer Y (2005) Genome duplication and the origin of angiosperms. Trends Ecol Evol 20:591–597

Drummond CS, Eastwood RJ, Miotto STS, Hughes CE (2012) Multiple continental radiations and correlates of diversification in *Lupinus* (Leguminosae): testing for key innovation with incomplete taxon sampling. Syst Biol 61:443–460

Ellwood S, Phan H, Jordan M, Hane J, Torres A, Avila C, Cruz-Izquierdo S, Oliver R (2008) Construction of a comparative genetic map in faba bean (*Vicia faba* L.); conservation of genome structure with *Lens culinaris*. BMC Genom 9:380

FAO (2011) FAOSTAT. Food and Agriculture Organization of the United Nations, Rome

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 Biol 11:59

Fredslund J, Madsen LH, Hougaard BK, Sandal N, Stougaard J, Bertioli D, Schauser L (2006) GeMprospector—online design of cross-species genetic marker candidates in legumes and grasses. Nucleic Acids Res 34:W670–W675

Gao L–L, Hane J, Kamphuis L, Foley R, Shi B-J, Atkins C, Singh K (2011) Development of genomic resources for the narrow-leafed lupin (*Lupinus angustifolius*): construction of a bacterial artificial chromosome (BAC) library and BAC-end sequencing. BMC Genom 12:521

Hecht V, Foucher F, Ferrandiz C, Macknight R, Navarro C, Morin J, Vardy ME, Ellis N, Beltran JP, Rameau C, Weller JL (2005) Conservation of *Arabidopsis* flowering genes in model legumes. Plant Physiol 137:1420–1434

Hughes C, Eastwood R (2006) Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. Proc Natl Acad Sci 103:10334–10339

International_Brachypodium_Initiative (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463:763–768

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

Jiao Y, Leebens-Mack J, Ayyampalayam S, Bowers J, McKain M, McNeal J, Rolf M, Ruzicka D, Wafula E, Wickett N, Wu X, Zhang Y, Wang J, Zhang Y, Carpenter E, Deyholos M, Kutchan T, Chanderbali A, Soltis P, Stevenson D, McCombie R, Pires J, Wong G, Soltis D, dePamphilis C (2012) A genome triplication associated with early diversification of the core eudicots. Genome Biol 13:R3

Kasprzak A, Šafar J, Janda J, Dolezel J, Wolko B, Naganowska B (2006) The bacterial artificial chromosome (BAC) library of the narrow-leafed lupin (*Lupinus angustifolius* L.). Cell Mol Biol Lett 11:396–407

Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, Caig V, Heller-Uszynska K, Jaccoud D, Hopper C, Aschenbrenner-Kilian M, Evers M, Peng K, Cayla C, Hok P, Uszynski G (2012) Diversity arrays technology: a generic genome profiling technology on open platforms. In: Pompanon F, Bonin A (eds) Data production and analysis in population genomics. Methods in Molecular Biology, vol 888. Humana Press, New York, pp 67–89. doi:10.1007/978-1-61779-870-2_5

Krzywinski M, Schein J, Birol İ, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645

Lagercrantz U, Lydiate DJ (1996) Comparative genome mapping in *Brassica*. Genetics 144:1903–1910

Lambers H, Clements JC, Nelson MN (2013) How a phosphorus-acquisition strategy based on carboxylate exudation powers the success and agronomic potential of lupines (Lupinus, Fabaceae). Am J Bot 100:263–288

Lavin M, Herendeen P, Wojchiechowski M (2005) Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. Syst Biol 54:575–594

Leitch AR, Leitch IJ (2008) Genomic plasticity and the diversity of polyploid plants. Science 320:481–483

Lesniewska K, Ksiazkiewicz M, Nelson MN, Mahé F, Aïnouche A, Wolko B, Naganowska B (2011) Assignment of 3 genetic linkage groups to 3 chromosomes of narrow-leafed lupin. J Hered 102:228–236

Li X, Yang H, Buirchell B, Yan G (2011) Development of a DNA marker tightly linked to low-alkaloid gene iucundus in narrow-leafed lupin (*Lupinus angustifolius* L.) for marker-assisted selection. Crop Pasture Sci 62:218–224

Mester DI, Ronin YI, Hu Y, Peng J, Nevo E, Korol AB (2003) Efficient multipoint mapping: making use of dominant repulsion-phase markers. Theor Appl Genet 107:1102–1112

Naganowska B, Wolko B, Sliwinska E, Kaczmarek Z (2003) Nuclear DNA content variation and species relationships in the genus *Lupinus* (Fabaceae). Ann Bot 92:349–355

Neff MM, Turk E, Kalishman M (2002) Web-based primer design for single nucleotide polymorphism analysis. Trends Genet 18:613–615

