INDEL variation in the regulatory region of the major flowering time gene *LanFTc1* is associated with vernalisation response and flowering time in narrow-leafed lupin (*Lupinus angustifolius* L.)

Candy M. Taylor[1], Lars G. Kamphuis[2,3,5], Weilu Zhang[1], Gagan Garg[2], Jens D. Berger[2], Mahsa Mousavi-Derazmahalleh[1], Philipp E. Bayer[4], David Edwards[4,5], Karam B. Singh[2,3,5], Wallace A. Cowling[1,5,*] & Matthew N. Nelson[1,5,6]

[1]UWA School of Agriculture and Environment, The University of Western Australia, Perth, Western Australia, Australia 6009, [2]Agriculture and Food, Commonwealth Scientific and Industrial Research Organisation, Floreat, Western Australia, Australia 6014, [3]Centre for Crop and Disease Management, Curtin University, Bentley, Western Australia, Australia 6102, [4]School of Biological Sciences, The University of Western Australia, Perth, Western Australia, Australia 6009, [5]The UWA Institute of Agriculture, The University of Western Australia, Perth, Western Australia, Australia 6009, [6]Natural Capital and Plant Health, Royal Botanic Gardens, Kew, Ardingly, West Sussex, United Kingdom RH17 6TN

** Correspondence: W. A. Cowling. e-mail: wallace.cowling@uwa.edu.au*

Short running title: INDEL variation in the regulatory region of *LanFTc1*

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/pce.13320

ABSTRACT

Narrow-leafed lupin (*Lupinus angustifolius* L.) cultivation was transformed by two dominant vernalisation-insensitive, early flowering time loci known as *Ku* and *Julius* (*Jul*), which allowed expansion into shorter season environments. However, reliance on these loci has limited genetic and phenotypic diversity for environmental adaptation in cultivated lupin. We recently predicted that a 1,423 bp deletion in the *cis*-regulatory region of *LanFTc1*, a *FLOWERING LOCUS T* (*FT*) homologue, de-repressed expression of *LanFTc1* and was the underlying cause of the *Ku* phenotype. Here, we surveyed diverse germplasm for *LanFTc1 cis*-regulatory variation, and identified two further deletions of 1,208 bp and 5,162 bp in the 5' regulatory region, which overlap the 1,423 bp deletion. Additionally, we confirmed that no other polymorphisms were perfectly associated with vernalisation responsiveness. Phenotyping and gene expression analyses revealed that *Jul* accessions possessed the 5,162 bp deletion and that the *Jul* and *Ku* deletions were equally capable of removing vernalisation requirement and up-regulating gene expression. The 1,208 bp deletion was associated with intermediate phenology, vernalisation responsiveness and gene expression, and therefore may be useful for expanding agronomic adaptation of lupin. This insertion/deletion (INDEL) series may also help resolve how the vernalisation response is mediated at the molecular level in legumes.

Key-words: Narrow-leafed lupin; *LanFTc1*; flowering time; vernalisation response; *cis*-regulation; insertion/deletion (INDEL); variant series

Brief Summary:

"Greater genetic and phenotypic diversity for flowering time would enhance the adaptation of narrow-leafed lupin to Australian and northern European agricultural environments. We studied the major floral integrator gene in narrow-leafed lupin, LanFTc1, and found a series of three overlapping independent deletions within the promoter region which reduce vernalisation responsiveness, shorten flowering time and heighten expression of LanFTc1. The smallest deletion results in a unique intermediate phenotype which may be valuable for breeding purposes and expanding the agronomic adaptation of lupin to new and existing production regions. Additionally, the insertion/deletion (INDEL) series may also help to increase knowledge of how signalling within the vernalisation pathway is mediated at a molecular level in legumes."

INTRODUCTION

Narrow-leafed lupin (*Lupinus angustifolius* L.) is one of three fully domesticated Old World *Lupinus* species originating from the Mediterranean and northern Africa (Gladstones, 1974). It is predominantly grown in Australia and several northern European countries, including Poland, Russia, Germany, Ukraine and Belarus, as a winter and summer annual pulse crop, respectively (FAO, 2014; Gladstones, 1970). The grain is most commonly marketed as a high protein livestock and aquaculture feed. However, in light of the nutritional and metabolomics properties of narrow-leafed lupin seed (Lima-Cabello et al., 2017) and the associated benefits to human health and disease prevention (Foyer et al., 2016; Kouris-Blazos & Belski, 2016), it is also being promoted in the human food market. In addition to high-protein grain production, narrow-leafed lupin has great agricultural value as a superior break crop. Lupin crops mobilise soil-bound phosphorus through carboxylate exudation (Lambers et al., 2013) and improve soil nitrogen through symbiosis, which, together with disease and weed control, is beneficial to the performance of subsequent crop rotations (Seymour et al., 2012).

Following domestication, one of the most important achievements in narrow-leafed lupin breeding has been the manipulation of phenology. In most wild populations, a prolonged period of exposure to cold winter temperatures, known as vernalisation, is required to promote the transition from vegetative to reproductive growth (Rahman & Gladstones, 1972). However, this trait caused great difficulty in the global transitioning of the species from a green manure and fodder crop into a broad acre grain crop. Within key production zones of Australia, winter temperatures are often mild and incapable of reliably saturating the vernalisation requirement from year to year, leading to delayed phenology, susceptibility to terminal drought stress and reduced final yields (Berger et al., 2012b; Gladstones, 1977). Similarly, strong vernalisation requirement was also a problem for summer cropping of lupin in northern Europe, where cool and wet autumn conditions hindered grain

maturation, particularly if sowing was delayed by the late arrival of spring or growing regions receive high amounts of summer rainfall (Kubok, 1988; Mikołajczyk, 1966). Two naturally-occurring dominant mutations were discovered independently during the 1960's in Australia and northern Europe and effectively removed the requirement for vernalisation, thereby improving the adaptation and yield stability of narrow-leafed lupin in short season environments. The most widely adopted of these is *Ku*, which arose as a spontaneous mutant in the cultivar 'Borre' and is capable of advancing flowering time by up to five weeks in Australia (Gladstones & Hill, 1969). It has been selected in almost all elite Australian cultivars released since the 1970's (Cowling, 1999; Stefanova & Buirchell, 2010) and has also been commonly used in European breeding programs (Boersma et al., 2007). The second mutation, *Julius* (*Jul*), was first discovered in the Russian bred cultivar, Krasnolistny, and became an important source of early phenology in Polish breeding during the 1980's (Kubok, 1988; Mikołajczyk, 1966). Due to striking similarity in photoperiodic as well as vernalisation responsiveness, *Ku* and *Jul* were thought to be controlled by the same gene (Rahman & Gladstones, 1972).

Global production of narrow-leafed lupin grain has stagnated in recent years with a leading cause being the limited genetic and adaptive diversity available within domesticated breeding pools of narrow-leafed lupin (Berger et al., 2012a). This lack of diversity largely stems from the species' recent domestication, which was based on a small number of founding individuals, and subsequent strong and persistent selection for key traits, such as phenology, all of which have resulted in severe genetic bottlenecks (Berger et al., 2012a; Cowling, 1999; Stefanova & Buirchell, 2010). Among the consequences for these bottlenecks is a reduction in genetic variation for flowering time within a domestic background. Furthermore, any variation that is present is hidden by the dominant overriding effect of the *Ku* and *Jul* loci.

A lack of phenological diversity is problematic for two main reasons. Firstly, despite the profound influence *Ku* and *Jul* have had in adapting narrow-leafed lupin to short-season production environments, many regions still lack options for cultivars adequately matched to their respective climates (Berger et al., 2008a; Berger et al., 2012b). The southwest of Western Australia and the eastern states of Australia are two such regions, where longer growing seasons coupled with reduced evapotranspiration could support later flowering cultivars and achieve higher yields (Chen et al., 2017). Secondly, the reliance on *Ku* and lack of genetic diversity due to domestication bottlenecks will reduce the ability of breeders to select stress tolerant cultivars adapted to climate change, which threatens future yield potential of lupins and crop harvestability (Nelson et al., 2010a). New genetic variation for phenology is required to increase the adoption of narrow-leafed lupin in longer-season environments and to maintain or extend feasible production zones in the face of climate change.

Recently, the genetic identify of *Ku* was revealed as a *FLOWERING LOCUS T* (*FT*) homologue, *LanFTc1* (Nelson et al., 2017). In *Arabidopsis*, *FT* has a well-defined role as a floral integrator gene, coordinating signals from the vernalisation, photoperiod and circadian clock pathways to promote flowering at an opportune time (Turck et al., 2008). A 1,423 bp deletion in the promoter region of *LanFTc1* was implicated as the casual mutation for the *Ku* allele, as its presence was perfectly predictive of vernalisation responsiveness in 216 wild and domesticated accessions and was associated with de-repressed expression of *LanFTc1* in the absence of vernalisation (Nelson et al., 2017). Given the demonstrated capacity for mutations in the non-coding sequence of this gene to modify its expression and plant phenology, we endeavoured to find other polymorphisms that may affect *cis*-regulation of *LanFTc1* and provide alternative sources of flowering time variation. Here, we report the discovery of a series of insertion/deletions (INDELs) in the 5' regulatory region of this gene which are associated with altered vernalisation responsiveness, flowering time and *LanFTc1* gene expression in the absence of vernalisation.

MATERIALS AND METHODS

Screening for polymorphisms within the genomic region of *LanFTc1* in diverse germplasm

Polymorphisms in the *LanFTc1* genomic region were explored in a panel of 48 narrow-leafed lupin accessions (Table S1) comprising: (1) the species reference genome cultivar, Tanijl (Hane et al., 2017); (2) 43 accessions, including 30 genetically diverse wild accessions representing the natural geographic range of the species throughout the Mediterranean Basin (Mousavi-Derazmahalleh et al., 2017) and 13 fully- or semi-domesticated accessions from Australia and Europe, all for which short-read sequencing data were generated to assemble the *LanFTc1* region; (3) Krasnolistny, the first cultivar described as carrying the *Julius* early flowering time locus (Mikołajczyk, 1966); (4) and three Polish cultivars (Kazan, Mirela and Sur) with Krasnolistny in their pedigree (Kubok, 1988).

The genomic sequence encompassing roughly 7 Kb upstream and 2 Kb downstream of the *LanFTc1* coding region was extracted from the Tanjil narrow-leafed lupin reference genome (Hane et al., 2017) plus the 43 accessions with short read sequencing data by aligning Illumina Paired End reads from each accession to the Tanjil reference using Bowtie2 v2.2.9 (--sensitive) (Langmead & Salzberg, 2012). Variants were called using samtools and bcftools (Li, 2011; Li et al., 2009), which were then filtered to remove artefactual sequence variant calls arising from misalignments close to large (> 1,000 bp) insertion/deletions (INDELs) and polymorphisms that were physically disrupted by others.

