Received: 6 November 2024 | Accepted: 24 February 2025

DOI: 10.1002/tpg2.70021

The Plant Genome

ORIGINAL ARTICLE

Genetic and transcriptomic analysis of lentil seed imbibition and dormancy in relation to its domestication

Azalea Guerra-García[1] | **Jana Balarynová**[2] | **Petr Smykal**[2] | **Eric J von Wettberg**[3] | **Scott D. Noble**[4] | **Kirstin E. Bett**[5]

[1]Departamento de Biotecnología y Bioquímica, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN), Irapuato, México

[2]Department of Botany, Palacký University, Olomouc, Czech Republic

[3]Department of Agriculture, Landscape, and Environment, Gund Institute for the Environment, University of Vermont, Burlington, Vermont, USA

[4]Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

[5]Department of Plant Sciences, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

Correspondence
Azalea Guerra-García, Departamento de Biotecnología y Bioquímica, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN), Irapuato, Guanajuato, México.
Email: azalea.guerra@cinvestav.mx

Kirstin E. Bett, Department of Plant Sciences, University of Saskatchewan, Saskatoon, SK, Canada.
Email: k.bett@usask.ca

Assigned to Associate Editor Agnieszka Golicz.

Funding information
Palacky University Grant Agency, Grant/Award Numbers: PrF-2023-001, PrF-2024-001; Ministry of Education, Youth and Sports of the Czech Republic, Grant/Award Number: LUAUS25035; Western Grains Research Foundation, Grant/Award Number: GC1903; Genome Canada, Grant/Award Number:

Abstract

Seed dormancy is an adaptation that delays germination to prevent the start of this process during unsuitable conditions. It is crucial in wild species but its loss was selected during crop domestication to ensure a fast and uniform germination. Water uptake, or imbibition, is the first step of germination. In the Fabaceae family, seeds have physical dormancy, in which seed coats are impermeable to water. We used an interspecific cross between an elite lentil line (*Lens culinaris*) and a wild lentil (*L. orientalis*) to investigate the genetic basis of imbibition capacity through quantitative trait locus (QTL) mapping and by using RNA from embryos and seed coats at different development stages, and phenotypic data of seed coat thickness (SCT) and proportion of imbibed seeds (PIS). Both characteristics were consistent throughout different years and locations, suggesting a hereditary component. QTL results suggest that they are each controlled by relatively few loci. Differentially expressed genes (DEGs) within the QTL were considered candidate genes. Two glycosyl-hydrolase genes (a β-glucosidase and a β-galactosidase), which degrade complex polysaccharides in the cell wall, were found among the candidate genes, and one of them had a positive correlation (β-glucosidase) between gene expression and imbibition capacity, and the other gene (β-galactosidase) presented a negative correlation between gene expression and SCT.

Abbreviations: ABA, abscisic acid; DAP, days after pollination; DEGs, differentially expressed genes; GA, gibberellins; GO, gene ontology; KEGG, Kyoto encyclopedia of genes and genomes; LG, linkage group; LOD, logarithm of the odds; PC, principal component; PCA, principal component analysis; PIS, proportion of imbibed seeds; QTL, quantitative trait locus; RIL, recombinant inbred line; SCT, seed coat thickness; WGCNA, weighted gene co-expression network analysis.

LSP18-16302; Government of Saskatchewan, Grant/Award Number: 20200026

Plain Language Summary

Seed dormancy prevents sprouting until the conditions are ideal, a crucial adaptation for wild plants. However, during the domestication of crops, farmers favored seeds that germinated quickly and uniformly, leading to a loss of dormancy. The first step in germination is imbibition, where seeds absorb water. Legume seeds have a hard outer coating that prevents water absorption. In our study, we compared a domesticated lentil variety with a wild one to investigate how genes influence imbibition. We measured seed coat thickness and water absorption rates and examined the genetic differences between the two types. Our findings showed that these traits remained consistent across different years and locations, indicating they are inherited. We identified several key genes involved, including one that helps break down complex sugars in the cell wall (β-glucosidase), which was associated with improved water uptake.

1 | INTRODUCTION

Seed germination and its developmental transitions have to be properly timed with the beginning of favorable growing seasons to allow seedlings to successfully establish (Willis et al., 2014). This process is controlled by seed dormancy, which is the inability of viable seeds to complete germination under favorable conditions (Bradford & Nonogaki, 2007). Loss of seed dormancy is one of the key traits in the domestication syndrome and has long been thought to be an essential step in domesticating annual seed crops like cereals and grain legumes (Hammer, 1984; Smýkal et al., 2018). Because dormant seeds would be lost to an early farmer, it is assumed that strong selection for non-dormant seeds occurred to ensure fast and uniform germination after sowing and, therefore, increase yields and easier harvesting (Abbo et al., 2014). As a result, low dormancy has been selected during the domestication of many crop species (Shu et al., 2016). The first step, and a critical one on the path to germination, is water uptake of seeds, called imbibition, which is an indicator of loss of dormancy (Bewley & Black, 1994). Furthermore, in legumes, imbibition is also related to cooking time (Cichy et al., 2015, 2019).

Seed dormancy encompasses a range of physiological and structural changes that can be divided into dormancy classes. In the Fabaceae family, physical dormancy, in which the seed coat is "hard" and impermeable to water, dominates (Smýkal et al., 2014). However, the genetics behind seed dormancy in grain legumes is not well known despite the efforts made in some domesticated members of the family, such as cowpea (Kongjaimun et al., 2012), rice bean (Isemura et al., 2012), pea (Balarynová et al., 2023; Hradilová et al., 2019; Sedláková et al., 2023; Smýkal et al., 2014; Williams et al., 2024; Zablatzká et al., 2021), soybean (Sedivy et al., 2017), chickpea (Sedláková et al., 2021, 2023), and common bean (Laosatit et al., 2022; Soltani et al., 2021). Contributing to seed impermeability are traits such as seed coat thickness (SCT), its texture, and its chemical composition (Hradilová et al., 2017; Janská et al., 2019; Williams et al., 2024; Zablatzká et al., 2021).

Seed dormancy is established during seed development, and environmental factors during seed maturation influence the level of dormancy (Bewley, 1997). Further changes in dormancy continue after seeds have been shed from the maternal plant, and these changes persist until imbibition occurs and seeds can germinate (Finch-Savage & Leubner-Metzger, 2006; Lamont & Pausas, 2023).

Lentil (*Lens culinaris* Medik) is among the founding set of domesticated crops in the Fertile Crescent, alongside wheat, barley, chickpea, peas, flax, and faba beans (Coyne et al., 2020; Guerra-García et al., 2021, 2024; Guerra-Garcia et al., 2022), and nowadays its use is increasing as a good source of plant-based protein. The wild lentil *Lens orientalis* (Boiss.), the closest wild relative of the domesticated lentil, occurs in disturbed habitats in Southeastern Turkey, the center of the Fertile Crescent (Abbo et al., 2009; Murphy & Fuller, 2017). Among the important traits that distinguish wild from cultivated lentil are the lack of pod dehiscence and the loss of dormancy in the latter, as also occurs in other legume crops. Although dormancy is difficult to infer from paleobotanical remains, seed shape and thickness have been inferred to correlate with domestication status (Abbo et al., 2009; Murphy & Fuller, 2017).

Using an interspecific cross between a wild lentil (*L. orientalis*) and an elite breeding line (*L. culinaris*), we investigated the genetic mechanisms associated with variation in seed imbibition in this cool-season grain legume. We measured SCT and the extent of imbibition in the interspecific population, allowing us to look at the relationship between these two traits and to identify quantitative trait loci (QTL) associated with them. Because dormancy is set during seed development, we searched for differentially expressed genes (DEGs) in the

embryos and seed coats of the parental genotypes at different stages by using RNAseq. Genes found both within the QTL and as DEGs were further explored and their expression was examined in the recombinant inbred lines (RILs).

Core Ideas

- No correlation between seed coat thickness and imbibition rates, suggesting they are controlled by different loci.
- An interspecific population between elite and wild lentil parents was used to identify quantitative trait loci.
- Differentially expressed genes between embryos and seed coats of wild and domesticated lentil were identified.
- Among the candidate genes identified, two glycosyl-hydrolases were highlighted.
- A β-glucosidase expression positively correlated with imbibition, while β-galactosidase negatively correlated with seed coat thickness.

2 | MATERIALS AND METHODS

2.1 | Biological material

To study the differences in seed imbibition and SCT in lentil, the interspecific population LR-68 was used. This population consists of 121 RILs derived from the interspecies cross wild lentil *L. orientalis* (IG 72643) × cultivated lentil *L. culinaris* (3339-3) (https://knowpulse.usask.ca/study/Lentil-Seed-Dormancy-Imbibition-Analysis). Seeds for this study originated from field trials grown around Saskatoon, SK, Canada, in 2018–2020, named Preston (Pres) 2018, Sutherland (Suth) 2019, and Saskatchewan Pulse Growers field (SPG) 2019. Each site-year was set up as a randomized complete block with three replications. Plants were grown and maintained using standard methods of the USask lentil breeding program. An additional set was grown at the Palacký University in Olomouc, the Czech Republic (CRep) as described in Hradilová et al. (2017), Janská et al. (2019), Williams et al. (2024), and Zablatzká et al. (2021). In brief, plants were grown in a peat-sand (90:10) substrate mix (Florcom Profi, BB Com Ltd.) in a greenhouse from January to May 2020. The photoperiod was extended to 14 h using light (Sylvania Grolux 600 W, Hortilux Schreder). All trials included all 121 RILs and the parents.

2.2 | Phenotyping imbibition capacity and seed coat thickness

SCT was estimated using a non-destructive optical coherence tomography system (Figure 1a) previously shown to estimate SCT of lentil seeds successfully (Nguyen et al., 2021). Seeds from the Pres-2018, Suth-2019, SPG-2019, and CRep 2020 site-years were sampled. The seed coats of 40–60 seeds from each genotype and site-year (replicates) were measured, and means were calculated. Phenotypic data are available at https://knowpulse.usask.ca/study/Lentil-Seed-Dormancy-Imbibition-Analysis.

Imbibition was evaluated in terms of the proportion of imbibed seeds (PIS) using seeds from three site-years: Suth-2019, SPG-2019, and one greenhouse-grown from the Czech Republic (CRep-2020). Seeds from Suth-2019 and SPG-2019 were tested in 2020; therefore, they were 1-year old. Seeds from Pres-2018 were excluded due to insufficient biological material for destructive testing and to minimize the effect of seed age as these were a year older than the others when the imbibition experiment was run. Seeds from CRep-2020 were grown and tested in 2020 when they were around 3 months old. Twenty-five seeds from each RIL and parents were placed on damp germination paper (Figure 1b) on baking sheets and covered with another damp sheet. Each baking sheet was placed inside a plastic bag to maintain humidity. After 24, 48, and 72 h, the number of imbibed (swollen) seeds was recorded. Three replicates from each site-year were tested and means were calculated on a site-year basis. For the estimation of PIS, the 48 h time point was selected since it captured the most contrasting behavior between wild and cultivated parental genotypes (Figure 1d), and is similar to previous studies in pea (Hradilová et al., 2017) and chickpea (Sedláková et al., 2021).

The Fligner–Killeen test of homogeneity of variances and Shapiro–Wilk normality test were applied to both phenotypic traits using R base (R Core Team, 2021). According to the results, analysis of variance (ANOVA) and Tukey tests were performed with R base (R Core Team, 2021) to test if there were differences in SCT between different site-years. In the case of PIS, the non-parametric Kruskal–Wallis and pairwise Wilcoxon tests were applied. Spearman correlation tests were then used to investigate if PIS and SCT were associated, and if they were consistent across years and locations. Furthermore, a principal component analysis (PCA) was performed in R combining the SCT and PIS data from all the site-years.