Nelson M, Phan H, Ellwood S, Moolhuijzen P, Hane J, Williams A, O'Lone C, Fosu-Nyarko J, Scobie M, Cakir M, Jones M, Bellgard M, Ksiazkiewicz M, Wolko B, Barker S, Oliver R, Cowling W (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

Nelson MN, Moolhuijzen PM, Boersma JG, Chudy M, Lesniewska K, Bellgard M, Oliver RP, Swiecicki W, Wolko B, Cowling WA, Ellwood SR (2010) Aligning a new reference genetic map of *Lupinus angustifolius* with the genome sequence of the model legume, *Lotus japonicus*. DNA Res 17:73–83

Parra-Gonzalez L, Aravena-Abarzua G, Navarro-Navarro C, Udall J, Maughan J, Peterson L, Salvo-Garrido H, Maureira-Butler I (2012) Yellow lupin (*Lupinus luteus* L.) transcriptome sequencing: molecular marker development and comparative studies. BMC Genom 13:425

Phan HTT, Ellwood SR, Adhikari K, Nelson MN, Oliver RP (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

Pierre J-B, Bogard M, Herrmann D, Huyghe C, Julier B (2011) A *CONSTANS*-like gene candidate that could explain most of the genetic variation for flowering date in *Medicago truncatula*. Mol Breed 28:25–35

Raman H, Raman R, Nelson MN, Aslam MN, Rajasekaran R, Wratten N, Cowling WA, Kilian A, Sharpe AG, Schondelmaier J (2012) Diversity array technology markers: genetic diversity analyses

and linkage map construction in rapeseed (*Brassica napus* L.). DNA Res 19:51–65

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang X-C, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183

Tomato_Genome_Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635–641

Turck F, Fornara F, Coupland G (2008) Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. Annu Rev Plant Biol 59:573–594

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer, New York

Wolko B, Weeden N (1989) Estimation of *Lupinus* genome polyploidy on the basis of isozymic loci number. Genetica Polonica 30:165–171

Wolko B, Clements JC, Naganowska B, Nelson MN, Yang H (2011) *Lupinus*. In: Kole C (ed) Wild crop relatives: genomic and breeding resources. Springer, Heidelberg, pp 153–206

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.). Mol Breed 14:145–151

Yang H, Tao Y, Zheng Z, Zhang Q, Zhou G, Sweetingham MW, Howieson JG, Li C (2013) Draft genome sequence, and a sequence-defined genetic linkage map of the legume crop species *Lupinus angustifolius* L. PLoS One 8:e64799

Young ND, Bharti AK (2012) Genome-enabled insights into legume biology. Annu Rev Plant Biol 63:283–305

Young ND, Debelle F, Oldroyd GE, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KF, Gouzy J, Schoof H, Van de Peer Y, Proost S, Cook DR, Meyers BC, Spannagl M, Cheung F, De Mita S, Krishnakumar V, Gundlach H, Zhou S, Mudge J, Bharti AK, Murray JD, Naoumkina MA, Rosen B, Silverstein KA, Tang H, Rombauts S, Zhao PX, Zhou P, Barbe V, Bardou P, Bechner M, Bellec A, Berger A, Berges H, Bidwell S, Bisseling T, Choisne N, Couloux A, Denny R, Deshpande S, Dai X, Doyle JJ, Dudez AM, Farmer AD, Fouteau S, Franken C, Gibelin C, Gish J, Goldstein S, Gonzalez AJ, Green PJ, Hallab A, Hartog M, Hua A, Humphray SJ, Jeong DH, Jing Y, Jocker A, Kenton SM, Kim DJ, Klee K, Lai H, Lang C, Lin S, Macmil SL, Magdelenat G, Matthews L, McCorrison J, Monaghan EL, Mun JH, Najar FZ, Nicholson C, Noirot C, O'Bleness M, Paule CR, Poulain J, Prion F, Qin B, Qu C, Retzel EF, Riddle C, Sallet E, Samain S, Samson N, Sanders I, Saurat O, Scarpelli C, Schiex T, Segurens B, Severin AJ, Sherrier DJ, Shi R, Sims S, Singer SR, Sinharoy S, Sterck L, Viollet A, Wang BB, Wang K, Wang M, Wang X, Warfsmann J, Weissenbach J, White DD, White JD, Wiley GB, Wincker P, Xing Y, Yang L, Yao Z, Ying F, Zhai J, Zhou L, Zuber A, Denarie J, Dixon RA, May GD, Schwartz DC, Rogers J, Quetier F, Town CD, Roe BA (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480:520–524

Zhu H, Choi H-K, Cook DR, Shoemaker RC (2005) Bridging model and crop legumes through comparative genomics. Plant Physiol 137:1189–1196