Both Tanjil and the P27255 wild-type *LanFTc1* sequence (GenBank ID KT862491) served as references to genotype the 1,423 bp insertion/deletion (INDEL) polymorphism previously identified by Nelson et al. (2017) in the 5' regulatory region and to survey for other alternative INDEL variations in this same region. After discovering additional INDELs in the 7 Kb sequence upstream of

LanFTc1, PCR primers were designed in the immediately adjacent sequences to screen for presence/absence of these INDELs for those accessions (Krasnolistny, Kazan, Mirela and Sur) for which no re-sequencing data were obtained. A summary of the INDELs the different PCR primers assay and the conditions for PCR amplification are provided in Table S2. To confirm the size and boundary of the newly identified INDELs, as determined from alignment of the short read sequences to the P27255 and Tanjil references, the nucleotide sequence of PCR products was determined by Sanger sequencing.

Additional INDELs and single nucleotide polymorphisms (SNPs) were also assessed in the coding and remaining non-coding sequences within the extracted *LanFTc1* genomic region in the 44 wild and domestic accessions using Tanjil as the reference genome. The P27255 wild-type *LanFTc1* sequence (GenBank ID KT862491), encompassing approximately 5 Kb upstream and 800 bp downstream of the *LanFTc1* coding region, was used as the reference to call variants within the 1,423 bp sequence in the 5' regulatory region that is deleted in the Tanjil reference genome.

Measuring degree-days to flowering and vernalisation responsiveness in diverse germplasm

The panel of 48 diverse accessions was phenotyped for time to flowering in two partially replicated trials (n = 1-3, as outlined in Table S1) for preliminary assessment of *LanFTc1* polymorphism genotype effect on vernalisation responsiveness and time to flowering. The first trial included 40 accessions representing the 0 bp, 1,208 bp and 1,423 bp INDEL variants. Data were gathered for all accessions except for P22603, which failed to germinate. In the second trial, six accessions carrying the 5,162 bp INDEL were compared with ten representatives of the three other INDEL variants.

All seeds were germinated in Jiffy-7® peat pellets within a controlled environment room (CER) at The University of Western Australia (Perth, Australia), which was maintained at 20°C constant temperature and with a 14 hour photoperiod. Two vernalisation treatments were provided: (1) a full vernalisation treatment in which two-day old seedlings were transferred to a 4 °C room (14 hour photoperiod) for 32 days before transferring to the CER for a further 140 days; and, (2) a mild, partial vernalisation treatment, whereby seven-day old seedlings were transferred to the 4 °C room for a total of eight days before transferring to the CER for a further 140 days. A mild vernalisation treatment was preferred to a fully non-vernalising treatment where accessions with a strong vernalisation requirement would not flower at all. Thus, the mild vernalisation treatment was designed to allow the degree of vernalisation responsiveness to be measureable in the most strongly vernalisation responsive accessions. All plants were transferred to the controlled environment room on the same day with approximately the same accumulated degree-days (~190 degree-days, with 0 °C as the baseline temperature) and were placed in a randomised block design.

Flowering was scored immediately after anthesis, which was indicated by an erect standard petal (i.e. open flower) or the changing colour of petals. To accommodate the few accessions which did not flower within the allocated time of the experiment, even with mild vernalisation treatment, flowering time was transformed to rate of flowering by taking the reciprocal of the degree-days to flowering. Three-way ANOVA was performed on the rate to flowering data using Genstat V.18, with vernalisation treatment and deletion category as main effects, and accessions nested within deletion category to subdivide variance among and within categories. Category effects were compared using orthogonal contrasts by least significant difference (LSD). ANOVA was performed separately for each phenotyping trial, and residual plots used to identify outliers and check that errors were randomly and independently distributed.

Assessing the relationship of INDEL variation in the 5' regulatory region and *LanFTc1* gene expression

Following the preliminary analysis of *LanFTc1* polymorphism effect on phenotype, a subset of accessions representing four variants of prominent INDELs in the 5' regulatory region of *LanFTc1* was selected from the original panel of 48 accessions to phenotype more precisely flowering time and vernalisation responsiveness, and to correlate these traits with *LanFTc1* gene expression. The representative subset included the following narrow-leafed lupins: (1) P27255, a wild Moroccan accession that is highly responsive to vernalisation; (2) 83A:476, a vernalisation-insensitive Australian breeding line; (3) P22660, a wild Israeli accession with mild sensitivity to vernalisation; (4) P29039, a vernalisation-insensitive Belarussian breeding line; and (5) Russian cultivar, Krasnolistny, also insensitive to vernalisation. P27255 and 83A:476 are the parents for a wild x domestic F_8 recombinant inbred line (RIL) mapping population (Boersma et al., 2005; Kroc et al., 2014; Nelson et al., 2010b; Nelson et al., 2006).

Seeds were scarified and imbibed in Milli-Q water for six hours before being immediately sown (non-vernalised treatment) or incubated in a darkened room at 4°C for 21 days in petri dishes (vernalised treatment). On the day of sowing the vernalised seeds, both treatments had accumulated approximately equal degree-days (calculated using base-line temperature of 0 °C). All plants were grown in a phytotron located at The University of Western Australia (Perth, Australia) with a diurnal temperature range of 18 ± 0.5 °C (day) to 14 ± 0.5 °C (night) and exposed to natural photoperiod (10-12 hours daylight during May to October 2017). Flowering time and degree-days to flowering was scored immediately after anthesis and the data analysed as outlined above.

The four uppermost fully emerged leaves were harvested for gene expression analyses from three biological replicates per treatment per accession at five growth stages: 4-leaf, 8-leaf, 12-leaf, 16-leaf and flowering. Samples were harvested between 12:00-14:00 hours and immediately snap-frozen in liquid nitrogen. RNA isolation, cDNA synthesis and quantitative reverse transcription PCR (qRT-PCR) were conducted according to the methods of Taylor et al. (2016) and Nelson et al. (2017). Briefly, the relative expression of *LanFTc1* was calculated as the average cycle threshold (C_T) for two primer pairs, which had previously been designed by Nelson et al. (2017) using transcript sequences from the draft Tanjil reference genome assembly (Kamphuis et al., 2015) to be specific to *LanFTc1* and to target unique portions of the coding sequence for gene-wide transcription assessment. The average *LanFTc1* C_T value was then normalised against *Ubiquitin C* (*UBC*), which had been validated as a robust reference gene under the same experimental conditions (Taylor et al., 2016), and relative expression was finally expressed as 40-ΔC_T. This method reports relative transcript abundance on a Log_2 scale, where a value of 40 represents the mean level of expression of *UBC* and the fold difference between treatments is calculated as $2^{\Delta\Delta CT}$ (where $\Delta\Delta C_T$ is equal to difference in average ΔC_T between vernalised and non-vernalised treatments) when primer efficiency is approximately 2.0 (Bari et al., 2006).

Nested ANOVA and polynomial linear regression of relative gene expression over degree-days to flowering were performed in Genstat V.18 as described above. Nested polynomial linear regression demonstrated non-significant differences between accessions within INDEL size categories, and therefore only the INDEL size category results are presented (Figs. 3a,b). The regressions were compared using orthogonal contrasts. The regression equations generated by Genstat were then used to plot smooth fitted curves, and we included the biological replicate data points for context (Figs. 3a,b).

Characterising the *LanFTc1* promoter region

Relative to P27255, representing the wild-type *LanFTc1* sequence, three large, distinct INDELs were identified in the 5' regulatory region. Two of the INDELs, which were associated with modified gene expression and phenology, were assessed relative to the wild-type for the presence/absence of transcription factor binding site motifs previously identified by Nelson *et al.* (2017). Putative binding site motifs were identified in that study using two open-access web-interface platforms, including: JASPAR 2014, which contains CORE Plantae matrix models (Mathelier et al., 2014); and PLACE, which contains *cis*-acting regulatory DNA elements in plants (Higo et al., 1999).

Assessment of linkage disequilibrium in the *LanFTc1* genomic region

To determine the likelihood of other polymorphisms within the *LanFTc1* genomic region being involved in modifying the response to vernalisation, we measured the association of each polymorphism with vernalisation responsiveness among the 44 accessions previously genotyped for polymorphisms in the *LanFTc1* genomic region. This was done by measuring pairwise linkage disequilibrium (r^2) of SNP and INDEL variants identified relative to the P27255 wild-type *LanFTc1* reference sequence (see above) with the vernalisation responsiveness phenotype, which was scored as a multi-allelic genotype (unresponsive, mildly responsive, and responsive). An r^2 value was also calculated for the four INDEL variants as a single multi-allelic polymorphism (0 bp, 1,208 bp, 1,423 bp and 5,162 bp). All pairwise r^2 values (--ld-window-r2 0) were calculated using PLINK v1.9 (Purcell et al., 2007). Default filtering settings in PLINK were used to remove markers with low quality or that were almost monomorphic from analysis. A linear adjusted association analysis was also conducted in PLINK v1.9 to determine the significance and strength of the association between the sequence variants and vernalisation responsiveness phenotype.

RESULTS

Screening for polymorphisms within the genomic region of *LanFTc1* in diverse germplasm

A 1,423 bp deletion between 4,248 and 2,826 bp upstream of the ATG start codon of *LanFTc1* was previously hypothesised as the causal sequence variant modifying vernalisation responsiveness in the breeding line, 83A:476 (*Ku*), compared to the wild-type *LanFTc1* sequence represented by P27255, and no polymorphisms were observed between accessions in the coding sequence (Nelson et al., 2017). To dissect this further, we surveyed the *LanFTc1* genomic region, from approximately 7 Kb upstream to 2 Kb downstream of the coding region in the Tanjil reference genome (Hane et al., 2017), for polymorphisms in 44 accessions of narrow-leafed lupin. This analysis revealed a total of 260 SNPs and 56 INDELs relative to Tanjil, excluding the 1,423 bp INDEL (Table 1; Table S3a). Importantly, no polymorphisms were found in the coding sequence. The genomic region features containing the most SNPs and INDELs (all less than 30 bp in length) were the large third intron and the 5' regulatory region, which contained approximately 50% and 32% of all polymorphisms. Variant calling using the P27255 wild-type *LanFTc1* sequence revealed a further 17 SNP and eight small INDELs (less than 20 bp in length) within the 1,423 bp sequence in the 5' regulatory region which is deleted in the Tanjil reference genome (Table S3b).

(Table 1 preferred position).

We then used P27255 and Tanjil as references to genotype the 1,423 bp INDEL and determine if there were any other major INDEL variations in the 5' regulatory region in 42 additional wild and domestic accessions. The wild-type *LanFTc1* sequence was present in a total of 29 wild accessions from the Mediterranean and two vernalisation responsive Australian cultivars (all known to carry the *ku* allele), while the 1,423 bp deletion was present in 10 Australian and Polish *Ku* cultivars.