2.3 | QTL mapping of the proportion of imbibed seeds and seed coat thickness

A previously generated genetic map for the LR-68 population (https://knowpulse.usask.ca/Geneticmap/2691123) was used along with the PIS and SCT phenotypic data to map genetic regions associated with those traits. The genetic

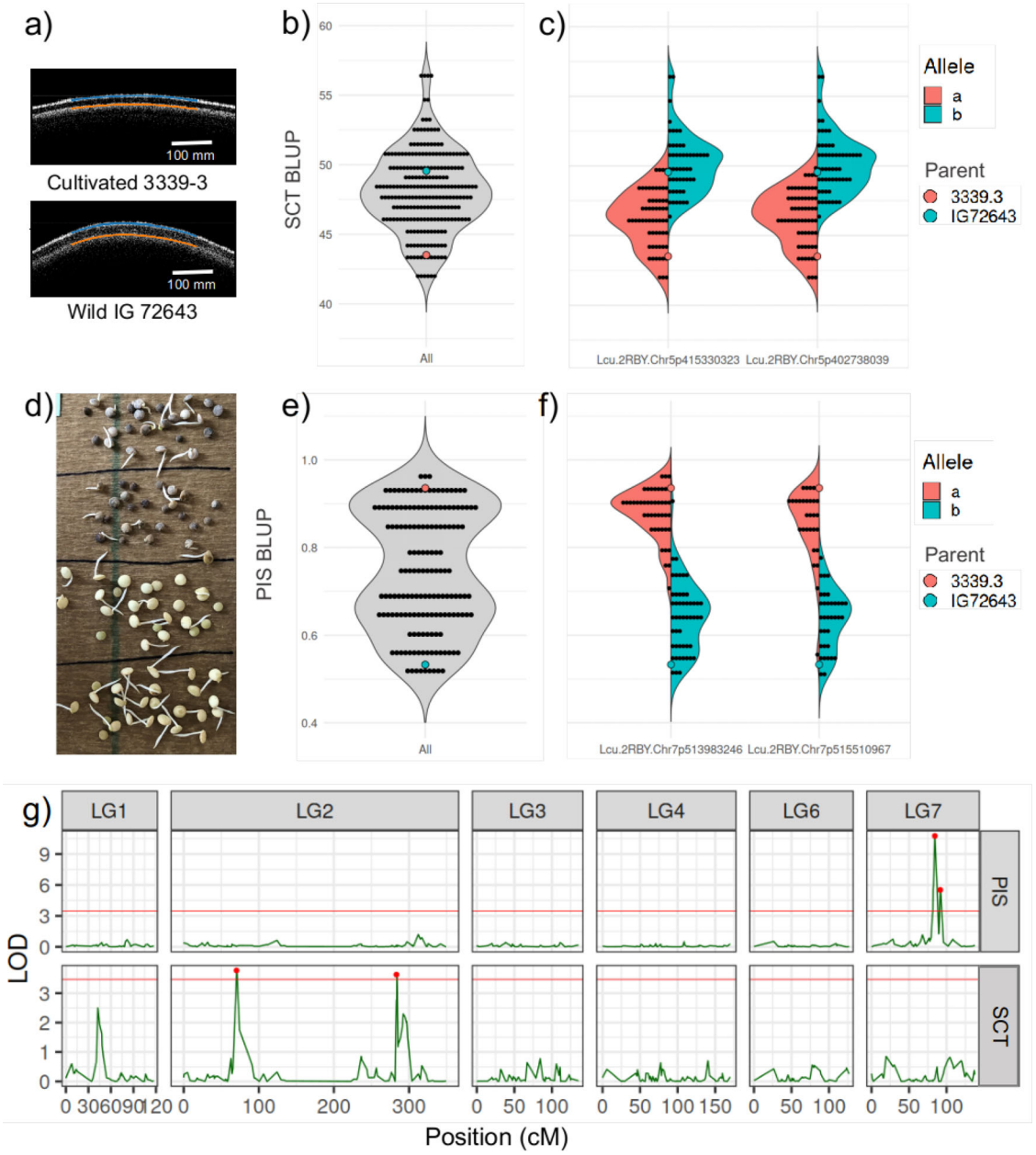

FIGURE 1 (a) Images based on optical coherence tomography of seed coats of a cultivated and a wild parent genotype. (b) Distribution of best linear unbiased prediction (BLUP) of seed coat thickness (SCT) from the LR-68 recombinant inbred lines (RILs), and (c) the same data separated according to the two quantitative trait locus (QTL) markers. (d) Lentil seeds from four different RILs after 48 h on damp germination paper. (e) Distribution of BLUP of the proportion of imbibed seeds (PIS) and (f) distribution of the same data separated based on their allele at the two QTL markers associated with this trait. (g) QTL plots for PIS (upper) and SCT (bottom). Red circles show the QTL regions identified, and the red line indicates the LOD threshold. LG, linkage group.

map has 616 genetic markers across six LGs. There is a 2–5 reciprocal translocation in the wild parent relative to the cultivated one (Cao et al., 2024), resulting in pseudolinkage of markers associated with these two chromosomes. Since the single nucleotide polymorphisms (SNPs) were called relative to an annotated *L. culinaris* genome assembly, the marker names reveal their chromosomal location in Lcu.2RBY (Ramsay et al., 2021). Chromosomes Lcu.2RBY.Chr2 and Lcu.2RBY.Chr5 combine to form LG2 (https://knowpulse.usask.ca/study/Lentil-Seed-Dormancy-Imbibition-Analysis).

Best linear unbiased predictions (BLUPs) across the environments (site-years) were estimated with the metan R package (Olivoto & Lúcio, 2020) and were then used for the QTL analysis to integrate the variation within and across the site-years. QTL mapping was performed with the qtl2 package (Broman et al., 2019). To establish the logarithm of odds (LOD) threshold, 100,000 permutations were performed for

each trait, setting $\alpha = 0.01$. The find_peaks function was used to identify the LOD peaks and to determine a 95% Bayes credible interval.

2.4 | RNA sequencing

To compare the gene expression during seed developmental stages that might determine PIS, RNA from seed coats and embryos at different developmental stages was analyzed. For this experiment, the parents of the LR-68 population, and selected 58 RILs representing the range of imbibition scores, and the cultivar Eston (*L. culinaris*) were grown in a glasshouse at Palacky University, Olomouc, Czech Republic, in January–April 2020 (Table S1).

The developmental stages were established according to days after pollination (DAP). The four time points sampled were 13, 17, 21, and 28 DAP (Table S1), following the protocol used by Zablatzká et al. (2021) and Balarynová et al. (2023). Three replicates for each treatment were included for the parental genotypes at the different stages. For the RILs, only seed coat tissue at 21 DAP and two replicates were included. This stage was selected based on previous experiments showing that the entirety of seed development is accomplished in about 26 days from pollination (Yu et al., 2023). It also corresponds to a similar analysis done on pea (Klčová et al., 2024) and chickpea (Sedláková et al., 2021, 2023). Seed coats and embryos were dissected on ice using a scalpel and tweezers and immediately frozen in liquid nitrogen. Frozen tissues were ground to a fine powder with liquid nitrogen using a sterile mortar and pestle. Total RNA was isolated using PureLink Plant RNA Reagent (Invitrogen). The Baseline-ZERO DNase (Epicenter) treatment removed residual DNA, followed by a phenol/chloroform extraction. Yield/quantity and purity were determined by using a NanoDrop 2000 spectrophotometer (Thermo Scientific). The integrity of the RNA samples was examined with an Agilent 2100 Bioanalyzer (Agilent Technologies).

The RNA sequencing for the RILs and parents was performed at the NovoGene facilities in Cambridge, UK, using an Illumina Novaseq platform, pair-end 150 bp, targeting 20 million pair reads per sample.

2.4.1 | Profile expressions, identification of differentially expressed genes

Reads were trimmed and adapters were removed using Trimmomatic 0.39 (Bolger et al., 2014) scanning the reads with a 4-base sliding window, cutting when the average quality <15, and keeping reads with a minimum length = 80. Reads were then aligned to the *L. culinaris* v2 reference genome (Ramsay et al., 2021) with HISAT 2.2.1 (Kim et al., 2019) and sorted with SAMtools (Danecek et al., 2021). To assemble the read alignments, merge the transcripts, and estimate their abundances, StringTie 2.2.0 (Kovaka et al., 2019) was used.

To identify DEGs, the R package Ballgown (Frazee et al., 2015) was implemented contrasting the cultivated (3339-3 and Eston) and wild (IG 72643) expression profiles from the same tissue and DAP (e.g., Cult-SeedCoat-13DAP vs. Wild-SeedCoat-13DAP; Table S2). In the case of the seed coat samples, wild and cultivated genotypes were compared at 13, 17, and 21 DAP, and for the embryos, the differences at 17, 21, and 28 DAP were evaluated. Differences in the seed coat at 28 DAP and embryos at 13 DAP were not tested since insufficient replicates were kept after filtering (see Results). Genes with FPKM (fragments per kilobase of transcript per million fragments mapped) <2 were filtered out to remove genes with very low expression for each comparison. The threshold for the DEGs discovery was set at q-value < 0.01 and a fold change >1.6 or <−1.6.

2.5 | Candidate genes and their expression in the RILs

Genes that were differentially expressed and residing within QTL for any of the two phenotypic traits were considered candidate genes. The expression of these genes was then traced in the RNA data of the RILs (seed coat at 21 DAP) and tested for correlation between the gene expression and SCT or PIS. For the correlations, non-annotated candidate genes were removed and Spearman tests were applied followed by the adjustment of p-values with the false discovery rate method because of the multiple tests.

2.6 | Gene ontology and KEGG enrichment analysis

The topGO package (Alexa & Rahnenfuhrer, 2022) and the annotation maps database from the GO.db package (Carlson, 2019) were used for testing enriched gene ontology (GO) terms in the expression profiles of the LR-68 parental genotypes and Eston elite cultivar. The *weight01* algorithm was applied, which combines the *weight* and *elim* GO graph methods, along with the Kolmogorov–Smirnov test ($p < 0.01$). Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was conducted with the KEGGREST package (Tenenbaum & Maintainer, 2022) applying a Wilcoxon test ($p < 0.01$).

2.7 | Co-expression networks

A co-expression network analysis was applied to the RNA data from the LR-68 parents (3336-3 and 72643) and the

Eston elite line. The co-expression networks were constructed with the weighted gene co-expression network analysis (WGCNA) algorithm using the WGCNA R package (Langfelder & Horvath, 2008). The analysis was applied to the normalized expression according to the tissue (embryos and seed coat). The pickSoftThreshold function, from the WGCNA package, was implemented, and the minimum module size was set to 30 to pick a soft-thresholding power for network construction.

After identifying the gene modules whose expression differs between the wild and domesticated genotypes, a GO was performed to identify gene terms enriched in the modules. In this case, the weight01 algorithm and Fisher tests were applied.

3 | RESULTS

3.1 | Phenotypic seed traits and QTL mapping

The parental genotypes are contrasting in seed traits, namely, SCT and PIS. These were evaluated in the LR-68 lentil population across different site-years (Figure 1). SCT exhibited a normal distribution (Shapiro–Wilk normality test, $p = 0.19$) and homogeneity of variances (Fligner–Killeen test, $p = 0.98$), while PIS did not follow a normal distribution ($p < 0.01$) nor homogeneity of variances ($p < 0.01$). Significant differences between the site-years for SCT were found with an ANOVA ($F = 4.01$, $p < 0.01$) and Tukey tests, specifically between Pres-2018 and the rest of the site-years (Figure S1). For PIS, the Kruskal–Wallis and Pairwise Wilcoxon tests found significant differences among the three site-years ($p < 0.01$; Figure S2).

