

A transcriptomic atlas facilitating systems biology approaches in pea

Gatepe Cedoine Kodjovi, Mathias Coulon, Lisa Couturier, Amélie Morin, Ingrid Goma-Louamba, Caroline Artault, Guillaume Tcherkez, Cécile Vriet, Sylvain La Camera, Nathalie Pourtau, et al.

▶ **To cite this version:**

Gatepe Cedoine Kodjovi, Mathias Coulon, Lisa Couturier, Amélie Morin, Ingrid Goma-Louamba, et al.. A transcriptomic atlas facilitating systems biology approaches in pea. Crop Science, 2025, 65 (6), pp.e70194. ⟨10.1002/csc2.70194⟩. ⟨hal-05375785⟩

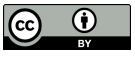
HAL Id: hal-05375785

https://hal.science/hal-05375785v1

Submitted on 24 Nov 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Check for updates

Received: 23 June 2025 | Accepted: 8 October 2025

DOI: 10.1002/csc2.70194

Crop Science

ORIGINAL ARTICLE

Genomics, Molecular Genetics & Biotechnology

A transcriptomic atlas facilitating systems biology approaches in pea

Gatepe Cedoine Kodjovi[1] | **Mathias Coulon**[1] | **Lisa Couturier**[1] | **Amélie Morin**[1] | **Ingrid Goma-Louamba**[1] | **Caroline Artault**[1] | **Guillaume Tcherkez**[2,3] | **Cécile Vriet**[1] | **Sylvain La Camera**[1] | **Nathalie Pourtau**[1] | **Bouziane Moumen**[1] | **Joan Doidy**[1]

[1]UMR CNRS 7267, EBI "Ecologie et Biologie des Interactions,", Université de Poitiers, Poitiers, France

[2]Research School of Biology, ANU College of Sciences, Australian National University, Canberra, Australian Capital Territory, Australia

[3]Institut de recherche en horticulture et semences, Université d'Angers, Beaucouzé, France

Correspondence
Joan Doidy, UMR CNRS 7267, EBI "Ecologie et Biologie des Interactions," Université de Poitiers, Poitiers, France.
Email: joan.doidy@univ-poitiers.fr

Assigned to Associate Editor Yuhui Wang.

Funding information
Centre National de la Recherche Scientifique; Université de Poitiers

Abstract

Pea is a model organism in biology and an agronomic crop, entering the genomics era. Although RNA-sequencing (RNA-seq) has become the main transcriptomic approach in biology, there is currently no genome-wide expression atlas covering a wide range of biological conditions for this species. Here, we generate a transcriptomic atlas in *Pisum sativum* by integrating 149 publicly available RNA-seq libraries, covering 10 cultivars, and a comprehensive collection of plant organs and several environmental stress conditions (heat, low temperature, nutrient, and water deficit). As proof of concept, we first verified the expression profiles of key gene families, such as sugar transporters and transcription factors (TF), across this transcriptomic atlas. Using a systems biology approach, we then inferred a regulatory network of genes responsive to water deficit, from which we predicted putative TF-target interactions, including genes encoding monosaccharide transporter *PsSTP13.2* and sugar facilitator *PsSWEET6*. Finally, we exploited our meta-analysis to identify new reference genes with stable expression based on their coefficient of variation. Ten reference genes were validated by quantitative PCR across various biological samples, including different pea varieties, organs, and stress conditions (water deficit and fungal pathogen infection). Three reference genes (*Psat6g102320*, *Psat6g163160*, and *Psat7g253080*) outperformed the expression stability of common housekeeping genes (*TFIIA*, *PPIIA*, and *β-tubulin*). Altogether, this atlas opens new avenues of integrative research in the genomics and systems biology era of legume crops.

Plain Language Summary

Pea (*Pisum sativum*) is a cultivated crop since the dawn of agriculture and a model organism in biology with the work of geneticist, Gregor Mendel. Nowadays, pea is

Abbreviations: bZIP, basic leucine zipper; Ct, cycle threshold; CV, coefficient of variation; ERF, Ethylene Response element binding Factors; GO, gene ontology; MST, monosaccharide transporter; PDA, potato dextrose agar; PP2A, protein phosphatase 2A; RNA-seq, RNA sequencing; RT-qPCR, real-time quantitative reverse transcription polymerase chain reaction; SUT, sucrose transporter; SWEET, sugar will eventually be exported transporter; TF, transcription factor; WD, water deficit.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2025 The Author(s). *Crop Science* published by Wiley Periodicals LLC on behalf of Crop Science Society of America.

entering the genomics era with the release of multiple omics metadata. However, there is currently no transcriptomic atlas (i.e., a collection of genome-wide expression studies in a variety of biological conditions) for this species. Here, we publish an atlas assembling seven transcriptomic studies that covers a comprehensive collection of pea varieties, plant organs, and diverse modalities of environmental stress. This atlas will facilitate systems biology approaches to identify key genes of agronomical interest, thus yielding potential benefits for agriculture and the environment.

1 | INTRODUCTION

Pea (*Pisum sativum*) is a cultivated crop since the dawn of agriculture, and still is one of the most important crop with an annual cultivation area over 10 million ha worldwide (FAOSTAT; http://www.fao.org/faostat/), providing a valuable source of dietary proteins, mineral nutrients, complex starch, and fibers with demonstrated health benefits (Dahl et al., 2012). Pea is also an agroecological crop, as the nitrogen-fixing ability of legumes benefits soil fertility and increases crop productivity without using fertilizers. *Pisum sativum* is also a model organism since the 18th century with the work of geneticist Gregor Mendel uncovering the basis of the laws of genetic inheritance. Although field pea is a widespread crop in agriculture and a model species in biology, it has only recently entered the "omics era" (Kreplak et al., 2019; Smýkal et al., 2012).

The development of omics technology has greatly benefited plant biology, with notable implementation of the first gene expression atlas in Arabidopsis (Winter et al., 2007) and in the model legume *Medicago truncatula* (Benedito et al., 2008) using microarray chips. With the recent advances in next-generation sequencing technologies, transcriptomic atlases have been upgraded to RNA-sequencing (RNA-seq) (Carrere et al., 2021; Sullivan et al., 2019; Waese et al., 2017). Thus, RNA-seq has become increasingly popular to study gene expression and transcriptomic atlas has been generated across multiple plant lineages (Perroud et al., 2018; Ramírez-González et al., 2018; Shinozaki et al., 2018; Sreedasyam et al., 2023), including in legume crops such as soybean (Severin et al., 2010; X. Zhang et al., 2025).

In pea, a reference genome has been released in the Caméor cultivar (Kreplak et al., 2019), followed by two genomes in the ZW1 and ZW6 cultivars (Liu et al., 2024; T. Yang et al., 2022). The release of reference genomes has provided valuable resources, facilitating the mapping of RNA-seq data. An increasing number of transcriptomes has been produced in recent years in legume crops. For instance, several studies have generated RNA-seq data in pea, enabling the identification of gene modules in response to changing climate conditions, such as heat (Huang et al., 2021) or water deficit (WD) (Morin et al., 2023). Besides, two transcriptomic studies have been published before the release of the pea genome (using de novo assembly), presenting gene expression atlas across a wide range of organs in pea (Alves-Carvalho et al., 2015; Sudheesh et al., 2015). However, to our knowledge, no gene atlas combining multi-experimental datasets has yet been attempted in this species.

Here, we present an integrative analysis of transcriptomic metadata in *P. sativum*. We generate a transcriptomic atlas from a comprehensive collection of RNA-seq studies performed in this species. We then exploit and validate this atlas through systems biology approaches. First, using heat map visualization, we analyze the expression profile of sugar transporter and transcription factor (TF) gene families. Then, we build a regulatory network of genes responsive to WD, highlighting hub candidates with co-expression patterns. Finally, we identify new reference genes with stable expression level across the transcriptomic atlas. The development of such integrative omics metadata combined with systems approaches opens new opportunities for genomics-assisted breeding strategies to enhance genetic gain in legume crops.

2 | MATERIALS AND METHODS

2.1 | Generation of the transcriptomic atlas

RNA-seq data (sequenced reads and fastq files) generated from *P. sativum* were downloaded from the Sequence Read Archive publicly available at NCBI (bioproject numbers listed in Table S1) using SRAtools v3.0.1 (SRA toolkit). A fastp v0.22.0 (S. Chen et al., 2018) analysis was performed to trim the adapters and filter out reads with a low-quality score, followed by a quality assessment performed using FastQC v0.12.1 (Babraham Bioinformatics). The RNA-seq reads were mapped to the *P. sativum* reference genome v1a (Kreplak et al., 2019) available at URGI (https://urgi.versailles.inra.fr/Species/Pisum/Pea-Genome-project) using STAR v2.7.10b (Dobin et al., 2012). Gene expression table counts were generated using FeatureCounts v2.0.1 (Liao et al., 2013) and the DESeq2 R-package was used with its median of ratios normalization method to generate the count data used for the atlas and for downstream analyses. The transcriptomic atlas com-

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

prising expression data of 44,756 genes of the pea genome across 81 biological conditions has been deposited in the open data repository Zenodo (https://doi.org/10.5281/zenodo.15635249).

