


## RESEARCH ARTICLE

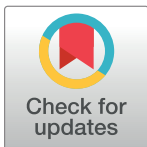
# Construction of a high-density interspecific (*Lens culinaris* x *L. odemensis*) genetic map based on functional markers for mapping morphological and agronomical traits, and QTLs affecting resistance to *Ascochyta* in lentil

Carlos Polanco <sup>\*</sup>, Luis Enrique Sáenz de Miera <sup>\*</sup>, Ana Isabel González, Pedro García, Richard Fratini, Francisca Vaquero, Francisco Javier Vences, Marcelino Pérez de la Vega

Área de Genética, Departamento de Biología Molecular, Universidad de León, León, Spain

 These authors contributed equally to this work.

\* [carlos.polanco@unileon.es](mailto:carlos.polanco@unileon.es)



## OPEN ACCESS

**Citation:** Polanco C, Sáenz de Miera LE, González AI, García P, Fratini R, Vaquero F, et al. (2019) Construction of a high-density interspecific (*Lens culinaris* x *L. odemensis*) genetic map based on functional markers for mapping morphological and agronomical traits, and QTLs affecting resistance to *Ascochyta* in lentil. PLoS ONE 14(3): e0214409. <https://doi.org/10.1371/journal.pone.0214409>

**Editor:** Roberto Papa, Università Politecnica delle Marche, ITALY

**Received:** November 16, 2018

**Accepted:** March 12, 2019

**Published:** March 27, 2019

**Copyright:** © 2019 Polanco et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are available within the paper and its Supporting Information files. The nucleotide sequences obtained in this work have been submitted to the NCBI Sequence Read Archive with accession numbers from SRR8608157 to SRR8608168 (*L. culinaris* cv. Alpo biological replicates), from SRR8608778 to SRR8608789 (*L. odemensis* ILWL235 biological replicates), and from SRR8613937 to SRR8614014 (RIL population).

## Abstract

Usage of high-throughput sequencing approaches allow for the generation and characterization of reference transcriptome datasets that support gene-based marker discovery, which in turn can be used to build genetic maps among other purposes. We have obtained a transcriptome assembly including 49,453 genes for the lentil (*Lens culinaris* Medik.) cultivar Alpo using RNAseq methodology. This transcriptome was used as reference to obtain 6,306 quality polymorphic markers (SNPs and short indels) analyzing genotype data from a RIL population at F<sub>7</sub> generation derived from the interspecific cross between *L. culinaris* cv. Alpo and *L. odemensis* accession ILWL235. *L. odemensis* is a wild species included in the secondary gene pool and can be used as a source for gene introgression in lentil breeding programs. Marker data were used to construct the first genetic interspecific map between these two species. This linkage map has been used to precisely identify regions of the CDC-Redberry lentil draft genome in which the candidate genes for some qualitative traits (seed coat spotting pattern, flower color, and stem pigmentation) could be located. The genome regions corresponding to a significant single quantitative trait locus (QTL) controlling “time to flowering” located in chromosome 6 and three QTLs regulating seed size and positioned in chromosomes 1 and 5 (two QTLs) were also identified. Significant QTLs for *Ascochyta* blight resistance in lentil were mapped to chromosome 6 in the genome region or close to it where QTLs for *Ascochyta* blight resistance have previously been reported.

## Introduction

In agriculture, the main goal of hybridization between crops and their wild relatives (wide-crosses) is the introgression of adaptive traits from the wild relatives into the cultivated forms as part of breeding programs [1]. Wide-crosses have extensively been used in plant breeding to

The lentil draft genome assembly v1.2 can be accessed through the JBrowse or BLAST function at <http://knowpulse.usask.ca>.

**Funding:** This project was financed by the Consejería de Educación, Junta de Castilla y León, Spain (grant LE005G18) and the Spanish Ministerio de Economía y Competitividad (grant AGL2013-44714-R, co-financed with FEDER funds).

**Competing interests:** The authors have declared that no competing interests exist.

broaden the gene pool of many crop species and to transfer desirable particular genes, mainly resistance genes to biotic or abiotic stresses. Response to biotic stresses is the main reason alleged to use wild crop relatives in breeding programs among the 10 main crops [2]. Wild germplasm provides the raw material used for resistance breeding in what is designated as the “smart breeding” strategy which facilitates exploration and utilization of the available natural genetic variation, thus expanding the genetic base of crop plants while providing more adaptive flexibility with respect to the future [3,4]. Wide-crosses implicate not only the primary gene pool, that is the wild form of the cultigen together with other species producing fertile hybrids, but also the secondary and tertiary gene pools. Wide-crosses have largely been used in cereal breeding [5–9], primarily in wheat (resistance has been introgressed from at least 52 species from 13 genera [10]), barley and rice, and in the dicot species tomato [11]. Although less extensively used than in cereal species, wide-crosses have also been used in breeding legume species [12–15]. Usually, wide-hybridization depends upon embryo rescue techniques to obtain interspecific hybrids since specific isolation often results in weak or completely unviable embryos. Thus, availability of efficient embryo rescue protocols based on *in vitro* culture techniques remains a bottleneck to systematic use of interspecific hybridization applied to breeding purposes [16].

Lentil (*Lens culinaris* Medik. subsp. *culinaris*) is a diploid ( $2n = 14$ ), self-pollinating annual cool season grain legume normally grown in temperate semi-arid regions, usually in rotation with cereals. Lentils are traditionally valued as a source of energy, proteins and iron in human nutrition. In addition, they are an important dietary source of fiber, minerals, vitamins and antioxidants. Likewise, lentils contribute to improve soil quality by replenishing soil nitrogen levels and can play a significant role in environmentally sustainable agricultural systems. The crop is now extensively cultivated throughout Western Asia, Northern Africa, the Indian sub-continent, Australia and North America, in particular in Canada as the current first world producer and exporter [17,18]. Wild lentil relatives are being broadly used to develop genetic maps, to analyze quantitative trait loci (QTLs) related to yield and other plant characters, as a source of resistance genes to both biotic and abiotic stresses, and in general to broadening the genetic base of lentil cultivars [19–29]. Singh et al. [30] have recently reviewed and summarized the available wild *Lens* germplasm, the characterization and evaluation of wild species for useful traits of interest to solve production-related problems, as well as the progress achieved to widen the genetic base of the cultivated lentil varieties through wide hybridization and explorative lentil genomics.

Interspecific *Lens* seeds most often fail to reach full development or else are unable to germinate, therefore *in vitro* embryo-rescue techniques are usually required to obtain hybrid plants between *Lens* species [31,32]. In a genotype-by-sequencing (GBS) characterization of *Lens* species, Wong et al. [33] included in the primary gene pool, in addition to the lentil cultigen, the wild ancestor *L. c.* subsp. *orientalis* (Boiss.) Ponert and the species *L. tomentosus* Ladiz., the secondary gene pool included *L. odemensis* Ladiz. and *L. lamottei* Czefr., in the tertiary *L. ervoides* Grande, and in the quaternary *L. nigricans* (M. Bieb.) Godr., while Ladizinsky and Abbo [34] included *L. odemensis*, *L. ervoides* and *L. tomentosus* in the secondary gene pool, and *L. nigricans* and *L. lamottei* in the tertiary.

Genetic maps based on recombination frequency distances between genes are classical elements of the formal genetic analysis, allowing researchers to locate genome regions of interest, hence remaining key breeding tools. The new high throughput technologies in genomics and transcriptomics provide an unprecedented level of insight into the structural diversity across crop genomes and afford thousands of genetic markers, therefore allowing to obtain highly saturated genetic maps providing a better location of QTLs and other genes of interest. Genetic maps based on markers derived from high throughput sequencing techniques have been

developed in lentil [35,36,19,37,38,24]. Several reviews on lentil genetic maps obtained from morphological to molecular markers have been published [18,36]. Excluding intersubspecific maps between the cultivated lentil and its wild ancestor (*L. c.* subsp. *orientalis*), true interspecific genetic maps between the cultivated lentil and other *Lens* species have so far been limited to *L. ervoides* [24].

Introgressed chromosome segments from alien species often include favorable genes that result in selection, structural rearrangements that hinder or prevent recombination in the affected chromosomal segments, or both. As a consequence, sets of relatively close alleles from the same parent tend to remain together from generation to generation and frequently lead to segregation distortion; which in turn can decrease the capacity to determine correct marker orders in genetic linkage maps while impeding accurate QTL detection [39]. Different allele selective values generate linkage drag, the unintentional co-selection of undesirable gene variants that are closely linked to selected loci of interest [40]. High-resolution genome analysis technologies enable description of crop diversity at an unprecedented chromosome-scale resolution and they are highly suited for genomic selection strategies. Likewise, genome-wide diversity data sets enable the dissection of linkage disequilibrium to characterize loci underlying selective sweeps. On the other hand, high-density genome data may help to increase the use of distant gene pools by enabling more efficient genomic selection strategies that efficiently predict performance based on genome wide marker combinations [40].

Chromosomal rearrangements between *Lens* species have been described based on karyotype and genetic mapping analyses [41–43]. Differences between *Lens* species karyotypes using fluorescent in situ hybridization (FISH) techniques have been identified [44–46]. Furthermore, Galasso [46] using highly repeated DNA sequences established a typical FISH karyotype for each *Lens* species with very little variation within species, except in the case of *L. odemensis* in which a karyotype polymorphism was found.

We present here the first genetic interspecific map between *L. culinaris* and *L. odemensis*, from the primary and the secondary gene pools [33] respectively, based on 6,306 SNPs and short indels. Markers located in expressed genes that were genotyped in both parents and in an F<sub>7</sub> RIL population which included a total of 78 lines. The genetic map was compared with the draft assembly of the *L. culinaris* (cv. CDC-Redberry) genome [47] and was used for mapping morphological and agronomical traits as well as QTLs related to resistance against the pathogen *Ascochyta lentis*.

## Materials and methods

### Plant and fungal materials

A bi-parental RIL population was developed from a single F<sub>1</sub> hybrid plant derived from an interspecific cross between *Lens culinaris* cultivar Alpo and *L. odemensis* accession ILWL235, followed by single seed descent from the F<sub>2</sub> population (120 F<sub>2</sub> seeds were sown). A total number of 78 surviving RILs at the F<sub>7</sub> generation were genotyped and phenotyped, as well as the two parental accessions. The *L. culinaris* cultivar Lupa was used as the susceptible control for inoculations with *Ascochyta lentis*.