The PCA from the combined SCT and PIS datasets showed that the first principal component (PC1) primarily reflected SCT and explained 55.8% of the variation; meanwhile, PC2 was mainly related to PIS and explained 31.4% of the variation (Figure S3). Significant and positive correlations were found among the SCT from different site-years (Figure S4). This was also observed for PIS but the data spread was greater than for SCT (Figure S4). The differences between the site-years suggest that both traits were influenced by the environment, particularly PIS. Nevertheless, the correlations support a genetic component kept under different environmental conditions. Regarding the relation between PIS and SCT, a weak correlation between their averages was detected (Spearman coefficient = −0.19, $p = 0.04$), but no correlation was found when the BLUPs, which integrate the variation between and across the evaluated site-years, were used (Spearman coefficient = −0.11, $p = 0.24$) (Table S3).

QTL mapping was applied to the two phenotypic traits for each site-year (Figure S5; Table S4) and to the BLUPs (Figure 1g). The QTL regions found with both approaches were consistent (Figure S5; Table S4). The results obtained with the BLUPs were used for downstream analyses because they had shorter Bayes intervals (0.95) and all the QTL identified with the BLUPs were also found with each site-year phenotypic data (Figure S5; Table S4; Supporting Information 2). For SCT, two QTLs were identified (LOD threshold = 3.30), both located on LG2, which includes chromosomes 2 and 5. The first QTL explained 13.7% of the variance, while the second explained 13.2%. Together, the two QTLs encompassed 1143 genes. Two more QTLs were found in chromosome 7 (LG7) associated with PIS (LOD threshold = 3.51), and one of them explained 34.3% of the variance and the other explained 19.4%, and 268 were present genes within their intervals (Figure 1g; Supporting Information 2). There was a clear phenotypic differentiation between the genotypes carrying the different alleles for the QTL (Figure 1c,f).

3.2 | Gene expression profiles differ between wild and domesticated lentil

The RNA reads were aligned against the *L. culinaris* reference genome Lcu.2RBY (Ramsay et al., 2021). Samples of both embryos and seed coats of the cultivated lentil had a mapping rate of >92% (Figure S6). This percentage decreased for reads from the wild lentil (*L. orientalis*) samples (>83%), particularly from the embryos (Figure S6).

After filtering and removing the genes with very low expression across the samples, ~12,000 genes were expressed in the seed coats from the LR-68 parents and Eston at different stages (Figure S7). This number decreased at later developmental stages. For example, on average ~11,500 expressed genes were present at 17 DAP and ~9000 genes at 28 DAP (Figure S7). In embryo samples, there was a clear differentiation between the expression profile of the domesticated and the wild genotypes when all expressed genes were considered (Figure S8). This was not the case in the seed coat samples, and no grouping according to the domesticated and wild lines was observed but rather according to the genotype (Figure S8).

The gene expression of the wild and cultivated genotypes from the same tissue (seed coat or embryo) at the same stage was compared to identify DEGs (Table S2). More DEGs were found in the embryo samples than in seed coat samples (Figures S9 and S10). In the seed coats, the number of DEGs ranged between 35 and 63 for each treatment, and between 249 and 633 for the embryos (Figure S9; Table S2). The gene expression of the DEGs differentiated the wild and domesticated genotypes (Figure S11).

3.3 | GO and KEGG enrichment analysis

GO analysis was performed to identify groupings of genes among the identified DEGs (Supporting Information 3). The enriched terms in the seed coat samples with the lowest *p* values were *oxidoreductase activity* at both 13 and 17 DAP and *translation* at 21 DAP. In the embryos were *translation* at 17 DAP, *photosynthesis* at 21 DAP, and *abscisic acid binding* at 28 DAP (Figures S12–S14; Supporting Information 3). Among the GO terms that were found in both embryos and seed coats were *molybdate ion transport* and *molybdenum ion binding*, *glycine cleavage system*, *translation*, and *structural constituent of ribosome* (Figures S12–S14).

The terms photosynthesis, light harvesting, defense response, abscisic acid (ABA)-activated signaling pathway, and ABA binding were among the enriched terms present only in the embryos. In the seed coat, pigment biosynthesis, catechol oxidase activity, monooxygenase activity, structural component of the cell wall, and extrinsic components of the membrane were among the terms with the highest significance values (Figures S12–S14).

In the KEGG pathway enrichment analysis, *legumain* showed the lowest *p*-value in the seed coats at 13 and 21 DAP, while at 17 DAP was *mitochondrial chaperone BCS1*. In the embryos at 17 and 28 DAP, *O-palmitoleoyl-L-serine hydrolase* pathway presented the most significant *p*-value, and *histone H2B* pathway at 21 DAP. The *glycine cleavage system H protein*, *soyasapogenol B glucuronide galactosyltransferase*, *palmitoleoyl-L-serine hydrolase*, and *histone* (H2B, H3, and H4) pathways were enriched in the embryo samples at different stages (Figure S15). Furthermore, *Xyloglucan-related* pathways were enriched in the seed coat at 13 DAP and 17 DAP, and the *pectinesterase* pathway in the embryos at 17 DAP showed a significant number of annotated terms.

3.4 | Co-expression network analysis

A coexpression networks analysis was performed to identify gene modules whose expression was correlated and differed between the wild and cultivated genotypes. In the seed coat, two gene modules stand out ("darkred" and "darkslateblue") (Figure S16). The "darkred" module was upregulated in the wild genotype (Figure S16), and the GO analysis showed *regulation of salicylic acid*, *transmembrane transport*, *cellulose synthesis*, *hydrogen peroxide catabolism* (Figure S17), *monooxygenase activity* (Figure S18), and *membrane* (Figure S19) were the enriched terms. Among the significant terms of the "darkslateblue" module, which was down-regulated in the wild genotype, were *cellular response to calcium ion* (Figure S17), *structural constituent of the cell wall* (Figure S18), and *Arp2/3 protein complex* (Figure S19).

Four gene modules differentiated the wild and domesticated genotypes in the embryos ("red," "magenta," "lightyellow," and "tan"; Figure S20). The expression of the "magenta" gene module was higher in the cultivated genotypes, particularly at 17 DAP, and the opposite pattern was observed in the other three modules (Figure S20). The terms with the lowest *p* values were *defense response to fungus*, *regulation of salicylic acid* and *hydrolase activity* in the red module; *positive regulation of transcription*, *endosperm development*, *cysteine endopeptidase inhibitor activity*, and *extracellular region* in the "magenta"; *cell wall modification*, *transmembrane transport*, *structural constituent of cytoskeleton*, and microtubules in the "lighyellow"; and in the tan module, the most significant enriched terms were *xylan biosynthetic process*, *xyloglucan metabolic process*, *abscisic acid binding*, *cellulose synthase activity*, and *membrane* (Figures S21–S23).

3.5 | Nineteen candidate genes were associated with seed coat thickness and 15 with proportion of imbibed seeds

The genes that were identified as DEGs and resided within the QTL were considered candidate genes and were further explored. Forty genes met this condition and after removing non-annotated genes, 30 candidate genes were kept. Only three were differentially expressed in seed coats, and the rest were differentially expressed in embryos. In terms of the phenotypic traits, 18 candidate genes were associated with SCT and 12 with PIS met this condition (Supporting Information 2).

Among the candidate genes identified, there were eight encoding ABA-responsive proteins, five embryonic abundant proteins, two 17.9 kDa heat shock proteins, and genes associated with pathogen response and/or with plant defense (Table 1; Supporting Information 2). The genes *Lcu.2RBY.5g046490* (*β*-galactosidase), *Lcu.2RBY.5g046640* (COL gene), *Lcu.2RBY.5g051280* (B3 domain transcription repressor VAL2), *Lcu.2RBY.5g055640* (peroxidase), and *Lcu.2RBY.7g071580* (*β*-glucosidase) were highly expressed in the cultivated elite lentil lines (Eston and 3339-3) (Figure 2). A candidate gene that showed higher expression in the wild lentil genotype was *Lcu.2RBY.5g053770* (hydroxymethylglutaryl-CoA lyase) in both seed coat and embryos at all stages (Figure 2 and Figure S24).

3.6 | Expression of candidate genes in the seed coat of RILs

Correlations between the expression of the candidate genes in the seed coat RILs at 21 DAP and the BLUP of the two phenotypic traits (SCT and PIS) were tested. After

TABLE 1 Relevant candidate genes that were differentially expressed in wild and domesticated lentil and within a quantitative trait locus (QTL) region. The "Exp. fold change" column shows the expression fold change between the wild and cultivated lentil at the tissue and time point indicated in the "tissue-days after pollination (DAP)" column. If a significant correlation was detected between the expression of the candidate gene in the recombinant inbred lines (RILs) and the seed coat thickness (SCT) or proportion of imbibed seeds (PIS), the Spearman coefficient is shown.

Gene	Gene description	QTL trait	Exp. fold change	Tissue-DAP	Exp~SCT Spearman coefficient	Exp~PIS Spearman coefficient	Notes	References
Lcu.2RBY.5g046490	*β*-Galactosidase	SCT	−1.590	Emb-28	−0.461	–	Required for the production of seed coat mucilage with correct hydration properties.	Dean et al., 2007
Lcu.2RBY.5g051280	B3 domain transcription repressor VAL2	SCT	−2.670	Emb-21	–	–	Repress DELAY OF GERMINATION1 (DOG1) expression.	Chen et al., 2020
Lcu.2RBY.5g051550	Cytochrome P450 family 71 protein	SCT	2.321	Emb-17	0.500	–	*CYP71* genes are associated with light responsiveness, hormonal regulation, and stress-related signaling	Feldmann, 2001; Sahoo et al., 2023
Lcu.2RBY.5g052860	Avr9/Cf-9 rapidly elicited protein	SCT	2.825	Coat-17	–	–	Cf-9, a receptor-like protein (RLP), confers resistance to the fungal pathogen *Cladosporium fulvum* that expresses the Avr9 avirulence gene. It is related to the initial development of a defense response.	Nekrasov et al., 2006; Romeis et al., 2000; Rowland et al., 2005
Lcu.2RBY.5g053770	Hydroxymethylglutaryl-CoA lyase	SCT	4.076 3.344 3.436 2.265	Emb-17 Emb-21 Emb-28 Coat-13	0.647	–	Induced expression of a cytosolic HMGL35 version in *A. thaliana*, which delays germination and leads to rapid wilting and chlorosis.	Hemmerlin et al., 2019
Lcu.2RBY.5g055640	Peroxidase	SCT	−2.521	Emb-28	−0.680	–	Secretory plant peroxidases are correlated with lignification and suberization.	Henriksen et al., 2001
Lcu.2RBY.5g055700	Subtilisin-like serine endopeptidase family protein	SCT	3.3151	Emb-28	0.655	–	Subtilisin-like proteases are a diverse family of serine peptidases present mostly in plants. They are involved in a broad spectrum of biological functions: protein turnover, plant development, interaction with the environment and plant defense response.	Figueir et al., 2018; Tornero et al., 1997; Vartapetian et al., 2011; Yamagata et al., 2000
Lcu.2RBY.7g070530	Pathogenesis-related protein 10b	PIS	2.617 2.664	Emb-21 Emb-28	–	−0.777	The protein family plays a role in growth development and in biotic stress response	Lopes et al., 2023
Lcu.2RBY.7g071580	*β*-Glucosidase	PIS	−1.674	Emb-21	–	0.552	Abundant in non-dormant dry seeds in *Arabidopsis*. Up-accumulation of *β*-1,3 glucanase is correlated with the release of seed coat-imposed dormancy in tobacco	Chibani et al., 2006; Leah et al., 1995

