



HAL
open science

Genome-wide identification of invertases in Fabaceae , focusing on transcriptional regulation of *Pisum sativum* invertases in seed subjected to drought

Amélie Morin, Fadia Kadi, Benoit Porcheron, Cécile Vriet, Laurence
Maurousset, Rémi Lemoine, Nathalie Pourtau, Joan Doidy

► To cite this version:

Amélie Morin, Fadia Kadi, Benoit Porcheron, Cécile Vriet, Laurence Maurousset, et al.. Genome-wide identification of invertases in Fabaceae , focusing on transcriptional regulation of *Pisum sativum* invertases in seed subjected to drought. *Physiologia Plantarum*, 2022, 174 (2), 10.1111/ppl.13673 . hal-04027277

HAL Id: hal-04027277

<https://hal.science/hal-04027277>

Submitted on 15 Mar 2023

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 Attribution - NonCommercial - NoDerivatives 4.0
International License

SPECIAL ISSUE ARTICLE

Genome-wide identification of invertases in *Fabaceae*, focusing on transcriptional regulation of *Pisum sativum* invertases in seed subjected to drought

Amélie Morin  | Fadia Kadi | Benoit Porcheron | Cécile Vriet  |
 Laurence Maurousset | Rémi Lemoine | Nathalie Pourtau  | Joan Doidy 

Université de Poitiers, UMR CNRS 7267, EBI
 “Ecologie et Biologie des Interactions”,
 Poitiers, France

Correspondence

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

Funding information

Centre National de la Recherche Scientifique;
 Grand Poitiers; Ministère de l'Enseignement
 Supérieur et de la Recherche Scientifique;
 Université de Poitiers

Edited by R. Le Hir

Abstract

Invertases are key enzymes for carbon metabolism, cleaving sucrose into energy-rich and signaling metabolites, glucose and fructose. Invertases play pivotal roles in development and stress response, determining yield and quality of seed production. In this context, the repertoire of invertase gene families is critically scarce in legumes. Here, we performed a systematic search for invertase families in 16 *Fabaceae* genomes. For instance, we identified 19 invertase genes in the model plant *Medicago* and 17 accessions in the agronomic crop *Pisum sativum*. Our comprehensive phylogenetic analysis sets a milestone for the scientific community as we propose a new nomenclature to correctly name plant invertases. Thus, neutral invertases were classified into four clades of cytosolic invertase (CINV). Acid invertases were classified into two cell wall invertase clades (CWINV) and two vacuolar invertase clades (VINV). Then, we explored transcriptional regulation of the pea invertase family, focusing on seed development and water stress. Invertase expression decreased sharply from embryogenesis to seed-filling stages, consistent with higher sucrose and lower monosaccharide contents. The vacuolar invertase *PsVINV1.1* clearly marked the transition between both developmental stages. We hypothesize that the predominantly expressed cell wall invertase, *PsCWINV1.2*, may drive sucrose unloading towards developing seeds. The same candidates, *PsVINV1.1* and *PsCWINV1.2*, were also regulated by water deficit during embryonic stage. We suggest that *PsVINV1.1* along with vacuolar sugar transporters maintain cellular osmotic pressure and *PsCWINV1.2* control hexose provision, thereby ensuring embryo survival in drought conditions. Altogether, our findings provide novel insights into the regulation of plant carbon metabolism in a challenging environment.

1 | INTRODUCTION

In plants, invertases catalyze the irreversible hydrolysis of sucrose into glucose and fructose, probably the three most important forms of sugar. Sucrose, being the principal end product of photosynthesis, is translocated from source leaf to sink organs, while glucose and fructose act both as energy-rich and signaling metabolites (Rolland

et al., 2006; Wind et al., 2010). Therefore, sucrose-cleaving enzymes play major roles in carbon partitioning, metabolism, plant nutrition, yield, signaling and stress tolerance (Ruan et al., 2010).

Invertases (EC 3.2.1.26) are generally classified based on their subcellular localization, solubility, and pH activity. The first group refers to “neutral/alkaline invertase,” with an optimal pH of 6.8–9.0, also called “cytosolic invertase” (CINV). Beyond cytosolic localization,

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Physiologia Plantarum* published by John Wiley & Sons Ltd on behalf of Scandinavian Plant Physiology Society.

CINV have also been reported in mitochondria, plastid and nucleus compartments (Murayama & Handa, 2007; Ruan, 2014). In contrast to acid invertases (see below), CINV has been much less documented. They seem to be involved in the accumulation of sugars in storage tissues (Wang et al., 2017) and overall control cytosolic sugar homeostasis for cellular function (Wan et al., 2018).