A suspension of the *A. lentis* isolate AL-84 in a concentration of  $1 \times 10^6$  conidia ml<sup>-1</sup> was used to spray-inoculate different replicates of three 15 day-old plantlets grown together under sterile conditions in capped glass tubes. Nine replicates of each of the two parental accessions and six replicates of each of the 78 RILs were inoculated. Another identical set of replicates per each line or parental sample was treated without the presence of the pathogen. Following 24 hours after the treatment, the aerial parts of the plantlets drawn from six replicates of the pathogen-inoculated and the pathogen-free sets were collected independently for each parental

accession and frozen for total RNA extraction. Three replicates of the pathogen-inoculated plants were collected and pooled for RNA extraction from each of the 78 RIL lines. Total RNA was isolated using the method of Chang et al. [48]. The remaining replicates from the pathogen-inoculated and the pathogen free sets were used to score disease severity. Disease symptoms were checked 14, 21, and 28 days after infection [49]. There were no clear signs of infection after 14 days, either in the susceptible genotypes used to check the infection progression. The plants started to deteriorate after 28 days into capped glass-tubes, so the scores at 21 days were used in further analyses, using the 1–9 categorical scale described by Ford et al. [50].

The mapping population and the parental accessions were also reared under greenhouse conditions. Later they were visually phenotyped for morphological traits such as the presence of seed coat spotting (*Scp*), flower color (purple or white), and anthocyanin pigmentation on the plantlet stem (*Gs*), as well as the quantitative traits seed size in mm<sup>2</sup> and time to flowering (scored as early, medium or late flowering). The average seed size was obtained from scanned images of 30 seeds using ImageJ software [51], and time to flowering was considered when 50% of the plants within a RIL had at least one flower [52].

### Libraries construction and Illumina sequencing

Total RNA quality and quantity of the samples to be sequenced was determined in the Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) and the Qubit 3.0 fluorometer (Thermo Fisher Scientific, Wilmington, DE, USA). The poly(A)+ mRNA fraction was isolated from the total RNA and strand-specific cDNA libraries were obtained following Illumina's recommendations. Briefly, poly(A)+ RNA was isolated on poly-T oligo-attached magnetic beads and chemically fragmented prior to reverse transcription and cDNA generation. The cDNA fragments then went through an end repair process, the addition of a single adenine base to the 3' end followed by ligation of the adapters. Finally, the products were purified and enriched using PCR to create the indexed final double stranded and strand-specific cDNA libraries. The quality of the libraries was analyzed by the High Sensitivity RNA assay in a 4200 TapeStation (Agilent Technologies) and the quantity of the libraries was determined by real-time PCR using a LightCycler 480 (Roche Applied Science, Penzberg, Germany) according to the manufacturer's protocols.

Twenty-four cDNA libraries were obtained using RNA from each of six replicates of both the inoculated and the non-inoculated plants from the two parental accessions. In addition, 78 cDNA libraries were obtained, one originating from each RIL line, using the pooled RNA from three inoculated replicates. Equimolar pools of cDNA libraries were sequenced by paired-end sequencing (100 x 2) in an Illumina HiSeq 2500 sequencer (Illumina Inc., San Diego, CA, USA).

### Transcriptome assembly, variant calling, and marker filtering

The quality of the raw sequence reads was assessed using the FASTQC tool [53]. The reads were filtered using Trimmomatic [54] with a sliding window size of 4 and leading and trailing parameters were set at a quality of 28. All the filtered reads obtained from the Alpo cultivar libraries were mapped after normalization onto the *L. culinaris* cv. CDC Redberry v1.2 genome assembly (<http://knowpulse.usask.ca>) using the STAR aligner [55]. A coordinate-sorted bam file was obtained using the SAMtools suite [56], which was used to generate a genome-guided *de novo* transcriptome assembly of the Alpo cultivar using the Trinity package [57] using the lentil v1.2 genome as reference. Duplicate and near identical contigs in the assembly were eliminated using the Dedupe BBTool from the Joint Genome Institute (University of California) with parameters  $e = 10$  and  $\text{minidentity} = 95$  in separate steps. The Magic-BLAST tool

available online at the NCBI webpage was used to map each transcript onto the pseudomolecules (chromosomes) and contigs of the lentil v1.2 genome assembly.

Before SNPs and short indels calling, the abundance of the ALPO transcripts was quantified using the Kallisto program [58] in order to filter the transcripts to exclude isoforms that were lowly expressed. A Trinity script was used to retain the single most highly expressed isoform for each gene. Such transcripts were used as the reference for variant calling using SAMtools mpileup [56]. A single variant call file (VCF) file was generated from the 102 bam files that were previously obtained using the STAR aligner in order to map independently the filtered reads derived from each library onto the reference Alpo transcriptome.

The VCF variants were filtered for read depth and low quality using the mpileup software filter options. A custom Perl script was then used to retain high-quality variants analysing the genotype QUAL scores and to remove multiallelic variants, missing calls for 28 or more RILs, variants showing different genotypes in the parental replicates, or fixed variants in the RILs. Finally, a single VCF variant was selected for each transcript of the ALPO reference transcriptome.

### Genetic linkage mapping, comparative mapping, and QTL detection

The high-quality markers retained were used to construct a genetic linkage map using the MSTMap software tool [59], which determines the correct marker order by computing the minimum spanning tree of an associated graph, making use of the following settings: Kosambi's distance function, cut-off p value =  $1 \times 10^{-12}$  for clustering the markers into linkage groups, 15 cM gap size, and the maximum likelihood objective function.

The locations of the markers in the physical (lentil v1.2 genome assembly) and genetic maps were compared and graphically represented using custom R scripts [60]. A circular comparative visualization of both maps was obtained using the R/Shiny application shinyCircos [61].

The interactive environment R/qtl [62] was used for mapping quantitative trait loci (QTL) onto the genetic linkage map by making use of Haley-Knott regression [63]. Genome-wide LOD significance thresholds were determined by a permutation test with 1,000 iterations at a significance level of 0.05. Confidence intervals were calculated using a 1.8-LOD drop method [64].

## Results

### Reference transcriptome, variant discovery, and segregation distortion

Sequencing of the 12 cDNA libraries from *L. culinaris* cv. Alpo yielded a total of 242,680,763 100-bp paired-end reads after the quality filtering. The normalization step retained 16,452,752 reads with an average input length of 200 bp, which were mapped onto the *L. culinaris* cv. CDC Redberry v1.2 genome assembly. The percentages of uniquely mapped, multimapped, and unmapped reads were 87.1%, 3.3%, and 9.6%, respectively. The normalized reads were *de novo* assembled using such mapping information, following the genome-guided method of the Trinity package, generating 61,194 'genes' (Trinity's definition) and 120,879 transcripts with an N50 of 2,521 bp. Duplicate and near identical contigs were removed in order to obtain a transcriptome from cv. ALPO that included 49,453 Trinity 'genes' and 102,892 transcripts with an N50 of 2,501 bp, an average length of 1,748.4 bp and a total of 179,931,148 assembled bases. When using only the longest isoform per gene the statistics amounted to an N50 of 2,219 bp, an average length of 1,308.2 bp and a total of 64,693,989 assembled bases. A total of 96.2% of the initially filtered reads aligned as proper pairs into the ALPO transcriptome. Transcript abundance was quantified and the single most highly expressed isoform for each gene

was retained in order to obtain the reference transcriptome, containing 49,453 assembled sequences (S1 File) that were used later for the variant calling procedure.

A total number of 1,139,915 variants (SNPs and short indels) were initially identified, and 120,483 were retained as quality polymorphic markers in the mapping population after undertaking the filtration applying the parameters described in the materials and methods section. These markers mapped onto 15,506 expressed genes. A single variant was chosen for each of them, applying the criteria of best adjustment to the expected 1:1 segregation considering the RIL genotypes. Most of these expressed sequences (13,128) were evenly distributed into the seven pseudomolecules (named as LcChr#) of the lentil v1.2 genome assembly, which correspond to the seven lentil chromosomes (Table 1; Fig 1). Another 1,996 sequences mapped onto minor contigs not yet assembled to the pseudomolecules and 382 sequences were not located in the lentil v1.2 genome assembly. Large regions were detected in LcChr2, LcChr4, and LcChr7 in which most of the marker frequencies scored near to the maximum for the allele derived from cv. ALPO (Fig 1). A higher than expected level of residual heterozygosity (average of 5.37%) was detected for the F<sub>7</sub> interspecific RIL population (1.56%).

The chi-square test revealed that 5,676 markers did not deviate significantly ( $p > 0.05$ ) from the expected Mendelian 1:1 ratio regarding parental allele contribution. The number of markers selected to construct the linkage map was increased up to 6,306 when the distorted markers that did not deviate significantly at  $p > 0.01$  were included. These 6,306 markers were located on all of the seven chromosomes (pseudomolecules) of the lentil v1.2 genome assembly yet were unevenly distributed between and within them (Table 1). High densities of non-distorted markers were found all along of the LcChr3 and LcChr6, while they were concentrated at several regions of LcChr1, LcChr2, and LcChr4 or only at very specific regions of LcChr5 and LcChr7 (Fig 1).

## High-density linkage map and comparative analysis

Data from the 6,306 SNP and short indels markers belonging to expressed sequences across of the 78 RILs were used to construct a high density linkage map (Fig 2). A total of 6,153 markers were mapped onto 10 linkage groups (LGs) while 153 markers remained unlinked to other markers (Table 2). These 6,153 markers grouped into 4,682 unique bins. The linkage groups were named using numbers and consecutive lowercase letters to match the respective lentil

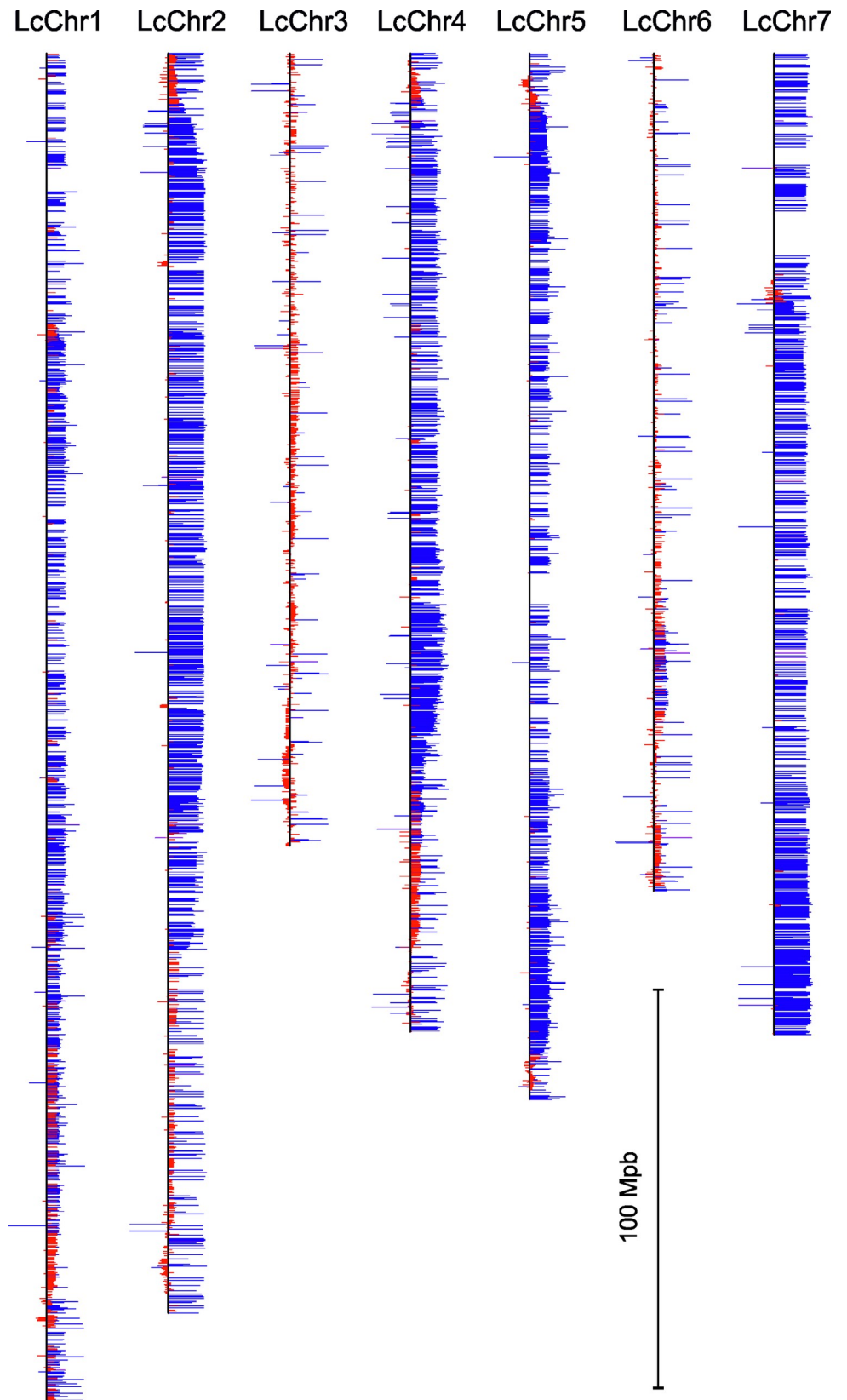
**Table 1. Chromosomal distribution of the functional markers identified.**