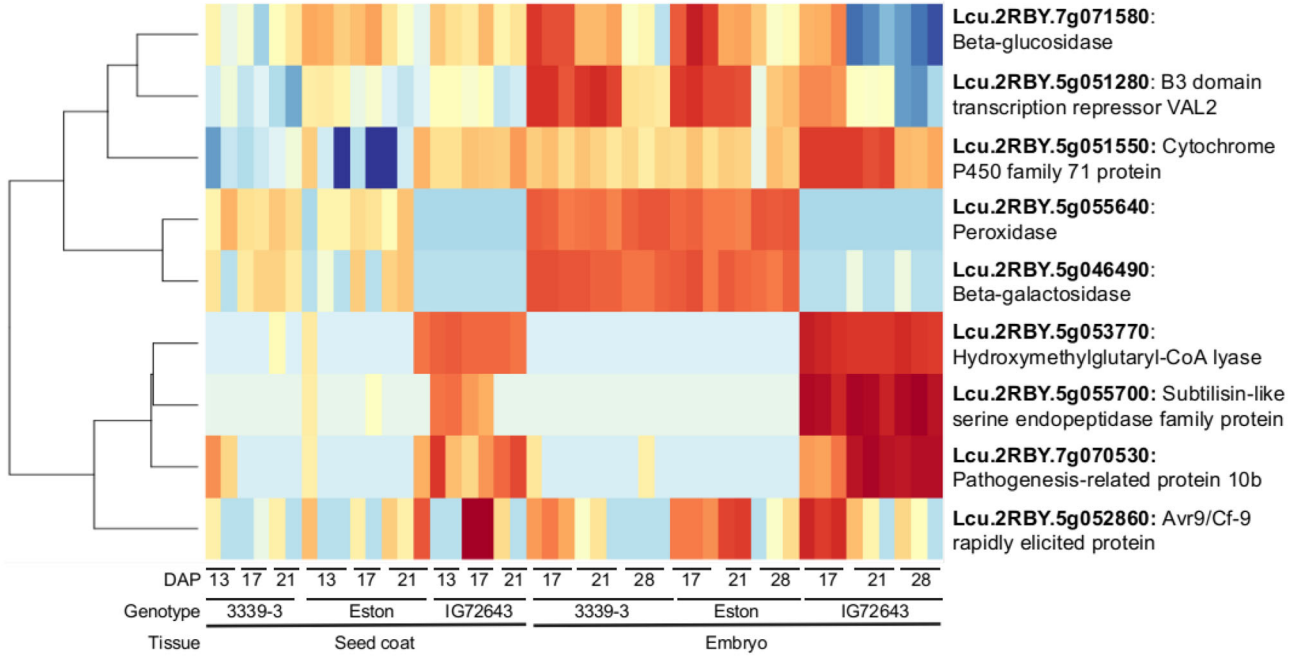

FIGURE 2 Relative gene expression levels of the relevant candidate in two domesticated lentil genotypes (Eston and 3339-3) and one wild lentil genotype (IG 72643) in embryos and seed coats sampled at different time points in terms of days after pollination (DAP).

correcting the *p*-value for the multiple tests, eight correlations were significant ($p < 0.05$) for SCT and two for PIS (Figures S25 and S26; Supporting Information 2). Candidate genes whose expression was positively correlated with SCT were *Lcu.2RBY.5g051550* (cytochrome P450 family 71 protein), *Lcu.2RBY.5g053770*, a hydroxymethylglutaryl-CoA lyase; *Lcu.2RBY.5g055700*, a subtilisin-like serine endopeptidase family protein; and the 17.9 kDa heat shock protein, *Lcu.2RBY.5g056140*. The expression of the candidate genes *Lcu.2RBY.5g046490* (*β*-galactosidase), *Lcu.2RBY.5g055640* (peroxidase), *Lcu.2RBY.5g050180* (embryonic abundant protein), and *Lcu.2RBY.7g070450* (ABA-responsive protein) was negatively correlated with SCT (Table 1; Supporting Information 2; Figure S25).

Only two significant correlations were found between expression in the RILs and PIS: a positive one for *Lcu.2RBY.7g071580* (*β*-glucosidase) and a negative correlation for *Lcu.2RBY.7g070530* (pathogenesis-related protein 10b) (Table 1; Supporting Information 2; Figure S26).

4 | DISCUSSION

Seed dormancy and imbibition capacity are two of the phenotypic traits modified during domestication and contribute to the phenomenon known as domestication syndrome (Hammer, 1984). Nonetheless, not all the phenotypic changes observed between wild and cultivated genotypes are true domestication traits, and Abbo et al. (2014) proposed the concept of "crucial domestication traits" to differentiate the domestication traits that are critical for cultivation and show clear dimorphism between wild and cultivated plants. In grain legumes, pod shattering, seed size, and dormancy are considered domestication traits, but only the latter seems to be crucial in cool-season legumes, such as lentil (Abbo et al., 2014).

4.1 | Seed coat thickness has a weak effect on the proportion of imbibed seeds

The induction and release of dormancy are controlled by intrinsic factors, such as phytohormones and specific regulators, that are influenced by environmental variables during seed maturation and storage (Née et al., 2017). In legumes, the prevalent type of dormancy is physical, in which the seed coat is impermeable to water, preventing imbibition until it is breached by factors in the soil (Smýkal et al., 2014; Wyse & Dickie, 2018).

The rate of water uptake by the dry seeds is controlled by the chemical composition, structural characteristics of the seed coat, and the water availability in the environment (Smýkal et al., 2014). In this study, we evaluated whether imbibition capacity, in terms of PIS, and SCT from an interspecific lentil population (LR-68) were consistent across different years and locations, and if these two traits were correlated. Our results showed that both traits are consistent throughout years and locations (Figure S4), particularly SCT,

suggesting they both are highly heritable. Nevertheless, PIS data had a higher data spreading, indicating that it is more affected by the environment than SCT, which has been previously reported in other species (Kissing Kucek et al., 2020; Penfield & MacGregor, 2017). The high heritability of these two traits and the relatively few QTL associated with them indicate that a few loci might be controlling them.

In other legumes, seed coat permeability is highly affected by its thickness. For example, seeds with thickened coats reduce their imbibition capacity in soybean (Jang et al., 2015; Noodén et al., 1985), in common bean (Balasubramanian et al., 2004; Wyatt, 1977), *Medicago*, and *Trifolium* (Russi et al., 1992). Our results show that this is not the case in lentil as only two significant negative correlations were detected between SCT and PIS from different site years (SCT-Pres-2018~Imb-CRep-2020 and SCT-Pres-2018~Imb-SPG-2019; Figure S4), and when BLUPs of PIS and SCT were tested, no correlation was found. The PCA applied to the phenotypic data was consistent with the results of the correlations since PC1 mainly explained SCT and PC2 explained PIS. A similar pattern of independence of the two traits was observed in pea, another cool-season legume (Hradilová et al., 2019; Williams et al., 2024).

4.2 | More genes were expressed in seed coats but more were differentially expressed in the embryos

More genes were expressed in the seed coat of the wild and cultivated genotypes at the three stages tested compared to the embryo samples (Figure S7). Despite this, a higher number of DEGs were detected in embryos. Furthermore, the number of expressed genes tended to decrease in the embryos at later developmental stages while the number of DEGs increased (Figure S9). Since seed development involves complex changes and wide phenotypic changes that occur during lentil domestication, it is expected that the identified DEGs are involved not only in PIS and SCT but also reflect other processes occurring during seed development. It is also worth noting that developmental stages were established according to DAP, and it is plausible that the development dynamics of the cultivated lentil were modified during the domestication. For example, in pea the seed development in cultivated genotypes is faster than in the wild ones (Balarynová et al., 2023). Furthermore, the communication between embryos and seed coats is crucial in the coordination of the developmental processes (Robert, 2019; W. Wang et al., 2022).

An interesting pattern was that upregulated DEGs were more frequent in the wild genotype (Figure S10; Table S2). This suggests that during domestication the suppression of the expression of some genes might have occurred. This has been reported in crops such as tomato (Sauvage et al., 2017), and a reduction in gene expression diversity was found in rice, cotton (Liu et al., 2019), common bean (Bellucci et al., 2014), and pea (Klčová et al., 2024).

The identification of DEGs to pinpoint candidate genes associated with phenotypic traits is a common approach (e.g., Guan et al., 2023; Verma et al., 2021; K. Wang et al., 2024; Yow et al., 2023; Zhao et al., 2022). However, phenotypic changes resulting from domestication may not involve changes in gene expression. In such cases, alterations in the gene product, rather than changes in expression levels, could be responsible for the observed phenotypic variations. For example, in soybean, a SNP caused an amino acid substitution that modified the action of endo-1,4-β-glucanase, affecting coat permeability without altering expression levels in the cultivar compared to the wild genotype (Jang et al., 2015). Nevertheless, it has been suggested that alleles influencing gene expression are among the most common genetic changes during crop evolution (Meyer & Purugganan, 2013).

4.3 | Glycosyl-hydrolases may influence both SCT and PIS capacity in lentil

Evidence from some legumes about the genetic mechanisms of physical dormancy suggests that changes in the level of dormancy are controlled by few loci (e.g., Isemura et al., 2012; Kongjaimun et al., 2012; Sedivy et al., 2017; Soltani et al., 2021), which is consistent with the early report of major-effect loci controlling seed dormancy in lentil by Ladizinsky et al. (1985). In the interspecific populations evaluated (LR-68), only two QTLs associated with each of the evaluated traits were identified (Figure 1; Supporting Information 2), supporting relatively few genes with significant contribution to PIS and SCT.

From a physiological perspective, impermeable seeds often have a hard, pectinaceous outer layer (Soltani et al., 2021; Werker et al., 1979), higher lignin content (Kannenberg & Allard, 1964), high amount of hydroxylated fatty acids (Shao et al., 2007), and/or the absence of pits in the seed coat (Soltani et al., 2021; Sun & Gong, 2024). The cell wall also plays an important role in determining permeability. It is constituted by a network composed mainly of cellulose microfibrils within a hydrated gel-type matrix of pectin and hemicelluloses (xyloglucan, mixed-linkage glucan, mannans, and xylans), and the degradation of these complex polysaccharides requires glycosyl-hydrolases (Fry, 2004; Holland et al., 2020). For example, in *M. truncatula* homologous glycosyl hydrolase genes have been associated with seed dormancy (Renzi et al., 2020) affects seed coat permeability in soybean by controlling the amount of β-1,4-glucans in the seed coat (Jang et al., 2015); in tobacco seeds, an accumulation of β-1,3 glucanase and the resulting hydrolysis of β-glucans correlated with the release of seed-coat

imposed dormancy (Leubner-Metzger, 2005) and the expression of β-1,3 glucanase is also elevated during pea seed germination (Petruzzelli et al., 1999). In our study, several glycosyl-hydrolases were among DEGs in the enriched GO terms and KEGG pathways (Figures S13 and S15), and in the gene modules from the coexpression networks (Figures S17 and S18).

