ORIGINAL ARTICLE

Genome-wide association study reveals white lupin candidate gene involved in anthracnose resistance

Joris A. Alkemade[1] · Nelson Nazzicari[2] · Monika M. Messmer[1] · Paolo Annicchiarico[2] · Barbara Ferrari[2] · Ralf T. Voegele[3] · Maria R. Finckh[4] · Christine Arncken[1] · Pierre Hohmann[1]

Abstract

Key message GWAS identifies candidate gene controlling resistance to anthracnose disease in white lupin.

Abstract White lupin (*Lupinus albus* L.) is a promising grain legume to meet the growing demand for plant-based protein. Its cultivation, however, is severely threatened by anthracnose disease caused by the fungal pathogen *Colletotrichum lupini*. To dissect the genetic architecture for anthracnose resistance, genotyping by sequencing was performed on white lupin accessions collected from the center of domestication and traditional cultivation regions. GBS resulted in 4611 high-quality single-nucleotide polymorphisms (SNPs) for 181 accessions, which were combined with resistance data observed under controlled conditions to perform a genome-wide association study (GWAS). Obtained disease phenotypes were shown to highly correlate with overall three-year disease assessments under Swiss field conditions ($r > 0.8$). GWAS results identified two significant SNPs associated with anthracnose resistance on gene *Lalb_Chr05_g0216161* encoding a RING zinc-finger E3 ubiquitin ligase which is potentially involved in plant immunity. Population analysis showed a remarkably fast linkage disequilibrium decay, weak population structure and grouping of commercial varieties with landraces, corresponding to the slow domestication history and scarcity of modern breeding efforts in white lupin. Together with 15 highly resistant accessions identified in the resistance assay, our findings show promise for further crop improvement. This study provides the basis for marker-assisted selection, genomic prediction and studies aimed at understanding anthracnose resistance mechanisms in white lupin and contributes to improving breeding programs worldwide.

Introduction

White lupin (*Lupinus albus* L., $2n = 50$) is a grain legume with a high-protein content and various health benefits that shows great potential to complement soybean and meet the growing demand for plant-based protein (Abraham et al. 2019; Annicchiarico 2008; Lucas et al. 2015). White lupin is

✉ Monika M. Messmer
monika.messmer@fibl.org

[1] Department of Crop Sciences, Research Institute of Organic Agriculture (FiBL), Frick, Switzerland

[2] Research Centre for Animal Production and Aquaculture, CREA, Lodi, Italy

[3] Institute of Phytomedicine, University of Hohenheim, Stuttgart, Germany

[4] Department of Ecological Plant Protection, University of Kassel, Witzenhausen, Germany

believed to originate from the North-Eastern Mediterranean, where wild forms (var. *graecus*) still persist, and has been cultivated for more than 4000 years across the Mediterranean and Eastern Africa (Wolko et al. 2011). Domestication of white lupin has been slow and systematic breeding efforts scarce. In general, lupin species have long been considered valuable assets in crop rotations due to their unique symbiosis with *Bradyrhizobium lupini,* making them highly efficient nitrogen fixers (Fernández-Pascual et al. 2007; Peix et al. 2015). Specifically, white lupin is one of the few crops that form specialized cluster roots, which increase phosphorus availability by carboxylate secretion, significantly increasing soil fertility (Gallardo et al. 2020; Lambers et al. 2013). Since the development of sweet low alkaloid varieties (Kroc et al. 2017), white lupin has become increasingly interesting for the food and feed industry (Lucas et al. 2015).

Cultivation of lupins, however, is severely compromised by the seed- and air-borne fungal pathogen *Colletotrichum lupini*, causing lupin anthracnose (Nirenberg et al. 2002; Talhinhas et al. 2016). Infected seeds (primary infection)

and rain-splash dispersal (secondary infection) can cause total yield loss under favorable conditions (Thomas and Sweetingham 2004; White et al. 2008). Typical symptoms are stem twisting and bending and necrotic lesions on stems and pods (Alkemade et al. 2021b). *Colletotrichum lupini* is a member of the *C. acutatum* species complex (clade 1), which contains numerous important plant pathogens (Damm et al. 2012). Contrary to the broad host range seen for most members of this complex, *C. lupini* is host specific to members of the genus *Lupinus* (Baroncelli et al. 2017; Talhinhas et al. 2016). The current lupin anthracnose outbreak started in the 1970s and coincided with a decrease in lupin production worldwide, especially in Europe (FAOSTAT 2021). The pandemic is caused by a globally dispersed and genetically uniform group (II) of highly aggressive strains originating from South America (Alkemade et al. 2021b; Dubrulle et al. 2020a). Little is known about the interaction between *C. lupini* and its host, but a hemibiotrophic lifestyle is considered likely (De Silva et al. 2017; Dubrulle et al. 2020b).

Disease management of anthracnose in white lupin is currently focused on planting pathogen-free seed and chemical control, although the latter strategy is not available for the organic sector and is considered problematic due to adverse environmental effects (Thomas et al. 2008; White et al. 2008). The dispersal of infected symptomless seeds is believed to be the most likely cause of the rapid spread of *C. lupini* strains across the globe. Advanced molecular diagnostics to determine infection levels are developed, but are not yet routinely available (Kamber et al. 2021; Pecchia et al. 2019). Breeding for resistance is therefore likely to be the most sustainable solution. However, no complete resistance has yet been found in white lupin and the trait is considered polygenic (Adhikari et al. 2009; Alkemade et al. 2021a; Jacob et al. 2017). Quantitative trait locus (QTL) mapping of anthracnose resistance using a recombinant inbred line population formed with the highly resistant Ethiopian landrace (P27174) and the susceptible cultivar Kiev Mutant revealed three major resistance QTLs in an Australian experiment (Książkiewicz et al. 2017; Phan et al. 2007; Yang et al. 2010). Unfortunately, accessions selected based on these QTLs did not show increased resistance either under controlled or Swiss field conditions (Alkemade et al. 2021a). The development of a high-throughput phenotyping system for field-relevant anthracnose resistance, together with the availability of a high-quality white lupin reference genome (Hufnagel et al. 2020), allows for more in-depth genomic studies. Genome-wide association studies (GWAS) have proved to be a valuable tool to determine the underlying genetics of quantitative traits in diverse populations and have led to the discovery of single-nucleotide polymorphism (SNP) markers and candidate genes associated with traits of interest for numerous crops (Liu and Yan 2019). As an example, GWAS was recently used with the closely related blue lupin (*L. angustifolius* L.) to identify SNP markers for pod shattering (Mousavi-Derazmahalleh et al. 2018b) and climatic adaptation (Mousavi-Derazmahalleh et al. 2018a).

The aim of this study was to identify SNP markers and candidate genes associated with anthracnose resistance in white lupin. A collection of 200 white lupin cultivars, breeding lines and landraces, originating from across the Mediterranean and important cultivation regions, was genotyped-by-sequencing (GBS) and phenotyped for anthracnose resistance under controlled conditions. These accessions were shown to be variable for key agronomic traits (Annicchiarico et al. 2010) such as drought tolerance (Annicchiarico et al. 2018) and grain yield (Annicchiarico et al. 2019). Understanding the genetic architecture of anthracnose resistance in white lupin will provide crucial information to support further crop improvement.

Material and methods

Germplasm collection

White lupin (*Lupinus albus* L.) accessions were collected across the Mediterranean region, Atlantic islands, Eastern Africa, Europe, Chile and Australia from seed genebanks and local partners. The accessions are described in Electronic Supplemental Material 1 (ESM_1). In total, the 200 accessions include commercial cultivars, breeding lines and traditional land races. The collection includes a large number of landraces from CREA's white lupin world collection (Annicchiarico et al. 2010), widely studied accessions such as Amiga, Feodora, Kiev Mutant and P27174 (Adhikari et al. 2009; Hufnagel et al. 2021), and recently discovered anthracnose-resistant lines (Alkemade et al. 2021a).

Disease phenotyping

The white lupin collection was phenotyped for anthracnose resistance under controlled conditions (25 ± 2 °C, 16 h light and 70% relative humidity) using the high-throughput protocol described by Alkemade et al. (2021a). Stem wound inoculations were performed with the highly virulent *Colletotrichum lupini* strain JA01 (genetic group II; Alkemade et al. (2021b)). Disease was assessed at 3, 7 and 10 days post-inoculation (dpi) with a 1 to 9 disease score index (DSI), with 1 being healthy and 9 completely diseased (Alkemade et al. 2021a). At 10 dpi, lesion size, stem length and shoot fresh weight were determined. The overall disease score (based on the three DSI assessments) is expressed as the standardized area under the disease progress curve (sAUDPC, Jeger and Viljanen-Rollinson (2001)), the lesion size as relative to overall stem length (LS_{rel}), and the shoot fresh weight is expressed relative to

a control (SFW$_{rel}$). All of the experiments were performed in a randomized complete block design with a minimum of 8 replicates per accession.