Chromosome <sup>1</sup>	Number of markers	Number of non-distorted markers <sup>2</sup>
LcChr1	2,137	688 (32.2)
LcChr2	2,227	797 (35.8)
LcChr3	1,619	1,462 (90.3)
LcChr4	2,014	736 (36.5)
LcChr5	2,089	470 (22.5)
LcChr6	1,284	961 (74.8)
LcChr7	1,758	131 (7.5)
LcContigs	1,996	924 (46.3)
Non-mapped	382	137 (35.8)
Total	15,506	6,306 (40.7)

<sup>1</sup> Lentil genome assembly v1.2

<sup>2</sup> Chi-square test,  $p < 0.01$ ; percentages in parentheses

<https://doi.org/10.1371/journal.pone.0214409.t001>



**Fig 1. Location of the 13,128 polymorphic markers onto the seven pseudomolecules (LcChr#) of the *L. culinaris* cv. CDC Redberry v1.2 genome assembly.** Size of horizontal lines represents the increased frequency deviation from the expected value of 0.5 (at vertical line) of the parental alleles derived from cv. ALPO (to the right) or *L. odemensis* (to the left). Blue horizontal lines denote distorted markers at  $p < 0.01$  (chi-square test) and red horizontal lines denote markers used for the linkage analysis.

<https://doi.org/10.1371/journal.pone.0214409.g001>

chromosomes of the lentil v1.2 assembly as best as possible. Marker name, linkage group and position (cM), reference transcript, and genome chromosome/contig location are provided in the [S2 File](#). Mapping quality was checked using a heat map plotted for all the linkage groups wherein all the 6,153 markers were lined up against each other in order to evaluate the assignment to the linkage groups and the marker order ([Fig 3](#)).

The comparison of the high-density genetic map of the RIL population obtained from the interspecific cross *L. culinaris* x *L. odemensis* against the *L. culinaris* v1.2 genome assembly revealed a high level of collinearity for the location of most of the analyzed functional markers ([Fig 4](#)). Unlinked markers were distributed along all of the seven chromosomes and markers from genes that were not located onto the lentil v1.2 genome were present in different LGs that were assigned to all of the seven chromosomes according to the collinearity shown ([Fig 5](#)).

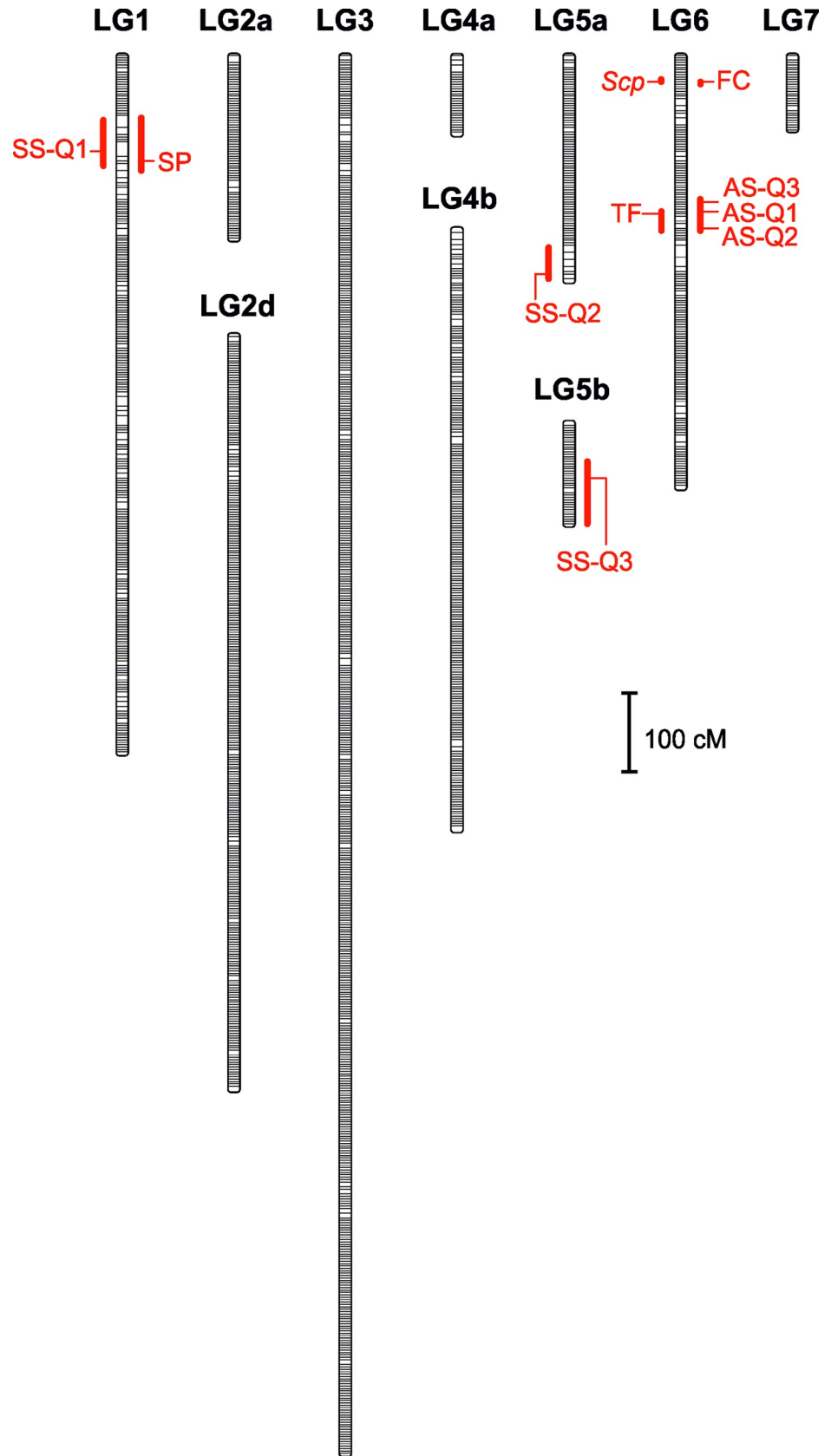
### Mapping morphological and agronomical traits

Two approaches were used to map three qualitative morphological characters: presence of seed coat spotting, flower color, and stem pigmentation. Firstly, by mapping the presence of related QTLs in the above linkage map ([Fig 2](#)) using the ‘scanone’ function of the R/qt software with parameter model = “binary” to locate the binary characters as QTLs, in order to obtain a confidence region where the candidate genes could be located. Secondly, by constructing a new linkage map using the MSTMap software tool considering 6,309 markers by adding the binary encoded phenotypic data of the RIL population for each of the three qualitative morphological markers.

The character presence of seed coat spotting (*Scp*) showed a significant QTL located on LG6 ([Table 3](#)), accounting for 85.07% of the observed variability, with a peak at marker S044140 which was located in the gene Lc25388. This gene encodes a MYB transcription factor MIXTA-like protein which is annotated on LcChr6 at position 11,352,212. The markers included in the 1.8-LOD support interval ([S2 File](#)) belong to genes annotated in the genome between positions 11,352,212 and 14,634,141 of LcChr6, with the exception of gene Lc26377 (see below, as peak marker for flower color). This region of LcChr6 expands 3.3 Mb and currently includes 46 annotated genes. Character *Scp* was flanked by the QTL peak marker S044140 (gene Lc25388) and the S061401 marker (above-mentioned as gene Lc26377), when the obtained linkage map included data of the morphological markers.

The character flower color showed a significant QTL overlapping the one previously described for the *Scp* character on LG6 ([Table 3](#)) which accounted for 84.20% of the observed variability. The support interval included 37 markers distributed in six bins ([S2 File](#)). The corresponding genes for 25 of these markers are annotated in a 7.7 Mb region of the genome in which 106 genes are located between positions 11,352,212 and 19,096,974 of LcChr6 (which includes the complete 3.3 Mb region of the *Scp* QTL), nonetheless the peak marker S061401 is annotated at position 157,341,084 of LcChr6. This marker is located in the Lc26377 gene encoding for a phosphate transporter PHO1-like protein. Other nine markers from bins included in the support interval belong to genes also annotated out of the above range at LcChr6, including an anthocyanin 5-aromatic acyltransferase (gene Lc27646) at position 105,439,424. The trait flower color was flanked by the same markers than *Scp* (see above) in the linkage map obtained which included the data of the morphological markers.





**Fig 2. Linkage map of the RIL population derived from *L. culinaris* cultivar ALPO x *L. odemensis* accession ILWL235.** Identified QTLs are shown as red horizontal lines (LOD peak) with solid vertical bars (with a 1.8-LOD drop confidence interval).

<https://doi.org/10.1371/journal.pone.0214409.g002>

The stem pigmentation character showed a significant QTL located on LG1 (Table 3) which accounted for 33.96% of the observed variability, with a peak at marker S006811 located in the Lc01979 gene (a Sulfite exporter TauE/SafE family protein) part of LcChr1 at position 75,787,823. The support interval included nine markers distributed in nine bins but two markers are currently annotated on LcChr5, one is positioned next to the peak marker, while another is located on LcChr6 of the draft genome assembly. Five out of the six LcChr1 markers are annotated between positions 46,253,696 to 96,580,293. This region of LcChr1 expands 50.3 Mb and counts with a total of 844 currently annotated genes. The character stem pigmentation was not included in any linkage group of the genetic map obtained considering also the data of the morphological markers.