Importantly, two new large INDEL variants were identified. The first of these was a 1,208 bp deletion observed between 3,970 and 2,763 bp upstream of the ATG start codon in the Israeli accession, P22660, relative to the P27255 wild-type *LanFTc1* sequence. The majority of this smaller deletion overlapped with that of the 1,423 bp variant. However, as determined from Sanger sequencing (GenBank ID MH166758), the first 277 bp at the 5' end of the 1,423 bp deletion was retained in P22660, whilst a further 62 bp immediately downstream of the 1,423 bp variant had been deleted (Fig. 1). The second INDEL variant was a prominent deletion of 5,162 bp in P29039 (a Belarussian breeding line) and Emir (a Polish cultivar) positioned between 6,209 and 1,048 bp upstream of the ATG start codon relative to the P27255 reference. Sanger sequencing supported a shared origin of the 5,162 bp deletion as both P29039 and Emir had identical deletion breakpoints (GenBank ID MH166759). This large deletion completely spanned the 1,423 bp and 1,208 bp deletions (Fig. 1).

(Figure 1 preferred position).

To determine if the narrow-leafed lupin cultivars carrying *Jul* contain the wild-type sequence or one of the three deletion genotypes, a PCR-marker (Table S2) approach was used. This confirmed that Krasnolistny, the original *Jul* cultivar, and three Polish cultivars known to descend from it (Kazan, Mirela and Sur) contain the 5,162 bp deletion as detected in P29039 and Emir (Fig S1).

Measuring degree-days to flowering and vernalisation responsiveness in diverse germplasm

To explore whether the four prominent INDEL variants in the 5' regulatory region of *LanFTc1* may affect vernalisation responsiveness and phenology in narrow-leafed lupin, we phenotyped rate to flowering (reciprocal of degree-days to flowering) in the full panel of 48 accessions under both mildly and fully vernalising conditions across two trials, with 39 accessions in the first trial (Fig. 2a) and 17 accessions in the second trial (Fig. 2b) (Table S1). In both trials, there were strong INDEL variant by

vernalisation treatment interactions ($P < 0.001$), in which vernalisation response was consistently proportional to flowering time (Fig. 2a,b). Thus, the strongest vernalisation response was observed in the late flowering wild-type accessions (0 bp INDEL), followed by the intermediate flowering 1,208 bp INDEL accession, and finally the early flowering 1,423 bp and 5,162 bp accessions, both of which had a small response to vernalisation (Fig. 2b). Accordingly, the large differences in rate to flowering observed between the 0 bp, 1,208 bp, and the 1,423 bp and 5,162 bp variants under mild vernalisation were greatly reduced under full vernalisation. Thus, under fully vernalising conditions, the strongly vernalisation responsive wild-types (0 bp deletions) flowered at the same rate as the intermediate 1,208 bp INDEL variant, and only marginally slower than the weakly vernalisation responsive accessions with the 1,423 bp and 5,162 bp deletions.

(Figure 2 preferred position).

Assessing the relationship of INDEL variation in the 5' regulatory region upon LanFTc1 *gene expression*

Based on the association between phenology and vernalisation responsiveness with INDEL variation in the 5' regulatory region of *LanFTc1*, we measured *LanFTc1* gene expression in five representative accessions, grown with and without vernalisation treatment. Four of the accessions included P27255, P22660, 83A:476 and P29039, representing the 0 bp (*ku*), 1,208 bp, 1,423 bp (*Ku*) and 5,162 bp deletions, respectively. Krasnolistny was also included to gain further evidence implicating the 5,162 bp deletion in the 5' regulatory region of *LanFTc1* as the causal mutation for the *Jul* locus.

Flowering time is earlier and vernalisation responsiveness is reduced or effectively lost in accessions with a deletion genotype

Orthogonal contrasts revealed that the flowering behaviour of the representative subset was consistent with the larger association study (Fig. 2). INDEL genotype by vernalisation treatment interactions were highly significant ($P < 0.001$), and ranked in the same order as previously. Thus, the wild-type 0 bp deletion was much more vernalisation responsive than the 1,208 bp deletion ($P_{diff} < 0.001$), which in turn was more responsive than the 1,423 bp and 5,162 bp INDELs ($P_{diff} < 0.001$) which had a similar low response to vernalisation ($P_{diff} = 0.883$) (Table 2). The two accessions with the 5,162 bp deletion genotype also had a similar low response to vernalisation ($P_{diff} = 0.450$), although P29039 was always slightly later flowering than Krasnolistny ($P_{diff} < 0.001$) (Table 2).

(Table 2 preferred position).

LanFTc1 *is expressed to varying degrees in accessions with a deletion genotype independently of vernalisation treatment*

All accessions had high gene expression under vernalising conditions during vegetative growth, starting from a common high basal level at the 4-leaf stage (approximately 277 degree-days, Fig. 3a). The deletion size categories (1,208 bp, 1,423 bp and 5,162 bp) had similar curvilinear increases in relative transcript abundance in the late vegetative stage (approximately 750 degree-days), while the wild-type (0 bp deletion) showed a flat slope, with no change in transcript abundance over time ($P = 0.851$).

Relative *LanFTc1* expression varied greatly among the accessions in the absence of vernalisation (Fig. 3b). Those with the 1,423 bp and 5,162 bp deletions behaved similarly to their respective vernalised treatments, with similar curvilinear increases in transcript abundance in the late vegetative stage. By contrast, relative expression levels in the accession (P22660) with the smaller 1,208 bp deletion rose rapidly from an intermediate basal level and reached similar levels as the larger deletion

categories by the mid-vegetative stage (Fig. 3b). Finally, the wild-type (0 bp deletion) had a slow linear increase in relative gene expression throughout the vegetative phase, starting from the lowest level of expression at the 4-leaf stage, and reached similar levels to the three deletion categories by the onset of flowering (Fig. 3b).

(Figure 3 preferred position).

Characterising the *LanFTc1* promoter region

The gene expression profiles of the three deletion variants indicate that sections of the 5' regulatory region are critical for regulating flowering time via the vernalisation pathway and *LanFTc1*. It appears that the 1,423 bp and 5,162 bp deletions are functionally equivalent, as they both result in insensitivity to vernalisation and similar *LanFTc1* expression profiles (Figs. 3a,b). Therefore, crucial regulatory elements responsible for full gene suppression in the absence of vernalising conditions should reside within the 1,423 bp INDEL sequence (Fig. 1). The functional activity of the 1,208 bp INDEL further refines this critical region. As the first 277 bp at the 5' end of the 1,423 bp INDEL are not also deleted within the 1,208 bp INDEL, it suggests that this 277 bp region is critical for establishing complete de-repression of *LanFTc1* (C_H; critical region for high early expression and vernalisation insensitivity). Additionally, as the 62 bp sequence at the 3' end of the 1,208 bp deletion is conserved in the wild-type, this indicates that the 1,146 bp sequence common to the 1,208 bp, 1,423 bp and 5,162 bp INDELs is responsible for establishing a moderate level of early gene activity without vernalisation (C_M; critical region for moderate early expression and moderate vernalisation responsiveness).

We next explored variation in the C_H and C_M critical regions that may explain differences in vernalisation responsiveness and *LanFTc1* expression. Both critical regions were screened for candidate transcription factor binding motifs that may have roles in the repression of *LanFTc1* within the P27255 representative wild-type sequence from the comprehensive list compiled by Nelson et al. (2017). A total of 31 individual motifs were found within the C_H region, including six in which the motif overhangs the C_H region and the adjacent 5' wild-type sequence and/or contains an alternative SNP allele in one or more vernalisation responsive accession(s) (Tables S4 and S5). Meanwhile as many as 168 individual motifs were identified within the C_M region, including seven for which one or more vernalisation responsive accession(s) have SNPs (Tables S4 and S5). Among all of the motifs identified upstream of the coding sequence within the 5' regulatory region and 5' UTR, the binding sites for three and five types of transcription factors were unique to the C_H and C_M regions, respectively, including several reported to have roles in determining flowering time (Table 3). Only a single type of transcription factor, named BRI1-EMS-SUPPRESSOR 1 (BES1), was common to both critical regions yet absent in the remainder of the adjacent 5' UTR and 5' regulatory region.

(Table 3 preferred position).

Assessment of linkage disequilibrium in the *LanFTc1* genomic region

To rule out the involvement of other polymorphisms in the *LanFTc1* gene region being involved in modifying vernalisation responsiveness, we measured the association of each polymorphism with the vernalisation responsiveness phenotype. A total of 48 INDELs and 206 SNPs were used for pairwise calculation of linkage disequilibrium (r^2) with the vernalisation responsiveness genotype. Low r^2 values of 0.35 or less were observed for the vast majority of polymorphisms (Fig. 4). One INDEL point at 752 bp in the P27255 wild-type reference sequence, which represents the 1,423 bp deletion in the 5' regulatory region, was the only variant with an r^2 value of 1.0 and perfect linkage

with the vernalisation responsiveness phenotype (Fig. 4). An r^2 value of 1.0 was also achieved when grouping the 1,423 bp, 5,162 bp and 1,208 bp deletions together as a single multi-allelic INDEL. The major association between the 1,423 bp INDEL and vernalisation responsiveness phenotype was highly significant (PLINK linear regression, coefficient t-statistic = 32.58, Bonferroni adjusted *P* = $6.60e^{-27}$). There was an additional small association with vernalisation responsiveness for a 1 bp INDEL at position 3,215 bp relative to the P27255 wild-type sequence (PLINK linear regression, Bonferroni adjusted *P* = 0.04) (Fig. 4).

(Figure 4 preferred position).

DISCUSSION

A series of *cis*-regulatory variants in a legume *FT* homologue

A 1,423 bp deletion in the 5' regulatory region of *LanFTc1*, an *FT* homologue, was recently hypothesised as the causal mutation behind the *Ku* locus that has been significant in establishing narrow-leafed lupin as a viable pulse crop in Australia and northern Europe (Nelson et al., 2017). Here, we have shown that another two independently occurring mutations, namely 1,208 bp and 5,162 bp deletions, overlap the same region of the promoter, creating a series of *cis*-regulatory variants that de-repress *LanFTc1* expression to varying extents. Additionally, we have shown that the 1,423 bp deletion, plus all three deletions when scored as a singular multi-allelic variant at the same position, are the only sequence variants in the *LanFTc1* wild-type sequence that are perfectly associated with vernalisation response and, as a consequence, early flowering time under non-vernalising conditions. The remaining 206 SNPs and 48 small INDELs within the 13 Kb wild-type genomic region are not associated with vernalisation responsiveness ($r^2 \leq 0.35$; Fig. 4), with the exception of one small INDEL (1 bp), which is weakly linked ($r^2 = 0.35$; Fig. 4). This finding

compliments those of Nelson *et al.* (2017), where presence/absence of the 1,423 bp deletion was perfectly predictive of vernalisation responsiveness in 216 accessions, and provides compelling evidence that the INDEL series are the likely causal mutations modifying vernalisation responsiveness, and thus phenology. However, we have not yet ruled out the possibility that other variants outside of this 13kb genomic region are also strongly or perfectly associated with vernalisation responsiveness. Lastly, we found strong evidence that the 5,162 bp deletion corresponds to *Jul*, supporting previous speculation that these two dominant early flowering time genes discovered independently in Australian and European breeding programs (Rahman & Gladstones, 1972) are in fact different alleles of the same gene, *LanFTc1*. We are currently developing bi-parental genetic populations to confirm this conclusion.