The second group comprises “acid invertases” that cleave sucrose most effectively at a lower pH of 3.5–5, and refers to both “cell wall invertase” (CWINV) and “vacuolar invertase” (VINV). VINVs can also be called soluble acid invertase, while CWINVs are bound to the cell wall layer, thus referred to the insoluble type of invertase. In vascular plants, CWINVs have an essential role in driving carbon fluxes towards non-photosynthetic organs, defined as sink organs (Ruan, 2014). By regulating the sucrose/hexose ratio, they contribute to phloem unloading towards competitive sink organs (e.g. root and seed) in response to developmental needs and environmental conditions. For instance, CWINVs perform a key metabolic role in regulating available carbon sources in plant-microbe interactions and in responses to abiotic stress (Ruan, 2014). Regarding the vacuolar type of invertase, VINVs generally contribute to sugar reserves in intracellular compartments. VINVs are key modulators for hexose accumulation and cell expansion in sinks organs (Ruan et al., 2010), as for instance in fruits and seeds. Both CWINV and VINV were shown to be associated with seed size in several crops (Mals et al., 2009; Wang et al., 2008; Xu et al., 2019), including legumes (Weber et al., 1996).

In *Fabaceae*, seed development can be temporarily divided into three phases: (i) embryogenesis, (ii) seed-filling, and (iii) desiccation/maturation (Munier-Jolain et al., 2010; Wang et al., 2018; Weber et al., 2005). Embryogenesis is a crucial phase for carbon metabolism regulation, when invertases sustain the developing embryo with energy-rich and signaling metabolites (glucose and fructose). During embryogenesis, sucrose is mainly degraded through the catalytic action of cell wall bound invertases, leading to a high ratio of hexoses to sucrose, thereby promoting embryonic growth through cell division (Wang et al., 2018; Weber et al., 1997). During the later stage of nutrient filling, the invertase activity decreases, and sucrose released from the seed coat is taken up directly by the cotyledons and stored as starch. Thus, by controlling the sugar provision during seed development, these enzymes are associated with crucial milestones, such as abortion rates and seed size during the embryogenesis stage, seed nutritional quality during the seed filling stage, and seed vigor during the maturation stage (Ruan et al., 2012; Wang et al., 2018; Weber et al., 1996). As such, it has been shown that cell wall invertases (Wang et al., 2008) and sugar transporters (Sosso et al., 2015) have been shaped through natural selection to ensure optimal seed dispersal and population fitness, and by human-mediated domestication to improve the nutritional value of seed crops.

Invertase gene families have been identified in many plants, including the model plant species *Arabidopsis* (Ji, Van den Ende, et al., 2005a; Sherson et al., 2003; Xiang & Van den Ende, 2013), agronomic crops like maize (*Zea mays*, Juárez-Colunga et al., 2018), and even in *Populus* tree species (Chen et al., 2015). Also, their crucial roles in crop yield and in response to abiotic stress have been

highlighted (Ruan et al., 2010). For instance, the invertase gene family has been mined in sugar cane, the expression levels of INV genes and the variation of enzymatic activities were evaluated in response to environmental stress, such as drought (Wang et al., 2017). However, invertase families have not yet been uncovered in legumes, which include major agro-ecological crops, cultivated worldwide as a sustainable source of plant protein and starch food/feed. To our knowledge, the invertase gene family has only been studied in soybean (*Glycine max*, Su et al., 2018). Therefore, the complete invertase families from model *Fabaceae* species like *Medicago truncatula* or *Lotus japonicus*, and from agronomic crops such as *Pisum sativum*, are still lacking and their contribution under abiotic stress has been overlooked in legumes.

Here, we mainly focused on *Pisum sativum*, a cultivated legume with a high nutritional seed value. Pea is also a model organism in biology since the work of famous geneticist Gregor Mendel. Nowadays, pea is entering the genomic era (Smýkal et al., 2012) with the recent release of a reference genome (Burstin et al., 2020; Kreplak et al., 2019). Our team has recently annotated major gene families controlling sugar transport in *P. sativum* (Doidy et al., 2019). Here, we uncover invertase families in sequenced *Fabaceae* genomes. Then, we focus on a comprehensive transcriptional analysis of the 17 invertase genes in *P. sativum*, from both a developmental aspect (from embryogenesis to seed-filling stage) and an environmental aspect (in response to water deficit). We also explored the gene expression profile of major invertase inhibitors and relevant vacuolar sugar transporters. From this study, we identify key candidate genes putatively involved in driving the carbon demand towards the developing seeds and maintaining cellular osmotic pressure under water-deficient conditions.