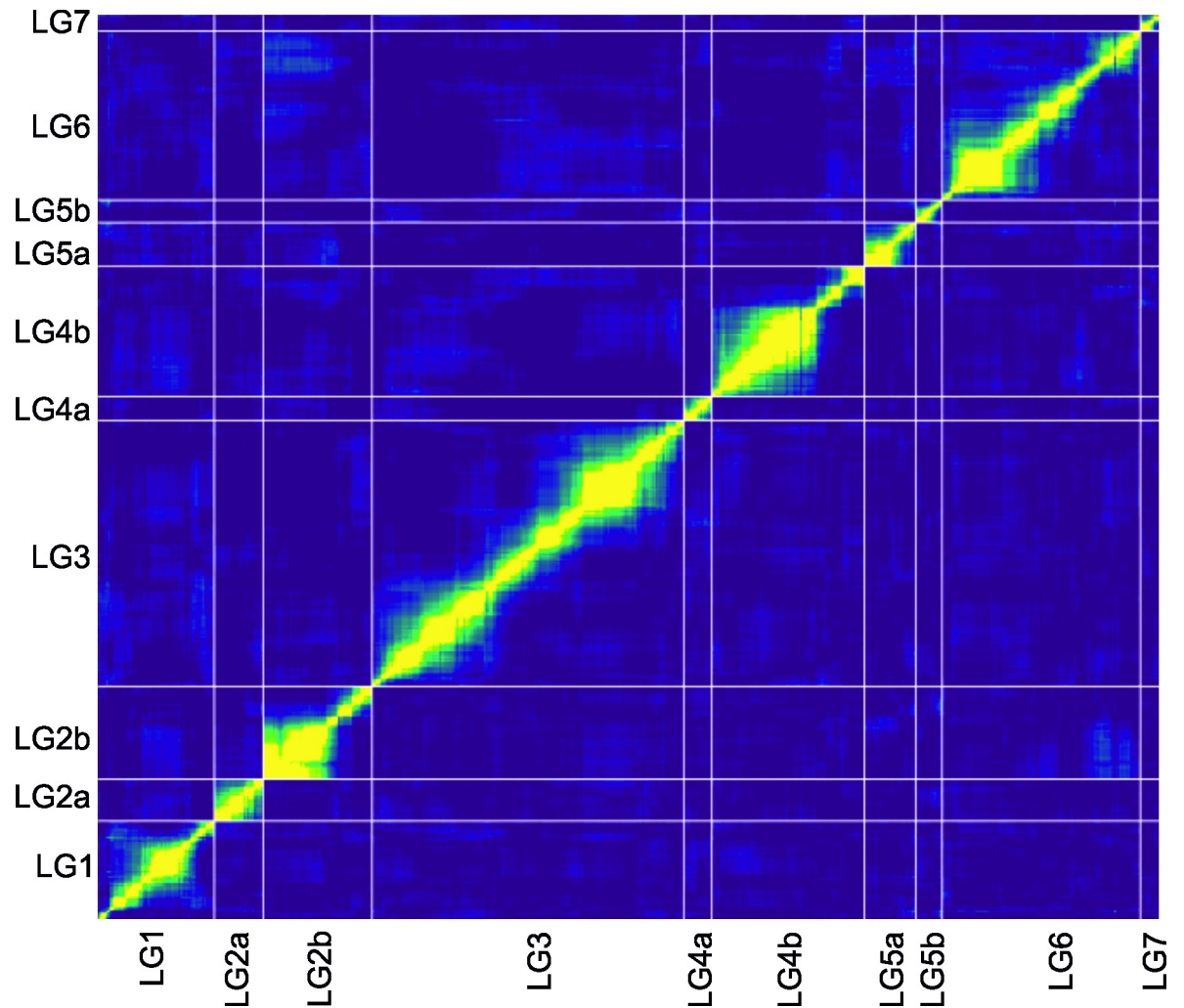
QTLs displaying a significant association with the quantitative agronomical characters time to flowering (TF) and seed size (SS) were identified in the RIL population (Table 3). A single QTL was uncovered for TF on LG6, accounting for 55.73% of the observed variability, comprising a peak between the bins mapped at positions 199.47 cM (two markers from LcChr6 genes annotated between 152,135,041 and 155,170,846) and 203.37 cM (six markers from LcChr6 genes, four of them annotated between 151,824,573 and 155,170,846). The support interval included another seven markers in five bins, with a total of 10 markers annotated in a region of LcChr6 that expands 8.8 Mb (151,824,573 to 160,668,788) in which 157 genes are currently annotated. The *L. odemensis* parental accession contributed positive alleles that increased the time to flowering (Table 3).

Three QTLs accounting globally for 35.48% of the observed variability in seed size were mapped onto LG1 (SS-Q1), LG5a (SS-Q2), and LG5b (SS-Q3) (Table 3). The SS-Q1 support interval fully overlaps with the QTL identified for stem pigmentation at LcChr1 (see above) and its peak was located between markers mapped at positions 104.68 cM and 110.61 cM. The corresponding genes, Lc03317 and Lc01979, are annotated flanking the LcChr1 region spanning from position 75,787,823 to 96,580,293 (20.8 Mb) in which a total of 366 genes are annotated. The SS-Q2 peak was located between the markers mapped to LG5a positions 277.53 cM and 284.76 cM, which correspond to the genes Lc21173 and Lc23619, respectively, flanking the LcChr5 region from position 18,739,851 to 22,169,473 (3.4 Mb) which counts with a total of 118 currently annotated genes. The support interval included another four LcChr5 markers that expand the region up to 31.2 Mb (positions 15,801,723 to 47,019,826) holding 726

**Table 2. Marker distributions and genetic lengths of the linkage groups.**

Linkage group	Number of markers	Number of unique bins	Length (cM)
LG1	678	597	889.49
LG2a; LG2b	283; 630	253; 623	232.78; 959.97
LG3	1,805	1,522	1,785.91
LG4a; LG4b	162; 884	79; 575	97.82; 760.53
LG5a; LG5b	297; 152	254; 122	284.76; 125.59
LG6	1,148	557	549.74
LG7	114	100	95.60
Unlinked	153		
Total	6,306	4,682	5,782.19

<https://doi.org/10.1371/journal.pone.0214409.t002>



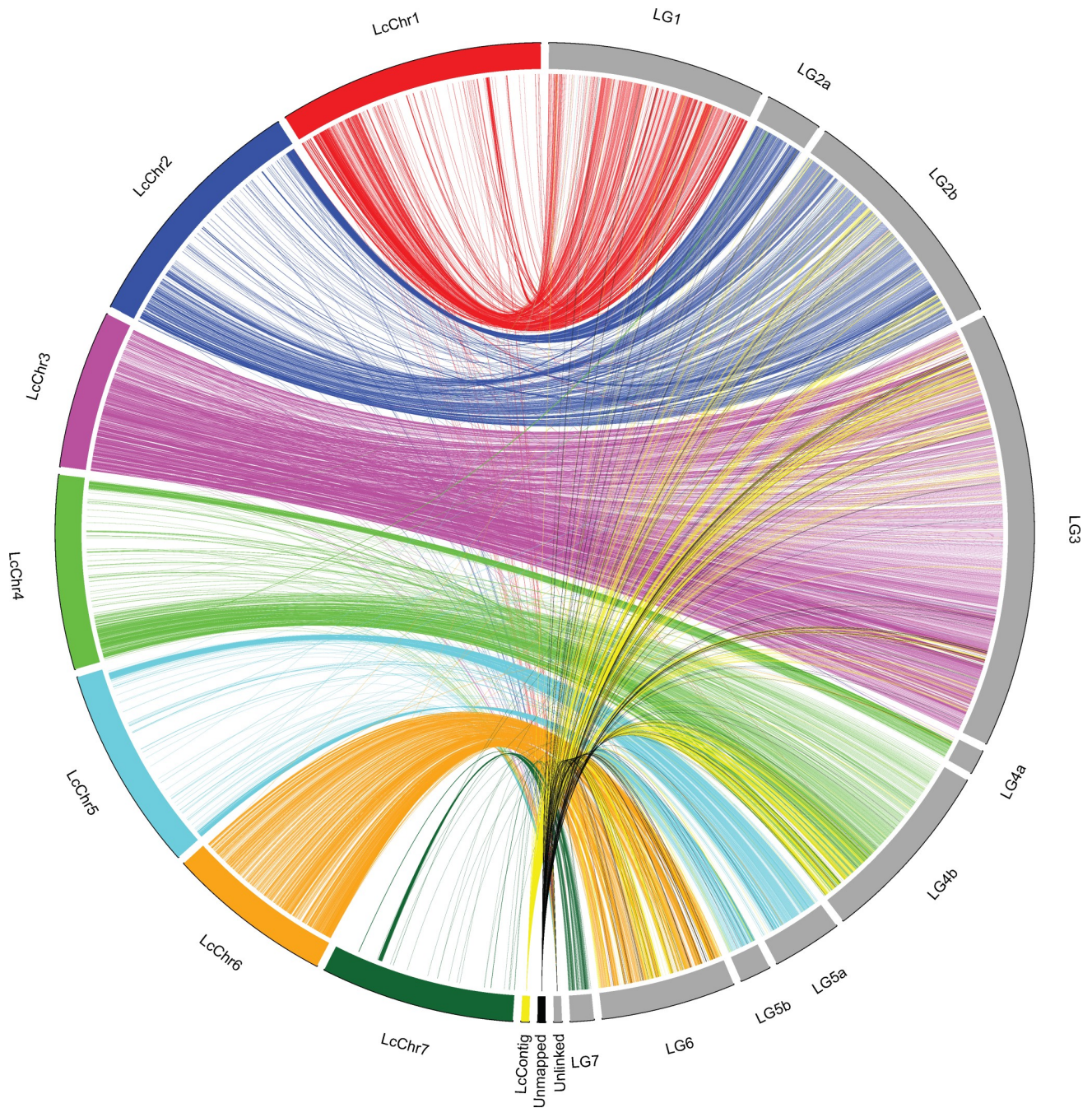
**Fig 3.** Heat plot for all of the linkage groups showing the pairwise comparison of recombination fractions (above of the diagonal) and the LOD score (below of the diagonal) for all of the markers uncovered. High recombination fraction estimates and low LOD scores are shaded in blue, while the warm colors represent the converse.

<https://doi.org/10.1371/journal.pone.0214409.g003>

currently annotated genes. The SS-Q3 peak marker S078816 was located in gene Lc23140, an EF hand protein with potential roles in Ca<sup>2+</sup> signal-mediated processes including growth and development [65], annotated at position 254,820,976 of LcChr5. The SS-Q3 support region included a total of 55 markers annotated in the LcChr5 region between positions 251,623,917 and 257,098,141 (5.5 Mb) with 204 currently annotated genes. The *L. odemensis* parental accession contributed negative alleles to seed size (Table 3).