2.2 | Heat map visualization

Expression values of sugar transporters (Doidy et al., 2019) and TFs (Kodjovi et al., 2023) genes were extracted from the transcriptomic atlas and heat maps were generated by an R package (pheatmap) with a Spearman hierarchical clustering applied to biological samples. Expression normalization was performed using z-scores of the mean expression data across gene rows.

2.3 | Gene regulatory network generation

The unique pea genome identifiers of the 3684 genes regulated by WD in root samples were retrieved from Morin et al. 2023. A weighted gene co-expression network was constructed using the WGCNA R package (Langfelder & Horvath, 2008) from a matrix of the pairwise correlation (Pearson) of the 3684 gene set computed across all experimental conditions of the transcriptomic atlas with a correlation coefficient threshold ($R > 0.9$ for positive and $R < -0.9$ for negative relations, respectively). The gene co-expression network was visualized on Cytoscape (v3.5) and analyzed using the network analyzer plugin (Shannon et al., 2003). A Cytoscape style adapted from the Virtual Plant platform (Katari et al., 2010) was applied to genes encoding TF, sugar transporters, and enzymes of the starch biosynthesis pathway.

2.4 | Identification of *cis*-regulatory motifs

The PlantPAN 4.0 (Chow et al., 2023) promoter analysis tool (https://plantpan.itps.ncku.edu.tw/plantpan4/promoter_analysis.php) was used to identify TF binding sites within the 1 kb promoter sequences.

2.5 | Identification and validation of reference genes

First, we filtered out genes presenting a null expression value in at least one sample in the raw expression data, retaining a list of 9242 genes with expression values across all samples. To assess the stability of gene expression across this dataset, the coefficient of variation ($CV = \sigma/\mu$) was calculated where σ and μ are standard deviation and mean, respectively. The top 10 genes with the lowest CV values were analyzed using gene ontology (GO) enrichment analysis (described in Morin et al., 2023) and by blastP comparison against *M. truncatula* and *Arabidopsis thaliana* genomes.

The top 10 genes with the lowest CV values were selected for further validation using real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR), in addition to the housekeeping genes *TFIIA*, *protein phosphatase 2A* (*PP2A*), and *β-tubulin* (Die et al., 2010). PCR primers were designed with the following criteria using AmplifX (v1.5.4): Tm of 56–59˚C (Tm delta inferior to 1˚C between primer pair), amplicon lengths of 80–160 bp and sequence length of 19–25 nucleotides with GC contents of 40%–60%. RNA extraction and cDNA synthesis were performed as described in Morin et al., 2022. Expression level was analyzed by RT-qPCR using GoTaq qPCR Master Mix (Promega) with a LightCycler 480 (Roche). The amplification program consisted of an initial denaturation step (2 min at 95˚C) followed by 40 amplification cycles (15 s at 95˚C, 1 min at 60˚C). The specificity of the amplification was checked with a melting curve analysis.

To analyze the stability of each reference gene, raw cycle threshold (Ct) values for all biological sample replicates were compared using the algorithms BestKeeper (Pfaffl et al., 2004), delta-Ct method (Silver et al., 2006), geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004) integrated on RefFinder (Xie et al., 2012) tool (https://www.ciidirsinaloa.com.mx/RefFinder-master/?type).

2.6 | Plant material used for RT-qPCR validation of reference genes

Plant material was produced from three independent experiments. The first experiment was performed to harvest multiple

Core Ideas

- This transcriptomic atlas generated in pea will facilitate systems biology approach in this model organism and legume crop.
- This atlas will facilitate the visualization of transcriptomic data in the scope of capturing key gene candidates with potential agronomic interest.
- Regulatory network of genes responsive to water deficit uncovers sugar transporter genes targeted by transcription factor (TF) candidates.
- New reference genes showing stable level of expression across various biological conditions were validated for quantitative reverse transcription polymerase chain reaction (qPCR) applications.

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

plant organs. Seeds of *P. sativum* (Kayanne, provided by KWS SAAT SE & Co. KGaA) were pregerminated and transferred individually in 3-L pots containing loam soil and grown in a greenhouse with a 14-h photoperiod complemented with artificial lighting when natural light was below 150 W/m^2. Roots, stems, and leaves at node 2 and node 9 ($n = 5$ for all samples) were harvested on plants at the vegetative stage (BBCH20); subsequently, roots and leaves at node 3, and the last vegetative leaf formed below the reproductive nodes, seeds (21 days after pollination), pods ($n = 5$ for all samples), and flowers ($n = 3$) were harvested on plants at the reproductive stage (BBCH33). The second experiment produced plants subjected to WD. Here, seeds of *P. sativum* (Captur, provided by Lemaire-Deffontaines) were sown in 3-L pots containing a mixture of loamy soil and sand (2:1 v:v). Plants were positioned as split-plot randomized blocks and grown in a greenhouse with a 14-h photoperiod complemented with artificial lighting when natural light was below 150 W/m^2. WD was applied at the beginning of the reproductive development by withdrawing water as soon as the first flower was pollinated and until the soil water content reached 40% of the field capacity for WD plants, while well-watered plants were maintained at 80% of the field capacity. WD was maintained at 40% of the field capacity during 14 days until the harvest of the plants, sampling the entire root system ($n = 5$) and the last source leaf ($n = 5$) formed during vegetative development. A last experiment was performed on detached pea leaves ($n = 4$; Kayanne, KWS) inoculated during 4 days in a closed tray under saturating humidity (22°C) with a potato dextrose agar (PDA) plug containing sporulating *Botrytis cinerea* strain B05.10 (Staats & van Kan, 2012), compared to control leaves with a sterile PDA plug. All samples were immediately frozen in liquid nitrogen and stored at −80°C.

3 | RESULTS AND DISCUSSION

3.1 | A transcriptomic atlas in *Pisum sativum*

We integrated a total of 149 publicly available RNA-seq libraries from seven international studies (Table 1). These transcriptomes were generated from 10 different pea varieties and covered a wide range of biological conditions, including a comprehensive collection of plant organs, various modalities of abiotic stress (mineral nutrition, water supply, and temperature), and biotic interactions (nodule) (Table S1).

The raw expression data from the source RNA-seq libraries were re-assembled to the reference genome (Kreplak et al., 2019) and the mean expression value was computed between biological replicates produced in individual studies (Table 1), thus providing a transcriptomic atlas of the 44,756 genes in the pea genome across 81 biological conditions. The resulting atlas was deposited in an open data repository (https://zenodo.org/records/15635249) publicly available to the research and breeding community. To our knowledge, this is the first transcriptomic meta-atlas generated in this model organism and legume crop.

3.2 | Visualizing gene expression profiles across the transcriptomic atlas

Visualization of transcriptomic data through heat maps is a valuable tool to provide a global view of gene expression in a wide range of biological conditions. As a first validation of this transcriptomic atlas, we explored the gene expression patterns at the family scale of sugar transporters and TFs, focusing on the expression profile of key gene candidates with potential agronomic interest.

3.2.1 | Clustering of the biological samples

To verify the data structure of the transcriptomic atlas, we conducted hierarchical clustering across all biological samples distributed on the heatmaps. Aside from a few outliers, all samples clearly clustered into four distinct groups (flower, seed/pod, shoot, and root). The heatmaps noticeably reveal that the biological samples clustered according to the nature of the plant organs for both the sugar transporter and TF expression datasets (see sample dendrograms presented in Figure 1 and Figure S1), validating the data assembly and biological structure of the transcriptomic atlas.

3.2.2 | Sugar transporter gene families

First, we focused on sugar transporters, as these genes have a significant impact on plant development, crop yield, and stress adaptation. Our lab has strong expertise on sugar transport in plants (Doidy et al., 2024; Lemoine et al., 2013), including in legume species (Doidy et al., 2012, 2019). In pea, we previously identified sugar transporter genes of the *sucrose transporter* (*SUT*), *monosaccharide transporter* (*MST*), and *sugar will eventually be exported transporter* (*SWEET*) families (Doidy et al., 2019). Here, we produced a heat map of the *PsSUT*, *PsMST*, and *PsSWEET* gene expression profiles across this transcriptomic atlas (Figure 1).

As proof of concept, we looked at the expression profile of well-characterized gene markers. For instance, SUT1 and SWEET12 family members are involved in phloem loading of sugars into the companion-cell sieve element complex (L.-Q. Chen et al., 2012; C. Zhang & Turgeon, 2009); therefore, such genes usually show a high expression profile in shoot samples. Here, we confirm that their gene homologs in pea,

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 List of RNA-sequencing (RNA-seq) studies used to generate the transcriptomic atlas.