A candidate gene that might be important in determining PIS is *Lcu.2RBY.7g071580*, which encodes the glycosyl hydrolases β-glucosidase (Figure 1; Supplementary Information 2). β-Glucosidases seem to be involved in both the breakdown of oligosaccharides produced during cell wall turnover and the release of monolignols from their glycosides, which enables lignification to strengthen secondary cell walls (Ketudat Cairns & Esen, 2010). β-Glucosidases have been associated with dormancy in plants. For example, a β-glucosidase is abundant in non-dormant dry seeds of *Arabidopsis* (Chibani et al., 2006), and in germinating barley seeds, a β-glucosidase is implicated in sugar liberation from β-glucans derived from cell wall polysaccharides (Leah et al., 1995). Interestingly, in barley, the β-glucosidase is abundant in the endosperm and is synthesized during the development of the grain and is stored until germination (Simos et al., 1994). In lentil, the domesticated genotypes showed higher expression compared to the wild lentil genotype (Figure 2), consistent with the hypothesis that β-glucosidase is involved in the hydrolytic breakdown of carbohydrates in the cell wall, leading to the release of dormancy imposed by the seed coat. Furthermore, the expression of β-glucosidase in the RILs was positively correlated with PIS (Figure S26).

Regarding SCT, a β-galactosidase, encoded by the *Lcu.2RBY.5g046490* gene, was associated with this trait. β-galactosidases, which are part of the glycosyl hydrolase family, are essential for the production of pectins in the cell wall in *Arabidopsis*. These pectins play a crucial role in providing its seed coat with the proper hydration properties for mucilage formation (Dean et al., 2007). No mucilage production occurs in the seed coat of legumes. Still, this protein is also involved in cell-to-cell adhesion (Šola et al., 2019), and in higher plants, β-galactosidase is the only enzyme capable of hydrolyzing galactosyl residues from the cell wall to release galactose during a variety of biological processes, including cell wall expansion and degradation, affecting the integrity and texture of tissues (Arunraj et al., 2020; Kishore & Kayastha, 2012). In species such as barley (Giannakouros et al., 1991), rice (Chantarangsee et al., 2007), and lupin (Buckeridge & Reid, 1994), the activity of this enzyme increases during germination. Consistent with this, this candidate gene was highly expressed in embryos of the domesticated genotypes at all stages examined (Figure 2).

Glycosyl-hydrolases, such as β-glucosidase, β-1,4-mannosyl-glycoprotein, and alpha-L-arabinofuranosidase, were also present in the enriched GO terms (Figure S13) and KEGG pathways (Figure S15). Although the degradation of pectins is complex and involves a larger number of enzymes, *pectinesterase* KEGG pathway was also enriched (Figure S15), and Soltani et al. (2021) reported a pectin acetylesterase 8 as a major contributor to seed dormancy in common bean. Additional enzymes and terms related to the synthesis of compounds of the cell wall were also enriched and differentially expressed. Among them were xyloglucan:xyloglucosyl transferase TCH4, xyloglucan 6-xylosyltransferase, and genes related to cellulose synthesis, along with other cell wall-related pathways (Figures S13, S15, S17, and S22; Supplementary Information 3). Notably, homologous genes were also found upregulated in pea seed coat (Hradilová et al., 2017).

Our results suggest that differential changes in cell wall composition between the wild and domesticated genotypes may be associated with variations in PIS and SCT in lentil, and that these changes might have been the result of the domestication. A limitation of our work is that expression changes associated with the differences in SCT and PIS could occur at different stages that were not sampled.

4.4 | Polyphenolic compounds and lentil seed properties

Lentil seeds are rich in secondary metabolites, including polyphenols, and most of these components are in the seed coat. Seeds with dark seed coats tend to have higher levels of phenolic compounds in comparison with pale seeds (Singh et al., 2017; Xu et al., 2007), and they might affect seed coat properties, including permeability (Penfield & MacGregor, 2017). These properties have been observed in wild pea (Balarynová et al., 2023) and chickpea seeds (Sedláková et al., 2021, 2023).

Peroxidases oxidize phenolic substrates and are involved in cell wall modulation, lignification and suberization (Henriksen et al., 2001). A peroxidase (*Lcu.2RBY.5g055640*) found within a QTL associated with SCT was differentially expressed and showed a negative correlation between expression in RILs and SCT (Table 1). Other three genes encoding peroxidases *(Lcu.2RBY.1g015620*, *Lcu.2RBY.1g044820*, and *Lcu.2RBY.4g042090*) were differentially expressed between wild and cultivated parents. *Hydrogen peroxide catabolism* (Figure S17) and *peroxidase activity* (Figure S18) were enriched in the "darkred" co-expression GO module in the seed coats.

Phenolic compounds are also components of cutin and suberin, along with glycerol and fatty acids (Pollard et al., 2008; Renard et al., 2021). The presence of cutin and/or suberin in seed coats (Chai et al., 2016; Fedi et al., 2017), and in the endosperm (De Giorgi et al., 2015; Panikashvili et al., 2009) is another factor that contributes to regulating seed

dormancy and longevity. Their occurrence is still a matter of debate in legumes (Janská et al., 2019) and has been demonstrated so far only in soybean (Ranathunge et al., 2010; Shao et al., 2007). GO terms associated with glycerol, another component of cutin and suberin, were detected in the "darkred" and "darksateblue" gene modules in seed coats. Specifically, *glycerol metabolic process* (Figure S17), *indole-3-glycerol-phosphate synthase activity*, and *1-acylglycerol-3-phosphate O-acyltransferase activity* (Figure S18) terms were enriched. These changes in the expression patterns of genes associated with the components of cutin and suberin might lead to modifications in seed coat properties, including water permeability.

4.5 | Other expression changes relevant for domesticated lentil

It is well known that plant hormones, such as ABA and gibberellins (GA), play a crucial role in dormancy and germination. ABA is required for the induction of dormancy during seed maturation and GA for germination (Gazzarrini & Tsai, 2015; Graeber et al., 2012). Recently, it was shown that ABA/GA has also been altered during pea domestication (Balarynová et al., 2023). Eleven ABA-related proteins were overexpressed in the wild embryos (Supplementary Information 2), and enriched GO terms related to ABA were also identified (Figures S12, S13, and S22). The candidate gene *Lcu.2RBY.5g051280* encodes a B3 domain transcription repressor VIVIPAROUS1/ABI3-LIKE 2 (VAL2) that silences seed dormancy by regulating DOG1 expression in *Arabidopsis* (Chen et al., 2020). *DOG1* is the major genetic regulator of dormancy in *Arabidopsis* by affecting the expression of ABA and GA metabolic genes (Li et al., 2022; Shu et al., 2016) and regulating properties of the endosperm, which acts as a germination barrier to control coat dormancy (Graeber et al., 2014). This gene might also play a role in determining dormancy in lentil.

5 | CONCLUSIONS

Analogous selection during the domestication process resulted in phenotypic convergence in common crop traits, and the loss of dormancy is one of the key traits that are part of the domestication syndrome. Evidence from different crops has shown that the molecular and genetic mechanisms that resulted in the loss of dormancy are not necessarily the same across different species. Furthermore, imbibition capacity is a quantitative and complex trait that is influenced by several genes with different contributions and by the environment.

Imbibition capacity in lentil is not affected by SCT, even though both traits are heritable. By using an interspecific population and a transcriptomic approach, thousands of DEGs were identified in the seed coat and embryos. However, QTL mapping suggested that relatively few genomic regions have a strong contribution to both traits. By combining the results of both approximations, we were able to detect two glycosyl-hydrolases that may have a high contribution to determining SCT and imbibition capacity in lentil: a β-galactosidase and a β-glucosidase, respectively. These enzymes might affect the evaluated phenotypic traits by altering the cell wall properties.

A better understanding of the genetic basis of important domestication traits and crop evolution might be translated into the practical use of wild relatives in breeding programs.

AUTHOR CONTRIBUTIONS

Azalea Guerra-García: Conceptualization; formal analysis; investigation; methodology; writing—original draft. **Jana Balarynová**: Formal analysis; methodology; validation; writing—review and editing. **Petr Smykal**: Conceptualization; investigation; writing—review and editing. **Eric J von Wettberg**: Conceptualization; investigation; supervision; writing—review and editing. **Scott D. Noble**: Software; writing—review and editing. **Kirstin E. Bett**: Conceptualization; funding acquisition; project administration; supervision; writing—review and editing.

ACKNOWLEDGMENTS

This work was conducted as part of the EVOLVES project funded by Genome Canada [grant: LSP18-16302] and managed by Genome Prairie. We are grateful for the matching financial support from Western Grains Research Foundation [grant: GC1903], the Government of Saskatchewan [ADF grant: 20200026], and the University of Saskatchewan. Petr Smykal and Jana Balarynová work was supported from Palacky University Grant Agency (PrF-2023-001 and PrF-2024-001), and the Ministry of Education, Youth and Sports (LUAUS25035). We thank the USask Pulse Crop Field Lab technical staff as well as Didier Socquet-Juglard, Derek Wright and Sandesh Neupane for their assistance with plant and seed production and management and Akiko Tomita for assistance with the imbibition tests and OCT imaging.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The genetic map of the interspecific population, gene expression and phenotypic data can be retrieved at https://knowpulse.usask.ca/study/Lentil-Seed-Dormancy-Imbibition-Analysis. The scripts are available at https://github.com/AzaleaGuerra. The README file contains further details about the script found in the repository.

ORCID

Azalea Guerra-García https://orcid.org/0000-0001-8132-467X
Petr Smykal https://orcid.org/0000-0002-6117-8510
Eric J von Wettberg https://orcid.org/0000-0002-2724-0317
Kirstin E. Bett https://orcid.org/0000-0001-7959-6959

REFERENCES

Abbo, S., Pinhasi van-Oss, R., Gopher, A., Saranga, Y., Ofner, I., & Peleg, Z. (2014). Plant domestication versus crop evolution: A conceptual framework for cereals and grain legumes. *Trends in Plant Science*, *19*, 351–360. https://doi.org/10.1016/j.tplants.2013.12.002

Abbo, S., Saranga, Y., Peleg, Z., Kerem, Z., Lev-Yadun, S., & Gopher, A. (2009). Reconsidering domestication of legumes versus cereals in the ancient near east. *The Quarterly Review of Biology*, *84*, 29–50. https://doi.org/10.1086/596462

Alexa, A., & Rahnenfuhrer, J. (2022). topGO: Enrichment Analysis for Gene Ontology (R package version 2.50.0) [Computer software]. CRAN.

Arunraj, R., Skori, L., Kumar, A., Hickerson, N. M., Shoma, N., & Samuel, M. A. (2020). Spatial regulation of alpha-galactosidase activity and its influence on raffinose family oligosaccharides during seed maturation and germination in. *Plant Signaling & Behavior*, *15*, 1709707. https://doi.org/10.1080/15592324.2019.1709707

Balarynová, J., Klčová, B., Tarkowská, D., Turečková, V., Trněný, O., Špundová, M., Ochatt, S., & Smýkal, P. (2023). Domestication has altered the ABA and gibberellin profiles in developing pea seeds. *Planta*, *258*, Article 25. https://doi.org/10.1007/s00425-023-04184-2

Balasubramanian, P., Vandenberg, A., & Hucl, P. (2004). Planting date and suboptimal seedbed temperature effects on dry bean establishment, phenology and yield. *Canadian Journal of Plant Science [Revue canadienne de phytotechnie]*, *84*, 31–36. https://doi.org/10.4141/p02-185

Bellucci, E., Bitocchi, E., Ferrarini, A., Benazzo, A., Biagetti, E., Klie, S., Minio, A., Rau, D., Rodriguez, M., Panziera, A., Venturini, L., Attene, G., Albertini, E., Jackson, S. A., Nanni, L., Fernie, A. R., Nikoloski, Z., Bertorelle, G., Delledonne, M., & Papa, R. (2014). Decreased nucleotide and expression diversity and modified coexpression patterns characterize domestication in the common bean. *The Plant Cell*, *26*, 1901–1912. https://doi.org/10.1105/tpc.114.124040

Bewley, J. (1997). Seed germination and dormancy. *The Plant Cell*, *9*, 1055–1066. https://doi.org/10.1105/tpc.9.7.1055

Bewley, J., & Black, M. (1994). *Seeds: Physiology of development and germination*. Springer US. https://doi.org/10.1007/978-1-4899-1002-8

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bradford, K. J., & Nonogaki, H. (Eds.). (2007). *Seed development, dormancy and germination*. Blackwell Publishing Ltd.