2 | MATERIAL AND METHODS

2.1 | Plant material and experimental design

Seeds of *Pisum sativum* L. (Safran genotype, Unisigma) were pre-germinated for 6 days at 22°C in the dark. The germinated seeds were transferred individually in 2.5 L pots containing a mixture of loamy soil and sand (2v–1v), and grown in a greenhouse with a 14 h photoperiod complemented with artificial lighting (120 $\mu\text{mol}/\text{m}^2/\text{s}$) when natural light was below 150 W/m^2 . The determination of soil water content was made at the beginning of the experimentation by saturating the substrate with water and drying for 48 h at 105°C. Seedlings were then irrigated to 80% of the water holding capacity every 2 days with tap water, and 50 ml of low nitrate nutrient solution (pH 6.5) was supplied once a week (Zuber et al., 2013). When the first flowering node was pollinated, irrigation was stopped for “water-deficient” plants (WD) until the soil water content reached 40% of the water holding capacity, while “well-watered” plants (WW) were continuously watered to 80% of the water holding capacity. Soil water content was maintained (to 80% of the water holding capacity for WW and 40% for WD plants) by weighing all pots daily. It took 2 days for plants subjected to water deficit to reach 40% of the water holding capacity, this

moment also concurred with the pollination of the first flower on the second flowering node. Water deficit was maintained until seed harvest 7 or 14 days after pollination (dap). At harvest, we measured stomatal conductance with a portable leaf porometer (SC-1, METER Group Inc.) on the last developed leaf prior to reproductive development. Volumetric soil water content was measured with a portable Fieldscout TDR 150 soil moisture instrument (Spectrum Technologies Inc.) with 12-cm long-rods to estimate available water for the plant and transpiration at midday, prior to sample collection.

Pod development was followed by a non-destructive approach, using a pea pod sizer developed by Arvalis-Institut du Végétal, to estimate whether the seeds had reached the final stage of seed abortion (FSSA), corresponding to a thickness exceeding 7 mm (Duthion & Pigeaire, 1991). Seeds of the second reproductive node were harvested at 7 dap (embryogenesis stage) or 14 dap (seed-filling stage, FSSA passed) by separating the seed from the maternal funiculus with a scalpel and were immediately placed in liquid nitrogen. A first experiment took place from September 26 to November 24, 2020 for molecular analyses and a second experiment repetition was performed from March 22 to May 25, 2021 for sugar quantification. All harvests were performed at 11:00 h. Experiments were carried out using a split-plot design with five biological replicates per condition for RNA extraction ($n = 5$) and four replicates of three pooled plants to obtain sufficient seed material at 7 dap while seeds of four individual plants were collected at 14 dap for sugar quantification ($n = 4$).

2.2 | Sequence identification of INV accessions in *Fabaceae* genomes

We mined *Fabaceae* genomes by sequence similarity to previously identified INV accessions using the blastP algorithm available at Phytozome v12 (<https://phytozome.jgi.doe.gov/pz/portal.html> Goodstein et al., 2012) for *Glycine max* Wm82.a2 (v1), *Medicago truncatula* (v4.0), *Phaseolus vulgaris* (v2.1), and *Trifolium pratense* (v2); the Pea Genome project for the recent release of the *Pisum sativum* (v1) genome (<https://urgi.versailles.inra.fr/Species/Pisum/Pea-Genome-project>; Kreplak et al., 2019), and the Legume Information System (<https://legumeinfo.org/home>, Dash et al., 2016) for *Arachis hypogaea*, *Arachis duranensis*, *Arachis ipaensis*, *Cajanus cajan*, *Cicer arietinum* (CDC Frontier kabuli type and ICC 4958 desi type), *Lotus japonicus* (v3.0), *Lupinus angustifolius* (v1.0), *Vigna angularis* (v3.0), *Vigna radiata*, and *Vigna unguiculata* (v1.0). A list of all accessions presented in this work is summarized in Table S1.

2.3 | Phylogenetic analysis of invertase families

After sequence retrieval from genomic databases, amino-acid alignment of respective INV families was performed using the ClustalW Multiple alignment algorithm on BioEdit (Hall, 1999) and then exported to MEGA7 software (Kumar et al., 2016) for evolutionary

analyses and construction of the phylogenetic trees. The evolutionary history was inferred using the Neighbor-Joining method. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. All ambiguous positions were removed for each sequence pair. There were a total of 828 and 989 positions in the final dataset for the alkaline/neutral INV and acid INV trees, respectively.