### Mapping QTLs affecting resistance to *Ascochyta*

Three significant QTLs (AS-Q1, AS-Q2, and AS-Q3) which accounted globally for 28.46% of the observed variability to *A. lentis* resistance in the RIL population, were mapped within a common 1.8-LOD support interval onto LG6 (Table 3). The bin which contained the AS-Q1 peak included two markers from LcChr6 genes that are annotated between 139,108,707 (gene Lc27513, ATP-dependent helicase BRM) and 147,095,064 (gene Lc25909, Plant/F1M20-13 protein). The AS-Q2 peak was located between the marker mapped to position 214.09 cM



**Fig 4. Circle comparative visualization of the markers locations in the lentil v1.2 genome and in the linkage groups of the RIL population genetic map.**

<https://doi.org/10.1371/journal.pone.0214409.g004>

(gene Lc27329, annotated at Chr6: 153,669,223) and the bin at 219.21 cM, which includes another three LcChr6 genes (Lc25078 at 108,049,280; Lc26336 at 152,733,327; Lc27664 at 154,983,742). The AS-Q3 peak was located in a bin that includes three LcChr6 genes at 79,397,962 (Lc25349, homeobox-leucine zipper ROC6), 140,610,179 (Lc25206, Methyltransferase PMT16), and 141,818,776 (Lc26436, uncharacterized protein). The 1.8-LOD support

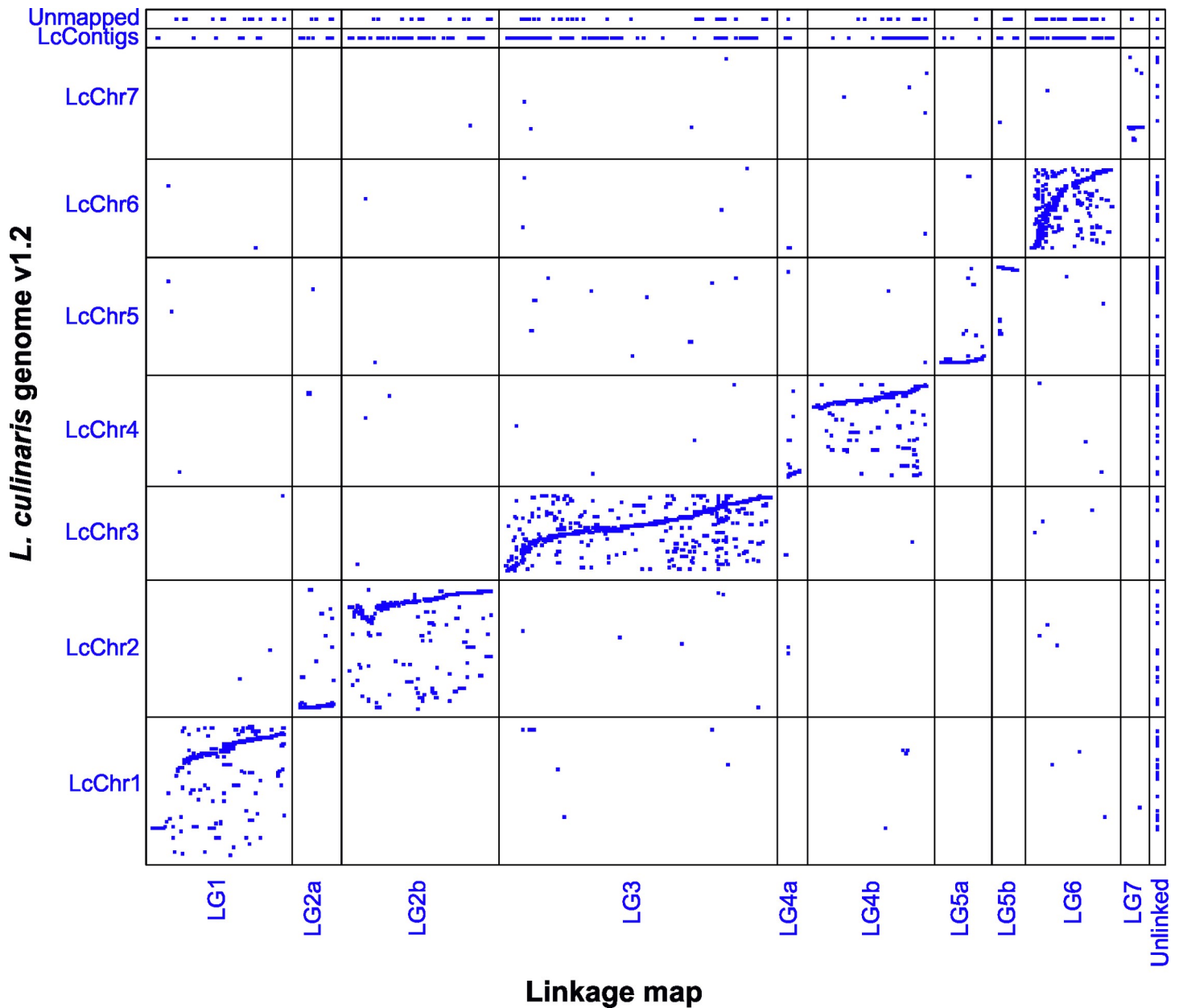


Fig 5. Dot plot representing marker collinearity between the lentil v1.2 genome chromosomes and the linkage groups (LGs) obtained from the genetic map constructed with the interspecific RIL population derived from *L. culinaris* cultivar ALPO x *L. odemensis* accession ILWL235.

<https://doi.org/10.1371/journal.pone.0214409.g005>

region included 22 bins containing a total of 75 markers located in LcChr6 genes plus 14 markers located in contigs not yet assigned to a chromosome. Most of the LcChr6 markers (69) are located in a 28.6Mb region between positions 132,032,337 and 160,668,788 which holds a total of 581 currently annotated genes; the remaining six markers were located in genes annotated in the following positions outside of those limits: 48,803,021, 53,893,775, 53,971,066, 79,267,292, 79,397,962, and 203,277,345. The *L. odemensis* parental accession contributed alleles that increased the resistance to *Ascochyta* infection as the scoring method [50] assigned low values to resistance and high values to susceptibility (Table 3).

**Table 3. List of significant QTLs detected in the interspecific *Lens culinaris* cv. ALPO x *L. odemensis* ILWL235 RIL population.**

Trait	QTL	LG <sup>1</sup>	Peak <sup>2</sup>	LOD	1.8-LOD <sup>3</sup>	Var. <sup>4</sup>	Add. <sup>5</sup>
Seed coat spotting ( <i>Scp</i> ; binary)	<i>Scp</i>	LG6	S044140	17.8	30.0–33.7	85.07	0.46
Flower color (FC; binary)	FC	LG6	S061401	16.5	31.3–34.6	84.20	0.46
Stem pigmentation (Gs; binary)	SP	LG1	S006811	6.0	79.0–145.7	33.96	0.32
Time to flowering (TF; range = 1–3)	TF	LG6	loc201	12.2	199.0–222.0	55.73	0.69
Seed size (SS; mm <sup>2</sup> )	SS-Q1	LG1	loc120	7.0	81.0–141.0	28.26	-1.28
range = 7.381–18.8	SS-Q2	LG5a	loc279	3.6	244.0–284.8	18.63	-0.97
mean = 14.89	SS-Q3	LG5b	S078816	3.2	47.0–125.6	23.13	-0.73
var = 4.02	Q1+Q2+Q3					35.48	
<i>Ascochyta</i> severity (AS; range = 1–7)	AS-Q1	LG6	S046029 bin	5.4	180.0–222.0	27.14	-0.81
mean = 4.49	AS-Q2	LG6	loc219	5.0	180.0–222.0	25.53	-0.80
var = 2.25	AS-Q3	LG6	S005268 bin	4.5	180.0–222.0	23.13	-0.73
	Q1+Q2+Q3					28.46	

<sup>1</sup> Linkage group.

<sup>2</sup> Marker or location (loc) in cM.

<sup>3</sup> 1.8-LOD support interval in cM.

<sup>4</sup> Phenotypic variance explained in percentage.

<sup>5</sup> Additive genetic effect of the alleles derived from the parental accession *L. odemensis* ILWL235

<https://doi.org/10.1371/journal.pone.0214409.t003>

## Discussion

High-throughput sequencing approaches, such as the RNAseq used in the present work, allow to generate and to characterize reference transcriptome datasets that support gene-based marker discovery useful and robust to undertake linkage mapping, trait dissection, or genetic diversity studies. Other transcriptome assemblies have previously been obtained with respect to other lentil cultivars using 454 Roche pyrosequencing [66,38], more recently also using the Illumina technology [67–69] and the Ion Proton system [70]. The assembly obtained in this work for cv. Alpo has displayed higher N50 values (2,501 bp), together with average (1,748.4 bp) and total lengths (179,931,148 bases) than those previously reported in transcriptomes obtained using Illumina sequencing platform. The N50 value is a metric widely used to assess the contiguity and the quality of an assembly [71]. Taking into account only the longest isoform of each gene, the total assembled bases amounted to 64,693,989 bp, corresponding to nearly half of the 130 Mbp of the whole lentil genome (4,063 Mbp) that has been identified as genic sequences [72]. The number of total assembled transcripts was also higher in the cv. Alpo transcriptome (102,879) than has been reported in previous publications (up to 84,069), while the number of different protein genes identified was lower (49,453) than the values reported by Singh et al. [68] which ranged from 52,565 to 55,667. Differences are a direct consequence of the distinct approaches that have been followed in the sequencing and read assembly processes, as well as a result of the different genotypes, plant tissues and development states of the materials studied.

Several genetic linkage maps of lentil (summarized recently by Ates et al. [36]) have been constructed using mapping populations mostly obtained by crossing cultivated lentil varieties or accessions. Intersubspecific maps developed from crosses using *L. culinaris* and the wild ancestor *L. culinaris* subsp. *orientalis* have likewise been published [73–75] while interspecific

linkage maps have been limited to the case of *L. ervoides* [24]. The map presented here (*LcxLo*-map) is the first for an interspecific cross between the primary and the secondary lentil gene pools [33] based on *L. culinaris* and *L. odemensis*. The *LcxLo*-map is composed of 6,306 SNPs and short indels spanning a total of 5,782.2 cM with an average distance of 1.23 cM between adjacent marker bins. To date, this is the *Lens* sp. map that incorporates the largest number of markers, excluding a consensus map for *L. culinaris* that contains 9,793 markers obtained by integration of three component maps with 5,385, 4,177, and 2,439 markers [36]. It is worth mentioning that all the markers included in the *LcxLo*-map are located in different functional regions that are transcribed, instead of being distributed along the whole genome. Therefore, a higher average distance value between markers is expected compared to the low value (0.1 cM) reported in the case of the consensus map for *L. culinaris* based on DArT markers [36], because only 3.2% of the lentil genome is considered to hold functional genic sequences and gene density is not evenly distributed along the lentil genome [72].