Broman, K. W., Gatti, D. M., Simecek, P., Furlotte, N. A., Prins, P., Sen, Ś., Yandell, B. S., & Churchill, G. A. (2019). R/qtl2: Software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics*, *211*, 495–502. https://doi.org/10.1534/genetics.118.301595

Buckeridge, M. S., & Reid, J. S. (1994). Purification and properties of a novel beta-galactosidase or exo-(1→4)-β-D-galactanase from the cotyledons of germinated *Lupinus angustifolius* L. seeds. *Planta*, *192*, 502–511. https://doi.org/10.1007/BF00203588

Cao, Z., Socquet-Juglard, D., Daba, K., Vandenberg, A., & Bett, K. E. (2024). Understanding genome structure facilitates the use of wild lentil germplasm for breeding: A case study with shattering loci. *The Plant Genome*, *17*(2). https://doi.org/10.1002/tpg2.20455

Carlson, M. (2019). GO.db: A Set of Annotation Maps Describing the Entire Gene Ontology (R package version 3.8.2) [Computer software]. CRAN.

Chai, M., Zhou, C., Molina, I., Fu, C., Nakashima, J., Li, G., Zhang, W., Park, J., Tang, Y., Jiang, Q., & Wang, Z.-Y. (2016). A class II KNOX gene, KNOX4, controls seed physical dormancy. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 6997–7002. https://doi.org/10.1073/pnas.1601256113

Chantarangsee, M., Tanthanuch, W., Fujimura, T., Fry, S. C., & Ketudat Cairns, J. (2007). Molecular characterization of β-galactosidases from germinating rice (*Oryza sativa*). *Plant Science: An International Journal of Experimental Plant Biology*, *173*, 118–134. https://doi.org/10.1016/j.plantsci.2007.04.009

Chen, N., Wang, H., Abdelmageed, H., Veerappan, V., Tadege, M., & Allen, R. D. (2020). HSI2/VAL1 and HSL1/VAL2 function redundantly to repress DOG1 expression in Arabidopsis seeds and seedlings. *The New Phytologist*, *227*, 840–856. https://doi.org/10.1111/nph.16559

Chibani, K., Ali-Rachedi, S., Job, C., Job, D., Jullien, M., & Grappin, P. (2006). Proteomic analysis of seed dormancy in Arabidopsis. *Plant Physiology*, *142*, 1493–1510. https://doi.org/10.1104/pp.106.087452

Cichy, K. A., Wiesinger, J. A., Berry, M., Nchimbi-Msolla, S., Fourie, D., Porch, T. G., Ambechew, D., & Miklas, P. N. (2019). The role of genotype and production environment in determining the cooking time of dry beans (*Phaseolus vulgaris* L.). *Legume Science*, *1*, e13. https://doi.org/10.1002/leg3.13

Cichy, K. A., Wiesinger, J. A., & Mendoza, F. A. (2015). Genetic diversity and genome-wide association analysis of cooking time in dry bean (*Phaseolus vulgaris* L.). *Theoretical and applied genetics [Theoretische und angewandte Genetik]*, *128*, 1555–1567. https://doi.org/10.1007/s00122-015-2531-z

Coyne, C. J., Kumar, S., Wettberg, E. J. B., Marques, E., Berger, J. D., Redden, R. J., Ellis, T. H. N., Brus, J., Zablatzká, L., & Smýkal, P. (2020). Potential and limits of exploitation of crop wild relatives for pea, lentil, and chickpea improvement. *Legume Science*, *2*, e36. https://doi.org/10.1002/leg3.36

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*, giab008. https://doi.org/10.1093/gigascience/giab008

Dean, G. H., Zheng, H., Tewari, J., Huang, J., Young, D. S., Hwang, Y. T., Western, T. L., Carpita, N. C., McCann, M. C., Mansfield, S. D., & Haughn, G. W. (2007). The Arabidopsis MUM2 gene encodes a beta-galactosidase required for the production of seed coat mucilage with correct hydration properties. *The Plant Cell*, *19*, 4007–4021. https://doi.org/10.1105/tpc.107.050609

De Giorgi, J., Piskurewicz, U., Loubery, S., Utz-Pugin, A., Bailly, C., Mène-Saffrané, L., & Lopez-Molina, L. (2015). An endosperm-associated cuticle is required for arabidopsis seed viability, dormancy and early control of germination. *PLoS Genetics*, *11*, e1005708. https://doi.org/10.1371/journal.pgen.1005708

Fedi, F., O'Neill, C. M., Menard, G., Trick, M., Dechirico, S., Corbineau, F., Bailly, C., Eastmond, P. J., & Penfield, S. (2017). Awake1, an

ABC-type transporter, reveals an essential role for suberin in the control of seed dormancy. *Plant Physiology*, *174*, 276–283. https://doi.org/10.1104/pp.16.01556

Feldmann, K. A. (2001). Cytochrome P450s as genes for crop improvement. *Current Opinion in Plant Biology*, *4*, 162–167. https://doi.org/10.1016/s1369-5266(00)00154-0

Figueiredo, J., Sousa Silva, M., & Figueiredo, A. (2018). Subtilisin-like proteases in plant defence: The past, the present and beyond. *Molecular Plant Pathology*, *19*, 1017–1028. https://doi.org/10.1111/mpp.12567

Finch-Savage, W. E., & Leubner-Metzger, G. (2006). Seed dormancy and the control of germination. *New Phytologist*, *171*(3), 501–523. https://doi.org/10.1111/j.1469-8137.2006.01787.x

Frazee, A. C., Pertea, G., Jaffe, A. E., Langmead, B., Salzberg, S. L., & Leek, J. T. (2015). Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nature Biotechnology*, *33*, 243–246. https://doi.org/10.1038/nbt.3172

Fry, S. C. (2004). Primary cell wall metabolism: Tracking the careers of wall polymers in living plant cells. *The New Phytologist*, *161*, 641–675. https://doi.org/10.1111/j.1469-8137.2004.00980.x

Gazzarrini, S., & Tsai, A. Y.-L. (2015). Hormone cross-talk during seed germination. *Essays in Biochemistry*, *58*, 151–164. https://doi.org/10.1042/bse0580151

Giannakouros, T., Karagiorgos, A., & Simos, G. (1991). Expression of β-galactosidase multiple forms during barley (*Hordeum vulgare*) seed germination. Separation and characterization of enzyme isoforms. *Physiologia Plantarum*, *82*, 413–418. https://doi.org/10.1111/j.1399-3054.1991.tb02926.x

Graeber, K., Linkies, A., Steinbrecher, T., Mummenhoff, K., Tarkowská, D., Turečková, V., Ignatz, M., Sperber, K., Voegele, A., de Jong, H., Urbanová, T., Strnad, M., & Leubner-Metzger, G. (2014). DELAY OF GERMINATION 1 mediates a conserved coat-dormancy mechanism for the temperature- and gibberellin-dependent control of seed germination. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, E3571–E3580. https://doi.org/10.1073/pnas.1403851111

Graeber, K., Nakabayashi, K., Miatton, E., Leubner-Metzger, G., & Soppe, W. J. J. (2012). Molecular mechanisms of seed dormancy. *Plant, Cell & Environment*, *35*, 1769–1786. https://doi.org/10.1111/j.1365-3040.2012.02542.x

Guan, M., Shi, X., Chen, S., Wan, Y., Tang, Y., Zhao, T., Gao, L., Sun, F., Yin, N., Zhao, H., Lu, K., Li, J., & Qu, C. (2023). Comparative transcriptome analysis identifies candidate genes related to seed coat color in rapeseed. *Frontiers in Plant Science*, *14*, 1154208. https://doi.org/10.3389/fpls.2023.1154208

Guerra-García, A., Gioia, T., von Wettberg, E., Logozzo, G., Papa, R., Bitocchi, E., & Bett, K. E. (2021). Intelligent characterization of lentil genetic resources: Evolutionary history, genetic diversity of germplasm, and the need for well-represented collections. *Current Protocols*, *1*, e134. https://doi.org/10.1002/cpz1.134

Guerra-Garcia, A., Haile, T., Ogutcen, E., Bett, K. E., & von Wettberg, E. J. (2022). An evolutionary look into the history of lentil reveals unexpected diversity. *Evolutionary Applications*, *15*, 1313–1325. https://doi.org/10.1111/eva.13467

Guerra-García, A., Trněný, O., Brus, J., Renzi, J. P., Kumar, S., Bariotakis, M., Coyne, C. J., Chitikineni, A., Bett, K. E., Varshney, R., Pirintsos, S., Berger, J., von Wettberg, E. J. B., & Smýkal, P. (2024). Genetic structure and ecological niche space of lentil's closest wild relative, *Lens orientalis* (Boiss.) Schmalh. *Plant Biology*, *26*(2), 232–244. https://doi.org/10.1111/plb.13615

Hammer, K. (1984). Das Domestikationssyndrom. In H. Böhme, A. Rieth, W. Müller-Stoll, R. Rieger, A. Rieth, H. Sagromsky, & H. Stubbe (Eds.), *Band 32* (pp. 269–294). De Gruyter. https://doi.org/10.1515/9783112642641-045

Hemmerlin, A., Huchelmann, A., Tritsch, D., Schaller, H., & Bach, T. J. (2019). The specific molecular architecture of plant 3-hydroxy-3-methylglutaryl-CoA lyase. *The Journal of Biological Chemistry*, *294*, 16186–16197. https://doi.org/10.1074/jbc.RA119.008839

Henriksen, A., Mirza, O., Indiani, C., Teilum, K., Smulevich, G., Welinder, K. G., & Gajhede, M. (2001). Structure of soybean seed coat peroxidase: A plant peroxidase with unusual stability and haem-apoprotein interactions. *Protein Science: A Publication of the Protein Society*, *10*, 108–115. https://doi.org/10.1110/ps.37301

Holland, C., Ryden, P., Edwards, C. H., & Grundy, M. M.-L. (2020). Plant cell walls: Impact on nutrient bioaccessibility and digestibility. *Foods*, *9*(2), 201. https://doi.org/10.3390/foods9020201

Hradilová, I., Duchoslav, M., Brus, J., Pechanec, V., Hýbl, M., Kopecký, P., Smržová, L., Štefelová, N., Vaclávek, T., Bariotakis, M., Machalová, J., Hron, K., Pirintsos, S., & Smýkal, P. (2019). Variation in wild pea (*Pisum sativum* subsp. elatius) seed dormancy and its relationship to the environment and seed coat traits. *PeerJ*, *7*, e6263. https://doi.org/10.7717/peerj.6263