To the best of our knowledge, this is the first report of a naturally occurring series of mutations in the non-coding region of a floral integrator gene in any legume species. However, it adds to a growing list of literature similarly reporting series of *cis*-regulatory variants of vernalisation and photoperiodic pathway genes in *Arabidopsis* and cereal crops. The largest and most widely published allelic series identified to date involves *VRN-1*, a MADS-box transcription factor that is orthologous to *APETALA1* in *Arabidopsis* (Yan et al., 2003) and which is involved in maintaining down-regulation of floral repressors following vernalisation within members of the Poaceae family (Chen & Dubcovsky, 2012). Within the promoter and intronic regions, a staggering number of INDELs ranging in size from 20 bp to 6,850 bp, in addition to a single SNP, have been identified in the A, B, D and G genomes of various di-, tetra- and hexaploid wild and domestic wheats and their progenitors (Fu et al., 2005; Golovnina et al., 2010; Konopatskaia et al., 2016; Milec et al., 2012; Muterko et al., 2015; Santra et al., 2009; Shcherban et al., 2012; Takumi et al., 2011; Yan et al., 2004; Zhang et al., 2015), plus the H genome of barley (*Hordeum vulgare* L.) (Fu et al., 2005). Our study also adds to others in *Arabidopsis* (Liu et al., 2014; Schwartz et al., 2009), perennial ryegrass (*Lolium perenne*) (Skøt et al., 2011), and wheats and barley (Chen et al., 2013; Yan et al., 2006) showing that *FT* orthologues have

similarly been a common target for the evolution of natural flowering time variation in a range of plant families. Lastly, a series of 7 bp tandem repeat INDELs has also been identified in *Arabidopsis* to modify the *cis*-regulation of *CONSTANS* (*CO*), a gene which encodes a zinc-finger transcription factor responsive to the photoperiodic and circadian clock flowering pathways (Rosas et al., 2014).

Implications for breeding and expansion of the adaptive range of narrow-leafed lupin

Discovery of an INDEL series in the promoter region of *LanFTc1* has major practical implications in light of its demonstrated capacity to modify vernalisation responsiveness and flowering time in narrow-leafed lupin. The *Ku* (1,423 bp deletion) and *Jul* (5,162 bp deletion) alleles have already been widely incorporated into domestic breeding programs in Australia and Europe. However, the 1,208 bp deletion present in the Israeli wild accession, P22660, represents a new form of valuable variation that has the potential to delay flowering time by approximately 2.5 weeks in the absence of vernalisation. Such variation would be extremely beneficial in expanding the production range of narrow-leafed lupin, plus increasing crop adaptation and yield potential in current environments with longer seasons, such as the southern Western Australian and eastern Australian growing regions. The predominance of the *Ku* and *Jul* alleles in breeding programs means that, without prior knowledge of the 1,208 bp INDEL variant, it would not be easily identified in the early stages of segregation from hybrids with breeding lines containing the dominant early alleles, *Ku* or *Jul*. The PCR marker designed by Nelson et al. (2017) will serve as a useful resource to screen for the 1,208 bp INDEL in future breeding (Table S2a).

Similar to the 1,208 bp deletion identified in wild germplasm from Israel, it is interesting to note that *Jul* is thought to have originated from the same region of the Middle East (Mikołajczyk, 1966). Evaluation studies of previous germplasm collection trips (Clements & Cowling, 1994; Gladstones & Crosbie, 1979) and a recent genetic and adaptive diversity analysis (Mousavi-Derazmahalleh et al.,

2017) have identified the Eastern Mediterranean and Northern Africa (including but not limited to parts of Morocco, the Middle East, and Aegean islands) as key geographic regions associated with early phenology in the natural habitat of *L. angustifolius*. The lower elevation and latitude of these regions, in combination with reduced, variable rainfall and increased seasonal temperatures, results in shorter growing seasons with heightened abiotic stresses that drive phenological evolution (Berger et al., 2008b; Berger et al., 2017). Therefore, it is possible that valuable *cis*-regulatory variations of *LanFTc1* or other genes regulating time to flowering exist in wild populations of narrow-leafed lupin from these origins. We are currently exploring this possibility in our research activities.

Understanding the regulation of *FT* homologues and the mediation of vernalisation responsiveness in the legume family

In addition to the potential benefits to narrow-leafed lupin breeding, the discovery of two new INDEL variants within the 5' regulatory region has also enabled us to explore which part of the promoter is critical for retaining normal repression of *LanFTc1* in the absence of vernalisation and, which if manipulated, is capable of modifying phenology. This critical region has been further divided into two zones, one of which is critical for establishing a medium level of de-repressed gene expression (C_M), while the second enables high and completely de-repressed transcriptional state (C_H). However, at this stage, it remains unclear as to why these regions are critical and what role(s) the deleted sequences play in the wild-type *ku* allele.

As evidenced from variant series in other species and flowering time genes, *cis*-regulatory changes are mediated by polymorphisms in a number of different ways. The classical *FT* promoter in *Arabidopsis* contains four major blocks (Adrian et al., 2010; Liu et al., 2014), comprising: firstly, the A block, positioned roughly 400 bp upstream of the ATG start codon and which contains a number of transcriptional elements, such as those bound by CO (Tiwari et al., 2010); secondly, the B block,

located approximately 1.8 kb upstream of the coding sequence and containing two conserved binding sequences for basic helix-loop-helix (bHLH) proteins (Adrian et al., 2010); the distal C block, located roughly 5.2 kb upstream of the start codon and which contains binding elements for proteins involved in delivering CO to motifs within Block A (Cao et al., 2014); and lastly, an intermittent sequence roughly 3.7 kb upstream of the ATG transcription start site that includes a block known as ID, which is involved in establishing physical proximity of the A and C blocks for photoperiod responsiveness. It is the latter in which two INDELs influencing promoter efficiency have evolved in *Arabidopsis* (Liu et al., 2014). The first of these includes an approximately 1.1 kb insertion near block ID that causes an excessive physical distance between blocks A and C, preventing their normal interaction. Meanwhile, a smaller deletion of approximately 250 bp contrastingly provides sufficient proximity of blocks A and C, such that the ID block is redundant. Previous research from Książkiewicz et al. (2016) has indicated a lack of sequence conservation between the 1,423 bp deletion (*Ku*) and ID block, therefore suggesting that the INDEL series in *LanFTc1* is unlikely to operate in a similar manner to that of the *FT* series in *Arabidopsis*. The discovery of the 5,162 bp deletion in *Jul* accessions in this study also suggests that this is not the case, as a significantly large proportion of the promoter that may correspond to other blocks has been deleted. However, further research to characterise the sequences either side of the 5,162 bp INDEL may be required to firmly eliminate the improvement of promoter efficiency by modification to regulatory element proximity as one possible consequence of the INDEL series in *LanFTc1*.

Alternative ways in which the INDEL series may instead modify *cis*-regulation of *LanFTc1* is through changing the profile of transcription factor binding sites within the promoter region or their capacity to be bound. A straightforward explanation is the complete removal of transcription factor binding site motifs from within the three deletions. We refined a list of candidate transcription factor motifs from Nelson et al. (2017), revealing a total of 168 and 31 individual motifs present in the wild-type promoter sequence yet which are absent in the C_M and C_H regions, respectively (Table S4 and S5).

However, this is still an extremely large number of candidate transcription factor motifs, and it will be very difficult to further resolve which may or may not have functional roles in *LanFTc1* regulation, especially if no further variants are found overlapping this region.

Copy number of transcription factor binding sites also represents another possibility of *cis*-regulatory modification. As has been demonstrated in *Arabidopsis*, increasing from three to four tandem repeats of a 7 bp motif for CYCLING DOF FACTOR 1 (CDF1) in the promoter of *CO* increases the day-time repression of this gene and significantly delays flowering time (Rosas et al., 2014). We identified motifs for the binding site of a single transcription factor, named BES1, present once within the C_H and twice within the C_M critical regions (Tables 2, S4 and S5), yet nowhere else in the 5' UTR and 5' regulatory region of *LanFTc1*. In *Arabidopsis*, BES1 interacts with EARLY FLOWERING 6 (EF6) and RELATIVE OF EARLY FLOWERING 6 (REF6) proteins to respectively repress the photoperiodic pathway, through unknown means, and *FLOWERING LOCUS C* (*FLC*), a repressor of *FT* that is itself repressed by vernalisation (Noh et al., 2004; Yu et al., 2008). Therefore, although both *FLC* and *REF6* are apparently absent from the narrow-leafed lupin genome (Hane et al., 2017), there is precedence for BES1 involvement in the regulation of flowering time and it is conceivable that it could be involved in the direct regulation of *LanFTc1* through partnership with other flowering time genes. In such a scenario, deletion of two copies of the BES1 binding site motif via the 1,208 bp INDEL genotype would be sufficient to elevate *LanFTc1* expression to an intermediate level, and deletion of all three motifs via the 1,423 bp or 5,162 bp INDELs would fully de-repress expression. With genome editing tools, such as the CRISPR/Cas-9 system (Bortesi & Fischer, 2015), and more efficient transformation protocols in narrow-leafed lupin (Barker et al., 2016), it may be feasible to modify BES1 binding site motifs in the wild-type sequence to test the validity of this hypothesis in future. If BES1 has a role in regulating *LanFTc1*, BES1's known involvement within the photoperiodic pathway could explain why cultivars with the 1,423 bp deletion are also less responsive to inductive long days

than wild-types without the large deletion in the promoter region of *LanFTc1* (J. D. Berger, unpublished data).