References	Varieties	Conditions	Organs	Data sources	Number of libraries	Number of conditions in Atlas (replicates)
Alves-Carvalho et al. (2015)	Caméor	Aeroponics, hydroponics with different levels of nitrate	Apical node, flowers, germinating seeds, leaves, nodules, peduncles, pods, roots, seeds, shoot, stems, tendrils	PRJNA267198	20	20 ($n = 1$)
Sudheesh et al. (2015)	Kaspa, Parafield	Glasshouse using a soil mixture	Pistil, flower, pod, seed, leaf, nodule, root, root-tip, seedling, stem, stamen, stipule, tendril	PRJNA277074 PRJNA277076	23	23 ($n = 1$)
Bahrman et al. (2019)	Champagne, Terese	Low temperature	Aerial parts	PRJNA543764	24	12 ($n = 2$)
Henriet et al. (2019)	Caméor	Water and/or sulfur stress	Seeds (9 DAP)	PRJNA517587	14	4 ($n = 3$ or 4)
Malovichko et al. (2020)	Spring-2	Growth chamber in peat soil	Seeds (10 and 20 DAP)	PRJNA533971	8	2 ($n = 4$)
Huang et al. (2021)	PR11-2, PR11-90, CDC Amarillo	Heat stress	Anthers, stipules	PRJNA757773	36	12 ($n = 3$)
Morin et al. (2023)	Safran	Water stress	Leaves, roots, seeds (7 and 14 DAP)	PRJNA991334	24	8 ($n = 3$)
					Total 149	Total 81

Abbreviation: DAP, days after pollination.

PsSUT1.1 (*Psat6g188160*) and *PsSWEET12* (*Psat7g038200*) show a particularly high expression level in the shoot sample cluster (Figure 1). These genes may have a conserved function for phloem loading in legume species (Doidy et al., 2019; Lu et al., 2020). We also verified the expression profiles of *PsSWEET15.1* (*Psat1g219320*), a putative sucrose facilitator highly expressed during seed development (Morin et al., 2023). The gene homolog in soybean (*GmSWEET15a*) was used as a marker of endosperm tissue in a multi-omics single-cell atlas (X. Zhang et al., 2025) and its knockout mutation led to higher seed abortion rates (S. Wang et al., 2019). Here, our heatmap visualization clearly confirmed the high expression level of *PsSWEET15.1* in all seed samples integrated in the atlas (Figure 1).

3.2.3 | Transcription factors

We then investigated TF families, since TFs are the primary regulators of genome-wide transcriptional changes. The pea genome comprises 1807 TF genes, distributed across 58 families (Kodjovi et al., 2023). Here, we built the heat maps showing the expression profiles of the three largest TF families in the pea genome (bHLH, ERF and MYB), from which we highlighted gene clusters with co-expression patterns across the transcriptomic atlas (see gene dendrograms presented in Figure S1).

We further analyzed the expression pattern of TF candidates involved in symbiotic nitrogen fixation, a key agroecological process of legumes. Pea, as a legume crop, establishes a root nodule symbiosis with nitrogen-fixing soil bacteria, increasing crop productivity and reducing the use of nitrogen fertilizers. Root nodule formation is accompanied by large transcriptional reprogramming orchestrated by a set of TFs (Mergaert et al., 2019). Here, we identified 335 TFs differentially regulated in nodule samples (Figure 2) and verified the regulation of gene homologs previously identified in the legume model *M. truncatula* (Table S2). For instance, NIN (NODULE INCEPTION) proteins are TFs that play key role during root nodule initiation and their mutation can lead to a non-nodulating phenotype (Schauser et al., 1999; Shen & Feng, 2024). Here, we confirm that the homologous *NIN*

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

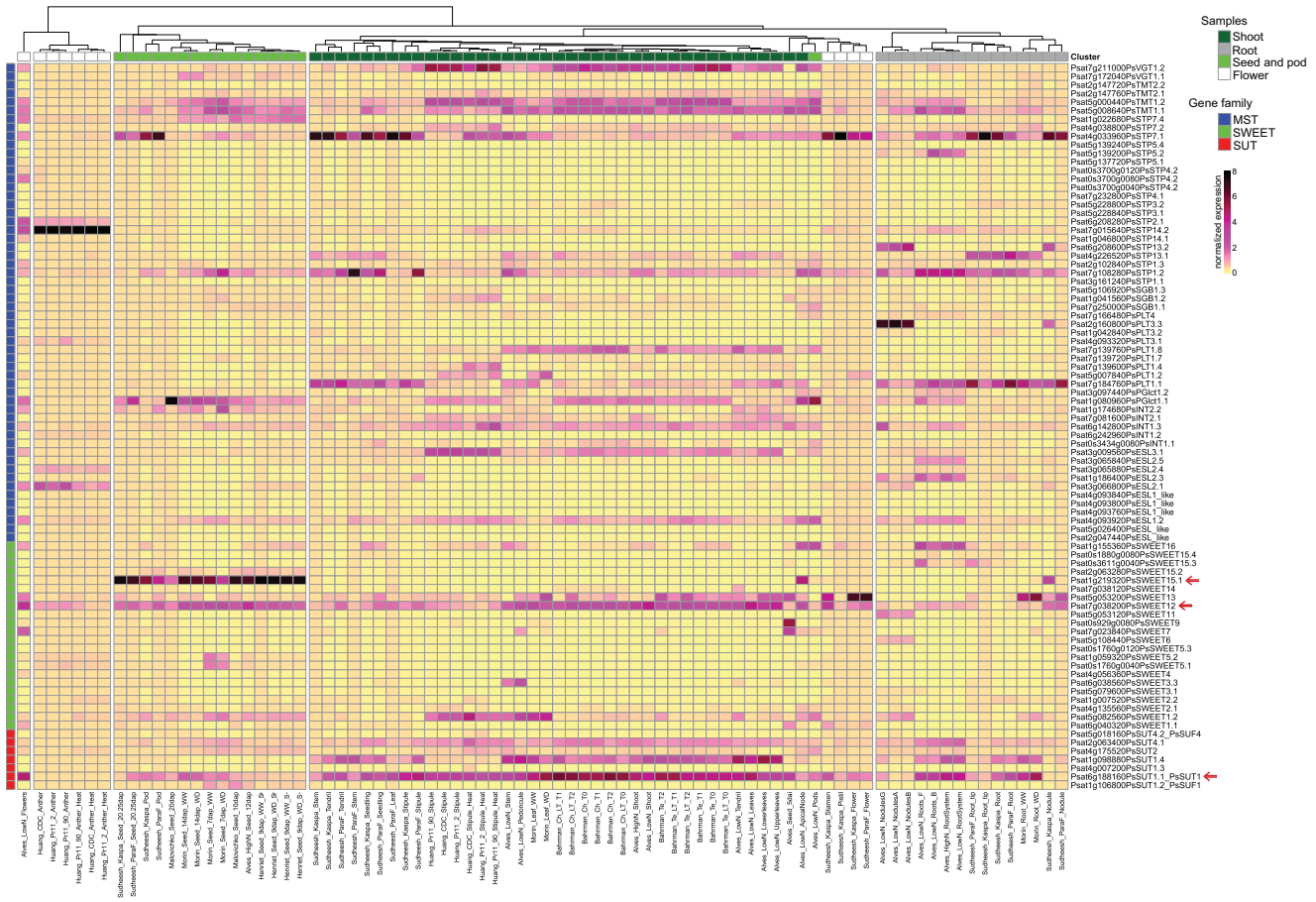


FIGURE 1 Heat map representation of sugar transporter gene expression across the atlas. Gene expression values of sugar transporter genes of the *SUT* (*sucrose transporter*), *MST* (*monosaccharide transporter*), and *SWEET* (*sugar will eventually be exported transporter*) families were normalized (*z*-scores of the mean expression data by gene row) and Spearman hierarchical clustering was applied to biological samples. Red arrows indicate gene examples with high expression profile in shoot (*PsSUT1.1* and *PsSWEET12*) and seed samples (*PsSWEET15.1*).

gene in pea (*Psat2g001120*) and a second *NIN-like* homolog (*Psat7g208800*) show specific induction of their gene expression in nodule samples integrated in the atlas (Figure 2). We also captured three TFs belonging to the ERF family, including *Psat7g259160*, a close homolog to *MtEFD* required for the formation of functional nitrogen-fixing nodules, which was found upregulated in nodule samples, while the expression of the tandem genes (*Psat4g225000* and *Psat4g225040*) homologous to *MtEDN3* and *MtEDN1* was highly repressed (Vernié et al., 2008). Lastly, a gene of the CAMTA family (calmodulin-binding transcription activators) *Psat7g054520* was downregulated in nodule samples; this gene is a close homolog to *MtCAMTA7* involved in calcium-regulated transcription during the early nodulation signaling pathway with their expression later suppressed upon infection (Y. Yang et al., 2015).

TFs serve as central nodes in the regulation of gene expression, orchestrating complex transcriptional networks, such as the regulation of developmental process, plant–microbe interaction, and stress adaptation. Advances in genomics and transcriptomics have illuminated their dynamic hub roles in signal integration (Li et al., 2015), making them valuable targets for next-generation breeding strategies. The dual roles of certain TF families, such as ERF (ethylene response factor) and NAC (NAM, ATAF1/2, and CUC2), in both nodulation and abiotic stress responses highlight their potential for developing climate-resilient legume varieties (de Zélicourt et al., 2012; Singh et al., 2025). Integrating TF hub markers into genomic-assisted selection can accelerate breeding programs by enabling the simultaneous improvement of key agronomical traits, such as nutrient-use efficiency and stress tolerance.