Field trials

Phenotypic data obtained under controlled conditions of twelve selected accessions, ranging from susceptible to resistance, were compared to phenotypic data of these twelve accessions acquired over three-year field trials in Switzerland (ESM_1). Field trials were performed within six row plots according to Alkemade et al. (2021a) at three distinct sites: Mellikon (47°34′05.3"N 8°21′19.3″E) in 2018 and 2019, Full-Reuenthal (47°36′02.8″N 8°11′35.2″E) in 2020 and Feldbach (47°14′20.0″N, 8°47′18.8″E) in 2018, 2019 and 2020. Trials performed in 2018 and 2019 are described in Alkemade et al. (2021a). In 2020, in Full-Reuenthal plot sizes were 1.32×3.5 m and in Feldbach plot sizes were 1.5×2.7 m with a seed density of 65 seed/m^2. Total field size was 304 m^2 in Full-Reuenthal and 259 m^2 in Feldbach. Trials were performed in a randomized complete block design consisting of 4 replicates. The field trials relied on natural infection and were scored 80, 100 and 135 days after sowing. The DSI ranged from 1 (healthy) to 9 (dead), as described in Alkemade et al. (2021a). The sAUDPC and yield (dt/ha) were determined.

Phenotypic data analysis

Statistical analyses of the phenotypic data were performed within R 4.0.3 (R Core Team 2020) using the packages *lme4* (Bates et al. 2015), *lmerTest* (Kuznetsova et al. 2017) and *emmeans* (Lenth et al. 2019), following a mixed model. The factors of interest (i.e., accession) were included as fixed effects, while environment, environment x accession and replicated block nested in environment were fitted as random factors, after confirming the assumptions of normality of residuals and homogeneity of variance. To achieve a normal distribution, data were transformed with a square root (yield), log10 (SFW$_{rel}$), square (sAUDPC$_{CC}$), or logit (LS$_{rel}$) transformation. The mean separation between accessions and the overall mean of all accessions combined was analyzed using Dunnett's test ($P \leq 0.05$). The data are presented as non-transformed estimated least-squares means obtained using the aforementioned mixed model. Estimated means of controlled and field conditions were correlated using the Pearson correlation coefficient. Broad sense heritability (H^2) was estimated as: genotypic variance/phenotypic variance (Toker 2004).

Genotyping and SNP calling

Genomic DNA was isolated from leaf tissue of three week old plants using DNeasy Plant Mini Kit (Qiagen, Hilden, D) and quantified with a Quant-iT™ PicoGreen™ dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). Samples were genotyped in four different batches referred to as TX2016-1 (88 samples), TX2016-2 (32 samples), EL2018 (40 samples) and EL2020 (40 samples), using slightly different procedures as follows:

Genotyping-by-sequencing (GBS) libraries for TX2016-1 and TX2016-2 were prepared with a modified Elshire et al. (2011) protocol. DNA samples (100 ng) were digested with restriction enzyme *ApeKI* (New England Biolabs, Ipswich, MA, USA) and ligated to unique barcodes and common adapters. Equal volumes of ligated products were pooled and purified with NucleoSpin Gel and PCR Clean-up (Macherey–Nagel, Düren, D). Template DNA (50 ng) was mixed with two primers (ESM_2) and KAPA Library Amplification Readymix (Roche, Basel, CH). Amplification steps were as follows: 5 min at 72 °C, 30 s at 98 °C and 10 cycles with 10 s at 98 °C, 30 s at 65 °C and 30 s at 72 °C. Sequencing was performed at the University of Texas (USA) on four Illumina HiSeq 2000 (Illumina Inc., San Diego, CA, USA) lanes, at 100 bp single end. DNA samples for libraries EL2018 and EL2020 were sent to The Elshire Group Ltd. (Palmerston North, New Zealand) for library preparation and sequencing. Library preparation was performed according to Elshire et al. (2011) as described above with the following changes: libraries were amplified with 14 PCR cycles and prepared using a KAPA HyperPrep Kit (Roche, Basel, CH) following the manufacturer's instructions. Sequencing was performed on a single Illumina HiSeq X lane, at 2X150 bp paired end.

GBS raw reads were demultiplexed using axe demultiplexer (Murray and Borevitz 2018). Trimming for restriction enzyme remnants, alignment on reference genome and SNP calling were performed using the dDocent pipeline (Puritz et al. 2014). For alignment we used the *L. albus* genome version 1.0 (Hufnagel et al. 2020) which was downloaded from https://www.whitelupin.fr/. The final genotype matrix, in the form of a vcf file, was further filtered for quality using the vcftools software (Danecek et al. 2011) with parameters − minQ 30 − max-non-ref-af 1 –non-ref-af 0.001. The resulting data set of 246,279 SNPS was filtered for monomorphic markers, minor allele frequency (MAF) < 5%, missing SNP marker rate > 10%, and a missing rate per individual > 20% (Pavan et al. 2020). Genotypes that deviated with 3 SD from the mean heterozygosity rate were removed (Marees et al. 2018). Missing data were imputed through Beagle (Browning and Browning 2016) within *statgenGWAS* (van Rossum et al. 2020) in *R*, resulting in 4611 high-quality SNPs for 181 accessions (ESM_3).

Linkage disequilibrium and population structure

Linkage disequilibrium (LD) of SNP markers was calculated as the pairwise squared correlation coefficient (r^2) between markers using *LD.decay* in *R* (Laido et al. 2014). Significant ($P \leq 0.05$) pair-wise LD estimates were used to calculate average LD decay within a sliding window of 5 kb. LD decay was visualized by plotting r^2 estimates against genetic distance (kb). A pairwise distance matrix derived from Euclidean distance of the full SNP dataset was calculated in R to construct a Ward Hierarchical clustering tree (Murtagh and Legendre 2014) with 1000 bootstraps using *pvclust* in R (Suzuki and Shimodaira 2006). The tree was generated with *ape* (Paradis and Schliep 2019) in R and modified in iTOL v 6.1 (Letunic and Bork 2007). Principal component analysis (PCA) was performed using the *prcomp* function in R based on 1,292 SNPs filtered for a MAF > 20 and physical distance > 2.5 kb. An Astle kinship matrix (Astle and Balding 2009) was generated using *statgenGWAS* in R using the pruned SNP dataset.

Genome-wide association mapping

A genome-wide association study (GWAS) was performed on 181 accessions and 4,611 SNPs using estimated least-square means for the traits disease score (sAUDPC), LS_{rel} and SFW_{rel}. The association between SNPs and phenotypes was determined by performing a single-trait GWAS following a single-locus mixed linear model (MLM) and multi-locus Bayesian information and linkage disequilibrium iteratively nested keyway (BLINK) model (Huang et al. 2019). MLM was performed within *statgenGWAS* in R following the method described in Kang et al. (2010). The first ten principal components (PCs) were included as covariates to control for population structure, and the Astle kinship matrix was included to account for cryptic relatedness (Astle and Balding 2009; Rincent et al. 2014). An efficient mixed model association (EMMA) algorithm was used to estimate the variance components (Kang et al. 2008). General least squares (GLS) were used to estimate effect size and P-value for each SNP. BLINK was performed within GAPIT 3.0 (Wang and Zhang 2021) using the first ten PCs. A Bonferroni corrected LOD threshold ($-\log10(0.05/\text{number of SNPs})$) was used to identify significant SNPs, and a fixed threshold of $-\log10(5.00\text{E}-05)$) was used to identify low peak SNPs otherwise missed by the highly conservative Bonferroni threshold (Storey and Tibshirani 2003). SNPs within 2.5 kb and ≥ 0.5 r^2 were considered linked. Manhattan and quantile–quantile (Q–Q) plots were generated within *statgenGWAS*.

Candidate gene selection

Candidate genes were considered when containing a significant SNP or a SNP in LD ($r^2 > 0.5$) with a significant SNP and when within 10 kb of significant SNPs. Candidate genes were located using the white lupin reference genome (v 1.0) browser (Hufnagel et al. 2020). Protein sequences were acquired and blasted (BLASTp) to find homologs in closely related species, i.e., blue lupin (*L. angustifolius*), peanut (*Arachis hypogea*), common bean (*Phaseolis vulgaris*), soybean (*Glycine max*) and model species *Medicago truncatula* (LPWG 2017). The potential function of each candidate gene was derived from annotations, literature and in silico analysis.

Results

Strong differentiation and heritability of anthracnose-related traits

Disease phenotyping of 200 white lupin accessions under controlled conditions revealed a range of resistant and susceptible accessions (Fig. 1). Strong differences between accessions ($P < 0.0001$) for all three anthracnose related traits, sAUDPC, LS_{rel} and SFW_{rel}, were observed with heritabilities of 0.77, 0.78 and 0.74, respectively. Significant ($P \leq 0.05$) Pearson correlations were found between the three traits ($r > (-)0.86$; ESM_4). For twelve selected accessions, strong correlations were observed between these three traits and the overall disease assessment means of three-year field trials in Switzerland ($r > (-)0.8$; ESM_5). We did not observe complete resistance against anthracnose, which we interpret as evidence that resistance in white lupin is quantitative. For sAUDPC, LS_{rel} and relative SFW_{rel}, 8, 13 and 2 accessions, respectively, were more resistant than the respective overall mean ($P \leq 0.05$, ESM_1). Six of these accessions originated from Ethiopia, two from Chile and one being the newly available commercial variety Frieda. In contrast, seven accessions were more susceptible than the overall mean. Remarkably, the Ethiopian landrace P27175 was resistant in one seed batch (FiBL 39) but susceptible in another (FiBL 19) and showed distinct heterozygosity rates of 0.77% and 32.2%, respectively (ESM_1). Resistant FiBL025 (R-6020) and susceptible LAP0119b (Aster) seeds showed a black-speckled morphology typical for wild *graecus* types.