The *culinaris* x *odemensis* map has been generated from a relative low number of available RILs and this fact could compromise the accuracy of the map, although a high collinearity level with respect to the location of the analyzed functional markers was found upon comparison against the lentil draft genome assembly v1.2.

The *LcxLo*-map has been used to find the genomic regions containing the candidate genes for three morphological characters. The seed coat spotting pattern (*Scp*) is an important characteristic that determines the market class and the end uses of the crop. A single gene controlling this trait has been proposed [41] and five alleles of the *Scp* gene have been reported, being the recessive homozygous genotype *scp scp* the non-spotted individuals [76]. The *Scp* locus was mapped onto LG6 by Fedoruk et al. [37] using an intraspecific linkage map with the same number of linkage groups as the seven lentil chromosomes, taken as the basis for the chromosome numbering of the lentil genome assembly [38]. The *Scp* locus was mapped by Eujayl et al. [77] to what they defined as LG3, with close linkage to the flower color locus. Kumar et al. [78] confirmed a monogenic complete dominance of purple flower color over white, fact previously proposed by Lal and Srivastava [79] who reported linkage with the mottled seed pattern.

We have found both traits, *Scp* and flower color (*FC*), in close linkage positioned in LG6 with overlapping support intervals for genomic location in the *LcxLo*-map. The *Scp* locus was mapped to a 3.3 Mb region of LcChr6 comprising 46 annotated genes including a candidate gene (Lc25388) encoding an MYB transcription factor MIXTA-like related with anthocyanin synthesis regulation and cell fate [80]. Most of the genes included in the *FC* support interval (25 out of 37) were annotated in a 7.7 Mb region that includes the whole 3.3 Mb region for *Scp*; nevertheless, another 10 genes which were included in the support interval are annotated outside of the 7.7 Mb region of the draft lentil genome, including gene Lc26377 corresponding to the flower color QTL peak marker. The Lc26377 gene encodes a phosphate transporter Pho-like that may have a pigmentation role as it has been reported that phosphate deficiency is related to anthocyanin accumulation [81,82].

Single gene inheritance of stem pigmentation in lentil has been reported with the purple color being controlled by a dominant allele [41,83–85]. Anthocyanin pigmentation of plantlet stems was present in the *L. odemensis* parental accession of the RIL population analyzed. This locus was mapped to LG1 using the QTL approach. However, the support interval included two genes currently annotated in LcChr5 and another gene annotated in LcChr6 of the draft genome assembly, all of them linked to five genes annotated in LcChr1 comprising a large 50.3 Mb region. The stem pigmentation locus *Gs* was mapped by Duran et al. [74] to their denominated LGIV, however an improved version of this map that our research group has obtained and which includes common markers with the map of Sharpe et al. [38], identified the named LGIV as part of the LG5/LcChr5 (unpublished data). The trait did not map to any linkage

group when the map was obtained comprising the morphological markers. Therefore, it has not been possible to determine its precise chromosomal location using the *LcxLo*-map data.

A single QTL was identified for the character flowering time located on LG6. All the genes included in the support interval are located at the LcChr6 of the lentil draft genome v1.2, in an 8.8 Mb region encompassing a total number of 157 annotated genes. Using an intersubspecific cross of *L. culinaris* ssp. *culinaris* and *L.c.* ssp. *orientalis*, Fratini et al. [22] described a QTL for flowering time close to the *Scp* locus in a linkage group named LGI. As the *Scp* locus was mapped onto LG6/LcChr6 of the *LcxLo*-map (see above) near to the QTL position, these two QTLs found respectively in intersubspecific and interspecific mapping populations may be related. Recently and through marker-trait association analysis, Kumar et al. [86] identified two SSR marker loci associated with flowering time, each one explaining 15% of total phenotypic variation observed for this trait in a panel of 96 genotypes of cultivated lentil. We tried to locate these two markers (ALD 16 and ALD P5) in the lentil genome v1.2 using their available PCR-primer sequences [25]. The ALD P5 primers were located in the LcChr3 at a distance of 256 bp (from 83,245,520 to 83,245,775) while it was not possible to find the chromosomal position of the ALD 16 locus.

A total of three QTLs were detected for the seed size trait in LG1, LG5a, and LG5b, corresponding to one genomic location in LcChr1 and two in LcChr5. QTLs for seed size in lentil have been previously mapped by Fratini et al. [22], Fedoruk et al. [37], and Verma et al. [87]. The phenotypic variances of 37% and 60% reported by Fratini et al. [22] and Fedoruk et al. [37] for seed size QTLs are higher than that reported in our study (35.5%) and also to the 27.5% reported by Verma et al. [87]. The comparison of the QTLs identified by Fratini et al. [22] and Verma et al. [87] is difficult with respect to their map positions in the *LcxLo*-map because no common markers or linkage groups can be resorted to. Fedoruk et al. [37] identified three QTLs for seed diameter, the most important was located in LG1/LcChr1 around the cotyledon color locus (*Yc*). One of the *Yc* flanking markers was LcC17238p3606 that is located in position 202,217,501 of LcChr1. Using association mapping analysis with SNP makers of 138 cultivated lentil accessions, Khazaei et al. [88] validated two SNPs closely associated to the seed diameter trait, located in LcChr7 and LcChr1 which respectively explained 11% and 14% of the observed variation. The location of the SNP in LcChr1 (212,502,660) is relatively close to the *Yc* linked marker reported by Fedoruk et al. [37] and to one of the markers included in the support interval of the QTL detected in LG1 (163,405,141) of the *LcxLo*-map, with 242 genes annotated between them.

Several genetic linkage maps have previously been used to identify QTLs for *Ascochyta* blight resistance in lentil (reviewed by Rodda et al. [89]), with magnitudes varying from 3% to 89% of the phenotypic variance evidenced. Common QTL locations between different studies are difficult to establish because of the differing nomenclature for linkage groups and the lack of common marker loci between the genetic maps. In spite of that, the QTL named AB\_NF1 on LG6 in the study of Sudheesh et al. [67] is comparable in position to QTL5 on LG1 of Rubeena et al. [85] and QTL1 on LG1 of Gupta et al. [49], based on a common SSR locus location [89]. The AB\_NF1 QTL accounted for 7% of the phenotypic variation in an F<sub>6</sub> RIL population from a cross of Indianhead x Northfield, being Northfield (ILL5588) a common parent in the majority of the studies published. We have located the SNP flanking markers of AB\_NF1 in the genes Lc28181 and Lc25002 located in the LcChr6 positions 197,394,513 and 198, 226,847, respectively. Most of the markers included in the common support interval of the QTLs detected in LG6 of the *LcxLo*-map are annotated in the region extending from 132,032,337 to 160,668,788, however one of them is located in position 203,277,345. The genomic regions of AB\_NF1 and the QTLs detected in this work are close enough to suggest that common genes may be involved in the *Ascochyta* resistance responses. Sudheesh et al. [67]



reported that AB\_NF1 linked markers were found to be associated with a minor gene that appeared to confer a partial resistance to an Australian field isolate (FT14125) evidencing on lentil genotypes a different pathogenicity pattern.

As new genetic diversity sources, wild relatives of crops have been used during many decades in plant breeding programs. Their utilization is expected to increase as a result of the ongoing improvements concerning species and diversity information and advances in breeding tools boosted by high throughput technologies. We have here reported the first genetic linkage map between *L. culinaris* and *L. odemensis* based on thousands of functional markers. This map has allowed to locate the genomic regions where the genes controlling different morphological and agronomical traits of lentil are present. Further studies in progress will allow for the precise identification of some of these genes.

## Supporting information

**S1 File. Reference transcriptome of cultivar ALPO.**  
(TXT)

**S2 File. List of 6,306 markers used to construct the genetic map, ordered according to their respective linkage groups.**  
(XLSX)

## Acknowledgments

The authors especially acknowledge access to the lentil genome data allowed by the Lentil Genome Sequencing LenGen Project granted by Dr. Bett (Univ. Saskatchewan, Canada).

## Author Contributions

**Conceptualization:** Carlos Polanco, Luis Enrique Sáenz de Miera, Pedro García, Marcelino Pérez de la Vega.

**Data curation:** Carlos Polanco, Luis Enrique Sáenz de Miera, Ana Isabel González, Pedro García.

**Formal analysis:** Carlos Polanco, Luis Enrique Sáenz de Miera, Pedro García, Francisca Vaquero, Francisco Javier Vences.

**Funding acquisition:** Carlos Polanco, Marcelino Pérez de la Vega.

**Investigation:** Carlos Polanco, Luis Enrique Sáenz de Miera, Ana Isabel González, Pedro García, Richard Fratini, Francisca Vaquero, Francisco Javier Vences.

**Methodology:** Carlos Polanco, Luis Enrique Sáenz de Miera, Ana Isabel González.

**Resources:** Marcelino Pérez de la Vega.

**Supervision:** Marcelino Pérez de la Vega.

**Writing – original draft:** Carlos Polanco, Marcelino Pérez de la Vega.

**Writing – review & editing:** Carlos Polanco, Ana Isabel González, Pedro García, Marcelino Pérez de la Vega.

## References

1. Warschefsky E, Varma-Penmetsa R, Cook DG, Wettberg EJB Von. Back to the wilds: Tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. *Am J*