2.4 | Analysis of transcriptomic data

Expression profiles of the pea invertase gene family were retrieved from the Pea Gene Atlas (Alves-Carvalho et al., 2015, https://urgi.versailles.inra.fr/download/pea/Pea_PSCAM_transcriptome/). Expression data for gene orthologs in *Medicago truncatula* were retrieved from the Medicago Gene Atlas (Benedito et al., 2008, hosted by LIPME [France] <https://lipm-browsers.toulouse.inra.fr/pub/expressionAtlas/app/mtgeav3/>). All transcriptomic data presented in this study are available in Table S3.

2.5 | Gene expression analysis by RT-qPCR

Total RNA from biological replicates ($n = 5$) was extracted using CTAB extraction method as described in Chang et al. (1993) modified by isopropanol precipitation of RNAs following the second step of chloroform:IAA. The pellet was then washed with ethanol (80°) and resuspended in water for a following purification using the RNeasy Mini Kit (Qiagen). RNA sample of 1 µg was treated with DNase I (Sigma-Aldrich) and then converted to cDNA using the M-MLV reverse transcriptase (Promega) with oligodT primers.

Gene 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 cycles of amplification with two stages (15 s at 95°C, 1 min at 60°C). The specificity of the amplification was checked with a melting curve analysis. Target gene expression was normalized to the reference gene Transcription Factor IIA (Psat5g118600; Die et al., 2010; Knopkiewicz & Wojtaszek, 2018). All primers are listed in Table S2.

2.6 | Soluble sugars quantification

Sucrose, glucose and fructose were extracted from approximately 10 mg of seed sample ($n = 4$) previously lyophilized and washed three times in methanol:chloroform:water (12:5:3 v/v). Supernatants containing soluble sugars were pooled and mixed with 0.6 volume of water and centrifuged. The upper aqueous phase was collected and evaporated at 50°C with a MiVac Quattro (Genevac). Soluble sugars were resuspended in 500 µl of water and quantified using sucrose/D-fructose/D-glucose Assay Kit (Megazyme).

TABLE 1 Comparative analysis of INV gene family and clade distribution in the 16 *Fabaceae* genomes and *Arabidopsis*

Invertase type Clade	Neutral/alkaline invertase				Acid invertase				Total/species
	CINV1	CINV2	CINV3	CINV4	CWINV1	CWINV2	VINV1	VINV2	
<i>Lupinus angustifolius</i>	4	1	4	1	5	2	3	1	21
<i>Arachis hypogaea</i> (4n)	6	4	2	2	10	6	4	2	36
<i>Arachis duranensis</i>	3	2	1	1	5	3	2	1	18
<i>Arachis ipaensis</i>	3	2	1	1	6	3	2	1	19
<i>Phaseolus vulgaris</i>	3	1	1	1	5	3	2	2	18
<i>Glycine max</i> (4n)	6	2	2	3	12	4	4	3	36
<i>Vigna angularis</i>	3	1	1	1	4	3	2	2	17
<i>Vigna radiata</i>	3	1	1	1	4	3	2	2	17
<i>Vigna unguiculata</i>	3	1	1	1	5	3	2	2	18
<i>Cajanus cajan</i>	3	1	1	2	5	1	2	1	16
<i>Lotus japonicus</i>	3	1	1	1	6	2	2	2	18
<i>Medicago truncatula</i>	3	1	2	2	5	4	1	1	19
<i>Pisum sativum</i>	3	1	1	2	5	2	2	1	17
<i>Cicer arietinum</i> (desi)	3	1	1	2	4	2	2	1	16
<i>Cicer arietinum</i> (kabuli)	3	1	1	2	5	2	2	1	17
<i>Trifolium pratense</i>	3	1	2	2	5	2	2	1	18
<i>Arabidopsis thaliana</i>	3	1	5	0	5	1	2	0	17

Note: *Fabaceae* species were ordered according to their phylogenetic relations among the Papilionoid lineage.