Weak population structure and fast LD decay

Genotyping by sequencing yielded 4611 high-quality SNPs for 181 accessions after filtering for monomorphic markers, minor allele frequency (MAF) < 5%, missing SNP marker
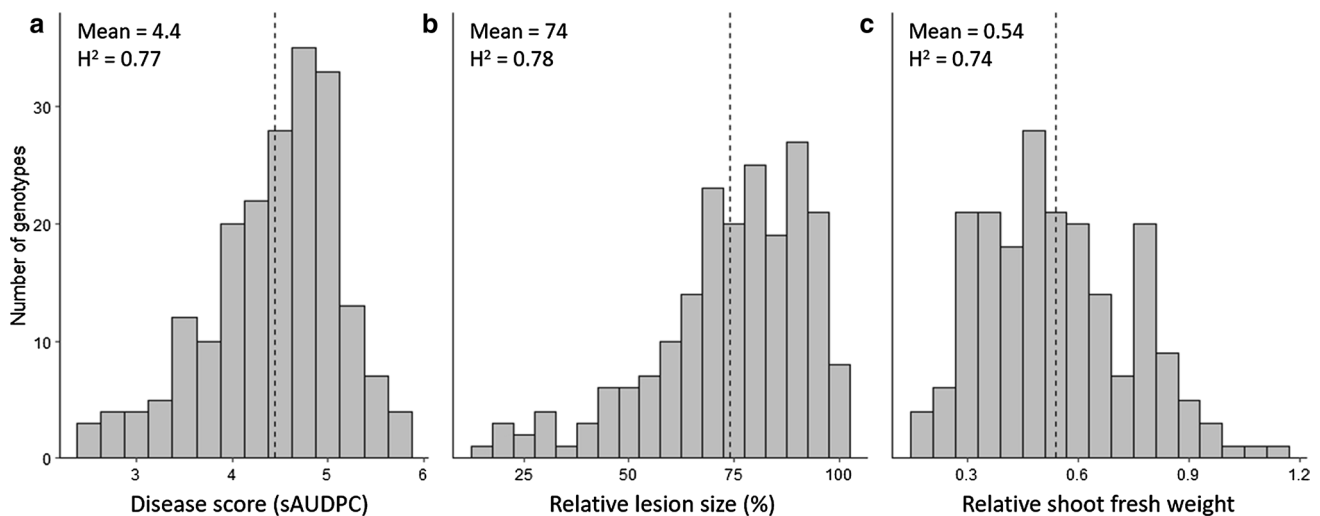
Fig. 1 Phenotypic variation in three anthracnose resistance traits among 200 white lupin accessions. **a** Disease score (standardized area under the disease progress curve (sAUDPC)). **b** Relative lesion size (%). **c** Relative shoot fresh weight. Dashed line indicates overall mean

rate > 10%, missing rate per accession > 20% and heterozygosity > 40% (mean Ho = 15%; ESM_1). LD decayed to half its maximum value at 2.9 kb ($r^2 = 0.45$), and SNPs were in linkage ($r^2 > 0.5$) over an average distance of 2.5 kb (ESM_6). Cluster analysis on the full SNP dataset distinguished 4 subgroups (I–IV) based on bootstrap support values (BS) > 90 and a branch length threshold of 10 (Fig. 2a). These subgroups could also be observed through PCA and Astle kinship analysis after pruning the SNP dataset (MAF > 20% and physical distance > 2.5 kb; Fig. 2b and c), which revealed overlap between groups II, III and IV, while group I was more clearly separated. Group I exclusively contained landraces originating from the South-East Mediterranean (Fig. 2a and ESM_1). Group II, which includes accessions from across the entire study area, encompasses most of the commercial cultivars and breeding lines used in this study (88%). A large proportion (27%) of group II includes landraces from North Africa, with half originating from Ethiopia. Group III consists mostly of Egyptian (64%) and Ethiopian (23%) landraces, with the Ethiopian landraces strongly clustering together (BS = 100). Group IV contains accessions from across the entire study area, including lines from the Iberian Peninsula (30%) and the Atlantic Isles (26%). Kinship between accessions showed relatively close relatedness among sampled accessions (Fig. 2c).

Three significant SNPs associated with anthracnose resistance

To map genetic variants associated with anthracnose resistance, we performed a genome-wide association study (GWAS) following a MLM and BLINK model. We included the first 10 principal components (PCs) and the

Astle kinship matrix to correct for population structure. The first 10 PCs explained 24% of the genetic variance. The resulting Q-Q plots revealed that the model was well calibrated, as we observed a good approximation between the expected and observed P-values (Fig. 3 and ESM_7). Following the MLM and applying a Bonferroni LOD-threshold of 4.96 ($P = 1.08E-05$), we identified two highly significant SNPs, Lalb_Chr05_2957601 and Lalb_Chr05_2957940, for both sAUDPC ($P = 6.11E-09$ and $2.41E-07$, respectively) and LS_{rel} ($P = 1.55E-06$ and $2.13E-06$, respectively; Fig. 3 and Table 1). Analysis with the BLINK model identified a significant association between Lalb_Chr05_2957601 and sAUDPC ($P = 6.58E-12$) and SFW_{rel} ($P = 2.72E-07$), and between Lalb_Chr05_2957940 and LS_{rel} ($P = 2.72E-07$; ESM_7 and Table 1). The two SNPs explained 12 to 16% of the observed variation (R^2_{LR}; Table 1), were not strongly linked with other SNPs and were found in exons of the same gene (*Lalb_Chr05g0216161*; Fig. 4). The minor allele frequencies (MAF) of these SNPs were 10 and 7%, respectively, and the non-reference alleles were significantly associated with increased anthracnose resistance when either heterozygous or homozygous (Fig. 4c). The non-reference alleles were only found homozygous for both SNPs in the Chilean accessions Fibl016 (Blu-25) and LAP0155a & b (Rumbo Baer), which were found to be highly resistant (sAUDPC = 2.55, 2.78 and 3.11, respectively).