Hradilová, I., Trněný, O., Válková, M., Cechová, M., Janská, A., Prokešová, L., Aamir, K., Krezdorn, N., Rotter, B., Winter, P., Varshney, R. K., Soukup, A., Bednář, P., Hanáček, P., & Smýkal, P. (2017). A combined comparative transcriptomic, metabolomic, and anatomical analyses of two key domestication traits: Pod dehiscence and seed dormancy in pea (*Pisum* sp.). *Frontiers in Plant Science*, *8*, 542. https://doi.org/10.3389/fpls.2017.00542

Isemura, T., Kaga, A., Tabata, S., Somta, P., Srinives, P., Shimizu, T., Jo, U., Vaughan, D. A., & Tomooka, N. (2012). Construction of a genetic linkage map and genetic analysis of domestication related traits in mungbean (*Vigna radiata*). *PLoS One*, *7*, e41304. https://doi.org/10.1371/journal.pone.0041304

Jang, S.-J., Sato, M., Sato, K., Jitsuyama, Y., Fujino, K., Mori, H., Takahashi, R., Benitez, E. R., Liu, B., Yamada, T., & Abe, J. (2015). A single-nucleotide polymorphism in an endo-1,4-β-glucanase gene controls seed coat permeability in soybean. *PLoS One*, *10*, e0128527. https://doi.org/10.1371/journal.pone.0128527

Janská, A., Pecková, E., Sczepaniak, B., Smýkal, P., & Soukup, A. (2019). The role of the testa during the establishment of physical dormancy in the pea seed. *Annals of Botany*, *123*, 815–829. https://doi.org/10.1093/aob/mcy213

Kannenberg, L. W., & Allard, R. W. (1964). An association between pigment and lignin formation in the seed coat of the lima bean. *Crop Science*, *4*, 621–622. https://doi.org/10.2135/cropsci1964.0011183X000400060021x

Ketudat Cairns, J. R., & Esen, A. (2010). β-Glucosidases. *Cellular and Molecular Life Sciences*, *67*, 3389–3405. https://doi.org/10.1007/s00018-010-0399-2

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*, 907–915. https://doi.org/10.1038/s41587-019-0201-4

Kishore, D., & Kayastha, A. M. (2012). A β-galactosidase from chick pea (*Cicer arietinum*) seeds: Its purification, biochemical properties

and industrial applications. *Food Chemistry*, *134*, 1113–1122. https://doi.org/10.1016/j.foodchem.2012.03.032

Kissing Kucek, L., Azevedo, M. D., Eagen, S. S., Ehlke, N. J., Hayes, R. J., Mirsky, S. B., Reberg-Horton, C., Ryan, M. R., Wayman, S., Wiering, N. P., & Riday, H. (2020). Seed dormancy in hairy vetch (*Vicia villosa* Roth) is influenced by genotype and environment. *Agronomy*, *10*, 1804. https://doi.org/10.3390/agronomy10111804

Klčová, B., Balarynová, J., Trněný, O., Krejčí, P., Cechová, M. Z., Leonova, T., Gorbach, D., Frolova, N., Kysil, E., Orlova, A., Ihling, C., Frolov, A., Bednář, P., & Smýkal, P. (2024). Domestication has altered gene expression and secondary metabolites in pea seed coat. *The Plant Journal: For Cell and Molecular Biology*, *118*, 2269–2295. https://doi.org/10.1111/tpj.16734

Kongjaimun, A., Kaga, A., Tomooka, N., Somta, P., Vaughan, D. A., & Srinives, P. (2012). The genetics of domestication of yard-long bean, *Vigna unguiculata* (L.) Walp. ssp. *unguiculata* cv.-gr. *sesquipedalis*. *Annals of Botany*, *109*, 1185–1200. https://doi.org/10.1093/aob/mcs048

Kovaka, S., Zimin, A. V., Pertea, G. M., Razaghi, R., Salzberg, S. L., & Pertea, M. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biology*, *20*, Article 278. https://doi.org/10.1186/s13059-019-1910-1

Ladizinsky, G., Cohen, D., & Muehlbauer, F. J. (1985). Hybridization in the genus Lens by means of embryo culture. *Theoretical and Applied Genetics*, *70*(1), 97–101. https://doi.org/10.1007/bf00264489

Lamont, B. B., & Pausas, J. G. (2023). Seed dormancy revisited: Dormancy-release pathways and environmental interactions. *Functional Ecology*, *37*(4), 1106–1125. https://doi.org/10.1111/1365-2435.14269

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1). https://doi.org/10.1186/1471-2105-9-559

Laosatit, K., Amkul, K., Chen, J., Lin, Y., Yuan, X., Wang, L., Chen, X., & Somta, P. (2022). A Class II KNOX Gene, *KNAT7-1*, regulates physical seed dormancy in mungbean [*Vigna radiata* (L.) Wilczek]. *Frontiers in Plant Science*, *13*, 852373. https://doi.org/10.3389/fpls.2022.852373

Leah, R., Kigel, J., Svendsen, I., & Mundy, J. (1995). Biochemical and molecular characterization of a barley seed beta-glucosidase. *The Journal of Biological Chemistry*, *270*, 15789–15797. https://doi.org/10.1074/jbc.270.26.15789

Leubner-Metzger, G. (2005). beta-1,3-Glucanase gene expression in low-hydrated seeds as a mechanism for dormancy release during tobacco after-ripening. *The Plant Journal: For Cell and Molecular Biology*, *41*, 133–145. https://doi.org/10.1111/j.1365-313X.2004.02284.x

Li, Q., Chen, X., Zhang, S., Shan, S., & Xiang, Y. (2022). DELAY OF GERMINATION 1, the master regulator of seed dormancy, integrates the regulatory network of phytohormones at the transcriptional level to control seed dormancy. *Current Issues in Molecular Biology*, *44*, 6205–6217. https://doi.org/10.3390/cimb44120423

Liu, G., Pei, W., Li, D., Ma, J., Cui, Y., Wang, N., Song, J., Wu, M., Li, L., Zang, X., Yu, S., Zhang, J., & Yu, J. (2019). A targeted QTL analysis for fiber length using a genetic population between two introgressed backcrossed inbred lines in upland cotton (*Gossypium hirsutum*). *The Crop Journal*, *7*(3), 273–282. https://doi.org/10.1016/j.cj.2018.11.005

Lopes, N. D. S., Santos, A. S., de Novais, D. P. S., Pirovani, C. P., & Micheli, F. (2023). Pathogenesis-related protein 10 in resistance to biotic stress: Progress in elucidating functions, regulation and modes of action. *Frontiers in Plant Science*, *14*, 1193873. https://doi.org/10.3389/fpls.2023.1193873

Meyer, R. S., & Purugganan, M. D. (2013). Evolution of crop species: Genetics of domestication and diversification. *Nature Reviews Genetics*, *14*, 840–852. https://doi.org/10.1038/nrg3605

Murphy, C., & Fuller, D. Q. (2017). Seed coat thinning during horsegram (*Macrotyloma uniflorum*) domestication documented through synchrotron tomography of archaeological seeds. *Scientific Reports*, *7*, Article 5369. https://doi.org/10.1038/s41598-017-05244-w

Née, G., Xiang, Y., & Soppe, W. J. J. (2017). The release of dormancy, a wake-up call for seeds to germinate. *Current Opinion in Plant Biology*, *35*, 8–14. https://doi.org/10.1016/j.pbi.2016.09.002

Nekrasov, V., Ludwig, A. A., & Jones, J. D. G. (2006). CITRX thioredoxin is a putative adaptor protein connecting Cf-9 and the ACIK1 protein kinase during the Cf-9/Avr9- induced defence response. *FEBS Letters*, *580*, 4236–4241. https://doi.org/10.1016/j.febslet.2006.06.077

Nguyen, B. D., Bett, K. E., & Noble, S. D. (2021). Optical coherence tomography applied to non-destructive seed coat thickness measurement. *Research Square*. https://doi.org/10.21203/rs.3.rs-723240/v1

Noodén, L. D., Blakley, K. A., & Grzybowski, J. M. (1985). Control of seed coat thickness and permeability in soybean : A possible adaptation to stress. *Plant Physiology*, *79*, 543–545. https://doi.org/10.1104/pp.79.2.543

Olivoto, T., & Lúcio, A. D. (2020). metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution*, *11*, 783–789. https://doi.org/10.1111/2041-210x.13384

Panikashvili, D., Shi, J. X., Schreiber, L., & Aharoni, A. (2009). The Arabidopsis DCR encoding a soluble BAHD acyltransferase is required for cutin polyester formation and seed hydration properties. *Plant Physiology*, *151*, 1773–1789. https://doi.org/10.1104/pp.109.143388

Penfield, S., & MacGregor, D. R. (2017). Effects of environmental variation during seed production on seed dormancy and germination. *Journal of Experimental Botany*, *68*, 819–825. https://doi.org/10.1093/jxb/erw436

Petruzzelli, L., Kunz, C., Waldvogel, R., Meins, F., Jr., & Leubner-Metzger, G. (1999). Distinct ethylene- and tissue-specific regulation of beta-1,3-glucanases and chitinases during pea seed germination. *Planta*, *209*, 195–201. https://doi.org/10.1007/s004250050622

Pollard, M., Beisson, F., Li, Y., & Ohlrogge, J. B. (2008). Building lipid barriers: Biosynthesis of cutin and suberin. *Trends in Plant Science*, *13*, 236–246. https://doi.org/10.1016/j.tplants.2008.03.003

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ramsay, L., Koh, C. S., Kagale, S., Gao, D., Kaur, S., Haile, T., Gela, T. S., Chen, L.-A., Cao, Z., Konkin, D. J., Toegelová, H., Doležel, J., Rosen, B. D., Stonehouse, R., Humann, J. L., Main, D., Coyne, C. J., McGee, R. J., Cook, D. R., … Bett, K. E. (2021). Genomic rearrangements have consequences for introgression breeding as revealed by genome assemblies of wild and cultivated lentil species. BioRxiv. https://doi.org/10.1101/2021.07.23.453237

Ranathunge, K., Shao, S., Qutob, D., Gijzen, M., Peterson, C. A., & Bernards, M. A. (2010). Properties of the soybean seed coat cuticle change during development. *Planta*, *231*, 1171–1188. https://doi.org/10.1007/s00425-010-1118-9

Renard, J., Martínez-Almonacid, I., Queralta Castillo, I., Sonntag, A., Hashim, A., Bissoli, G., Campos, L., Muñoz-Bertomeu, J., Niñoles, R., Roach, T., Sánchez-León, S., Ozuna, C. V., Gadea, J., Lisón, P., Kranner, I., Barro, F., Serrano, R., Molina, I., & Bueso, E. (2021). Apoplastic lipid barriers regulated by conserved homeobox transcription factors extend seed longevity in multiple plant species. *The New Phytologist*, *231*, 679–694. https://doi.org/10.1111/nph.17399

Renzi, J. P., Duchoslav, M., Brus, J., Hradilová, I., Pechanec, V., Václavek, T., Machalová, J., Hron, K., Verdier, J., & Smýkal, P. (2020). Physical dormancy release in seeds is related to environmental variations. *Plants*, *9*(4), 503. https://doi.org/10.3390/plants9040503

Robert, H. S. (2019). Molecular communication for coordinated seed and fruit development: What can we learn from auxin and sugars? *International Journal of Molecular Sciences*, *20*(4), 936. https://doi.org/10.3390/ijms20040936

Romeis, T., Tang, S., Hammond-Kosack, K., Piedras, P., Blatt, M., & Jones, J. D. (2000). Early signalling events in the Avr9/Cf-9-dependent plant defence response. *Molecular Plant Pathology*, *1*, 3–8. https://doi.org/10.1046/j.1364-3703.2000.00001.x