Lastly, the location of INDELs relative to transcription factor binding site motifs can influence the affinity for transcription factor binding, as demonstrated in the case of the *FT* homologue allelic series in perennial ryegrass (Skøt et al., 2011). Relative to the wild-type sequence designated as the C haplotype, a deletion of five nucleotides positioned seven to 11 bp downstream of a conserved motif (the A haplotype) resulted in a two day delay in flowering time, whereas a six nucleotide deletion positioned one to six bp directly 3' of the conserved motif (the B haplotype) resulted in a seven day flowering time delay. Here, we have identified three classes of transcription factor binding site motifs which are disrupted by the 5' end of the C_H region and which are completely absent in the C_M region. However, motifs for these same transcription factors are also found on several occasions elsewhere within the *LanFTc1* genomic region, and small INDEL polymorphisms present in vernalisation responsive accessions without the 1,208 bp, 1,423 bp or 5,162 bp INDELs can also be found disrupting some of these motifs. Therefore, it seems unlikely that any of these motifs are functional or critical to *LanFTc1*; however, further research will be required to better characterise other motifs adjacent to the large deletions.

Despite our lack of knowledge as to how they impact gene expression, the discovery of the INDEL variant series in the 5' regulatory region of *LanFTc1* has provided us with a rare opportunity to better explore possible ways in which *FT* homologues are regulated, and vernalisation responsiveness is mediated, at the molecular level outside of the Brassicaceae and Poaceae. At present, our greatest understanding concerning vernalisation response within the legume family comes from *Medicago truncatula* (Weller & Ortega, 2015). In this model species, an *FTa1* homologue is upregulated following the return of warm conditions post-vernalisation, and loss-of-function mutations within the coding sequence render plants insensitive to vernalisation (Laurie et al., 2011).

Three induced mutant lines with dominant early, vernalisation-insensitive flowering have been shown to contain transposon insertions in the large third intron or the 3' regulatory region that result in up-regulated expression of *FTa1*, suggesting that these genomic regions are important sites for conferring transcriptional repression in the wild-type (Jaudal et al., 2013). However, similar to the present story in narrow-leafed lupin, it is unknown what elements within these regions are important for the vernalisation pathway and *FTa1* transcription. Thus far, it appears that methylation in the *FTa1* genomic region is unlikely to play a role, with no differences observed between the mutants with induced transposon insertions and wild-type plants (Jaudal et al., 2013). The discovery of further INDEL variants in *LanFTc1* could provide further clues of how signalling mediated through the vernalisation pathway is centred on *FT* homologues at the molecular level in the Fabaceae.

ACKNOWLEDGEMENTS

This research was generously funded by grants DAW00238 and UWA00147 from the Grains Research and Development Corporation (GRDC), Australia and funding from The University of Western Australia. Our thanks go to Michał Książkiewicz (Polish Academy of Sciences) for the kind provision of seed for a number of Russian and Polish cultivars used in this study.

REFERENCES

Adrian J., Farrona S., Reimer J., Albani M., Coupland G. & Turck F. (2010) *cis*-regualtory elements and chromatin state coordinately control temporal and spatial expression of *FLOWERING LOCUS T* in *Arabidopsis*. *The Plant Cell*, **22**, 1425-1440.

Bari R., Pant B.D., Stint M. & Scheible W.-R. (2006) PHO_2, $MicroRNA_{399}$, and PHR_1 define a phoshate-signaling pathway in plants. *Plant Physiology*, **141**, 988-999.

Barker, S.J., Si, P., Hodgson, L., Ferguson-Hunt, M., Khentry, Y., Krishnamurthy, P., …., Erskine, W. (2016) Regeneration selection improves transformation efficiency in narrow-leaf lupin. *Plant Cell Tissue and Organ Culture*, **126**, 219–228.

Behringer C. & Schwechheimer C. (2015) B-GATA transcription factors - insights into their structure, regulation, and role in plant development. *Frontiers in Plant Science*, **6**, 90.

Berger J., Buirchell B., Palta J., Luckett D., Ludwig C. & Shrestha D. (2008a) G x E analysis of narrow-leafed lupin historical trials indicates little specific adaptation among Australian cultivars. In *Lupins for Health and Wealth*, (eds J.A. Palta & J.D.Berger), pp. 317-320. Proceedings of the 12th International Lupin Conference, 14-18 September 2008, Fremantle, Western Australia. International Lupin Association, Canterbury, New Zealand.

Berger J., Ludwig C. & Buirchell B. (2008b) Ecogeography of the Old World lupins: characterising the habitat range. In *Lupins for Health and Wealth*, (eds J.A. Palta & J.D.Berger), pp. 355-361. Proceedings of the 12th International Lupin Conference, 14-18 September 2008, Fremantle, Western Australia. International Lupin Association, Canterbury, New Zealand.

Berger J., Shrestha D. & Ludwig C. (2017) Reproductive strategies in Mediterranean legumes: trade-offs between phenology, seed size and vigor within and between wild and domesticated *Lupinus* species collected along aridity gradients. *Frontiers in Plant Science*, **8**, 548.

Berger J.D., Buirchell B.J., Luckett D.J. & Nelson M.N. (2012a) Domestication bottlenecks limit genetic diversity and constrain adaptation in narrow-leafed lupin (*Lupinus angustifolius* L.). *Theoretical and Applied Genetics*, **124**, 637-652.

Berger J.D., Buirchell B.J., Luckett D.J., Palta J.A., Ludwig C. & Liu D.L. (2012b) How has narrow-leafed lupin changed in its 1st 40 years as an industrial, broad-acre crop? A G×E-based characterization of yield-related traits in Australian cultivars. *Field Crops Research*, **126**, 152-164.

Boersma J.G., Buirchell B.J., Sivasithamparam K. & Yang H. (2007) Development of a sequence-specific PCR marker linked to the *Ku* gene which removes the vernalization requirement in narrow-leafed lupin. *Plant Breeding*, **126**, 306-309.

Boersma J.G., Pallotta M., Li C., Buirchell B.J., 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.). *Cellular and Molecular Biology Letters*, **10**, 331-344.

Borner R., Kampmann G., Chandler J., Gleißner R., Wisman E., Apel K. & Melzer S. (2000) A MADS domain gene involved in the transition to flowering in *Arabidopsis*. *The Plant Journal*, **24**, 591-599.

Bortesi L. & Fischer R. (2015) The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, **33**, 41-52.

Bowman J.L., Smyth D.R. & Meyerowitz E.M. (1989) Genes directing flower development in *Arabidopsis*. *The Plant Cell*, **1**, 37-52.

Cao S., Kumimoto R., Gnesutta N., Calogero A., Mantovani R. & Holt III B. (2014) A distal *CCAAT*/NUCLEAR FACTOR Y complex promotes chromatin looping at the *FLOWERING LOCUS T* promoter and regulates the timing of flowering in *Arabidopsis*. *The Plant Cell*, **26**, 1009-1017.

Chen A. & Dubcovsky J. (2012) Wheat TILLING mutants show that the vernalization gene *VRN1* down-regulates the flowering repressor *VRN2* in leaves but is not essential for flowering. *PLOS Genetics*, **8**, e1003134.

Chen C., Fletcher A., Lawes R., Berger J. & Roberston M. (2017) Modelling phenological and agronomic adaptation options for narrow-leafed lupins in the southern grainbelt of Western Australia. *European Journal of Agronomy*, **89**, 140-147.

Chen F., Gao M., Zhang J., Zuo A., Shang X. & Cui D. (2013) Molecular characterization of vernalization and response genes in bread wheat from the Yellow and Huai Valley of China. *BMC Plant Biology*, **13**, 199.

Clements J. & Cowling W. (1994) Patterns of morphological diversity in relation to geographical origins of wild *Lupinus angustifolius* from the Aegean region. *Genetic Resources and Crop Evolution*, **41**, 109-122.

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

Cui R., Han J., Zhao S., Su K., Wu F., Du X., ..., Meng Z. (2009) Functional conservation and diversification of class E floral homeotic genes in rice (*Oryza sativa*). *The Plant Journal*, **61**, 767-781.

De Smet I., Lau S., Erismann J.S., Axiotis I., Kolb M., Kientz M., ..., Jürgens G. (2013) Transcriptional repression of *BODENLOS* by HD-ZIP transcription factor HB5 in *Arabidopsis thaliana*. *Journal of Experimental Botany*, **64**, 3009-3019.

FAO (2014) FAOSTAT. Food and Agriculture Organization of the United Nations, Rome, Italy.

Foyer C.H., Lam H.-M., Nguyen H.T., Siddique K.H.M., Varshney R.K., Colmer T.D., ..., Considine M.J. (2016) Neglecting legumes has compromised human health and sustainable food production. *Nature Plants*, **2**, 16112.

Fu D., Szűcs P., Yan L., Helguera M., Skinner J.S., von Zitzewitz J., ..., Dubcovsky J. (2005) Large deletions within the first intron in *VRN-1* are associated with spring growth habit in barley and wheat. *Molecular Genetics and Genomics*, **273**, 54-65.

Gladstones J.S. (1970) Lupins as crop plants. *Field Crop Abstracts*, **23**, 123-148.

Gladstones J.S. (1974) Lupins of the Mediterranean region and Africa, Western Australian Department of Agriculture, Technical Bulletin No. 26.

Gladstones J.S. (1977) The narrow-leafed lupin in Western Australia. Western Australia Department of Agriculture, Bulletin 3990.

Gladstones J.S. & Crosbie G.B. (1979) Lupin wild types introduced into Western Australia to 1973. Department of Agriculture, Western Australia, Technical Bulletin No. 43.

Gladstones J.S. & Hill G.D. (1969) Selection for economic characters in *Lupinus angustifolius* and *L. digitatus*. 2. Time of flowering. *Australian Journal of Experimental Agriculture and Animal Husbandry*, **9**, 213-220.

Golovnina K., Kondratenko E., Blinov A. & Goncharov N. (2010) Molecular characterization of vernalization loci *VRN1* in wild and cultivated wheats. *BMC Plant Biology*, **10**, 168.

Hane J.K., Ming Y., Kamphuis L.G., Nelson M.N., Garg G., Atkins C.A., ..., Singh K.B. (2017) A comprehensive draft genome sequence for lupin (*Lupinus angustifolius*), an emerging health food: insights into plant-microbe interactions and legume evolution. *Plant Biotechnology Journal*, **15**, 318-330.

Higo K., Ugawa Y., Iwamoto M. & Korenaga T. (1999) Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research*, **27**, 297-300.

Hill J.P. & Lord E.M. (1989) Floral development in *Arabidopsis thaliana*: a comparison of the wild type and the homeotic pistillata mutant. *Canadian Journal of Botany*, **67**, 2922-2936.

Jaudal M., Yeoh C.C., Zhang L., Stockum C., Mysore K.S., Ratet P. & Putterill J. (2013) Retroelement insertions at the Medicago *FTa1* locus in *spring* mutants eliminate vernalisation but not long-day requirements for early flowering. *The Plant Journal*, **76**, 580-591.