Analysis with the BLINK model identified an additional SNP on chromosome 5, Lalb_Chr05_3706534 ($P = 1.91E-07$), associated with LS_{rel} (ESM_7 and Table 1). This SNP explained 10% of the observed variation and showed a MAF of 16%. The non-reference allele of this SNP was implicated in decreased anthracnose resistance (ESM_8). SNPs Lalb_Chr05_3688076 and Lalb_Chr05_3784474 were
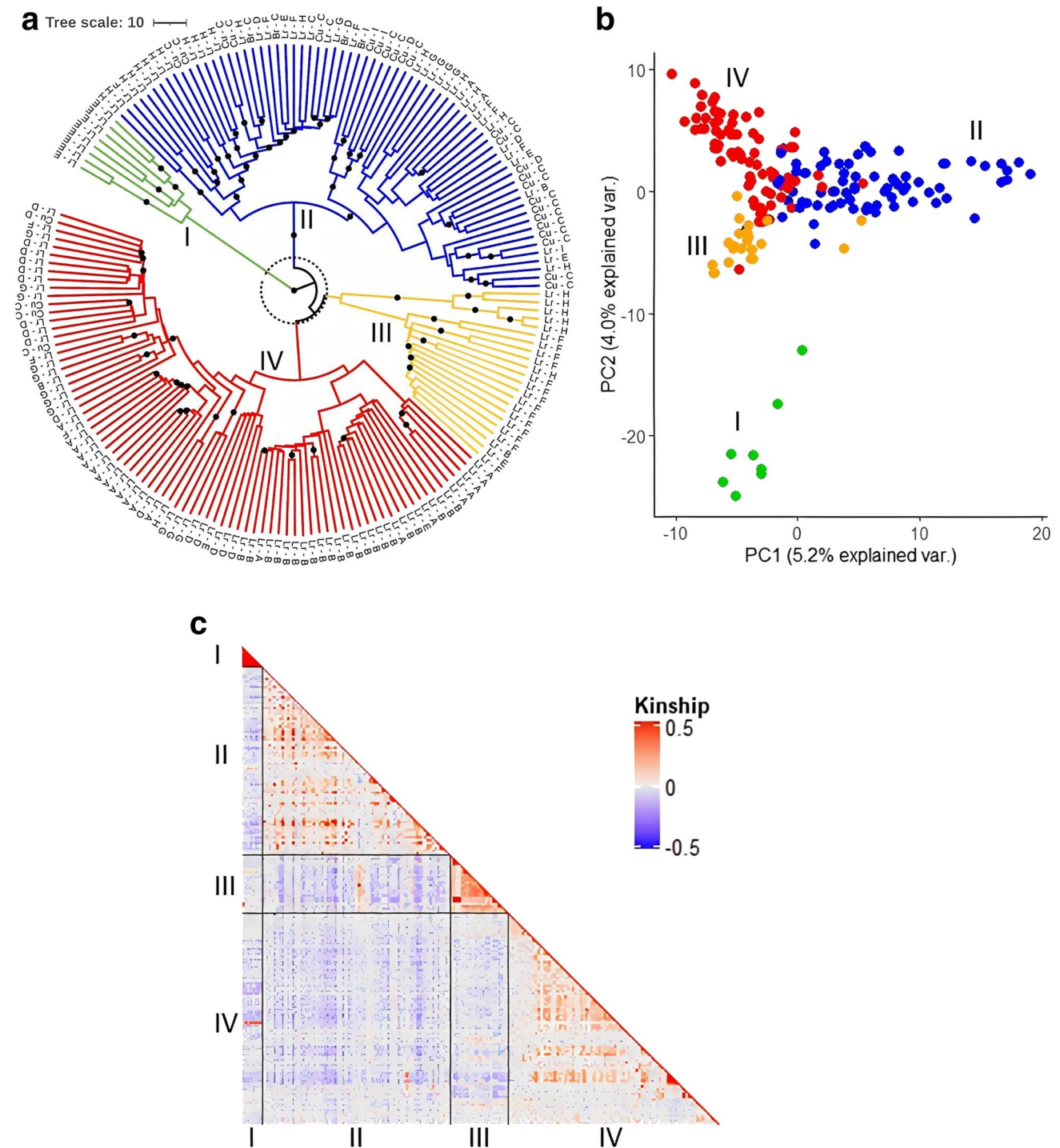
Fig. 2 Genetic diversity and population structure of 181 white lupin accessions. **a** Ward cluster analysis (1,000 bootstraps). Colors represent subgroups (I-IV). Black dots represent bootstrap support values (> 90), and dotted circle indicates branch length of 10. Letters indicate accession type (inner) and region of origin (outer) with Lr = landrace, Cu = cultivar, Br = breeding line, A = Atlantic Isles, B = Iberian Peninsula, C = West and Eastern Europe, D = North-East Mediterranean, E = South-East Mediterranean, F = Egypt, G = South-West Mediterranean, H = East Africa and I = Chile. **b** Principal component analysis (PCA). Each dot represents an accession and colors represent subgroups (I-IV). **c** Heatmap of the Astle kinship value among accessions
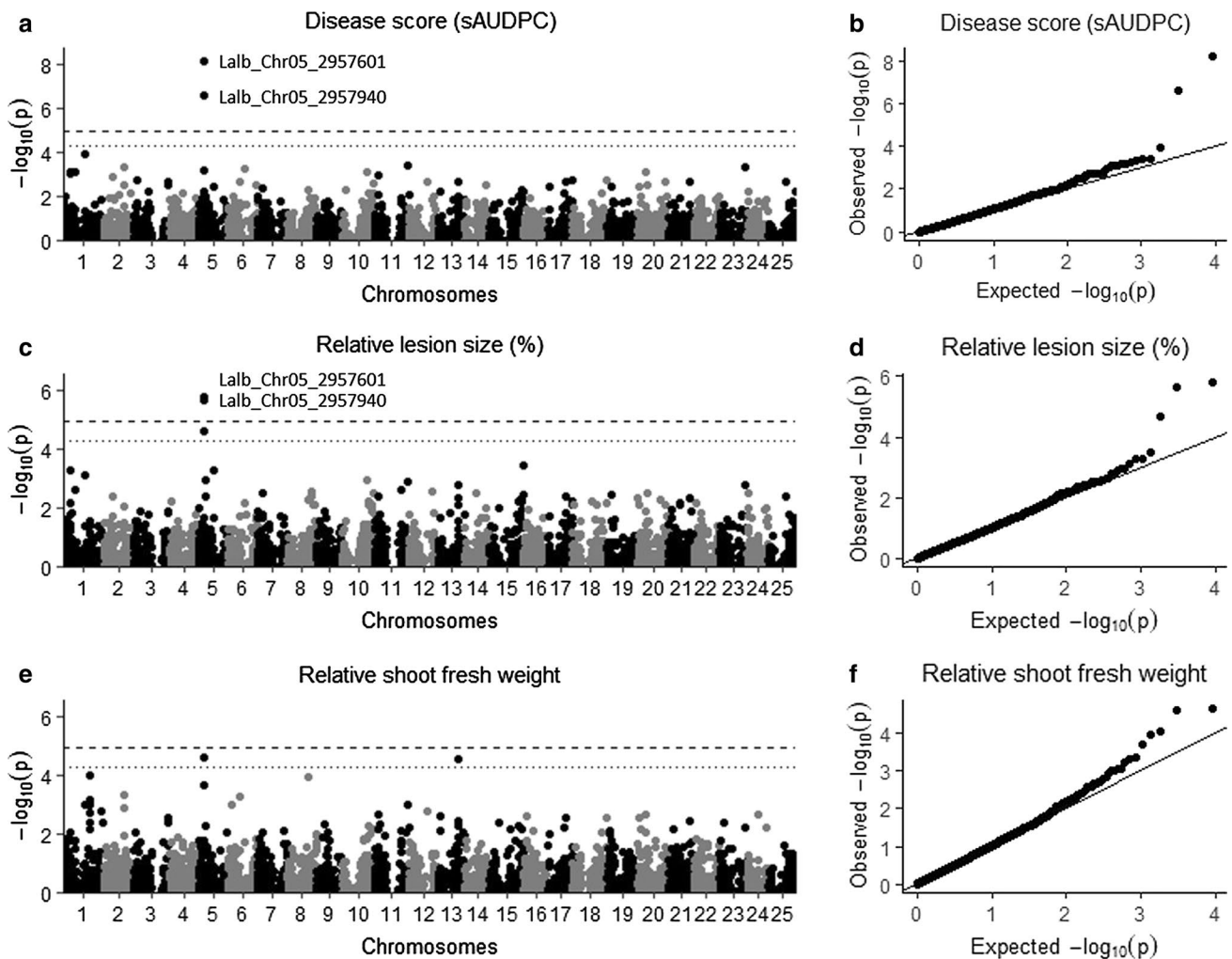
Fig. 3 Manhattan and corresponding Q–Q plots by MLM showing SNP association with anthracnose resistance. **a, b** Disease score (standardized area under the disease progress curve (sAUDPC)). **c, d** Relative lesion size (%). **e, f** Relative shoot fresh weight. Upper dashed line indicates Bonferroni corrected LOD threshold of 4.96 ($P = 1.08E{-}05$), and lower dotted line indicates fixed LOD threshold of 4.3 ($P = 5.00E{-}05$)

considered linked with Lalb_Chr05_3706534 (ESM_9). Other promising SNPs, above the fixed LOD threshold of 4.3 ($P = 5.00E{-}5$)), were found on chromosome 1 (Lalb_Chr01_3872625) for LS_{rel}, on chromosome 13 (Lalb_Chr13_12108967) for SFW_{rel} and on chromosome 6 (Lalb_Chr06_9655085) for sAUDPC (ESM_9).

Candidate genes involved in resistance pathways

Candidate genes were considered when containing a significant SNP, a SNP linked ($r^2 > 0.5$) to a significant SNP or when located within 10 kb of a significant SNP. The significant SNPs Lalb_Chr05_2957601 and Lalb_Chr05_2957940 are both located within an exon of the same gene: *Lalb_Chr05g0216161* (Table 1, Fig. 4d). This gene is annotated

as a putative chromatin regulator and encodes a protein containing a Von Willebrand factor type A (VWFA) as well as a RING zinc-finger domain (ZF; Fig. 4d). Homologs in closely related legume species encode for RING zinc-finger E3 ubiquitin ligases, which are widely associated with plant immunity (Marino et al. 2012). Lalb_Chr05_3706534 is located within an intron of gene *Lalb_Chr05g0217341* which encodes a putative transcription regulator, and homologs in related species encode a paired amphipathic helix protein Sin3. Linked to this SNP is Lalb_Chr05_3784474, which is located within an exon of the gene *Lalb_Chr05g0217471* which encodes for a non-specific serine/threonine protein kinase (ESM_9). Homologs in closely related species encode leucine-rich repeat (LRR) receptor-like serine/threonine-protein kinases, which are often implicated in plant

Table 1 Significant SNPs and associated candidate genes

Trait	SNP[a]	Allele	P value MLM	P value BLINK	R^2_{LR}	MAF	Candidate gene	Annotation
Disease Score (sAUDPC)	Lalb_Chr05_2957601	G/A	6.11E-09	6.58E-12	0.16	0.10	*Lalb_Chr05 g0216161*	E3 ubiquitin-protein ligase
	Lalb_Chr05_2957940	C/T	2.41E-07	0.016[b]	0.13	0.07		
Relative lesion size (%)	Lalb_Chr05_2957601	G/A	1.55E-06	0.042[b]	0.12	0.10		
	Lalb_Chr05_2957940	C/T	2.13E-06	3.98E-10	0.12	0.07		
Relative shoot fresh weight	Lalb_Chr05_2957601	G/A	2.29E-05[b]	2.72E-07	0.09	0.10		
	Lalb_Chr05_2957940	C/T	1.81E-04[b]	0.11[b]	0.07	0.07		
Relative lesion size (%)	Lalb_Chr05_3706534	A/G	2.21E-05[b]	1.91E-07	0.10	0.16	*Lalb_Chr05 g0217341*	Paired amphipathic helix protein Sin3-like 4

SNP single-nucleotide polymorphism, *MAF* minor allele frequency, R^2_{LR} likelihood-ratio-based R^2, *sAUDPC* standardized area under the disease progress curve.