Rowland, O., Ludwig, A. A., Merrick, C. J., Baillieul, F., Tracy, F. E., Durrant, W. E., Fritz-Laylin, L., Nekrasov, V., Sjölander, K., Yoshioka, H., & Jones, J. D. G. (2005). Functional analysis of Avr9/Cf-9 rapidly elicited genes identifies a protein kinase, ACIK1, that is essential for full Cf-9-dependent disease resistance in tomato. *The Plant Cell*, *17*, 295–310. https://doi.org/10.1105/tpc.104.026013

Russi, L., Cocks, P. S., & Roberts, E. H. (1992). Coat thickness and hard-seededness in some *Medicago* and *Trifolium* species. *Seed Science Research*, *2*, 243–249. https://doi.org/10.1017/s0960258500001434

Sahoo, B., Nayak, I., Parameswaran, C., Kesawat, M. S., Sahoo, K. K., Subudhi, H. N., Balasubramaniasai, C., Prabhukarthikeyan, S. R., Katara, J. L., Dash, S. K., Chung, S.-M., Siddiqui, M. H., Alamri, S., & Samantaray, S. (2023). A comprehensive genome-wide investigation of the cytochrome 71 (*OsCYP71P6*) gene family: Revealing the impact of promoter and gene variants (Ser33Leu) of *OsCYP71P6* on yield-related traits in *Indica rice* (*Oryza sativa* L.). *Plants*, *12*(17), 3035. https://doi.org/10.3390/plants12173035

Sauvage, C., Rau, A., Aichholz, C., Chadoeuf, J., Sarah, G., Ruiz, M., Santoni, S., Causse, M., David, J., & Glémin, S. (2017). Domestication rewired gene expression and nucleotide diversity patterns in tomato. *The Plant Journal: For Cell and Molecular Biology*, *91*, 631–645. https://doi.org/10.1111/tpj.13592

Sedivy, E. J., Wu, F., & Hanzawa, Y. (2017). Soybean domestication: The origin, genetic architecture and molecular bases. *The New Phytologist*, *214*, 539–553. https://doi.org/10.1111/nph.14418

Sedláková, V., Hanáček, P., Grulichová, M., Zablatzká, L., & Smýkal, P. (2021). Evaluation of seed dormancy, one of the key domestication traits in chickpea. *Agronomy*, *11*, 2292. https://doi.org/10.3390/agronomy11112292

Sedláková, V., Zeljković, S. Ć., Štefelová, N., Smýkal, P., & Hanáček, P. (2023). Phenylpropanoid content of chickpea seed coats in relation to seed dormancy. *Plants*, *12*, 2687. https://doi.org/10.3390/plants12142687

Shao, S., Meyer, C. J., Ma, F., Peterson, C. A., & Bernards, M. A. (2007). The outermost cuticle of soybean seeds: Chemical composition and function during imbibition. *Journal of Experimental Botany*, *58*, 1071–1082. https://doi.org/10.1093/jxb/erl268

Shu, K., Meng, Y. J., Shuai, H. W., Liu, W. G., Du, J. B., Liu, J., & Yang, W. Y. (2016). Dormancy and germination: How does the crop seed decide?. *Plant Biology*, *17*, 1104–1112. https://doi.org/10.1111/plb.12356

Simos, G., Panagiotidis, C. A., Skoumbas, A., Choli, D., Ouzounis, C., & Georgatsos, J. G. (1994). Barley beta-glucosidase: Expression during seed germination and maturation and partial amino acid sequences. *Biochimica et biophysica acta*, *1199*, 52–58. https://doi.org/10.1016/0304-4165(94)90095-7

Singh, B., Singh, J. P., Kaur, A., & Singh, N. (2017). Phenolic composition and antioxidant potential of grain legume seeds: A review. *Food Research International*, *101*, 1–16. https://doi.org/10.1016/j.foodres.2017.09.026

Smýkal, P., Nelson, M., Berger, J., & Von Wettberg, E. (2018). The impact of genetic changes during crop domestication. *Agronomy*, *8*, 119. https://doi.org/10.3390/agronomy8070119

Smýkal, P., Vernoud, V., Blair, M. W., Soukup, A., & Thompson, R. D. (2014). The role of the testa during development and in establishment of dormancy of the legume seed. *Frontiers in Plant Science*, *5*, 351. https://doi.org/10.3389/fpls.2014.00351

Šola, K., Gilchrist, E. J., Ropartz, D., Wang, L., Feussner, I., Mansfield, S. D., Ralet, M.-C., & Haughn, G. W. (2019). RUBY, a putative galactose oxidase, influences pectin properties and promotes cell-to-cell adhesion in the seed coat epidermis of Arabidopsis. *The Plant cell*, *31*, 809–831. https://doi.org/10.1105/tpc.18.00954

Soltani, A., Walter, K. A., Wiersma, A. T., Santiago, J. P., Quiqley, M., Chitwood, D., Porch, T. G., Miklas, P., McClean, P. E., Osorno, J. M., & Lowry, D. B. (2021). The genetics and physiology of seed dormancy, a crucial trait in common bean domestication. *BMC Plant Biology*, *21*, Article 58. https://doi.org/10.1186/s12870-021-02837-6

Sun, Y., & Gong, Y. (2024). Research advances on the hard seededness trait of soybean and the underlying regulatory mechanisms. *Frontiers in Plant Science*, *15*, 1419962. https://doi.org/10.3389/fpls.2024.1419962

Tenenbaum, D., & Maintainer, B. (2022). KEGGREST: Client-side REST access to the Kyoto Encyclopedia of Genes and Genomes (KEGG) (R package version 1.38.0) [Computer software]. CRAN.

Tornero, P., Conejero, V., & Vera, P. (1997). Identification of a new pathogen-induced member of the subtilisin-like processing protease family from plants. *The Journal of Biological Chemistry*, *272*, 14412–14419. https://doi.org/10.1074/jbc.272.22.14412

Vartapetian, A. B., Tuzhikov, A. I., Chichkova, N. V., Taliansky, M., & Wolpert, T. J. (2011). A plant alternative to animal caspases: Subtilisin-like proteases. *Cell Death and Differentiation*, *18*, 1289–1297. https://doi.org/10.1038/cdd.2011.49

Verma, S. K., Mittal, S., Gayacharan, Wankhede, D. P., Parida, S. K., Chattopadhyay, D., Prasad, G., Mishra, D. C., Joshi, D. C., Singh, M., Singh, K., & Singh, A. K. (2021). Transcriptome analysis reveals key pathways and candidate genes controlling seed development and size in ricebean (*Vigna umbellata*). *Frontiers in Genetics*, *12*, 791355. https://doi.org/10.3389/fgene.2021.791355

Wang, K., Nan, L.-L., Xia, J., Yao, Y.-H., Cheng, J., & Chen, J.-R. (2024). Transcriptome analysis reveals candidate genes for different root types of alfalfa (*Medicago sativa*) after water stress induced by PEG-6000. *Chemical and Biological Technologies in Agriculture*, *11*(1), Article 107. https://doi.org/10.1186/s40538-024-00640-7

Wang, W., Xiong, H., Sun, K., Zhang, B., & Sun, M.-X. (2022). New insights into cell-cell communications during seed development in flowering plants. *Journal of Integrative Plant Biology*, *64*, 215–229. https://doi.org/10.1111/jipb.13170

Werker, E., Marbach, I., & Mayer, A. M. (1979). Relation between the anatomy of the testa, water permeability and the presence of phenolics in the genus *Pisum*. *Annals of Botany*, *43*, 765–771. https://doi.org/10.1093/oxfordjournals.aob.a085691

Williams, O. R., Vander Schoor, J. K., Butler, J. B., Hecht, V. F. G., & Weller, J. L. (2024). Physical seed dormancy in pea is genetically separable from seed coat thickness and roughness. *Frontiers in Plant Science*, *15*, 1359226. https://doi.org/10.3389/fpls.2024.1359226

Willis, C. G., Baskin, C. C., Baskin, J. M., Auld, J. R., Lawrence Venable, D., Cavender-Bares, J., Donohue, K., & de Casas, R. R. (2014). The evolution of seed dormancy: Environmental cues, evolutionary hubs, and diversification of the seed plants. *The New Phytologist*, *203*, 300–309. https://doi.org/10.1111/nph.12782

Wyatt, J. E. (1977). Seed coat and water absorption properties of seed of near-isogenic snap bean lines differing in seed coat color. *Journal of the American Society for Horticultural Science*, *102*, 478–480. https://doi.org/10.21273/JASHS.102.4.478

Wyse, S. V., & Dickie, J. B. (2018). Ecological correlates of seed dormancy differ among dormancy types: A case study in the legumes. *The New Phytologist*, *217*, 477–479. https://doi.org/10.1111/nph.14777

Xu, B. J., Yuan, S. H., & Chang, S. K. C. (2007). Comparative analyses of phenolic composition, antioxidant capacity, and color of cool season legumes and other selected food legumes. *Journal of Food Science*, *72*, S167–S177. https://doi.org/10.1111/j.1750-3841.2006.00261.x

Yamagata, H., Uesugi, M., Saka, K., Iwasaki, T., & Aizono, Y. (2000). Molecular cloning and characterization of a cDNA and a gene for subtilisin-like serine proteases from rice (*Oryza sativa* L.) and *Arabidopsis thaliana*. *Bioscience, Biotechnology, and Biochemistry*, *64*, 1947–1957. https://doi.org/10.1271/bbb.64.1947

Yow, A. G., Laosuntisuk, K., Young, R. A., Doherty, C. J., Gillitt, N., Perkins-Veazie, P., Jenny Xiang, Q.-Y., & Iorizzo, M. (2023). Comparative transcriptome analysis reveals candidate genes for cold stress response and early flowering in pineapple. *Scientific Reports*, *13*, Article 18890. https://doi.org/10.1038/s41598-023-45722-y

Yu, B., Gao, P., Song, J., Yang, H., Qin, L., Yu, X., Song, H., Coulson, J., Bekkaoui, Y., Akhov, L., Han, X., Cram, D., Wei, Y., Zaharia, L. I., Zou, J., Konkin, D., Quilichini, T. D., Fobert, P., Patterson, N., … Xiang, D. (2023). Spatiotemporal transcriptomics and metabolic profiling provide insights into gene regulatory networks during lentil seed development. *The Plant Journal: For Cell and Molecular Biology*, *115*, 253–274. https://doi.org/10.1111/tpj.16205

Zablatzká, L., Balarynová, J., Klčová, B., Kopecký, P., & Smýkal, P. (2021). Anatomy and histochemistry of seed coat development of wild (*Pisum sativum* subsp. *elatius* (M. Bieb.) Asch. et Graebn. and domesticated pea (*Pisum sativum* subsp. *sativum* L.). *International Journal of Molecular Sciences*, *22*(9), 4602. https://doi.org/10.3390/ijms22094602

Zhao, J., Shi, X., Chen, L., Chen, Q., Tian, X., Ai, L., Zhao, H., Yang, C., Yan, L., & Zhang, M. (2022). Genetic and transcriptome analyses reveal the candidate genes and pathways involved in the inactive shade-avoidance response enabling high-density planting of soybean. *Frontiers in Plant Science*, *13*, 973643. https://doi.org/10.3389/fpls.2022.973643

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Guerra-García, A., Balarynová, J., Smykal, P., von Wettberg, E. J., Noble, S. D., & Bett, K. E. (2025). Genetic and transcriptomic analysis of lentil seed imbibition and dormancy in relation to its domestication. *The Plant Genome*, *18*, e70021. https://doi.org/10.1002/tpg2.70021