Johannesson H., Wang Y. & Engström P. (2001) DNA-binding and dimerization preferences of *Arabidopsis* homeodomain-leucine zipper transcription factors *in vitro*. *Plant Molecular Biology*, **45**, 63-73.

Kamphuis L.G., Hane J.K., Nelson M.N., Gao L., Atkins C.A. & Singh K.B. (2015) Transcriptome sequencing of different narrow-leafed lupin tissue types provides a comprehensive uni-gene assembly and extensive gene-based molecular markers. *Plant Biotechnology Journal*, **13**, 14-25.

Konopatskaia I., Vavilova V., Kondratenko E., Blinov A. & Goncharov N. (2016) *VRN1* genes variability in tetraploid wheat species with a spring growth habit. *BMC Plant Biology*, **16**, 244.

Kouris-Blazos A. & Belski R. (2016) Health benefits of legumes and pulses with a focus on Australian sweet lupins. *Asia Pacific Journal of Clinical Nutrition*, **25**, 1-17.

Kroc M., Koczyk G., Święcicki W., Kilian A. & Nelson M.N. (2014) New evidence of ancestral polyploidy in the Genistoid legume *Lupinus angustifolius* L. (narrow-leafed lupin). *Theoretical and Applied Genetics*, **127**, 1237-1249.

Książkiewicz M., Rychel S., Nelson M., Wyrwa K., Naganowska B. & Wolko B. (2016) Expansion of the phosphatidylethanolamine binding protein family in legumes: a case study of *Lupinus angustifolius* L. *FLOWERING LOCUS T* homologs, *LanFTc1* and *LanFTc2*. *BMC Genomics*, **17**, 820.

Kubok I. (1988) *The History of Lupine Breeding in Poland*. Plant Breeding and Acclimatization Institute, Radzików, Poland.

Lambers H., Clements J.C. & Nelson M.N. (2013) How a phosphorus-acquisition strategy based on carboxylate exudation powers the success and agronomic potential of lupines (*Lupinus*, Fabaceae). *American Journal of Botany*, **100**, 263-288.

Laurie R.E., Diwadkar P., Jaudal M., Zhang L., Hecht V., Wen J., ..., Macknight R.C. (2011) The Medicago *FLOWERING LOCUS T* homolog, *MtFTa1*, is a key regulator of flowering time. *Plant Physiology*, **156**, 2207-2224.

Lima-Cabello E., Alche V., Foley R., Andrikopoulos S., Morahan G., Singh K., ..., Jimenez-Lopez J. (2017) Narrow-leafed lupin (*Lupinus angustifolius* L.) β-conglutin proteins modulate the insulin signaling pathway as potential type 2 diabetes treatment and inflammatory-related disease amelioration. *Molecular Nutrition & Food Research*, **61**, 1600819.

Liu L., Adrian J., Pankin A., Hu J., Dong X., von Korff M. & Turck F. (2014) Induced and natural variation of promoter length modulates the photoperiodic response of *FLOWERING LOCUS T*. *Nature Communications*, **5**.

Luo X., Sun X., Liu B., Zhu D., Bai X., Cai H., ..., Zhu Y. (2013) Ectopic expression of a WRKY homolog from *Glycine soja* alters flowering time in *Arabidopsis*. *PLOS ONE*, **8**, e73295.

Mandaokar A., Thines B., Shin B., Lange B.M., Choi G., Koo Y.J., ..., Browse J. (2006) Transcriptional regulators of stamen development in Arabidopsis identified by transcriptional profiling. *The Plant Journal*, **46**, 984-1008.

Mandel M.A. & Yanofsky M.F. (1998) The *Arabidopsis AGL9* MADS box gene is expressed in young flower primordia. *Sexual Plant Reproduction*, **11**, 22-28.

Mathelier A., Zhao X., Zhang A.W., Parcy F., Worsley-Hunt R., Arenillas D.J., ..., Wasserman W.W. (2014) JASPAR 2014: an extensively expanded and update open-access database of transcription factor binding profiles. *Nucleic Acids Research*, **42**, D1423-D1147.

Mikołajczyk J. (1966) Genetic studies in *Lupinus angustifolius.* Part. III. Inheritance of the alkaloid content, seed hardness and length of the growing season in blue lupin. *Genetica Polonica*, **7**, 181-196.

Milec Z., Tomková L., Sumíková T. & Pánková K. (2012) A new multiplex PCR test for the determination of *Vrn-B1* alleles in bread wheat (*Triticum aestivum* L.). *Molecular Breeding*, **30**, 317-323.

Mola J., Grotewold E. & Koesa R. (1998) How genes paint flowers and seeds. *Trends in Plant Science*, **3**, 212-217.

Mousavi-Derazmahalleh M., Bayer P., Nevado B., Hurgobin B., Filatov D., Kilian A., ..., Nelson M. (2017) Exploring the genetic and adaptive divserity of a pan-Mediterranean crop wild relative: narrow-leafed lupin. *Theoretical and Applied Genetics*, **In press**.

Muterko A., Balashova I., Cockram J., Kalendar R. & Sivolap Y. (2015) The new wheat vernalization repsonse allele *Vrn-D1s* is caused by DNA transposon insertion in the first intron. *Plant Molecular Biology Reporter*, **33**, 294-303.

Nelson M.N., Berger J.D. & Erskine W. (2010a) Flowering time control in annual legumes: prospects in a changing global climate. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, **5**, 017.

Nelson M.N., Książkiewicz M., Rychel S., Besharat N., Taylor C.M., Wyrwa K., ..., Wolko B. (2017) The loss of vernalization requirement essential to domestication in narrow-leafed lupin is associated with a deletion in the promoter and de-repressed expression of an *FT* homologue. *New Phytologist*, **213**, 220-232.

Nelson M.N., Moolhuijzen P.M., Boersma J.G., Chudy M., Lesniewska K., Bellgard M., ..., Ellwood S.R. (2010b) Aligning a new reference genetic map of *Lupinus angustifolius* with the genome sequence of the model legume, *Lotus japonicus*. *DNA Research*, **17**, 73-83.

Nelson M.N., Phan H.T.T., Ellwood S.R., Moolhuijzen P.M., Hane J., Williams A., ..., Cowling W.A. (2006) The first gene-based map of *Lupinus angustifolius* L.-location of domestication genes and conserved synteny with *Medicago truncatula*. *Theoretical and Applied Genetics*, **113**, 225-238.

Noh B., Lee S.-H., Kim H.-J., Yi G., Shin E.-A., Lee M., ..., Noh Y.-S. (2004) Divergent roles of a pair of homologous jumonji/zinc-finger-class transcription factor proteins in the regulation of Arabidopsis flowering time. *The Plant Cell*, **16**, 2601-2613.

Ooka H., Satoh K., Doi K., Nagata T., Otomo Y., Murakami K., ..., Kikuchi S. (2003) Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Research*, **10**, 239-247.

Paz-Ares J., Ghosal D., Wienand U., Peterson P.A. & Saedler H. (1987) The regulatory *c1* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators. *The EMBO Journal*, **6**, 3553-3558.

Pelaz S., Ditta G.S., Baumann E., Wisman E. & Yanofsky M.F. (2000) B and C floral organ identity functions require *SEPALLATA* MADS-box genes. *Nature*, **405**, 200-203.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., ..., Sham P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, **81**, 559-575.

Rahman M.S. & Gladstones J.S. (1972) Control of lupin initiation by vernalization, photoperiod and temperature under controlled envrionment. *Australian Journal of Experimental Agriculture and Animal Husbandry*, **12**, 638-645.

Reyes J.C., Muro-Pastor M.I. & Florencio F.J. (2004) The GATA family of transcription factors in Arabdiopsis and rice. *Plant Physiology*, **134**, 1718-1732.

Rosas U., Mei Y., Xie Q., Banta J., Zhou R., Seufferheld G., ..., Purugganan M. (2014) Variation in *Arabidopsis* flowering time associated with *cis*-regulatory variation in *CONSTANS*. *Nature Communications*, **5**, 3651.

Sainz M.B., Grotewold E. & Chandler V.L. (1997) Evidence for direct activation of an anthocyanin promoter by the maize C1 protein and comparison of DNA binding by related Myb domain proteins. *The Plant Cell*, **9**, 611-625.

Santra D., Santra M., Allan R., Campbell K. & Kidwell K. (2009) Genetic and molecular characterization of vernalization genes *Vrn-A1*, *Vrn-B1*, and *Vrn-D1* in spring wheat germplasm from the Pacific Northwest region of the U.S.A. *Plant Breeding*, **28**, 576-584.

Schwartz C., Balasubramanian S., Warthmann N., Michael T., Lempe J., Sureshkumar S., ..., Weigel D. (2009) *Cis*-regulatory changes at *FLOWERING LOCUS T* mediate natural variation in flowering responses in *Arabidopsis thaliana*. *Genetics*, **183**, 723-732.

Seymour M., Kirkegaard J., Peoples M., White P. & French R. (2012) Break-crop benefits to wheat in Western Australia - insights from over three decades of research. *Crop and Pasture Science*, **63**, 1-16.

Shcherban A., Efremova T. & Salina E. (2012) Identification of a new *Vrn-B1* allele using two near-isogenic wheat lines with difference in heading time. *Molecular Breeding*, **29**, 675-685.

Skøt L., Sanderson R., Thomas A., Skøt K., Thorogood D., Latypova G., ..., Armstead I. (2011) Allelic variation in the perennial ryegrass *FLOWERING LOCuS T* gene is associated with changes in flowering time across a range of populations. *Plant Physiology*, **155**, 1013-1022.

Stefanova K.T. & Buirchell B. (2010) Multiplicative mixed models for genetic gain assessment in lupin breeding. *Crop Science*, **50**, 880-891.

Takumi S., Koyamam K., Fujiwara K. & Kobayashi F. (2011) Identification of a large deletion in the first intron of the *VRN-D1* locus, associated with loss of vernalization requirement in wild wheat progenitor *Aegilops tauschii* Coss. *Genes and Genetic Systems*, **86**, 183-195.

Tamaki S., Matsuo S., Wong H.L., Yokoi S. & Shimamoto K. (2007) Hd3a protein is a mobile flowering signal in rice. *Science*, **316**, 1033-1036.

Taylor C., Jost R., Erskine W. & Nelson M. (2016) Identifying stable reference genes for qRT-PCR normalisation in gene expression studies of narrow-leafed lupin (*Lupinus angustifolius* L.). *PLoS ONE*, **11**, e0148300.

Tiwari S., Shen Y., Chang H.-C., Hou Y., Harris A., Fong Ma S., ..., Ratcliffe O. (2010) The flowering time regulator CONSTANS is recruited to the *FLOWERING LOCUS T* promoter via a unique *cis*-element. *New Phytologist*, **187**, 57-66.