[a]*Lalb Lupinus albus*, *Chr* chromosome, number = position on chromosome. [b]Not significant
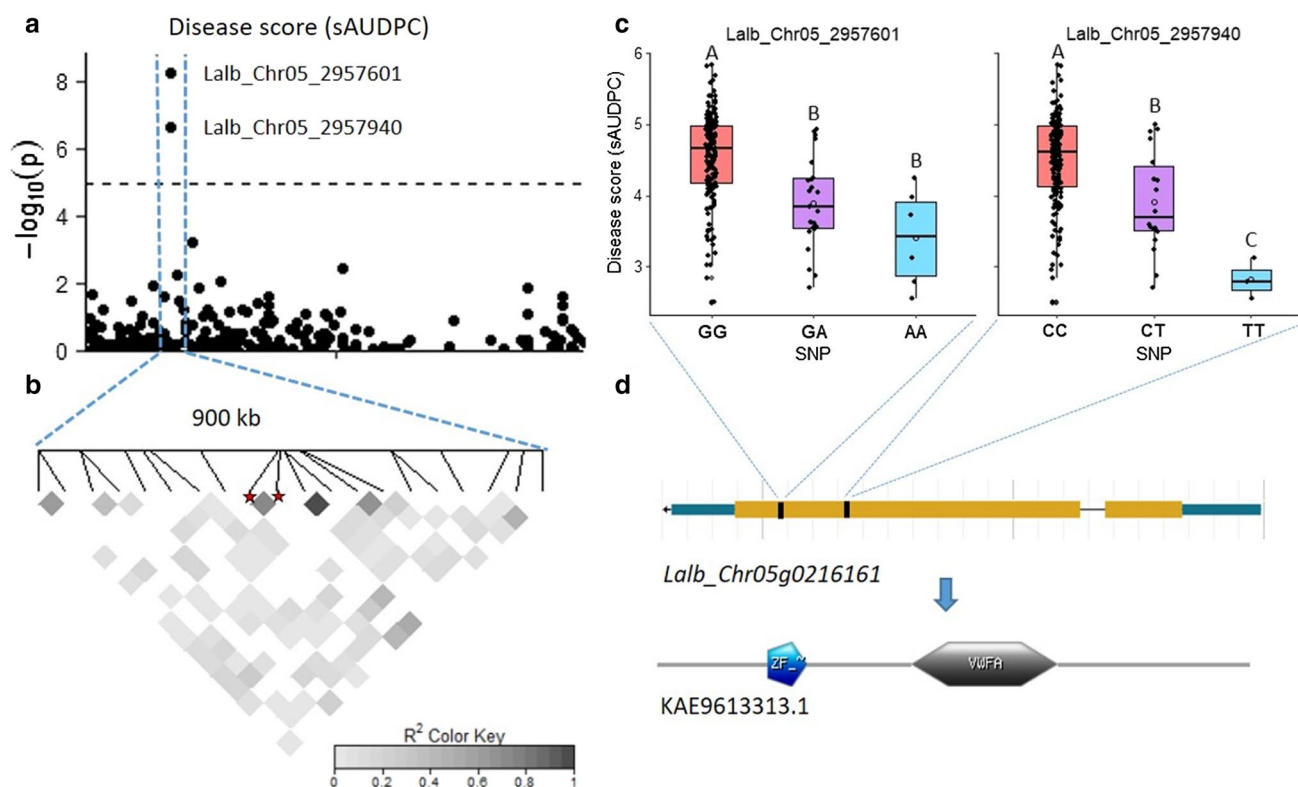

Fig. 4 Characterization of SNP Lalb_Chr05_2957601 & -940. **a** Manhattan plot of chromosome 5, dashed line indicates Bonferroni corrected LOD threshold of 4.96 (*P* = 1.08E-05). **b** Linkage disequilibrium (LD) heatmap of 20 SNPs within 500 kb of significant SNPs (stars). **c** Boxplots showing allele effect on disease score (standardized area under the disease progress curve (sAUDPC)). Capital letters within plot indicate significant difference (Tukey-HSD, *P* ≤ 0.05). **d** Candidate gene *Lalb_Chr05g0216161*, showing protein coding region (orange) and spliced non-coding RNA (blue; www.whitelupin.fr) and corresponding protein KAE9613313.1 with Von Willebrand factor type A (VWFA) and RING zinc-finger domain (ZF; De Castro et al. (2006))

defense against fungal pathogens (Afzal et al. 2008; Tang et al. 2017).

Discussion

Disease phenotypes of twelve selected accessions obtained under controlled conditions, strongly correlated ($r > 0.8$) to overall three-year field plot disease assessments in Switzerland, confirming field-relevance of high-throughput phenotyping under controlled conditions (Alkemade et al. 2021a). No complete resistance was observed, but based on sAUDPC, LS_{rel} and SFW_{rel}, a total of 15 different accessions showed to be significantly more resistant to anthracnose compared to the overall mean. High resistance was found for Chilean cultivar Rumbo Baer and the breeding line Blu-25, both of which appear to derive from a resistant landrace from the Azores (von Baer et al. 2009). Five resistant accessions originated from Ethiopia, which was previously shown to be a good source for white lupin anthracnose resistance (Adhikari et al. 2009; Cowling et al. 1999). However, Ethiopian landrace P27174 (Fibl020 & 38), used as a resistant parent for an anthracnose resistance QTL study (Książkiewicz et al. 2017; Yang et al. 2010), was not shown to be resistant in this study and in Alkemade et al. (2021a), suggesting the occurrence of cross-pollination or admixture. Resistance and heterozygosity rate in Ethiopian landrace P27175 differed among seed batches, suggesting differences in seed quality, outcrossing or that P27175 represents a mixture of genotypes. These observed irregularities for P27174 and P27175 should be further investigated.

White lupin has long been cultivated across the Mediterranean and North-Eastern Africa, with its primary center of origin believed to be in the Balkans up to Western Turkey where wild *graecus* types are still found (Wolko et al. 2011). This study, which contains accessions collected from across the traditional cultivation regions of white lupin, revealed an exceptionally fast LD decay (2.9 kb). This fast LD decay is consistent with an earlier study by Hufnagel et al. (2021) and the fact that white lupin has a modest rate of outcrossing (Green et al. 1980). Studies on other grain legume species, including pea (> 50 kb; Gali et al. (2019)), soybean (> 240 kb; Wen et al. (2018)), and common bean (> 1 Mb; Diniz et al. (2019)), as well as the closely related blue lupin (> 77 kb; Mousavi-Derazmahalleh et al. (2018b)), have all reported slower rates of LD decay. The analyzed population showed an average heterozygosity rate of 15%, which is similar to observed outcrossing rates of 10% (Green et al. 1980). The heterozygosity within the studied population could have reduced the power of the GWAS to detect major loci (Alqudah et al. 2020). Increasing sample size, marker coverage and creating an inbred population could improve the analysis.

In addition, we detected a weak population structure, finding four subgroups (I–IV). Principal component analysis showed overlap between Groups II, III and IV, while only Group I, exclusively containing landraces from the South-Eastern Mediterranean, formed a clearly distinct group. Group III consisted primarily of Egyptian and Ethiopian landraces, with Ethiopian accessions strongly grouping together. Landraces from Ethiopia were previously reported to form a distinct group within white lupin (Raman et al. 2014) and were shown to be most closely related to wild *graecus* types (Hufnagel et al. 2020, 2021), which suggests these landraces derived in isolation and are still little domesticated. In contrast with these results, 10 of the 16 Ethiopian landraces collected in this study were classified in Group II, containing commercial cultivars and landraces from all across the collection area. Taken together, we interpret the fast LD decay, weak population structure and the grouping of commercial varieties with landraces to indicate that there have been few recent breeding events in white lupin, which implies that there is great potential for further crop improvement of this re-emerging protein crop.

GWAS analysis by MLM and BLINK identified three significant SNPs, Lalb_Chr05_2957601, 2957940 and 3706534, on chromosome 5 associated with anthracnose resistance. Additionally, three SNPs were identified above the fixed LOD threshold of 4.3 on chromosomes 1, 6 and 13. Compared to MLM, BLINK improved statistical power and removed redundant genetically linked SNPs (Huang et al. 2019). It should be considered, however, that the genome-wide marker coverage in this study might not have been adequate to replace the polygenic effect of the kinship matrix and population structure by covariate markers as done in multi-locus methods such as BLINK (Tibbs Cortes et al. 2021). The identified SNPs do not correspond to previously reported QTLs associated with anthracnose or phomopsis (*Diaporthe toxica*) resistance in white lupin (Cowley et al. 2014; Książkiewicz et al. 2017). Corresponding candidate genes also do not reflect anthracnose resistance genes identified in blue lupin, including *Lanr1, Anman* and *LanrBo* (Fischer et al. 2015; Yang et al. 2004, 2008).