3.3 | Building gene regulatory networks

In the systems biology era, building gene regulatory networks is a modern approach to capture regulatory hubs pivotal to environmental changes. To further exploit our transcriptomic atlas, we performed a network analysis of genes differentially

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

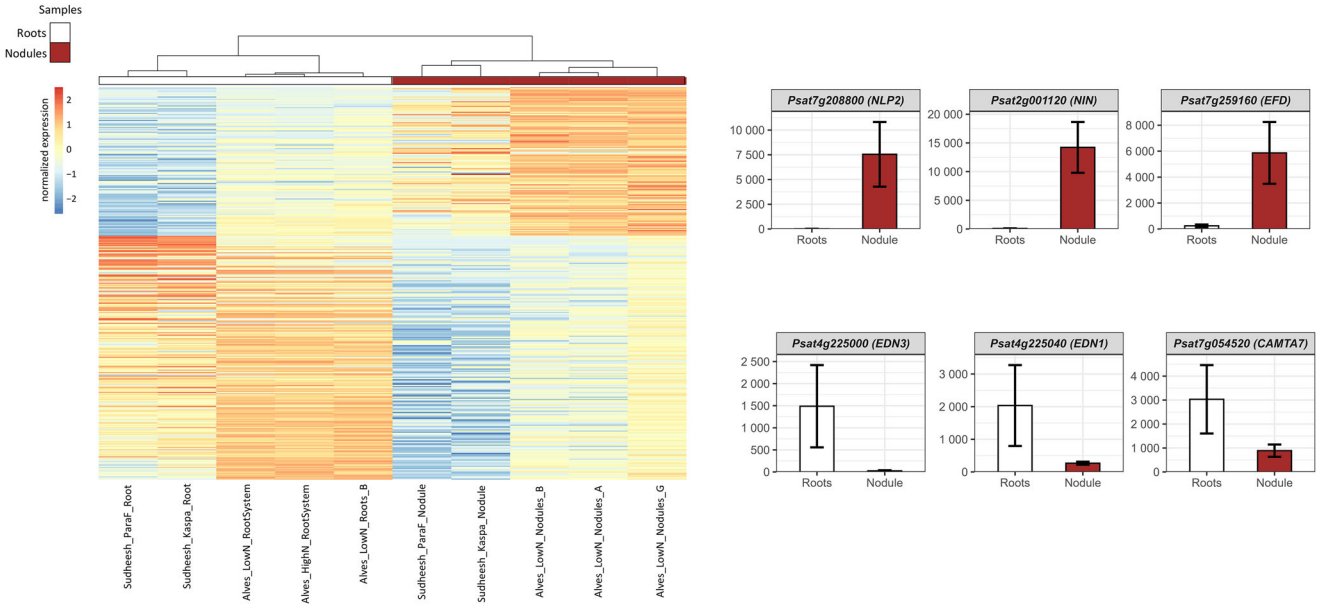


FIGURE 2 Heat map representation of transcription factor (TF) genes differentially expressed in nodules (left panel). In total, 335 TFs were found differentially expressed (127 up-regulated and 208 down-regulated) in nodule versus root samples. Gene expression values of TF genes were normalized (log2 of the mean expression data by gene row) and Spearman hierarchical clustering was applied to biological samples. Graphs of TF candidates differentially expressed in nodule compared to root samples (right panels), with gene annotation of the closest homolog in *Medicago truncatula* indicated in brackets: *Psat2g001120* (*NIN*, *NODULE INCEPTION*), *Psat7g208800* (*NLP2*, *NIN-like protein 2*), *Psat7g259160* (*EFD*, *ethylene response element binding factor [ERF] required for nodule differentiation*), *Psat4g225000* (*EDN3*, *ERF differentially regulated during nodulation 3*), *Psat4g225040* (*EDN1*, *ERF differentially regulated during nodulation 1*), and *Psat7g054520* (*CAMTA7*, *Calmodulin-binding transcription activator 7*).

regulated by WD. To do so, we selected the 3684 differentially expressed genes during WD in root (Morin et al., 2023; Figure S2) and computed a gene expression correlation network across the transcriptomic atlas (Figure S3). The inferred gene regulatory network shows 1605 WD-responsive genes interconnected by 44,364 edges (Figure S3; Table S3). We performed a GO-term analysis to identify the biological processes that were overrepresented in this gene network. This analysis notably showed a significant enrichment of terms associated with membrane transport, regulation of transcription, and nucleic acid metabolic processes (Table S4). In this network, two hub genes showed the highest degree of interconnection (both with 225 edges): *Psat1g112600*, a gene homolog to Kunitz trypsin inhibitor *AtKTI5* (*At1g17860*) and *Psat7g096040*, a gene homolog to *UMAMIT23* (*At1g68170*) in Arabidopsis.

TFs play essential roles in regulating the expression of target genes involved in crucial biological processes, such as water stress responses (Joshi et al., 2016). In this network, we predict the influence of 88 TFs governing the genome-wide transcriptional changes to WD in root (Table S5). The TF hub with the highest degree of connection (with 210 edges) is a basic leucine zipper (bZIP) TF (*Psat5g257080*) homolog to *TGA9* (*At1g08320*) in Arabidopsis, a positive regulator of autophagy induced by drought (Tang & Bassham, 2022; P. Wang et al., 2020).

We then focused on a subnetwork of putative TF targeting sugar transporter genes (Figure 3). We isolated the subnetwork with predicted direct interactions of two sugar transporter candidates, *PsSTP13.2* encoding a putative monosaccharide transporter (Pegler et al., 2024) and *PsSWEET6* encoding a putative sugar facilitator responsive to drought (Amoah & Adu-Gyamfi, 2025). These relevant sugar transporter candidates *PsSTP13.2* and *PsSWEET6* were previously identified as differentially expressed genes in an RNA-seq analysis of roots subjected to drought (Morin et al., 2023), and here, we were interested in identifying putative TF regulators involved in transcriptional reprogramming of root carbon allocation. The inferred network comprised three TFs, *PsGFR* (*Psat2g140880*), homolog to *G2-LIKE FLAVONOID REGULATOR* (*At5g45580*), a gene involved in the regulation of flavonoid synthesis (Petridis et al., 2016); *PsWRKY72* (*Psat3g196240*), homologs to *AtWRKY72* (*At5g15130*) and *GmWRKY72*, which is involved in the repression of phytoalexin synthesis in soybean (Lin et al., 2024); and *PsbZIP6* (*Psat0ss26165g0040*), homolog to *AtbZIP6* (*At2g22850*), which showed a high expression level in root tissues and an upregulation in roots in response to abiotic stress (Kilian et al., 2007) (Figure 3). We also performed a promoter analysis of *PsSTP13.2* and *PsSWEET6* to identify *cis*-regulatory elements using PlantPAN 4.0 (Chow et al., 2023). These genes harbored multiple copies of bZIP, G2-like, and WRKY

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

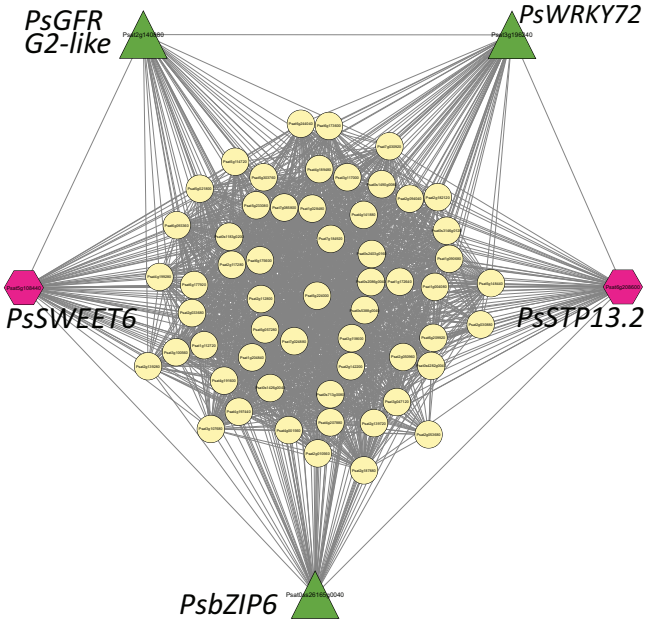


FIGURE 3 Regulatory subnetwork including sugar transporters *PsSTP13.2* and *PsSWEET6* as putative transcription factor (TF) target of bZIP, G2-like, and WRKY candidates. A Cytoscape style adapted from the Virtual Plant platform (Katari et al., 2010) was applied to sugar transporters represented by pink hexagons (*PsSTP13.2 Psat6g208600* and *PsSWEET6 Psat5g108440*); transcription factors are represented by green triangles (*PsbZIP6 Psat0ss26165g0040*, *PsGFR G2-like Psat2g140880*, and *PsWRKY72 Psat3g196240*) and other genes (yellow circles) of this network.

binding motifs in their promoter region (Figure S4), thus corroborating the TF-target predictions.