- Bot. 2014; 101: 1791–1800. [10.3732/ajb.1400116](https://doi.org/10.3732/ajb.1400116). <https://doi.org/10.3732/ajb.1400116> PMID: [25326621](https://pubmed.ncbi.nlm.nih.gov/25326621/)
2. Dempewolf H, Baute G, Anderson J, Kilian B, Smith C, Guarino L. Past and future use of wild relatives in crop breeding. *Crop Sci.* 2017; 57: 1070–1082. <https://doi.org/10.2135/cropsci2016.10.0885>
  3. McCouch S. Diversifying selection in plant breeding. *PLoS Biol.* 2004; 2: e347. <https://doi.org/10.1371/journal.pbio.0020347> PMID: [15486582](https://pubmed.ncbi.nlm.nih.gov/15486582/)
  4. Nelson R, Wiesner-Hanks T, Wisser R, Balint-Kurti P. Navigating complexity to breed disease-resistant crops. *Nat Rev Genet.* 2018; 19: 21–33. <https://doi.org/10.1038/nrg.2017.82> PMID: [29109524](https://pubmed.ncbi.nlm.nih.gov/29109524/)
  5. Brar DS, Khush GS. Alien introgression in rice. *Plant Mol Biol.* 1997; 35: 35–47. <https://doi.org/10.1023/A:1005825519998> PMID: [9291958](https://pubmed.ncbi.nlm.nih.gov/9291958/)
  6. Mujeeb-Kazi A, Kazi AG, Dundas I, Rasheed A, Ogonnaya F, Kishii M, et al. Genetic diversity for wheat improvement as a conduit to food security. *Advances in Agronomy* 2013; 122: 179–257. <https://doi.org/10.1016/B978-0-12-417187-9.00004-8>
  7. Piaskowski J, Murphy K, Kisha T, Jones S. Perennial wheat lines have highly admixed population structure and elevated rates of outcrossing. *Euphytica* 2017; 213: 171. <https://doi.org/10.1007/s10681-017-1961-x>
  8. Sharma HC. How wide can a wide cross be? *Euphytica* 1995; 82: 43–64. <https://doi.org/10.1007/BF00028709>
  9. Zaim M, El Hassouni K, Gamba F, Filali-Maltouf A, Belkadi B, Sourour A, et al. Wide crosses of durum wheat (*Triticum durum* Desf.) reveal good disease resistance, yield stability, and industrial quality across Mediterranean sites. *Field Crops Res.* 2017; 214: 219–227. <https://doi.org/10.1016/j.fcr.2017.09.007>
  10. Wulff BBH, Moscou MJ. Strategies for transferring resistance into wheat: from wide crosses to GM cassettes. *Front Plant Sci* 2014; 5: 692. <https://doi.org/10.3389/fpls.2014.00692> PMID: [25538723](https://pubmed.ncbi.nlm.nih.gov/25538723/)
  11. Zhang H, Mittal N, Leamy LJ, Barazani O, Song BH. Back into the wild—Apply untapped genetic diversity of wild relatives for crop improvement. *Evol Appl* 2017; 10: 5–24. <https://doi.org/10.1111/eva.12434> PMID: [28035232](https://pubmed.ncbi.nlm.nih.gov/28035232/)
  12. Abberton MT. Interspecific hybridization in the genus *Trifolium*. *Plant Breed.* 2007; 126: 337–342. <https://doi.org/10.1111/j.1439-0523.2007.01374.x>
  13. Muehlbauer FJ, Kaiser WJ, Simon CJ. Potential for wild-species in cool season food legume breeding. *Euphytica* 1994; 73: 109–114. <https://doi.org/10.1007/BF00027187>
  14. Ocampo B, Conicella C, Moss JP. Wide crossing: Opportunities and progress. In: Knight R, editor. Linking research and marketing opportunities for pulses in the 21<sup>st</sup> century. Kluwer Academic Publishers, Dordrecht. 2000. pp. 411–419.
  15. Sharma S. Prebreeding using wild species for genetic enhancement of grain legumes at ICRIASAT. *Crop Sci.* 2017; 57: 1132–1144. <https://doi.org/10.2135/cropsci2017.01.0033>
  16. Sharma DR, Kaur R, Kumar K. Embryo rescue in plants—A review. *Euphytica* 1996; 89: 325–337. <https://doi.org/10.1007/BF00022289>
  17. Fratini R, Pérez de la Vega M. Genetics of economic traits in lentil: seed traits and adaptation to climatic variations. *Grain Legum.* 2011; 57: 18–20.
  18. Pérez de la Vega M, Fratini R, Muehlbauer FJ. Lentil (*Lens culinaris* Medik.). In: Pérez de la Vega M, Torres AM, Cubero JI, editors. *Genetics, Genomics and Breeding of Cool Season Grain Legumes*. Enfield, Science Publishers, 2011. pp. 98–150.
  19. Bhadauria V, Ramsay L, Bett KE, Banniza S. QTL mapping reveals genetic determinants of fungal disease resistance in the wild lentil species *Lens ervoides*. *Sci Rep* 2017; 7: 3231. <https://doi.org/10.1038/s41598-017-03463-9> PMID: [28607439](https://pubmed.ncbi.nlm.nih.gov/28607439/)
  20. Dadu RHR, Ford R, Sambasivam P, Gupta D. A novel *Lens orientalis* resistance source to the recently evolved highly aggressive Australian *Ascochyta lentis* Isolates. *Front Plant Sci* 2017; 8: 1038. <https://doi.org/10.3389/fpls.2017.01038> PMID: [28659965](https://pubmed.ncbi.nlm.nih.gov/28659965/)
  21. Fiala JV, Tullu A, Banniza S, Séguin-Swartz G, Vandenberg A. Interspecies transfer of resistance to anthracnose in lentil (*Lens culinaris* Medik.). *Crop Sci.* 2009; 49: 825–830. <https://doi.org/10.2135/cropsci2008.05.0260>
  22. Fratini R, Durán Y, García P, Pérez de la Vega M. Identification of quantitative trait loci (QTL) for plant structure, growth habit and yield in lentil. *Spanish J Agricul Res.* 2007; 5(3): 348–356. <https://doi.org/10.5424/sjar/2007053-255>
  23. Gorim LY, Vandenberg A. Evaluation of wild lentil species as genetic resources to improve drought tolerance in cultivated lentil. *Front Plant Sci* 2017; 8: 1129. <https://doi.org/10.3389/fpls.2017.01129> PMID: [28706524](https://pubmed.ncbi.nlm.nih.gov/28706524/)

24. Gujaria-Verma N, Vail SL, Carrasquilla-Garcia N, Penmetsa RV, Cook DR, Farmer AD, et al. Genetic mapping of legume orthologs reveals high conservation of synteny between lentil species and the sequenced genomes of *Medicago* and chickpea. *Front Plant Sci* 2014; 5: 676. <https://doi.org/10.3389/fpls.2014.00676> PMID: 25538716
25. Kumar J, Srivastva E, Singh M, Kumar S, Nadarajan N, Sarker A. Diversification of indigenous gene-pool by using exotic germplasm in lentil (*Lens culinaris* Medikus subsp. *culinaris*). *Physiol Mol Biol Plants* 2014; 20(1): 125–132. <https://doi.org/10.1007/s12298-013-0214-2> PMID: 24554846
26. Singh M, Rana MK, Kumar K, Bisht IS, Dutta M, Gutam NK, et al. Broadening the genetic base of lentil cultivars through inter-sub-specific and interspecific crosses of *Lens* taxa. *Plant Breed.* 2013; 132: 667–675. <https://doi.org/10.1111/pbr.12089>
27. Singh M, Bisht IS, Kumar S, Dutta M, Bansal KC, Karale M, et al. Global wild annual *Lens* collection: A potential resource for lentil genetic base broadening and yield enhancement. *PLoS ONE* 2014; 9: e107781. <https://doi.org/10.1371/journal.pone.0107781> PMID: 25254552
28. Singh M, Rana JC, Singh B, Kumar S, Saxena DR, Saxena A, et al. Comparative agronomic performance and reaction to fusarium wilt of *Lens culinaris* × *L. orientalis* and *L. culinaris* × *L. ervoides* derivatives. *Front. Plant Sci* 2017; 8: 1162. <https://doi.org/10.3389/fpls.2017.01162> PMID: 28751897
29. Suvorova GN, Ikonnikov AV. Characteristics of recombinant lines of lentil *L. culinaris* × *L. orientalis*. *Russ Agric Sci* 2014; 40: 22–26. <https://doi.org/10.3103/S1068367414010182>
30. Singh M, Sharma SK, Singh B, Malhotra N, Chandora R, Sarker A, et al. Widening the genetic base of cultivated gene pool following introgression from wild *Lens* taxa. *Plant Breed.* 2018; 137: 470–485. <https://doi.org/10.1111/pbr.12615>
31. Fratini R, Ruiz ML. Interspecific hybridization in the genus *Lens* applying *in vitro* embryo rescue. *Euphytica* 2006; 150: 271–280. <https://doi.org/10.1007/s10681-006-9118-3C>
32. Fratini R, Ruiz ML. Wide crossing in lentil through embryo rescue. *Methods Mol Biol.* 2011; 710: 131–139. [https://doi.org/10.1007/978-1-61737-988-8\\_11](https://doi.org/10.1007/978-1-61737-988-8_11) PMID: 21207267
33. Wong MML, Gujaria-Verma N, Ramsay L, Yuan HY, Caron C, Diapari M, et al. Classification and characterization of species within the genus *Lens* using genotyping-by-sequencing (GBS). *PLoS ONE* 2015; 10: e0122025. <https://doi.org/10.1371/journal.pone.0122025> PMID: 25815480
34. Ladizinsky G, Abbo S. The *Lens* genus. In: *The Search for Wild Relatives of Cool Season Legumes*, Ladizinsky G, Abbo S, editors. Springer Int Publ Switzerland; 2015. pp. 1–28. [https://doi.org/10.1007/978-3-319-14505-1\\_1](https://doi.org/10.1007/978-3-319-14505-1_1)
35. Ates D, Sever T, Aldemir S, Yagmur B, Temel HY, Kaya HB, et al. Identification QTLs controlling genes for Se uptake in lentil seeds. *PLoS ONE* 2016; 11(3): e0149210. <https://doi.org/10.1371/journal.pone.0149210> PMID: 26978666
36. Ates D, Aldemir S, Alsaleh A, Erdogmus S, Nemli S, Kahrman A, et al. A consensus linkage map of lentil based on DArT markers from three RIL mapping populations. *PLoS ONE* 2018; 13(1): e0191375. <https://doi.org/10.1371/journal.pone.0191375> PMID: 29351563
37. Fedoruk MJ, Vandenberg A, Bett KE. Quantitative trait loci analysis of seed quality characteristics in lentil using single nucleotide polymorphism markers. *Plant Genome* 2013; 6: 3. <https://doi.org/10.3835/plantgenome2013.05.0012>
38. Sharpe AG, Ramsay L, Sanderson LA, Fedoruk MJ, Clarke WE, Li R, et al. Ancient orphan crop joins modern era: gene-based SNP discovery and mapping in lentil. *BMC Genomics* 2013; 14: 192. <https://doi.org/10.1186/1471-2164-14-192> PMID: 23506258
39. Zhang L, Wang S, Li H, Deng Q, Zheng A, Li S, et al. Effects of missing marker and segregation distortion on QTL mapping in F<sub>2</sub> populations. *Theor Appl Genet.* 2010; 121: 1071–1082. <https://doi.org/10.1007/s00122-010-1372-z> PMID: 20535442
40. Voss-Fels K, Snowdon RJ. Understanding and utilizing crop genome diversity via high-resolution genotyping. *Plant Biotechnol J.* 2016; 14: 1086–1094. <https://doi.org/10.1111/pbi.12456> PMID: 27003869
41. Ladizinsky G. The genetics of several morphological traits in the lentil. *J Hered.* 1979; 70: 135–137. <https://doi.org/10.1093/oxfordjournals.jhered.a109209>
42. Ladizinsky G. Wild lentils. *CRC Crit Rev Plant Sci* 1993; 12: 169–184. <https://doi.org/10.1080/07352689309701900>
43. Tadmor Y, Zamir D, Ladizinsky G. Genetic mapping of ancient translocation in the genus *Lens*. *Theor Appl Genet.* 1987; 73: 883–892. <https://doi.org/10.1007/BF00289394> PMID: 24241299
44. Balyan HS, Houben A, Ahan R. Karyotype analysis and physical mapping of 18S-5.8S-25S and 5S ribosomal RNA loci in species of genus *Lens* Miller (Fabaceae). *Caryologia* 2002; 55: 121–128. <https://doi.org/10.1080/00087114.2002.10589267>
45. Fernández M, Ruiz ML, Linares C, Fominaya A, Pérez de la Vega M. 5S rDNA genome regions of *Lens* species. *Genome* 2005; 48: 937–942. <https://doi.org/10.1139/g05-052> PMID: 16391700