SNPs Lalb_Chr05_2957601 and 2957940 are both located in the same coding region of *Lalb_Chr05g0216161* which encodes for a protein with a RING zinc-finger and VWFA domain. The MAF for these SNPs was low, and non-references alleles were only homozygous in Chilean, Ethiopian and Moroccan accessions, but were also present in wild *graecus* types (LD37, GR38, and Batsi; Hufnagel et al. (2021)). Homologs in closely related legume species encode RING zinc-finger E3 ubiquitin-protein ligases. E3 ubiquitin-ligases have frequently been shown to be involved in different steps of plant immunity (Duplan and Rivas 2014; Marino et al. 2012; Zhou and Zeng 2017). In pepper (*Capsicum annuum*), the RING finger protein gene, *CaRFP1*, containing

a VWFA domain, was shown to act as E3 ubiquitin ligase and was highly upregulated during *C. coccodes* infection (Hong et al. 2007). Other RING type E3 ubiquitin ligases were shown to influence resistance against *Magnaporthe oryzae* in rice (Park et al. 2016), *Xanthonomas* infection in *C. annuum* (Lee et al. 2011) and *Ralstonia solanacearum* in tobacco (Ghannam et al. 2016). Besides biotic stress, RING E3 ubiquitin ligases have shown to improve resistance against abiotic stresses (Cho et al. 2017; Lee and Kim 2011), such as drought (Cheng et al. 2012) and salt stress (Kim and Kim 2013), and were shown to be involved in various plant developmental processes (Shu and Yang 2017), such as root development (Sakai et al. 2012). In conclusion, the identified gene, *Lalb_Chr05g0216161*, might play an important role in anthracnose resistance in white lupin and should be further investigated.

SNP Lalb_Chr05_3706534 is located in the noncoding region of *Lalb_Chr05g0217341*. Homologs of *Lalb_Chr05g0217341* in closely related legume species encode for paired amphipathic helix protein Sin3-like 4 proteins, which are known to be involved in powdery mildew (*Podosphaera fusca*) resistance in cucumber (Liu et al. 2021). SNP Lalb_Chr05_3784474 is considered linked to Lalb_Chr05_3706534 and is located within the gene *Lalb_Chr05g0217471* which encodes a LRR containing receptor-like protein with a serine/threonine kinase. LRR receptor kinases are well known as resistance genes and for their role in plant immunity (Afzal et al. 2008; Ellis et al. 2000; Tang et al. 2017), and serine/threonine kinases were shown to be involved in signaling during pathogen recognition and subsequent activation of plant defense mechanisms (Afzal et al. 2008; Goff and Ramonell 2007). LRR receptor-like serine/threonine-protein kinases were shown to confer resistance against apple scab (*Venturia inaequalis*) in apple (Padmarasu et al. 2018) and against rice blast (*Magnaporthe grisea*) in rice (Song et al. 2008). The fact that the three significant SNPs, representing two loci (*Lalb_Chr05g0216161* and *Lalb_Chr05g0217341*), could only explain 16 to 34% of disease phenotypic variance, the identification of few low peak SNPs and the quantitative nature of the trait suggests that anthracnose resistance is under polygenic control as indicated by Książkiewicz et al. (2017), but the identification of one major loci (*Lalb_Chr05g0216161*) could also imply oligogenic control.

This study showed that GWAS, thanks to the weak population structure, fast LD decay and the availability of a high-quality reference genome, is a powerful tool to identify resistance loci in white lupin and provides the basis for further gene mapping. Further characterization of the E3 ubiquitin ligase encoding *Lalb_Chr05g0216161* gene, e.g., via gene expression studies, could shed first light on white lupin resistance mechanisms against anthracnose disease. The obtained dataset also provides a basis for marker-assisted selection and the development of genomic prediction models for anthracnose resistance. Overall, this study contributes to understanding the genetic make-up of anthracnose resistance in white lupin and supports future crop improvement.

Authors contribution statement JAA, PH and MMM conceived the original idea for this study. PH, MMM and PA acquired the funding for this project. JAA conducted the experiments and took the lead in manuscript writing. JAA analyzed the data with help from NN. BF contributed with providing lab assistance. CA planned and performed the field trials and provided seeds of accessions Fibl001-FiBL040. JAA designed the figures and tables with input from NN, PH, PA, MMM, RTV and MRF, who also significantly contributed to data interpretation and provided critical feedback that shaped the final version of the manuscript.

Data availability The datasets generated during the current study are available at: https://doi.org/10.5281/zenodo.5142130.

Declataions

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All experiments described in this manuscript comply with the current laws of the country in which they were performed.

References

Abraham EM, Ganopoulos I, Madesis P, Mavromatis A, Mylona P, Nianiou-Obeidat I, Parissi Z, Polidoros A, Tani E, Vlachostergios D (2019) The use of lupin as a source of protein in animal feeding: Genomic tools and breeding approaches. Int J Mol Sci 20:851. https://doi.org/10.3390/ijms20040851

Adhikari K, Buirchell B, Thomas GJ, Sweetingham MW, Yang H (2009) Identification of anthracnose resistance in *Lupinus albus* L. and its transfer from landraces to modern cultivars. Crop Pasture Sci 60:472–479. https://doi.org/10.1071/CP08092

Afzal AJ, Wood AJ, Lightfoot DA (2008) Plant receptor-like serine threonine kinases: roles in signaling and plant defense. Mol Plant-Microbe Interact 21:507–517. https://doi.org/10.1094/MPMI-21-5-0507

Alkemade JA, Messmer MM, Arncken C, Leska A, Annicchiarico P, Nazzicari N, Książkiewicz M, Voegele RT, Finckh MR, Hohmann P (2021a) A high-throughput phenotyping tool to identify field-relevant anthracnose resistance in white lupin. Plant Dis. https://doi.org/10.1094/PDIS-07-20-1531-RE

Alkemade JA, Messmer MM, Voegele RT, Finckh MR, Hohmann P (2021b) Genetic diversity of *Colletotrichum lupini* and its virulence on white and Andean lupin. Sci Rep 11:13547. https://doi.org/10.1038/s41598-021-92953-y

Alqudah AM, Sallam A, Baenziger PS, Börner A (2020) GWAS: Fast-forwarding gene identification and characterization in temperate cereals: lessons from barley—a review. J Adv Res 22:119–135. https://doi.org/10.1016/j.jare.2019.10.013

Annicchiarico P (2008) Adaptation of cool-season grain legume species across climatically-contrasting environments of southern Europe. Agron J 100:1647–1654. https://doi.org/10.2134/agronj2008.0085

Annicchiarico P, Harzic N, Carroni AM (2010) Adaptation, diversity, and exploitation of global white lupin (*Lupinus albus* L.) landrace genetic resources. Field Crops Res 119:114–124. https://doi.org/10.1016/j.fcr.2010.06.022

Annicchiarico P, Nazzicari N, Ferrari B, Harzic N, Carroni AM, Romani M, Pecetti L (2019) Genomic prediction of grain yield in contrasting environments for white lupin genetic resources. Mol Breed 39:142. https://doi.org/10.1007/s11032-019-1048-6

Annicchiarico P, Romani M, Pecetti L (2018) White lupin (*Lupinus albus*) variation for adaptation to severe drought stress. Plant Breed 137:782–789. https://doi.org/10.1111/pbr.12642

Astle W, Balding DJ (2009) Population structure and cryptic relatedness in genetic association studies. Stat Sci 24:451–471. https://doi.org/10.1214/09-STS307

Baroncelli R, Talhinhas P, Pensec F, Sukno SA, Le Floch G, Thon MR (2017) The *Colletotrichum acutatum* species complex as a model system to study evolution and host specialization in plant pathogens. Front Microbiol 8:2001. https://doi.org/10.3389/fmicb.2017.02001

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48. https://doi.org/10.18637/jss.v067.i01

Browning BL, Browning SR (2016) Genotype imputation with millions of reference samples. Am J Hum Genet 98:116–126. https://doi.org/10.1016/j.ajhg.2015.11.020

Cheng M-C, Hsieh E-J, Chen J-H, Chen H-Y, Lin T-P (2012) Arabidopsis RGLG2, functioning as a RING E3 ligase, interacts with AtERF53 and negatively regulates the plant drought stress response. Plant Physiol 158:363–375. https://doi.org/10.1104/pp.111.189738

Cho SK, Ryu MY, Kim JH, Hong JS, Oh TR, Kim WT, Yang SW (2017) RING E3 ligases: key regulatory elements are involved in abiotic stress responses in plants. BMB Rep 50:393–400. https://doi.org/10.5483/BMBRep.2017.50.8.128

Cowley R, Luckett DJ, Ash GJ, Harper JD, Vipin CA, Raman H, Ellwood S (2014) Identification of QTLs associated with resistance to Phomopsis pod blight (*Diaporthe toxica*) in *Lupinus albus*. Breed Sci 64:83–89. https://doi.org/10.1270/jsbbs.64.83

Cowling W, Buirchell B, Sweetingham M, Yang H, Thomas G, Luckett D, Brown A, Hamblin J (1999) Anthracnose resistance in lupins-an innovative Australian research effort 1996–1998. In: Lupin, an ancient crop for the new millennium: proceedings of the 9th international Lupin conference, Klink/Muritz, Germany, 20–24 June, 1999. International Lupin Association, pp 60–62

Damm U, Cannon PF, Woudenberg JHC, Crous PW (2012) The *Colletotrichum acutatum* species complex. Stud Mycol 73:37–113. https://doi.org/10.3114/sim0010

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330

De Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34:W362–W365. https://doi.org/10.1093/nar/gkl124

De Silva DD, Crous PW, Ades PK, Hyde KD, Taylor PW (2017) Life styles of *Colletotrichum* species and implications for plant biosecurity. Fungal Biol Rev 31:155–168. https://doi.org/10.1016/j.fbr.2017.05.001