Gene network inference is a valuable tool in systems biology approach, enabling the prediction of TF-target interactions. The TF candidates identified here are putatively involved in the genome-wide transcriptional response to WD, while sugar transporter targets could be implicated in the reallocation of carbon resources toward roots, remodeling the root architecture to efficiently forage water in the soil. Further validation of such gene candidates identified through systems biology approach may open new breeding strategies to develop legume crops resilient to climate change.

3.4 | Identification of reference genes

As a final proof of concept, we exploited this transcriptomic atlas to identify new reference genes. Universal reference genes are characterized by stable expression level that remains constant across various conditions. Usually, reference genes are housekeeping genes involved in basic cellular processes. For instance, in *P. sativum*, several studies use *transcription factor IIA*, *PP2A*, and *β-tubulin* as reference genes to normalize qPCR results (Die et al., 2010; Knopkiewicz & Wojtaszek, 2018). However, expression levels of such genes often vary in response to experimental conditions, forcing the researchers to validate the reference genes for every experimental study. For instance, Die et al. (2010) recommends using the *elongation factor1-α/β-tubulin* pair for the RT-qPCR normalization of leaf samples, while *PP2A* or the *TFIIA/β-tubulin* pair worked best for normalization of root samples.

In the omics era, we now have a wide range of transcriptomic metadata that can be exploited to identify reference genes with a stable level of expression across various and comprehensive sets of biological conditions. Here, we used a genome-wide driven approach to identify reference genes in *P. sativum*. To do so, we exploited this transcriptomic meta-analysis to capture genes with the highest stability based on their CV (Table S6), as previously achieved in *Gossypium hirsutum* (Smitha et al., 2019). Then, we selected the top 10 most stable genes based on their CV across the transcriptomic atlas (Table 2). The expression profiles of the top 10 genes with the lowest CV were verified through a heat map visualization (Figure S5). Surprisingly, GO-term analysis of the top 10 genes with the lowest CV values showed numerous biological terms related to basic nucleic acid processes (Table S7), indicating that such genes are also part of the basal housekeeping machinery of cellular processes. Comparison of homology by blast against Medicago and Arabidopsis genomes (Table S8) confirmed the correct annotation and putative function of the top 10 reference candidate genes.

3.4.1 | Validation of reference genes by RT-qPCR

We then tested the stability of the top 10 candidate reference genes and the three housekeeping genes (*TFIIA*, *PP2A*, and *β-tubulin*) by RT-qPCR (Table S9; Figure S6) across multiple organs and abiotic and biotic stress experiments. Briefly, a first experiment was conducted to produce biological samples from multiple plant organs (i.e., leaves, stem, root, flower, pod, and seeds); another experiment was performed to produce plants subjected to WD, and the last experiment produced leaf samples inoculated with the fungal pathogen *Botrytis cinerea*. The reference genes exhibited a mean Ct value ranging from 22.4 to 27.5 across all experiments, with *Psat0s3790g0040* showing the highest expression and *Psat5g227600* the lowest expression levels, respectively (Figure 4A). Based on this first analysis of raw Ct values, *Psat5g227600* and *Psat5g227680* seem to present the lowest variability (CV = 0.019), while *β-tubulin* shows the highest variability (CV = 0.036) across all experiments (noteworthy, these CV values are calculated from RT-qPCR Ct values, while CV values reported in Table 2 are calculated from RNA-seq data; hence these two metrics are on entirely dif-

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

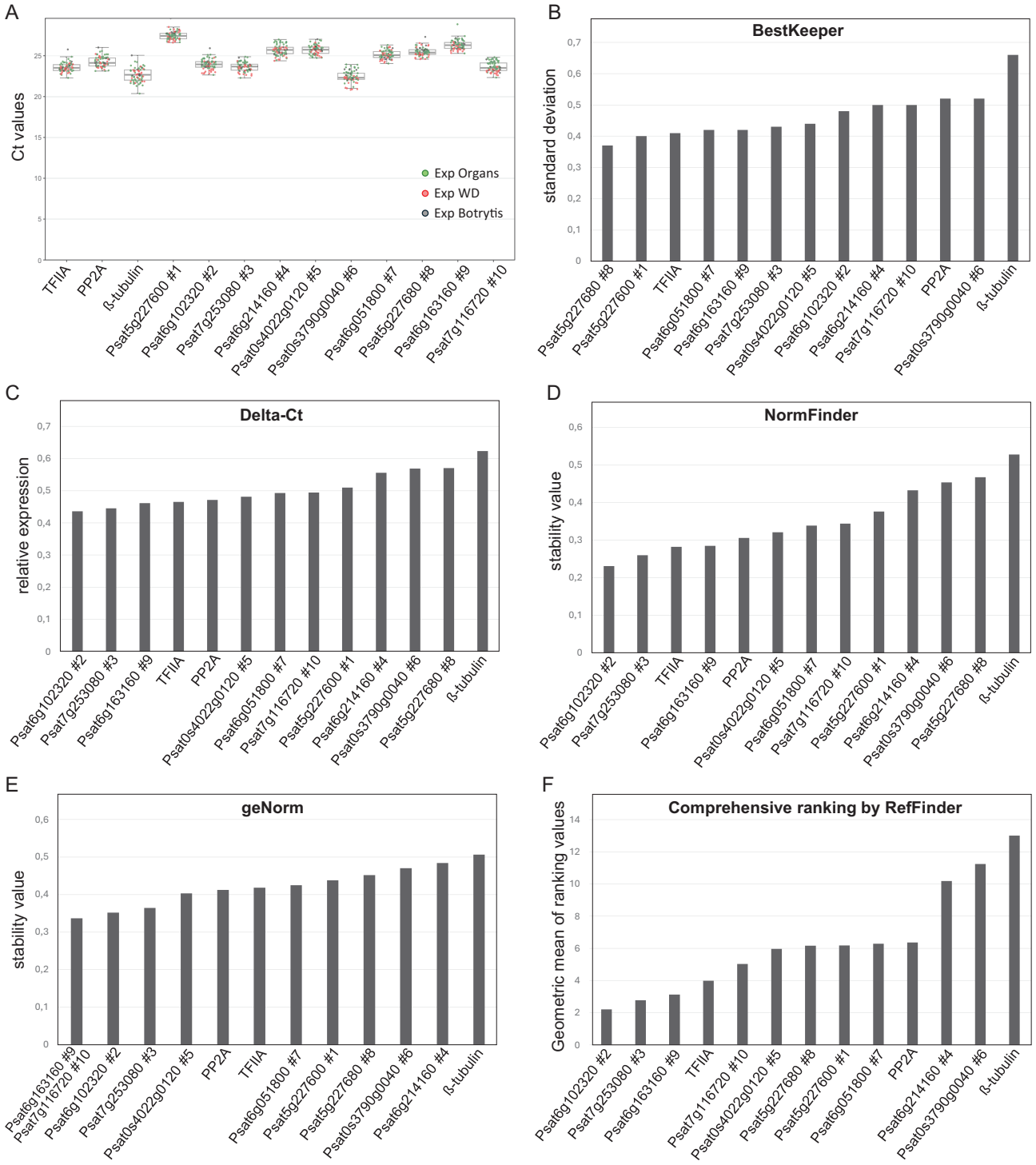


FIGURE 4 Experimental validation by real-time quantitative reverse transcription polymerase chain reaction (RT-qPCR) and stability analysis of the candidate reference genes. (A) Boxplots of raw cycle threshold (Ct) showing biological replicate distribution of experiment 1 (green dots) comprising various plant organs harvested at both vegetative (BBCH20) and reproductive (BBCH33) stages, experiment 2 (red dots) comprising leaf and root samples from plants subjected to water deficit (WD) compared to well-watered control, and experiment 3 (gray dots) comprising leaf samples inoculated by *Botrytis cinerea* compared to mock control. (B–F). Gene expression stability analyzed using BestKeeper (B), delta-Ct method (C), geNorm (D), NormFinder (E), and RefFinder (F) ranked in the order of least stable to most stable reference genes.

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 List of top 10 candidate reference genes ordered by coefficient of variation.

Ranking	CV	Gene ID	Gene description from *Pisum sativum* genome
#1	0.2942	*Psat5g227600*	PPR repeat family
#2	0.3401	*Psat6g102320*	Eukaryotic DNA topoisomerase I + DNA binding fragment
#3	0.3417	*Psat7g253080*	CASC3/Barentsz eIF4AIII binding
#4	0.3708	*Psat6g214160*	Domain of unknown function (DUF382)
#5	0.3739	*Psat0s4022g0120*	Unknown gene
#6	0.3777	*Psat0s3790g0040*	Superfamilies 1 and 2 helicase ATP-binding type-1 domain
#7	0.3784	*Psat6g051800*	Endonuclease/Exonuclease/phosphatase family
#8	0.3785	*Psat5g227680*	Unknown gene
#9	0.3785	*Psat6g163160*	HCNGP-like protein
#10	0.3789	*Psat7g116720*	Guanylate-binding protein + N-terminal domain

Abbreviation: CV, coefficient of variation.

ferent scales). Altogether, the biological replicate distribution (Figure 4A) confirms the potential stability of the candidate reference genes and the housekeeping genes (*TFIIA* and *PP2A*) in diverse experiments, including different pea varieties, plant organs at both the reproductive and vegetative stages, and under conditions of abiotic and biotic stress.