46. Galasso I. Distribution of highly repeated DNA sequences in species of the genus *Lens* Miller. *Genome* 2003; 46: 1118–1124. <https://doi.org/10.1139/g03-077> PMID: 14663530
47. Bett K, Ramsay L, Chan C, Sharp A, Cook D, Penmetsa R, et al. Lentil 1.0 and beyond. Plant and Animal Genomes XXIV Conference, San Diego, California, USA; 2016.
48. Chang S, Puryear J, Cairney J. A simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol Rep.* 1993; 11: 113–116. <https://doi.org/10.1007/BF02670468>
49. Gupta D, Taylor WJ, Inder P, Phan HTT, Ellwood SR, Mathur PN, et al. Integration of EST-SSR markers of *Medicago truncatula* into intraspecific linkage map of lentil and identification of QTL conferring resistance to ascochyta blight at seedling and pod stages. *Mol Breed.* 2012; 30(1): 429–439. <https://doi.org/10.1007/s11032-011-9634-2>
50. Ford R, Pang ECK, Taylor PWJ. Genetics of resistance to ascochyta blight (*Ascochyta lentis*) of lentil and the identification of closely linked RAPD markers. *Theor Appl Genet.* 1999; 98: 93–98. <https://doi.org/10.1007/s001220051044>
51. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 2012; 9: 671–675. <https://doi.org/10.1038/nmeth.2089> PMID: 22930834
52. Erskine W, Hussain A, Tahir M, Bahksh A, Ellis RH, Summerfield RJ, et al. Field evaluation of a model of photothermal flowering responses in a world lentil collection. *Theor Appl Genet.* 1994; 88: 423–428. <https://doi.org/10.1007/BF00223655> PMID: 24186029
53. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
54. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; 30(15): 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
55. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; 29: 15–21. <https://doi.org/10.1093/bioinformatics/bts635> PMID: 23104886
56. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 2009; 25(16): 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
57. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 2013; 8: 1494–512. <https://doi.org/10.1038/nprot.2013.084> PMID: 23845962
58. Bray NL, Pimentel H, Melsted P, Pachter P. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol.* 2016; 34: 525–527. <https://doi.org/10.1038/nbt.3519> PMID: 27043002
59. Wu Y, Bhat PR, Close TJ, Lonardi S. Efficient and Accurate Construction of Genetic Linkage Maps from the Minimum Spanning Tree of a Graph. *PLoS Genet.* 2008; 4(10): e1000212. <https://doi.org/10.1371/journal.pgen.1000212> PMID: 18846212
60. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. 2016. Available from <https://www.R-project.org/>.
61. Yu Y, Ouyang Y, Yao W. ShinyCircos: an R/Shiny application for interactive creation of Circos plot. *Bioinformatics* 2018; 34(7): 1229–1231. <https://doi.org/10.1093/bioinformatics/btx763> PMID: 29186362.
62. Broman KW, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 2003; 19(7): 889–890. <https://doi.org/10.1093/bioinformatics/btg112> PMID: 12724300
63. Haley CS, Knott SA. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 1992; 69: 315–324. <https://doi.org/10.1038/hdy.1992.131> PMID: 16718932
64. Manichaikul A, Dupuis J, Sen S, Broman KW. Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. *Genetics* 2006; 174(1): 481–9. <https://doi.org/10.1534/genetics.106.061549> PMID: 16783000
65. Zeng H, Zhang Y, Zhang X, Pi E, Zhu Y. Analysis of EF-Hand proteins in soybean genome suggests their potential roles in environmental and nutritional stress signaling. *Front Plant Sci* 2017; 8: 877. <https://doi.org/10.3389/fpls.2017.00877> PMID: 28596783
66. Kaur S, Cogan NO, Pembleton LW, Shinzuka M, Savin KW, Materne M, et al. Transcriptome sequencing of lentil based on second-generation technology permits large-scale unigenic assembly and SSR marker discovery. *BMC Genomics* 2011; 12: 265. <https://doi.org/10.1186/1471-2164-12-265> PMID: 21609489
67. Sudheesh S, Rodda MS, Davidson J, Javid M, Stephens A, Slater AT, et al. SNP-Based linkage mapping for validation of QTLs for resistance to Ascochyta blight in lentil. *Front Plant Sci* 2016; 7: 1604. <https://doi.org/10.3389/fpls.2016.01604> PMID: 27853461

68. Singh D, Singh CK, Taunk J, Tomar RSS, Chaturvedi AK, Gaikwad K, et al. Transcriptome analysis of lentil (*Lens culinaris* Medikus) in response to seedling drought stress. *BMC Genomics* 2017; 18: 206. <https://doi.org/10.1186/s12864-017-3596-7> PMID: 28241862
69. Sari E, Bhadauria V, Ramsay L, Borhan MH, Lichtenzweig J, Bett KE, et al. Defense responses of lentil (*Lens culinaris*) genotypes carrying non-allelic ascochyta blight resistance genes to *Ascochyta lentis* infection. *PLoS ONE* 2018; 13(9): e0204124. <https://doi.org/10.1371/journal.pone.0204124> PMID: 30235263
70. Khorramdelazad M, Bar I, Whatmore P, Smetham G, Bhaaskaria V, Yang Y, et al. Transcriptome profiling of lentil (*Lens culinaris*) through the first 24 hours of *Ascochyta lentis* infection reveals key defence response genes. *BMC Genomics* 2018; 19: 108. <https://doi.org/10.1186/s12864-018-4488-1> PMID: 29385986
71. Alhakami H, Mirebrahim H, Lonardi S. A comparative evaluation of genome assembly reconciliation tools. *Genome Biol.* 2017; 18: 93. <https://doi.org/10.1186/s13059-017-1213-3> PMID: 28521789
72. Ogutcen E, Ramsay L, Bishop von Wettberg E, Bett K. Capturing variation in *Lens* (Fabaceae): Development and utility of an exome capture array for lentil. *Appl Plant Sci* 2018; 6: e1165. <https://doi.org/10.1002/aps3.1165>
73. De la Puente R, García P, Polanco C, Pérez de la Vega M. An improved intersubspecific genetic map in *Lens* including functional markers. *Spanish J Agricul Res.* 2013; 11(1): 132–136. <https://doi.org/10.5424/sjar/2013111-3283>
74. Durán Y, Fratini R, García P, Pérez de la Vega M. An intersubspecific genetic map of *Lens*. *Theor Appl Genet.* 2004; 108: 1265–1273. <https://doi.org/10.1007/s00122-003-1542-3> PMID: 14676948
75. Havey MJ, Muehlbauer FJ. Linkages between restriction fragment length, isozyme, and morphological markers in lentil. *Theor Appl Genet.* 1989; 77: 395–401. <https://doi.org/10.1007/BF00305835> PMID: 24232618
76. Vandenberg A, Slinkard AE. Genetics of seed coat color and pattern in lentil. *J Hered.* 1990; 81: 484–488. <https://doi.org/10.1093/oxfordjournals.jhered.a111030>
77. Eujayl I, Baum M, Powell W, Erskine W, Pehu E. A genetic linkage map of lentil (*Lens* sp.) based on RAPD and AFLP markers using recombinant inbred lines. *Theor Appl Genet.* 1998; 97: 83–89. <https://doi.org/10.1007/s001220050869>
78. Kumar Y, Sharma B, Mishra SK, Tyagi MC. Linkage between genes for flower colour, seed coat pattern and seed coat colour in lentil (*Lens culinaris* Medik.). *Journal of Lentil Research* 2005; 2: 22–26.
79. Lal S, Srivastava RS. Inheritance of flower colour in lentils. *Indian J Genet Plant Breed* 1975; 35(1): 29–30.
80. Outchkourov NS, Karlova R, Hölscher M, Schrama X, Bilou I, Jongedijk E, et al. Transcription Factor-Mediated Control of Anthocyanin Biosynthesis in Vegetative Tissues. *Plant Physiol.* 2018; 176: 1862–1878. <https://doi.org/10.1104/pp.17.01662> PMID: 29192027
81. Poirier Y, Bucher M. Phosphate transport and homeostasis in *Arabidopsis*. *The Arabidopsis Book* 2002; 1: e0024. <https://doi.org/10.1199/tab.0024> PMID: 22303200
82. Liu TY, Huang TK, Tseng CY, Lai YS, Lin SI, Lin WY, et al. PHO2-Dependent degradation of PHO1 modulates phosphate homeostasis in *Arabidopsis*. *Plant Cell* 2012; 24: 2168–2183. <https://doi.org/10.1105/tpc.112.096636> PMID: 22634761
83. Emami MK, Sharma B. Linkage between three morphological markers in lentil. *Plant Breed.* 1999; 118: 579–581. <https://doi.org/10.1046/j.1439-0523.1999.00416.x>
84. Hoque ME, Mishra SK, Sarker A. Inheritance and linkage relationship between morphological and RAPD markers in lentil (*Lens culinaris* Medik.). *Indian J Genet Plant Breed* 2002; 62: 5–10.
85. Rubeena P, Taylor WJ, Ades PK, Ford R. QTL mapping of resistance in lentil (*Lens culinaris* ssp. *culinaris*) to ascochyta blight (*Ascochyta lentis*). *Plant Breed.* 2006; 125: 506–512. <https://doi.org/10.1111/j.1439-0523.2006.01259.x>
86. Kumar J, Gupta S, Gupta DS, Singh NP. Identification of QTLs for agronomic traits using association mapping in lentil. *Euphytica* 2018; 214: 75. <https://doi.org/10.1007/s10681-018-2155-x>
87. Verma P, Goyal R, Chahota RK, Sharma TR, Abdin MZ, Bhatia S. Construction of a genetic linkage map and identification of QTLs for seed weight and seed size traits in lentil (*Lens culinaris* Medik.). *PLoS One* 2015; 10: e0139666. <https://doi.org/10.1371/journal.pone.0139666> PMID: 26436554
88. Khazaei H, Fedoruk M, Caron CT, Vandenberg A, Bett KE. Single Nucleotide Polymorphism markers associated with seed quality characteristics of cultivated lentil. *Plant Genome* 2018; 11: 170051. <https://doi.org/10.3835/plantgenome2017.06.0051>
89. Rodda MS, Davidson J, Javid M, Sudheesh S, Blake S, Forster JW, et al. Molecular breeding for Ascochyta blight resistance in lentil: Current progress and future directions. *Front Plant Sci* 2017; 8: 1136. <https://doi.org/10.3389/fpls.2017.01136> PMID: 28706526