Diniz AL, Giordani W, Costa ZP, Margarido GR, Perseguini JMK, Benchimol-Reis LL, Chiorato AF, Garcia AAF, Vieira MLC (2019) Evidence for strong kinship influence on the extent of linkage disequilibrium in cultivated common beans. Genes 10:5. https://doi.org/10.3390/genes10010005

Dubrulle G, Pensec F, Picot A, Rigalma K, Pawtowski A, Nicolleau S, Harzic N, Nodet P, Baroncelli R, Le Floch G (2020a) Phylogenetic diversity and effect of temperature on pathogenicity of *Colletotrichum lupini*. Plant Dis 104:938–950. https://doi.org/10.1094/PDIS-02-19-0273-RE

Dubrulle G, Picot A, Madec S, Corre E, Pawtowski A, Baroncelli R, Zivy M, Balliau T, Le Floch G, Pensec F (2020b) Deciphering the infectious process of *Colletotrichum lupini* in lupin through transcriptomic and proteomic analysis. Microorganisms 8:1621. https://doi.org/10.3390/microorganisms8101621

Duplan V, Rivas S (2014) E3 ubiquitin-ligases and their target proteins during the regulation of plant innate immunity. Front Plant Sci 5:42. https://doi.org/10.3389/fpls.2014.00042

Ellis J, Dodds P, Pryor T (2000) Structure, function and evolution of plant disease resistance genes. Curr Opin Plant Biol 3:278–284. https://doi.org/10.1016/S1369-5266(00)00080-7

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6:e19379. https://doi.org/10.1371/journal.pone.0019379

FAOSTAT (2021) Statistics Division of the Food and Agriculture Organization of the United Nations. http://www.fao.org/faostat/en/#data/QC. (01–06–2021)

Fernández-Pascual M, Pueyo JJ, Felipe M, Golvano MP, Lucas MM (2007) Singular features of the *Bradyrhizobium-Lupinus* symbiosis. Dyn Soil Dyn Plant 1:1–16

Fischer K, Dieterich R, Nelson MN, Kamphuis LG, Singh KB, Rotter B, Krezdorn N, Winter P, Wehling P, Ruge-Wehling B (2015) Characterization and mapping of *LanrBo*: a locus conferring anthracnose resistance in narrow-leafed lupin (*Lupinus angustifolius* L.). Theor Appl Genet 128:2121–2130. https://doi.org/10.1007/s00122-015-2572-3

Gali KK, Sackville A, Tafesse EG, Lachagari V, McPhee K, Hybl M, Mikić A, Smýkal P, McGee R, Burstin J (2019) Genome-wide association mapping for agronomic and seed quality traits of field pea (*Pisum sativum* L.). Front Plant Sci 10:1538. https://doi.org/10.3389/fpls.2019.01538

Gallardo C, Hufnagel B, Soriano A, Divol F, Marquès L, Doumas P, Peret B (2020) Developmental atlas of white lupin cluster roots. BioRxiv. https://doi.org/10.1101/2020.03.26.009910

Ghannam A, Jacques A, de Ruffray P, Kauffmann S (2016) NtRING1, putative RING-finger E3 ligase protein, is a positive regulator of the early stages of elicitin-induced HR in tobacco. Plant Cell Rep 35:415–428. https://doi.org/10.1007/s00299-015-1893-7

Goff KE, Ramonell KM (2007) The role and regulation of receptor-like kinases in plant defense. Gene Regul Syst Bio 1:167–175. https://doi.org/10.1177/117762500700100015

Green A, Brown A, Oram R (1980) Determination of outcrossing rate in a breeding population of *Lupinus albus* L. (white lupin). Plant Breed 84:181–191

Hong JK, Choi HW, Hwang IS, Hwang BK (2007) Role of a novel pathogen-induced pepper C3–H–C4 type RING-finger protein gene, *CaRFP1*, in disease susceptibility and osmotic stress tolerance. Plant Mol Biol 63:571–588. https://doi.org/10.1007/s11103-006-9110-2

Huang M, Liu X, Zhou Y, Summers RM, Zhang Z (2019) BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. GigaScience 8:giy154. https://doi.org/10.1093/gigascience/giy154

Hufnagel B, Marques A, Soriano A, Marquès L, Divol F, Doumas P, Sallet E, Mancinotti D, Carrere S, Marande W (2020) High-quality genome sequence of white lupin provides insight into soil exploration and seed quality. Nat Commun 11:1–12. https://doi.org/10.1038/s41467-019-14197-9

Hufnagel B, Soriano A, Taylor J, Divol F, Kroc M, Sanders H, Yeheyis L, Nelson M, Péret B (2021) Pangenome of white lupin provides insights into the diversity of the species. Plant Biotechnol J. https://doi.org/10.1111/pbi.13678

Jacob I, Feuerstein U, Heinz M, Schott M, Urbatzka P (2017) Evaluation of new breeding lines of white lupin with improved resistance to anthracnose. Euphytica 213:236. https://doi.org/10.1007/s10681-017-2011-4

Jeger M, Viljanen-Rollinson S (2001) The use of the area under the disease-progress curve (AUDPC) to assess quantitative disease resistance in crop cultivars. Theor Appl Genet 102:32–40. https://doi.org/10.1007/s001220051615

Kamber T, Malpica-López N, Messmer MM, Oberhänsli T, Arncken C, Alkemade JA, Hohmann P (2021) A qPCR assay for the fast detection and quantification of *Colletotrichum lupini*. Plants 10:1548. https://doi.org/10.3390/plants10081548

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S-y, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. Nat Genet 42:348–354. https://doi.org/10.1038/ng.548

Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E (2008) Efficient control of population structure in model organism association mapping. Genetics 178:1709–1723. https://doi.org/10.1534/genetics.107.080101

Kim JH, Kim WT (2013) The Arabidopsis RING E3 ubiquitin ligase AtAIRP3/LOG2 participates in positive regulation of high-salt and drought stress responses. Plant Physiol 162:1733–1749. https://doi.org/10.1104/pp.113.220103

Kroc M, Rybiński W, Wilczura P, Kamel K, Kaczmarek Z, Barzyk P, Święcicki W (2017) Quantitative and qualitative analysis of alkaloids composition in the seeds of a white lupin (*Lupinus albus* L.) collection. Genet Resour Crop Evol 64:1853–1860. https://doi.org/10.1007/s10722-016-0473-1

Książkiewicz M, Nazzicari N, Ha Y, Nelson MN, Renshaw D, Rychel S, Ferrari B, Carelli M, Tomaszewska M, Stawiński S (2017) A high-density consensus linkage map of white lupin highlights synteny with narrow-leafed lupin and provides markers tagging key agronomic traits. Sci Rep 7:15335. https://doi.org/10.1038/s41598-017-15625-w

Kuznetsova A, Brockhoff PB, Christensen RH (2017) lmerTest package: tests in linear mixed effects models. J Stat Softw 82:1–26. https://doi.org/10.18637/JSS.V082.I13

Laido G, Marone D, Russo MA, Colecchia SA, Mastrangelo AM, De Vita P, Papa R (2014) Linkage disequilibrium and genome-wide association mapping in tetraploid wheat (*Triticum turgidum* L.). PLoS ONE 9:e95211. https://doi.org/10.1371/journal.pone.0095211

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. https://doi.org/10.3732/ajb.1200474

Lee DH, Choi HW, Hwang BK (2011) The pepper E3 ubiquitin ligase RING1 gene, *CaRING1*, is required for cell death and the salicylic acid-dependent defense response. Plant Physiol 156:2011–2025. https://doi.org/10.1104/pp.111.177568

Lee J-H, Kim WT (2011) Regulation of abiotic stress signal transduction by E3 ubiquitin ligases in Arabidopsis. Mol Cells 31:201–208. https://doi.org/10.1007/s10059-011-0031-9

Lenth R, Singmann H, Love J, Buerkner P, Herve M (2019) Emmeans: estimated marginal means, aka least-squares means (Version 1.3.4). https://CRAN.R-project.org/package=emmeans. (20 September 2020)

Letunic I, Bork P (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23:127–128. https://doi.org/10.1093/bioinformatics/btl529

Liu HJ, Yan J (2019) Crop genome-wide association study: a harvest of biological relevance. Plant J 97:8–18. https://doi.org/10.1111/tpj.14139

Liu X, Gu X, Lu H, Liu P, Miao H, Bai Y, Zhang S (2021) Identification of novel loci and candidate genes for resistance to powdery mildew in a resequenced cucumber germplasm. Genes 12:584. https://doi.org/10.3390/genes12040584

LPWG (2017) A new subfamily classification of the *Leguminosae* based on a taxonomically comprehensive phylogeny. Taxon 66:44–77. https://doi.org/10.12705/661.3

Lucas MM, Stoddard FL, Annicchiarico P, Frias J, Martinez-Villaluenga C, Sussmann D, Duranti M, Seger A, Zander PM, Pueyo JJ (2015) The future of lupin as a protein crop in Europe. Front Plant Sci 6:705. https://doi.org/10.3389/fpls.2015.00705

Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, Derks EM (2018) A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. Int J Methods Psychiatr Res 27:e1608. https://doi.org/10.1002/mpr.1608