3.4.2 | Stability analysis of candidate reference genes

Expression stability of the genes was further analyzed using several tools: BestKeeper (Pfaffl et al., 2004), delta-Ct method (Silver et al., 2006), geNorm (Vandesompele et al., 2002), and NormFinder (Andersen et al., 2004) algorithms, integrated on RefFinder (Xie et al., 2012) tool. The expression stability of the reference genes is ranked in the order of least stable to the most stable reference genes (Figure 4 B–F). BestKeeper analyzes expression stability based on standard deviation and CV, ranking *Psat5g227600*, *Psat5g227680*, and *TFIIA* as the most stable genes with the lowest variability (Figure 4B). The delta-Ct method compares the relative expression of gene pairs in each sample, ranking *Psat6g102320*, *Psat7g253080*, and *Psat6g163160* with the smallest mean standard deviation of delta-Ct (Figure 4C). NormFinder uses a model that calculates the stability value from both the total variation of the reference genes and the variation between subgroups in a sample set, ranking *Psat6g102320*, *Psat7g253080*, and *TFIIA* as the most stable genes (Figure 4D). The geNorm algorithm also calculates stability values by averaging the pairwise variations among all genes across all samples, positioning *Psat6g163160* and *Psat7g116720* (both ranked in first position) and *Psat6g102320* as the most stable reference genes (Figure 4E). Finally, a comprehensive stability rank of each gene is calculated by RefFinder using the geometric mean of stability position given by each individual method, ranking *Psat6g102320*, *Psat6g163160*, and *Psat7g253080* as the most stable reference genes across all tested samples (Figure 4F). In contrast, the housekeeping gene *β-tubulin* was determined to be the least stable gene through all analysis tools.

All 10 newly identified reference genes together with the two housekeeping genes (*TFIIA* and *PP2A*) yielded low Ct variability (Figure 4A) and high stability profiles (Figure 4B–F), however comprehensive stability ranking (Figure 4F) indicates that gene normalization of RT-qPCR results is best achieved using the reference gene *Psat6g102320*. Thus, our transcriptome-wide approach identifies new reference genes that outperform some of the commonly used housekeeping genes and are suitable for qPCR applications by the research community.

4 | CONCLUSION

Here, we present an integrative analysis of transcriptomic metadata in *P. sativum*, a model organism in biology and a legume crop of agroecological importance. The development of such omics metadata combined with systems biology approaches will advance our understanding of key genetic determinants (e.g., gene modules responding to environmental changes, hub TF candidates, and reference genes) with several applications to the research and breeding communities working on legume crop improvement. The bursting of next-generation sequencing data together with new breeding tools in pea represents a valuable resource for innovative omics breeding approaches (Padhy et al., 2025; Pandey et al., 2021). Such integrative genomics-assisted breeding approaches will accelerate genetic gain through development of improved and adaptive legume cultivars.

AUTHOR CONTRIBUTIONS

Gatepe Cedoine Kodjovi: Data curation; formal analysis; methodology. **Mathias Coulon**: Investigation. **Lisa Couturier**: Investigation. **Amélie Morin**: Formal anal-

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ysis; investigation; writing—review and editing. **Ingrid Goma-Louamba**: Data curation; formal analysis. **Caroline Artault**: Investigation. **Guillaume Tcherkez**: Writing—review and editing. **Cécile Vriet**: Methodology; supervision; writing—review and editing. **Sylvain La Camera**: Methodology; supervision; writing—review and editing. **Nathalie Pourtau**: Methodology; supervision; writing—review and editing. **Bouziane Moumen**: Data curation; supervision; writing—review and editing. **Joan Doidy**: Conceptualization; data curation; formal analysis; methodology; supervision; writing—original draft.

ACKNOWLEDGMENTS

This work was supported by the Université de Poitiers, the Centre National de la Recherche Scientifique (CNRS) and an internal collaborative grant (ACI) funded by the EBI laboratory. We thank students of the Master Gphy (Génie Physiologique, Biotechnologique et Informatique, Université de Poitiers) for their contributions to bioinformatics analyses. We thank all members of the SEVE team (EBI, Université de Poitiers) for their technical assistance and research contributions.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAIBILITY STATEMENT

The transcriptomic atlas generated in this study have been deposited in the open data repository Zenodo (10.5281/zenodo.15635249) available at https://zenodo.org/records/15635249.

ORCID

Mathias Coulon https://orcid.org/0009-0009-8935-0157
Amélie Morin https://orcid.org/0000-0003-3318-8778
Guillaume Tcherkez https://orcid.org/0000-0002-3339-956X
Cécile Vriet https://orcid.org/0000-0001-6625-268X
Bouziane Moumen https://orcid.org/0000-0002-1789-6160
Joan Doidy https://orcid.org/0000-0003-3966-3218

REFERENCES

Alves-Carvalho, S., Aubert, G., Carrère, S., Cruaud, C., Brochot, A.-L., Jacquin, F., Klein, A., Martin, C., Boucherot, K., Kreplak, J., Da Silva, C., Moreau, S., Gamas, P., Wincker, P., Gouzy, J., & Burstin, J. (2015). Full-length de novo assembly of RNA-seq data in pea (*Pisum sativum* L.) provides a gene expression atlas and gives insights into root nodulation in this species. *The Plant Journal*, *84*, 1–19. https://doi.org/10.1111/tpj.12967

Amoah, J. N., & Adu-Gyamfi, M. O. (2025). Effect of drought acclimation on sugar metabolism in millet. *Protoplasma*, *262*, 35–49. https://doi.org/10.1007/s00709-024-01976-5

Andersen, C. L., Jensen, J. L., & Ørntoft, T. F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research*, *64*, 5245–5250. https://doi.org/10.1158/0008-5472.CAN-04-0496

Bahrman, N., Hascoët, E., Jaminon, O., Dépta, F., Hû, J.-F., Bouchez, O., Lejeune-Hénaut, I., Delbreil, B., & Legrand, S. (2019). Identification of genes differentially expressed in response to cold in *Pisum sativum* using RNA sequencing analyses. *Plants*, *8*, 288. https://doi.org/10.3390/plants8080288

Benedito, V. A., Torres-Jerez, I., Murray, J. D., Andriankaja, A., Allen, S., Kakar, K., Wandrey, M., Verdier, J., Zuber, H., Ott, T., Moreau, S., Niebel, A., Frickey, T., Weiller, G., He, J., Dai, X., Zhao, P. X., Tang, Y., & Udvardi, M. K. (2008). A gene expression atlas of the model legume *Medicago truncatula*. *The Plant Journal*, *55*, 504–513. https://doi.org/10.1111/j.1365-313X.2008.03519.x

Carrere, S., Verdier, J., & Gamas, P. (2021). MtExpress, a comprehensive and curated RNAseq-based gene expression atlas for the model legume *Medicago truncatula*. *Plant and Cell Physiology*, *62*, 1494–1500. https://doi.org/10.1093/pcp/pcab110

Chen, L.-Q., Qu, X.-Q., Hou, B.-H., Sosso, D., Osorio, S., Fernie, A. R., & Frommer, W. B. (2012). Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science*, *335*, 207–211. https://doi.org/10.1126/science.1213351

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*, i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chow, C.-N., Yang, C.-W., Wu, N.-Y., Wang, H.-T., Tseng, K.-C., Chiu, Y.-H., Lee, T.-Y., & Chang, W.-C. (2023). PlantPAN 4.0: Updated database for identifying conserved non-coding sequences and exploring dynamic transcriptional regulation in plant promoters. *Nucleic Acids Research*, *52*, D1569–D1578. https://doi.org/10.1093/nar/gkad945

Dahl, W. J., Foster, L. M., & Tyler, R. T. (2012). Review of the health benefits of peas (*Pisum sativum* L.). *British Journal of Nutrition*, *108*(Suppl 1), S3–S10. https://doi.org/10.1017/S0007114512000852

de Zélicourt, A., Diet, A., Marion, J., Laffont, C., Ariel, F., Moison, M., Zahaf, O., Crespi, M., Gruber, V., & Frugier, F. (2012). Dual involvement of a *Medicago truncatula* NAC transcription factor in root abiotic stress response and symbiotic nodule senescence. *The Plant Journal*, *70*, 220–230.