3 | RESULTS

3.1 | A phylogenetic review of plant invertase, focusing on *Fabaceae* genomes

First, we mined all 16 available *Fabaceae* genomes to identify invertase gene families in model and crop species (Table 1). Thus, we identified *de novo* 21 LalNV in *Lupinus angustifolius*, 36 AhINV in *Arachis hypogaea*, 18 AdlINV in *Arachis duranensis*, and 19 AilINV in *Arachis ipaensis*; Lupin and peanut being the most basal species with a sequenced genome in the *Fabaceae* family (Bertioli et al., 2019; Hane et al., 2017). In species belonging to the Phaseoloid clade, we identified 18 PvINV in *Phaseolus vulgaris*, 36 GmlINV in *Glycine max*, 17 ValINV in *Vigna angularis*, 17 VrINV in *Vigna radiata*, 18 VulINV in *Vigna unguiculata*, and 16 CcINV in *Cajanus cajan*. We identified 18 LjINV in the model *Lotus japonicus*, which belong to the intermediate Robinoid clade. In species belonging to the “Inverted Repeat-Lacking Clade” (IRLC), we identified 16 CalINV in *Cicer arietinum* (desi type) and 17 CalINV in chickpea (kabuli type), 19 MtINV in *Medicago truncatula*, 17 PsINV in *Pisum sativum*, and 18 TpINV in *Trifolium pratense*. Finally, we also included the *Arabidopsis* invertase family, comprising 17 AtINV, in our phylogenetic analysis.

As neutral and acid invertases share no sequence homology, we aligned CINV independently of CWINV and VINV (Figures S1 and S2), from which we built two distinct phylogenetic trees (Figures 1 and 2). CINV distributed among four clades (CINV1, CINV2, CINV3, and CINV4) while CWINV and VINV both distributed in two clades (CWINV1, CWINV2 and VINV1, VINV2). Here, we propose a new and

simple nomenclature system for naming plant invertase and we renamed all legume and *Arabidopsis* accessions based upon phylogenetic grouping to our proposed nomenclature (Figures 1 and 2; see Section 4.1).

Most Phaseoloid species comprised six CINV genes and most IRLC species comprised seven CINV genes, except for *Medicago* and *Trifolium*, which comprise eight CINV because of gene duplication in the CINV3 clade (Figure 1 and Table 1). Diploid legume species possessed three gene copies in the CINV1 clade, a single accession in both the CINV2 and CINV3 clades, but the number of gene paralogs varies in the CINV4 clade. Indeed, IRLC species seem to harbor an additional accession in the CINV4 (e.g. PsCINV4.1 and PsCINV4.2), while a single gene copy is found in Phaseoloid species (e.g. PvCINV4.1).

Most Phaseoloid species comprised eight CWINV and four VINV genes, while most IRLC species comprised seven CWINV and three VINV genes (Figure 2 and Table 1). This is because diploid Phaseoloid species possess one additional CWINV accession in the CWINV2 clade (e.g. PvCWINV2.3) and one additional VINV accession in the VINV2 clade (e.g. PvVINV2.2) compared with IRLC species (Figure 2).

In total, diploid species possess an invertase family of relatively moderate size, comprising 16–21 gene loci, while both tetraploid species *Arachis hypogaea* and *Glycine max* showed 36 invertase genes (Table 1). Interestingly, *Lupinus angustifolius* showed the highest number of invertases among diploid species (Table 1), which seems the results of duplication events in the CINV (CINV1 and CINV3 clades Figure 1) and in the VINV family (VINV1 Figure 2).