Marino D, Peeters N, Rivas S (2012) Ubiquitination during plant immune signaling. Plant Physiol 160:15–27. https://doi.org/10.1104/pp.112.199281

Mousavi-Derazmahalleh M, Bayer PE, Nevado B, Hurgobin B, Filatov D, Kilian A, Kamphuis LG, Singh KB, Berger JD, Hane JK (2018a) Exploring the genetic and adaptive diversity of a pan-Mediterranean crop wild relative: narrow-leafed lupin. Theor Appl Genet 131:887–901. https://doi.org/10.1007/s00122-017-3045-7

Mousavi-Derazmahalleh M, Nevado B, Bayer PE, Filatov DA, Hane JK, Edwards D, Erskine W, Nelson MN (2018b) The western

Mediterranean region provided the founder population of domesticated narrow-leafed lupin. Theor Appl Genet 131:2543–2554. https://doi.org/10.1007/s00122-018-3171-x

Murray KD, Borevitz JO (2018) Axe: rapid, competitive sequence read demultiplexing using a trie. Bioinformatics 34:3924–3925. https://doi.org/10.1093/bioinformatics/bty432

Murtagh F, Legendre P (2014) Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? J Classif 31:274–295. https://doi.org/10.1007/s00357-014-9161-z

Nirenberg HI, Feiler U, Hagedorn G (2002) Description of *Colletotrichum lupini* comb. nov. in modern terms. Mycologia 94:307–320. https://doi.org/10.1080/15572536.2003.11833238

Padmarasu S, Sargent D, Patocchi A, Troggio M, Baldi P, Linsmith G, Poles L, Jänsch M, Kellerhals M, Tartarini S, Velasco R (2018) Identification of a leucine-rich repeat receptor-like serine/threonine-protein kinase as a candidate gene for *Rvi12 (Vb)*-based apple scab resistance. Mol Breed 38:1–14. https://doi.org/10.1007/s11032-018-0825-y

Paradis E, Schliep K (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528. https://doi.org/10.1093/bioinformatics/bty633

Park CH, Shirsekar G, Bellizzi M, Chen S, Songkumarn P, Xie X, Shi X, Ning Y, Zhou B, Suttiviriya P, Wang M, Umemura K, Wang G-L (2016) The E3 ligase APIP10 connects the effector AvrPiz-t to the NLR receptor Piz-t in rice. PLoS Path 12:e1005529. https://doi.org/10.1371/journal.ppat.1005529

Pavan S, Delvento C, Ricciardi L, Lotti C, Ciani E, D'Agostino N (2020) Recommendations for choosing the genotyping method and best practices for quality control in crop genome-wide association studies. Front Genet. https://doi.org/10.3389/fgene.2020.00447

Pecchia S, Caggiano B, Da Lio D, Cafà G, Le Floch G, Baroncelli R (2019) Molecular detection of the seed-borne pathogen *Colletotrichum lupini* targeting the hyper-variable IGS region of the ribosomal cluster. Plants 8:222. https://doi.org/10.3390/plants8070222

Peix A, Ramírez-Bahena MH, Flores-Félix JD, de la Vega PA, Rivas R, Mateos PF, Igual JM, Martínez-Molina E, Trujillo ME, Velázquez E (2015) Revision of the taxonomic status of the species *Rhizobium lupini* and reclassification as *Bradyrhizobium lupini* comb. nov. Int J Syst Evol Microbiol 65:1213–1219. https://doi.org/10.1099/ijs.0.000082

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. https://doi.org/10.1093/dnares/dsm009

Puritz JB, Hollenbeck CM, Gold JR (2014) dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ 2:e431. https://doi.org/10.7717/peerj.431

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Raman R, Cowley RB, Raman H, Luckett DJ (2014) Analyses using SSR and DArT molecular markers reveal that ethiopian accessions of white lupin (*Lupinus albus* L.) represent a unique genepool. Open J Genet 04:87–98. https://doi.org/10.4236/ojgen.2014.42012

Rincent R, Moreau L, Monod H, Kuhn E, Melchinger AE, Malvar RA, Moreno-Gonzalez J, Nicolas S, Madur D, Combes V, Dumas F, Altmann T, Brunel D, Ouzunova M, Flament P, Dubreuil P, Charcosset A, Mary-Huard T (2014) Recovering power in association mapping panels with variable levels of linkage disequilibrium. Genetics 197:375–387. https://doi.org/10.1534/genetics.113.159731

Sakai T, Mochizuki S, Haga K, Uehara Y, Suzuki A, Harada A, Wada T, Ishiguro S, Okada K (2012) The wavy growth 3 E3 ligase family controls the gravitropic response in *Arabidopsis* roots. Plant J 70:303–314. https://doi.org/10.1111/j.1365-313X.2011.04870.x

Shu K, Yang W (2017) E3 ubiquitin ligases: ubiquitous actors in plant development and abiotic stress responses. Plant Cell Physiol 58:1461–1476. https://doi.org/10.1093/pcp/pcx071

Song D, Li G, Song F, Zheng Z (2008) Molecular characterization and expression analysis of *OsBISERK1*, a gene encoding a leucine-rich repeat receptor-like kinase, during disease resistance responses in rice. Mol Biol Rep 35:275–283. https://doi.org/10.1007/s11033-007-9080-8

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proc Natl Acad Sci USA 100:9440–9445. https://doi.org/10.1073/pnas.1530509100

Suzuki R, Shimodaira H (2006) Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics 22:1540–1542. https://doi.org/10.1093/bioinformatics/btl117

Talhinhas P, Baroncelli R, Le Floch G (2016) Anthracnose of lupins caused by *Colletotrichum lupini*: a recent disease and a successful worldwide pathogen. J Plant Pathol 98:5–14

Tang D, Wang G, Zhou J-M (2017) Receptor kinases in plant-pathogen interactions: more than pattern recognition. Plant Cell 29:618–637. https://doi.org/10.1105/tpc.16.00891

Thomas G, Sweetingham M (2004) Cultivar and environment influence the development of lupin anthracnose caused by *Colletotrichum lupini*. Australas Plant Pathol 33:571–577. https://doi.org/10.1071/AP04060

Thomas G, Sweetingham M, Adcock K (2008) Application of fungicides to reduce yield loss in anthracnose-infected lupins. Crop Protect 27:1071–1077. https://doi.org/10.1016/j.cropro.2007.12.012

Tibbs Cortes L, Zhang Z, Yu J (2021) Status and prospects of genome-wide association studies in plants. The Plant Genome 14:e20077. https://doi.org/10.1002/tpg2.20077

Toker C (2004) Estimates of broad-sense heritability for seed yield and yield criteria in faba bean (*Vicia faba* L.). Hereditas 140:222–225. https://doi.org/10.1111/j.1601-5223.2004.01780.x

van Rossum B-J, Kruijer W, van Eeuwijk F, Boer M, Malosetti M, Bustos-Korts D, Millet E, Paulo J, Verouden M (2020) Package 'statgenGWAS'. https://CRAN.R-project.org/package=statgenGWAS

von Baer E, von Baer I, Riegel R (2009) Pecosa-Baer: a new cultivar of white lupin with determined bushy growth habit, sweet grain and high protein content. Chil J Agric Res 69:577–580. https://doi.org/10.4067/S0718-58392009000400013

Wang J, Zhang Z (2021) GAPIT Version 3: boosting power and accuracy for genomic association and prediction. Genom Proteom Bioinform. https://doi.org/10.1016/j.gpb.2021.08.005

Wen Z, Tan R, Zhang S, Collins PJ, Yuan J, Du W, Gu C, Ou S, Song Q, An YQC, Boyse J, Chilvers M, Wang D (2018) Integrating GWAS and gene expression data for functional characterization of resistance to white mould in soya bean. Plant Biotechnol J 16:1825–1835. https://doi.org/10.1111/pbi.12918

White P, French B, McLarty A, Grains Research and Development Corporation (2008) Producing lupins. Department of Agriculture and Food, Perth, Western Australia

Wolko B, Clements JC, Naganowska B, Nelson MN, Yang Ha (2011) *Lupinus*. In: Kole C (ed) Wild Crop Relatives: Genomic and Breeding Resources: Legume Crops and Forages. Springer, Berlin, Heidelberg, pp 153–206. https://doi.org/10.1007/978-3-642-14387-8_9

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. https://doi.org/10.1023/B:MOLB.0000038003.49638.97

Yang H, Lin R, Renshaw D, Li C, Adhikari K, Thomas G, Buirchell B, Sweetingham M, Yan G (2010) Development of sequence-specific PCR markers associated with a polygenic controlled trait for marker-assisted selection using a modified selective genotyping strategy: a case study on anthracnose disease resistance in white lupin (*Lupinus albus* L.). Mol Breed 25:239–249. https://doi.org/10.1007/s11032-009-9325-4

Yang H, Renshaw D, Thomas G, Buirchell B, Sweetingham M (2008) A strategy to develop molecular markers applicable to a wide range of crosses for marker assisted selection in plant breeding: a case study on anthracnose disease resistance in lupin (*Lupinus angustifolius* L.). Mol Breed 21:473–483. https://doi.org/10.1007/s11032-007-9146-2

Zhou B, Zeng L (2017) Conventional and unconventional ubiquitination in plant immunity. Mol Plant Pathol 18:1313–1330. https://doi.org/10.1111/mpp.12521

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.