Turck F., Fornara F. & Coupland G. (2008) Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. *Annual Review of Plant Biology*, **59**, 573-594.

Wang L., Yin H., Qian Q., Yang J., Huang C., Hu X. & Luo D. (2009) *NECK LEAF 1*, a GATA type transcription factor, modulates organogenesis by regulating the expression of multiple regulatory genes during reproductive development in rice. *Cell Research*, **19**, 598-611.

Wang Y., Henriksson E., Söderman E., Nordin Henriksson K., Sundberg E. & Engström P. (2003) The Arabidopsis homeobox gene, *ATHB16*, regulates leaf development and the sensitivity to photoperiod in Arabdiopsis. *Developmental Biology*, **264**, 228-239.

Weller J.L. & Ortega R. (2015) Genetic control of flowering time in legumes. *Frontiers in Plant Science*, **6**, 207.

Yan L., Fu D., Li C., Blechl A., Tranquilli G., Bonafede M., ..., Dubcovsky J. (2006) The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proceedings of the National Academy of Sciences, USA*, **103**, 19581-19586.

Yan L., Helguera M., Kato K., Fukuyama S., Sherman J. & Dubcovsky J. (2004) Allelic variation at the *VRN-1* promoter region in polyploid what. *Theoretical and Applied Genetics*, **109**, 1677-1686.

Yan L., Loukoianov A., Tranquilli G., Helguera M., Fahima T. & Dubcovsky J. (2003) Positional cloning of the wehat vernalization gene *VRN1*. *Proceedings of the National Academy of Sciences*, *USA*, **100**, 6263-6268.

Yoo S.K., Chung K.S., Kim J., Lee J.H., Hong S.M., Yoo S.J., ..., Ahn J.H. (2005) *CONSTANS* activates *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* through *FLOWERING LOCUS T* to promote flowering in Arabidopsis. *Plant Physiology*, **139**, 770-778.

Yu X., Li L., Li L., Guo M., Chory J. & Yin Y. (2008) Modulation of brassinosteroid-regulated gene expression by jumonji domain-containing proteins ELF6 and REF6 in *Arabidopsis*. *Proceedings of the National Academy of Sciences, USA*, **105**, 7618-7623.

Zhang Z., Gao M.W., S., Chen F. & Cui D. (2015) Allelic variation at the vernalization and photoperiod sensitivity loci in Chinese winter wheat cultivars (*Triticum aestivum* L.). *Frontiers in Plant Science*, **6**, 470.

Zhao Y., Medrano L., Ohashi K., Fletcher J.C., Yu H., Sakai H. & Meyerowitz E.M. (2004) HANABA TARANU is a GATA transcription factor that regulates shoot apical meristem and flower development in Arabdiopsis. *The Plant Cell*, **16**, 2586-2600.

TABLES

Table 1: A summary of SNP and INDEL polymorphisms (excluding the large INDEL within the 5' regulatory region) observed in the genomic region from approximately 7 Kb upstream to 2 Kb downstream of *LanFTc1* in 44 accessions of narrow-leafed lupin relative to the Tanjil narrow-leafed lupin reference genome. Genomic region features include regulatory regions adjacent to the coding sequence (CDS), and the untranslated regions (UTRs), exons (coding sequences) and introns (non-coding intra-genic sequences) of *LanFTc1*.

***LanFTc1* genomic region feature**	**Coordinates on Pseudochromosome NLL-10**	**Coordinates on Scaffold_276_44**	**Number of SNP polymorphisms**	**Number of INDEL polymorphisms**
5' regulatory region	8,016,843 – 8,023,566	3,823 – 10,546	91	11
5' UTR	8,023,567 – 8,023,842	10,547 – 10,822	1	1
Exon 1 (CDS)	8,023,843 – 8,024,040	10,823 – 11,020	0	0
Intron 1	8,024,041 – 8,024,162	11,021 – 11,142	1	0
Exon 2 (CDS)	8,024,163 – 8,024,225	11,143 – 11,205	0	0
Intron 2	8,024,226 – 8,024,381	11,206 – 11,361	2	3
Exon 3 (CDS)	8,024,382 – 8,024,420	11,362 – 11,400	0	0
Intron 3	8,024,421 – 8,030,939	11,401 – 17,919	127	29
Exon 4 (CDS)	8,030,940 – 8,031,155	17,920 – 18,135	0	0
3' UTR	8,031,156 – 8,031,424	18,136 – 18,404	3	2
3' regulatory region	8,031,425 – 8,033,155	18,405 – 20,135	35	10
Total			260	56

Table 2: Average days and degree-days to flowering in vernalised and non-vernalised treatments for narrow-leafed lupins representing the 5' regulatory region wild-type sequence (*ku*) for *LanFTc1* and three major deletion variants of 1,208 bp, 1,423 bp (*Ku*), and 5,162 bp (*Jul*). Differences between accession means within and across treatments greater than 1.9 days and 33.64 degree-days are significant (least significant difference $P < 0.05$).

Accession	**Deletion genotype**	**Days to flowering**		**Degree-days to flowering**	
		Vernalised	**Non-vernalised**	**Vernalised**	**Non-vernalised**
P27255	0 bp (*ku*)	53.7	134.3	1023.2	2434.8
P22660	1,208	49.0	67.0	941.5	1256.5
83A:476	1,423 bp (*Ku*)	51.0	49.0	976.5	941.5
P29039	5,162 (*Jul*)	58.3	55.7	1104.8	1058.2
Krasnolistny	5,162 (*Jul*)	52.3	50.7	999.8	970.7

Table 3: A list of candidate transcription factor (TF) binding site motifs unique to sequences within the 5' regulatory regions critical for establishing moderate (C_M) or high (C_H) levels of de-repressed *LanFTc1* expression during early vegetative growth.

Motifs present in the wild-type sequence	Motifs present in the C_H sequence	Motifs present in the C_M sequence	Role of TF in flowering within other angiosperms	References
AGL9	1[a]	0	AGAMOUS-LIKE 9 (AGL9), also known as SEPALLATA3 (SEP3), is a MADS-box TF that is involved in establishing identity of petals, stamens and carpels in *Arabidopsis*. In rice (*Oryza sativa*), knock out of two *AGL9/SEP3* homologues, *OsMADS7* and *OsMADS8*, also results in delayed flowering.	Mandel and Yanofsky (1998) Pelaz et al. (2000) Cui et al. (2009)
ATHB5	1	0	ARABIDOPSIS THALIANA HOMEOBOX PROTEIN 5 (ATHB5) is a homeodomain leucine zipper (HDZip) TF that forms a heterodimer with its family member, ARABIDOPSIS THALIANA HOMEOBOX PROTEIN 16 (ATHB16), which regulates photoperiodic responsiveness in *Arabidopsis.*	Johannesson et al. (2001) Wang et al. (2003) De Smet et al. (2013)
PI	1	0	PISTILLATA (PI) is a MADS-box TF that is involved in establishing identify of petals and stamens in *Arabidopsis.*	Hill and Lord (1989) Bowman et al. (1989)
AT3G20750	0	1	AT3G20750 (also known as GATA29) is a member of the GATA protein family and contains a HAN domain, which has roles in regulating cell differentiation and speciation, including for floral organs, in *Arabidopsis*. The rice (*Oryza sativa*) *AT3G20750* homologue, *NECK LEAF 1* (*NL1*), has a similar role in floral organ identity, and its overexpression is thought to affect regulation of *Hd3a*, a rice *FT* homologue, and delay flowering.	Reyes et al. (2004) Zhao et al. (2004) Behringer and Schwechheimer (2015) Wang et al. (2009) Tamaki et al. (2007)
C1	0	1	C1 is a MYB TF involved in anthocyanin biosynthesis, thus flower colouration, in maize (*Zea mays*).	Paz-Ares et al. (1987) Sainz et al. (1997) Mola et al. (1998)
MADSA	0	1	MADSA, also known as AGAMOUS-LIKE 20 (AGL20) and SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1) in *Arabidopsis*, is a MADS-box TF activated in the shoot apical meristems during the transition from vegetative to floral development, and which integrates signals from the giberellin (GA) pathway, and also the photoperiod pathway through CONSTANS (CO) via *FT.*	Borner et al. (2000) Yoo et al. (2005)
NAC6	0	1	NAC6 (also known as NAC2 or AtNAC2) is a member of the NAC TF family, which has roles in regulating morphogenesis and stress responses, and is itself involved in the regulation of stamen development in *Arabidopsis*.	Ooka et al. (2003) Mandaokar et al. (2006)
WRKY	0	1	WRKY20 is a member of the WRKY TF family that has roles in plant stress responses and development. Wild soybean (*Glycine soja*) homologue *WRKY20* is abundantly expressed in flowers and floral meristems, and is thought to be involved in positive regulation of the autonomous pathway. Its overexpression in *Arabidopsis* results in early flowering and is associated with upregulation of floral integrator genes, including *FT* and *SOC1*.	Luo et al. (2013)
BES1	1	2	BES1 is a member of the BES1/BZR1 TF family and interacts with EARLY FLOWERING 6 (EF6) and RELATIVE OF EARLY FLOWEIRNG 6 (REF6) in *Arabidopsis* to repress the photoperiodic pathway and *FLOWERING LOCUS C* (*FLC*), a repressor of *FT*.	Noh et al. (2004) Yu et al. (2008)

[a] Note that six of seven nucleotides forming this motif are located within the C_H region and that one or more vernalisation responsive accessions contain a SNP within the motif.