Die, J. V., Román, B., Nadal, S., & González-Verdejo, C. I. (2010). Evaluation of candidate reference genes for expression studies in *Pisum sativum* under different experimental conditions. *Planta*, *232*, 145–153. https://doi.org/10.1007/s00425-010-1158-1

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2012). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635

Doidy, J., van Tuinen, D., Lamotte, O., Corneillat, M., Alcaraz, G., & Wipf, D. (2012). The *Medicago truncatula* sucrose transporter family. Characterization and implication of key members in carbon partitioning towards arbuscular mycorrhizal fungi. *Molecular Plant*, *5*, 1346–1358. https://doi.org/10.1093/mp/sss1079

Doidy, J., Vidal, U., & Lemoine, R. (2019). Sugar transporters in Fabaceae, featuring SUT MST and SWEET families of the model plant *Medicago truncatula* and the agricultural crop *Pisum sativum*. *PLoS One*, *14*, e0223173. https://doi.org/10.1371/journal.pone.0223173

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Doidy, J., Wang, Y., Gouaille, L., Goma-Louamba, I., Jiang, Z., Pourtau, N., Le Gourrierec, J., & Sakr, S. (2024). Sugar transport and signaling in shoot branching. *International Journal of Molecular Sciences*, *25*, 13214. https://doi.org/10.3390/ijms252313214

Henriet, C., Aimé, D., Térézol, M., Kilandamoko, A., Rossin, N., Combes-Soia, L., Labas, V., Serre, R.-F., Prudent, M., Kreplak, J., Vernoud, V., & Gallardo, K. (2019). Water stress combined with sulfur deficiency in pea affects yield components but mitigates the effect of deficiency on seed globulin composition. *Journal of Experimental Botany*, *70*, 4287–4304. https://doi.org/10.1093/jxb/erz114

Huang, S., Gali, K. K., Lachagari, R. V. B., Chakravartty, N., Bueckert, R. A., Tar'an, B., & Warkentin, T. D. (2021). Identification of heat responsive genes in pea stipules and anthers through transcriptional profiling. *PLoS One*, *16*, e0251167. https://doi.org/10.1371/journal.pone.0251167

Joshi, R., Wani, S. H., Singh, B., Bohra, A., Dar, Z. A., Lone, A. A., Pareek, A., & Singla-Pareek, S. L. (2016). Transcription factors and plants response to drought stress: Current understanding and future directions. *Frontiers in Plant Science*, *7*, 1029. https://doi.org/10.3389/fpls.2016.01029

Katari, M. S., Nowicki, S. D., Aceituno, F. F., Nero, D., Kelfer, J., Thompson, L. P., Cabello, J. M., Davidson, R. S., Goldberg, A. P., Shasha, D. E., Coruzzi, G. M., & Gutiérrez, R. A. (2010). VirtualPlant: A software platform to support systems biology research. *Plant Physiology*, *152*, 500–515. https://doi.org/10.1104/pp.109.147025

Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., D'Angelo, C., Bornberg-Bauer, E., Kudla, J., & Harter, K. (2007). The AtGenExpress global stress expression data set: Protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *The Plant Journal*, *50*, 347–363. https://doi.org/10.1111/j.1365-313X.2007.03052.x

Knopkiewicz, M., & Wojtaszek, P. (2018). Validation of reference genes for gene expression analysis using quantitative polymerase chain reaction in pea lines (*Pisum sativum*) with different lodging susceptibility. *Annals of Applied Biology*, *174*, 86–91. https://doi.org/10.1111/aab.12475

Kodjovi, G. C., Pourtau, N., Moumen, B., & Doidy, J. (2023). Transcription factor dataset in the pea genome. *Social Science Research Network*. https://doi.org/10.2139/ssrn.4420182

Kreplak, J., Madoui, M.-A., Cápal, P., Novák, P., Labadie, K., Aubert, G., Bayer, P. E., Gali, K. K., Syme, R. A., Main, D., Klein, A., Bérard, A., Vrbová, I., Fournier, C., d'Agata, L., Belser, C., Berrabah, W., Toegelová, H., Milec, Z., & Burstin, J. (2019). A reference genome for pea provides insight into legume genome evolution. *Nature Genetics*, *51*, 1411–1422. https://doi.org/10.1038/s41588-019-0480-1

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

Lemoine, R., Camera, S. L. A., Atanassova, R., Dédaldéchamp, F., Allario, T., Pourtau, N., Bonnemain, J.-L., Laloi, M., Coutos-Thévenot, P., Maurousset, L., Faucher, M., Girousse, C., Lemonnier, P., Parrilla, J., & Durand, M. (2013). Source-to-sink transport of sugar and regulation by environmental factors. *Frontiers in Plant Science*, *4*, Article 272. https://doi.org/10.3389/fpls.2013.00272

Li, Y., Varala, K., & Coruzzi, G. M. (2015). From milliseconds to lifetimes: Tracking the dynamic behavior of transcription factors in gene networks. *Trends in Genetics*, *31*, 509–515. https://doi.org/10.1016/j.tig.2015.05.005

Liao, Y., Smyth, G. K., & Shi, W. (2013). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656

Lin, J., Monsalvo, I., Kwon, H., Pullano, S., & Kovinich, N. (2024). The WRKY family transcription factor GmWRKY72 represses glyceollin phytoalexin biosynthesis in soybean. *Plants*, *13*, 3036. https://doi.org/10.3390/plants13213036

Liu, N., Lyu, X., Zhang, X., Zhang, G., Zhang, Z., Guan, X., Chen, X., Yang, X., Feng, Z., Gao, Q., Shi, W., Deng, Y., Sheng, K., Ou, J., Zhu, Y., Wang, B., Bu, Y., Zhang, M., Zhang, L., & Gong, Y. (2024). Reference genome sequence and population genomic analysis of peas provide insights into the genetic basis of Mendelian and other agronomic traits. *Nature Genetics*, *56*, 1964–1974. https://doi.org/10.1038/s41588-024-01867-8

Lu, M.-Z., Snyder, R., Grant, J., & Tegeder, M. (2020). Manipulation of sucrose phloem and embryo loading affects pea leaf metabolism, carbon and nitrogen partitioning to sinks as well as seed storage pools. *The Plant Journal*, *101*, 217–236. https://doi.org/10.1111/tpj.14533

Malovichko, Y. V., Shtark, O. Y., Vasileva, E. N., Nizhnikov, A. A., & Antonets, K. S. (2020). Transcriptomic insights into mechanisms of early seed maturation in the garden pea (*Pisum sativum* L.). *Cells*, *9*, 779. https://doi.org/10.3390/cells9030779

Mergaert, P., Kereszt, A., & Kondorosi, E. (2019). Gene expression in nitrogen-fixing symbiotic nodule cells in *Medicago truncatula* and other nodulating plants. *The Plant Cell*, *32*, 42–68. https://doi.org/10.1105/tpc.19.00494

Morin, A., Kadi, F., Porcheron, B., Vriet, C., Maurousset, L., Lemoine, R., Pourtau, N., & Doidy, J. (2022). Genome-wide identification of invertases in Fabaceae, focusing on transcriptional regulation of *Pisum sativum* invertases in seed subjected to drought. *Physiologia Plantarum*, *174*, e13673. https://doi.org/10.1111/ppl.13673

Morin, A., Porcheron, B., Kodjovi, G. C., Moumen, B., Vriet, C., Maurousset, L., Lemoine, R., Pourtau, N., & Doidy, J. (2023). Genome-wide transcriptional responses to water deficit during seed development in *Pisum sativum*, focusing on sugar transport and metabolism. *Physiologia Plantarum*, *175*, e14062. https://doi.org/10.1111/ppl.14062

Padhy, A. K., Singh, A., Chaurasia, S., Parida, S. K., Tripathi, K., & Bhatia, S. (2025). Key determinants of seed size for enhancing genetic gain in legumes. *Plant, Cell & Environment*. https://doi.org/10.1111/pce.15632

Pandey, A. K., Rubiales, D., Wang, Y., Fang, P., Sun, T., Liu, N., & Xu, P. (2021). Omics resources and omics-enabled approaches for achieving high productivity and improved quality in pea (*Pisum sativum* L.). *Theoretical and Applied Genetics*, *134*, 755–776. https://doi.org/10.1007/s00122-020-03751-5

Pegler, J. L., Patrick, J. W., McDermott, B., Brown, A., Oultram, J. M. J., Grof, C. P. L., & Ward, J. M. (2024). *Phaseolus vulgaris* STP13.1 is an H+-coupled monosaccharide transporter, present in source leaves and seed coats, with higher substrate affinity at depolarized potentials. *Plant Direct*, *8*, e585. https://doi.org/10.1002/pld3.585

Perroud, P.-F., Haas, F. B., Hiss, M., Ullrich, K. K., Alboresi, A., Amirebrahimi, M., Barry, K., Bassi, R., Bonhomme, S., Chen, H., Coates, J. C., Fujita, T., Guyon-Debast, A., Lang, D., Lin, J., Lipzen, A., Nogué, F., Oliver, M. J., Ponce De León, I., & Rensing, S. A. (2018). The *Physcomitrella patens* gene atlas project: Large-scale RNA-seq based expression data. *The Plant Journal*, *95*, 168–182. https://doi.org/10.1111/tpj.13940

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Petridis, A., Döll, S., Nichelmann, L., Bilger, W., & Mock, H.-P. (2016). *Arabidopsis thaliana* G2-LIKE FLAVONOID REGULATOR and BRASSINOSTEROID ENHANCED EXPRESSION1 are low-temperature regulators of flavonoid accumulation. *New Phytologist*, *211*, 912–925. https://doi.org/10.1111/nph.13986

Pfaffl, M. W., Tichopad, A., Prgomet, C., & Neuvians, T. P. (2004). Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnology Letters*, *26*, 509–515. https://doi.org/10.1023/B:BILE.0000019559.84305.47

Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., Davey, M., Jacobs, J., van Ex, F., Pasha, A., Khedikar, Y., Robinson, S. J., Cory, A. T., Florio, T., Concia, L., Juery, C., Schoonbeek, H., Steuernagel, B., Xiang, D., … Tan, Y. (2018). The transcriptional landscape of polyploid wheat. *Science*, *361*, eaar6089.