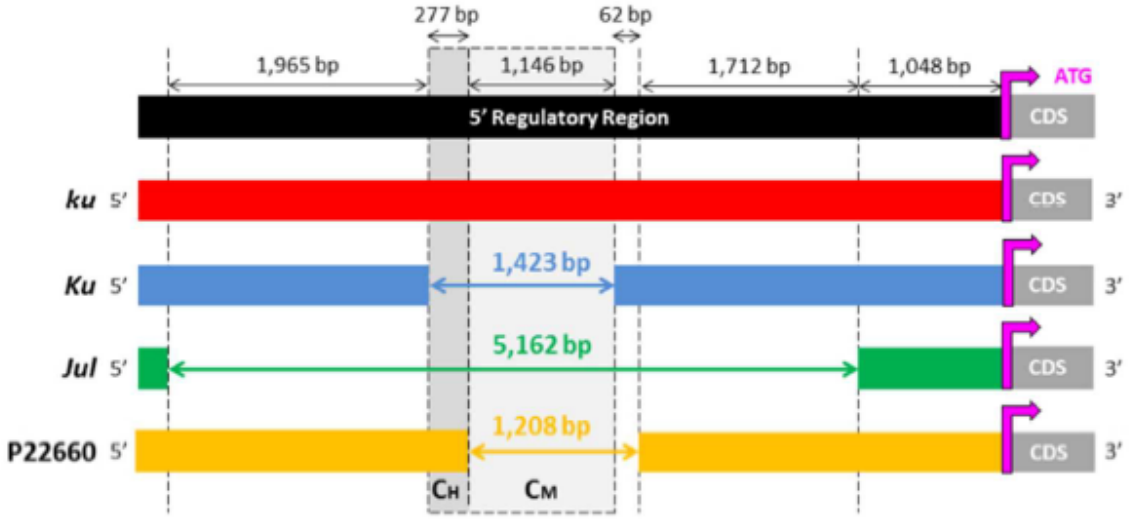

Figure 1. A schematic illustrating the positions of INDEL genotypes in the 5' regulatory region of *LanFTc1* relative to the start codon (ATG) of the coding sequence (CDS). The wild-type *LanFTc1* sequence (*ku*) was obtained from P27255, a wild Moroccan accession, and the 1,423 bp deletion (*Ku*) from the Tanjil reference genome. A 5,162 bp deletion (*Julius* or *Jul*) was found in several European breeding lines and cultivars, including Krasnolistny (Russian cultivar), P29039 (Belarussian breeding line) and Emir (Polish cultivar). A 1,208 bp deletion was identified in P22660, a wild accession from Israel. Critical regions of the regulatory region, if deleted, enable high (C_H, shaded dark grey) and moderately high (C_M, shaded light grey) levels of LanFTc1 expression, respectively, relative to the wild-type sequence.

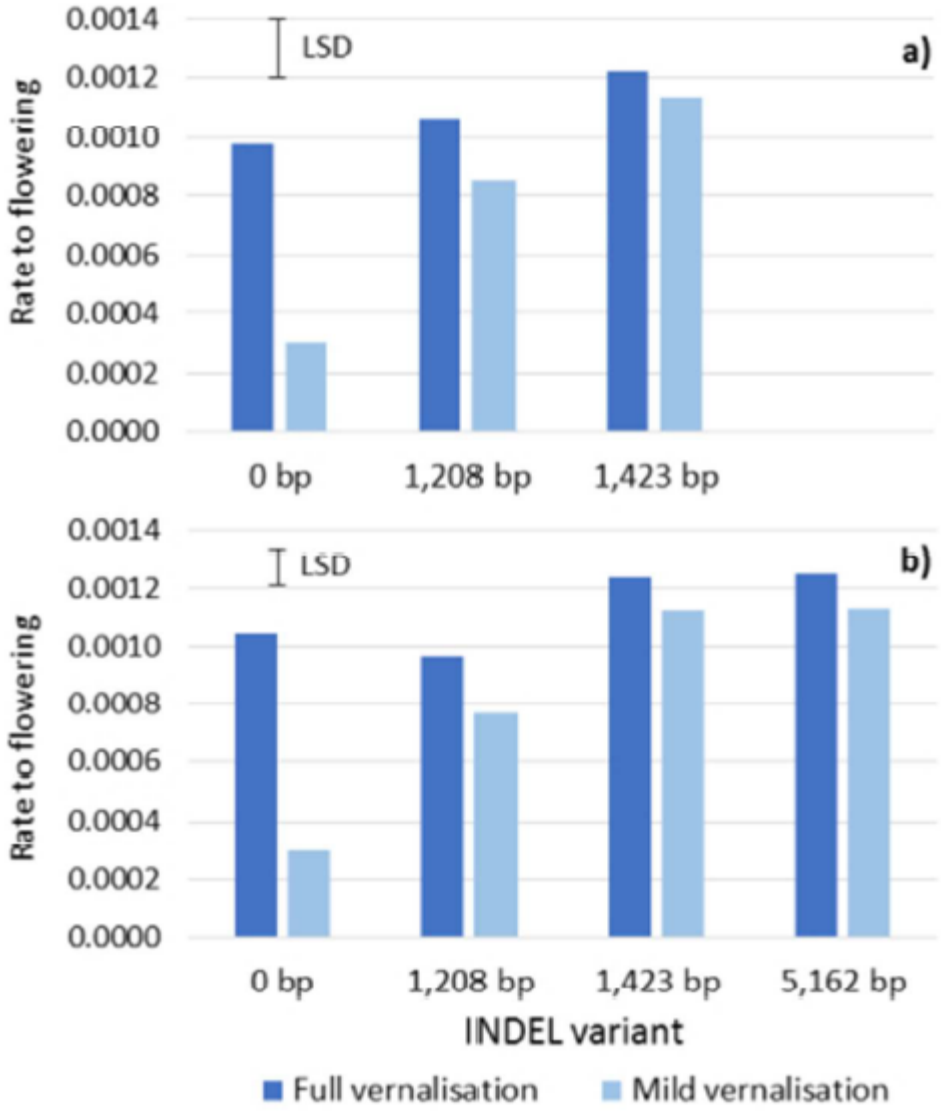

Figure 2. Average rate to flowering (reciprocal of degree-days to flowering) in mildly and fully vernalising conditions for narrow-leafed lupins possessing various insertion/deletions (INDELs) in the 5' regulatory region of *LanFTc1* in two phenotyping trials. (a) Trial 1 included narrow-leafed lupins carrying the 0 bp deletion (n = 29), 1,208 bp deletion (n = 1), and 1,423 bp deletion (n = 9). (b) Trial 2 included narrow-leafed lupins with the 0 bp deletion (n = 2), 1,208 bp deletion (n = 1), 1,423 bp deletion (n = 8), and 5,162 bp deletion (n = 6). The least significant difference (LSD) value is provided to compare responses within and between vernalisation treatments in each phenotyping trial (a: LSD = 0.00020; b: LSD = 0.00011).

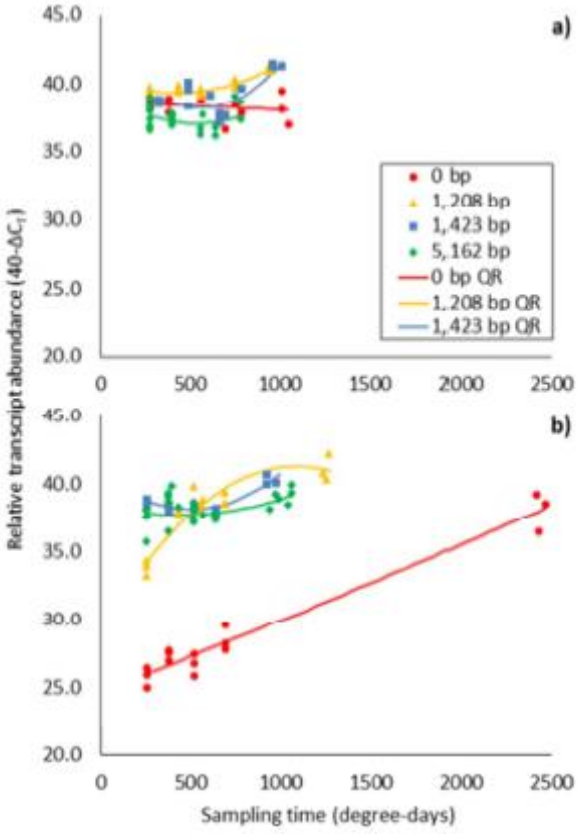

Figure 3. Relative expression of *LanFTc1* at various degree-days from 4-leaf stage (approximately 277 degree-days) to first flowering in deletion categories of narrow-leafed lupin (red circle, 0 bp deletion wild-type; yellow triangle, 1,208 bp deletion; blue square, 1,423 bp deletion; green diamond, 5,162 bp deletion) with (a) and without (b) vernalisation. The quadratic regression model (QR) captured 94.7% of variance, and indicated significant intercept, linear and quadratic slope differences ($P < 0.001$) between category/vernalisation treatment combinations. The final sampling time is at first flower, and therefore varies widely between treatment combinations.

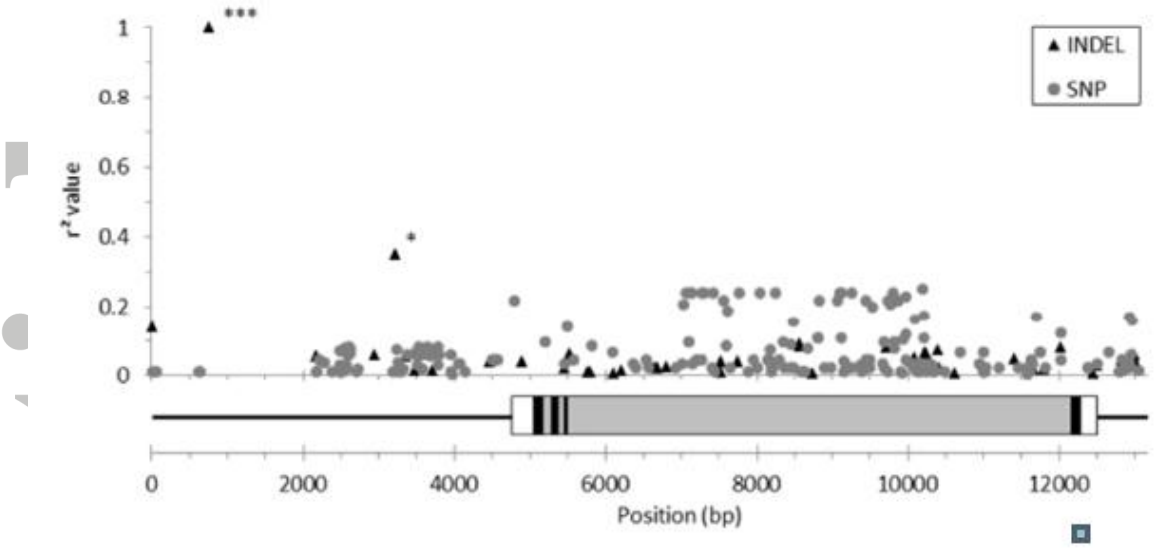

Figure 4. Linkage (represented as r^2) of insertion/deletions (INDELs; black triangles) and single nucleotide polymorphisms (SNPs; grey circles) identified in the narrow-leafed lupin *LanFTc1* wild-type genomic sequence, represented by Moroccan accession P27255, with vernalisation responsiveness. The positions of polymorphisms are indicated relative to the base pair (bp) position along the wild-type *LanFTc1* genomic sequence (GenBank ID KT862491) and a schematic of the *LanFTc1* genomic features, including: regulatory regions (RRs; solid black line), untranslated regions (UTRs; solid white bar with black border), exons (solid black bar with black border), and introns (solid grey bar with black border). An r^2 value of 1.0 represents perfect linkage with vernalisation response phenotype. Asterisks denote significant associations between polymorphisms and the vernalisation responsiveness phenotype (Bonferroni adjusted *P* values: * $0.01 < P < 0.05$; ** $0.001 < P < 0.01$; *** $P < 0.001$).