Schauser, L., Roussis, A., Stiller, J., & Stougaard, J. (1999). A plant regulator controlling development of symbiotic root nodules. *Nature*, *402*, 191–195. https://doi.org/10.1038/46058

Severin, A. J., Woody, J. L., Bolon, Y.-T., Joseph, B., Diers, B. W., Farmer, A. D., Muehlbauer, G. J., Nelson, R. T., Grant, D., Specht, J. E., Graham, M. A., Cannon, S. B., May, G. D., Vance, C. P., & Shoemaker, R. C. (2010). RNA-seq Atlas of *Glycine max*: A guide to the soybean transcriptome. *BMC Plant Biology*, *10*, Article 160. https://doi.org/10.1186/1471-2229-10-160

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome research*, *13*, 2498–2504. https://doi.org/10.1101/gr.1239303

Shen, L., & Feng, J. (2024). NIN—At the heart of nitrogen-fixing Nodule symbiosis. *Frontiers in Plant Science*, *14*, 1284720. https://doi.org/10.3389/fpls.2023.1284720

Shinozaki, Y., Nicolas, P., Fernandez-Pozo, N., Ma, Q., Evanich, D. J., Shi, Y., Xu, Y., Zheng, Y., Snyder, S. I., Martin, L. B. B., Ruiz-May, E., Thannhauser, T. W., Chen, K., Domozych, D. S., Catalá, C., Fei, Z., Mueller, L. A., Giovannoni, J. J., & Rose, J. K. C. (2018). High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening. *Nature Communications*, *9*, Article 364. https://doi.org/10.1038/s41467-017-02782-9

Silver, N., Best, S., Jiang, J., & Thein, S. L. (2006). Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Molecular Biology*, *7*, Article 33. https://doi.org/10.1186/1471-2199-7-33

Singh, A., Sutradhar, M., Prasad, S. S., Ramteke, P. W., Rajeshwar, J. R., & Vagvolgyi, C. (2025). Plant novel transcription factors and their roles in gene regulation. *Preprints.org*. 10.20944/preprints202502.1341.v1

Smitha, P. K., Vishnupriyan, K., Kar, A. S., Anil Kumar, M., Bathula, C., Chandrashekara, K. N., Dhar, S. K., & Das, M. (2019). Genome wide search to identify reference genes candidates for gene expression analysis in *Gossypium hirsutum*. *BMC Plant Biology*, *19*, Article 405. https://doi.org/10.1186/s12870-019-1988-3

Smýkal, P., Aubert, G., Burstin, J., Coyne, C. J., Ellis, N. T. H., Flavell, A. J., Ford, R., Hýbl, M., Macas, J., Neumann, P., McPhee, K. E., Redden, R. J., Rubiales, D., Weller, J. L., & Warkentin, T. D. (2012). Pea (*Pisum sativum* L.) in the genomic era. *Agronomy*, *2*, 74–115.

Sreedasyam, A., Plott, C., Hossain, M. S., Lovell, J. T., Grimwood, J., Jenkins, J. W., Daum, C., Barry, K., Carlson, J., Shu, S., Phillips, J., Amirebrahimi, M., Zane, M., Wang, M., Goodstein, D., Haas, F. B., Hiss, M., Perroud, P.-F., Jawdy, S. S., … Schmutz, J. (2023). JGI Plant Gene Atlas: An updateable transcriptome resource to improve functional gene descriptions across the plant kingdom. *Nucleic Acids Research*, *51*, 8383–8401. https://doi.org/10.1093/nar/gkad616

Staats, M., & van Kan, J. A. (2012). Genome update of *Botrytis cinerea* strains B05.10 and T4. *Eukaryotic Cell*, *11*, 1413–1414. https://doi.org/10.1128/EC.00164-12

Sudheesh, S., Sawbridge, T. I., Cogan, N. O. I., Kennedy, P., Forster, J. W., & Kaur, S. (2015). De novo assembly and characterisation of the field pea transcriptome using RNA-seq. *BMC Genomics*, *16*, Article 611. https://doi.org/10.1186/s12864-015-1815-7

Sullivan, A., Purohit, P. K., Freese, N. H., Pasha, A., Esteban, E., Waese, J., Wu, A., Chen, M., Chin, C. Y., Song, R., Watharkar, S. R., Chan, A. P., Krishnakumar, V., Vaughn, M. W., Town, C., Loraine, A. E., & Provart, N. J. (2019). An 'eFP-Seq Browser' for visualizing and exploring RNA sequencing data. *The Plant Journal*, *100*, 641–654. https://doi.org/10.1111/tpj.14468

Tang, J., & Bassham, D. C. (2022). Autophagy during drought: Function, regulation, and potential application. *The Plant Journal*, *109*, 390–401.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, *3*, Article research0034.1. https://doi.org/10.1186/gb-2002-3-7-research0034

Vernié, T., Moreau, S., De Billy, F. ranç., Plet, J., Combier, J.-P., Rogers, C., Oldroyd, G., Frugier, F., Niebel, A., & Gamas, P. (2008). EFD is an ERF transcription factor involved in the control of nodule number and differentiation in *Medicago truncatula*. *The Plant Cell*, *20*, 2696–2713. https://doi.org/10.1105/tpc.108.059857

Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., Cumming, M., Kelley, L. A., Sternberg, M. J., Krishnakumar, V., Ferlanti, E., Miller, J., Town, C., Stuerzlinger, W., & Provart, N. J. (2017). ePlant: Visualizing and exploring multiple levels of data for hypothesis generation in plant biology. *The Plant Cell*, *29*, 1806–1821. https://doi.org/10.1105/tpc.17.00073

Wang, P., Nolan, T. M., Yin, Y., & Bassham, D. C. (2020). Identification of transcription factors that regulate ATG8 expression and autophagy in Arabidopsis. *Autophagy*, *16*, 123–139. https://doi.org/10.1080/15548627.2019.1598753

Wang, S., Yokosho, K., Guo, R., Whelan, J., Ruan, Y.-L., Ma, J. F., & Shou, H. (2019). The soybean sugar transporter GmSWEET15 mediates sucrose export from endosperm to early embryo. *Plant Physiology*, *180*, 2133–2141. https://doi.org/10.1104/pp.19.00641

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., & Provart, N. J. (2007). An "electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One*, *2*, e718. https://doi.org/10.1371/journal.pone.0000718

Xie, F., Xiao, P., Chen, D., Xu, L., & Zhang, B. (2012). miRDeepFinder: A miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Molecular Biology*, *80*, 75–84. https://doi.org/10.1007/s11103-012-9885-2

Yang, T., Liu, R., Luo, Y., Hu, S., Wang, D., Wang, C., Pandey, M. K., Ge, S., Xu, Q., Li, N., Li, G., Huang, Y., Saxena, R. K., Ji, Y., Li, M., Yan, X., He, Y., Liu, Y., Wang, X., & Zong, X. (2022). Improved pea reference genome and pan-genome highlight genomic features and evolutionary characteristics. *Nature Genetics*, *54*, 1553–1563. https://doi.org/10.1038/s41588-022-01172-2

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Yang, Y., Sun, T., Xu, L., Pi, E., Wang, S., Wang, H., & Shen, C. (2015). Genome-wide identification of *CAMTA* gene family members in *Medicago truncatula* and their expression during root nodule symbiosis and hormone treatments. *Frontiers in Plant Science*, *6*, 459.

Zhang, C., & Turgeon, R. (2009). Downregulating the sucrose transporter *VpSUT1* in *Verbascum phoeniceum* does not inhibit phloem loading. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 18849–18854. https://doi.org/10.1073/pnas.0904189106

Zhang, X., Luo, Z., Marand, A. P., Yan, H., Jang, H., Bang, S., Mendieta, J. P., Minow, M. A. A., & Schmitz, R. J. (2025). A spatially resolved multi-omic single-cell atlas of soybean development. *Cell*, *188*, 550–567.E519. https://doi.org/10.1016/j.cell.2024.10.050

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: Kodjovi, G. C., Coulon, M., Couturier, L., Morin, A., Goma-Louamba, I., Artault, C., Tcherkez, G., Vriet, C., La Camera, S., Pourtau, N., Moumen, B., & Doidy, J. (2025). A transcriptomic atlas facilitating systems biology approaches in pea. *Crop Science*, *65*, e70194. https://doi.org/10.1002/csc2.70194

14350653, 2025, 6, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/csc2.70194 by Sylvain LA CAMERA - Cochrane France , Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License