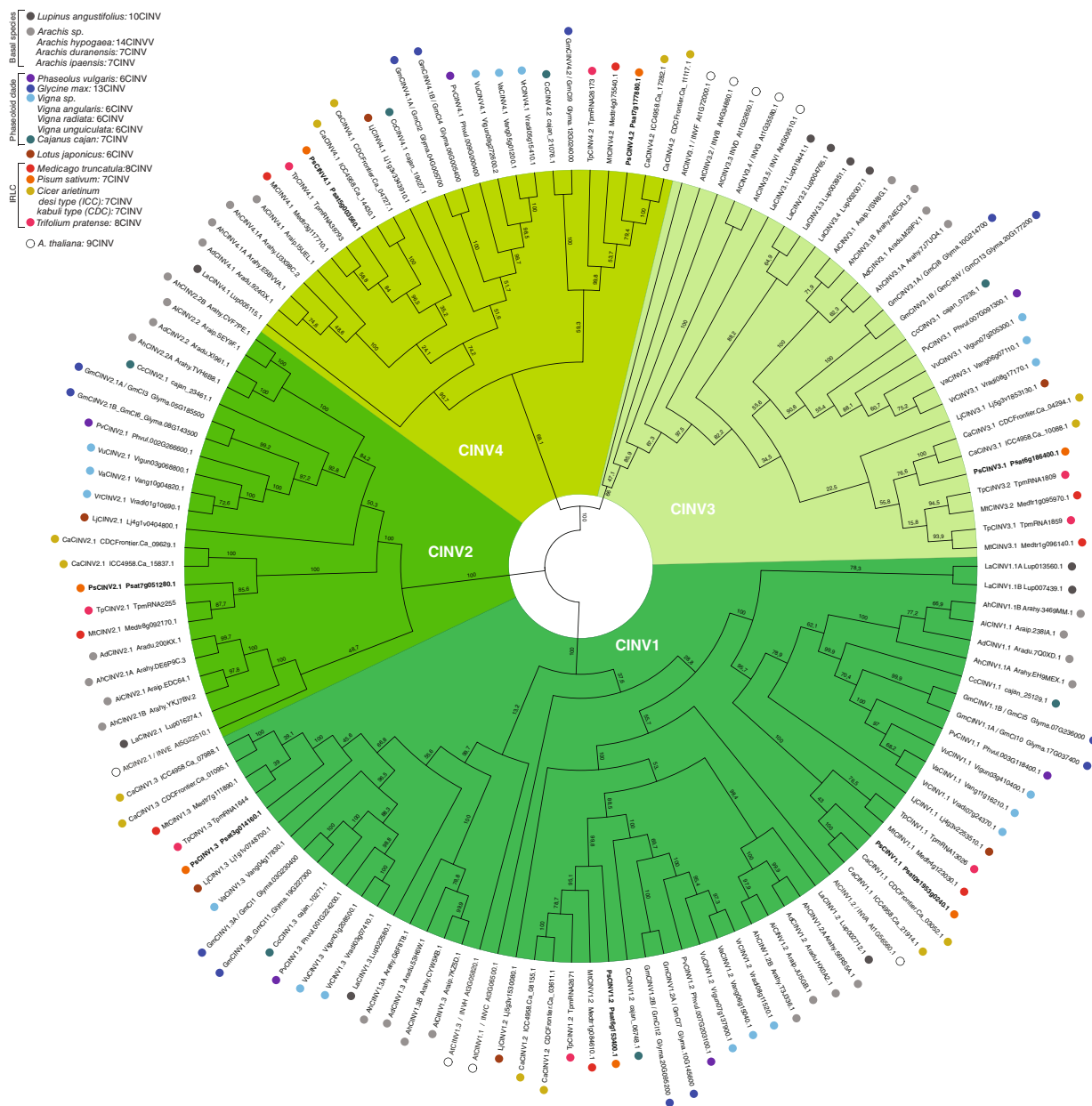


FIGURE 1 Phylogenetic tree of the *Fabaceae* alkaline/neutral invertase gene families. A total of 125 alkaline/neutral invertase accessions were retrieved from 16 legume genomes, along with nine *Arabidopsis* accessions. CINV were named upon phylogenetic grouping into particular clades (CINV1, CINV2, CINV3, and CINV4)

3.2 | Protein sequence motif of acid invertase

Plant invertases harbor conserved motifs in their sequence for sucrose hydrolysis: a β -fructosidase motif (NDPD/NG) close to their N-terminus, a cysteine-containing catalytic domain (WECPD for CWINs and WECVD for VINs), as well as an aspartate residue (D-239) necessary for binding and hydrolyzing sucrose. Invertases lacking either the β -fructosidase motif or the catalytic domain are considered as non-functional enzymes (Wan et al., 2018). Based on this assumption, PsCWINV1.3 (lacking NDPD/NG) and PsCWINV1.5 (with an aberrant MECPD) may be non-functional invertases (Figure S3). The

orthologous proteins MtCWINV1.3 and MtCWINV1.5 from the legume model *M. truncatula* also seem to harbor motifs of non-functional invertases (Figure S3).

Interestingly, CWINV2 accessions cluster with AtCWINV6, a fructan exohydrolase (Figure 2). Putative fructan exohydrolases seem to exhibit specific motifs: (i) a different cysteine-containing catalytic domain (VIECPD in *P. sativum* and VCECPD in *M. truncatula*), (ii) the aspartate necessary for binding and hydrolyzing sucrose, Asp-239 is substituted by Gln (D239Q) in *P. sativum* and by Leu (D239L) in *M. truncatula*, and (iii) amino acid substitutions of the hydrophobic residues Trp-47 (W47F) and Trp-82 (W82F, PsCWINV2.2, and

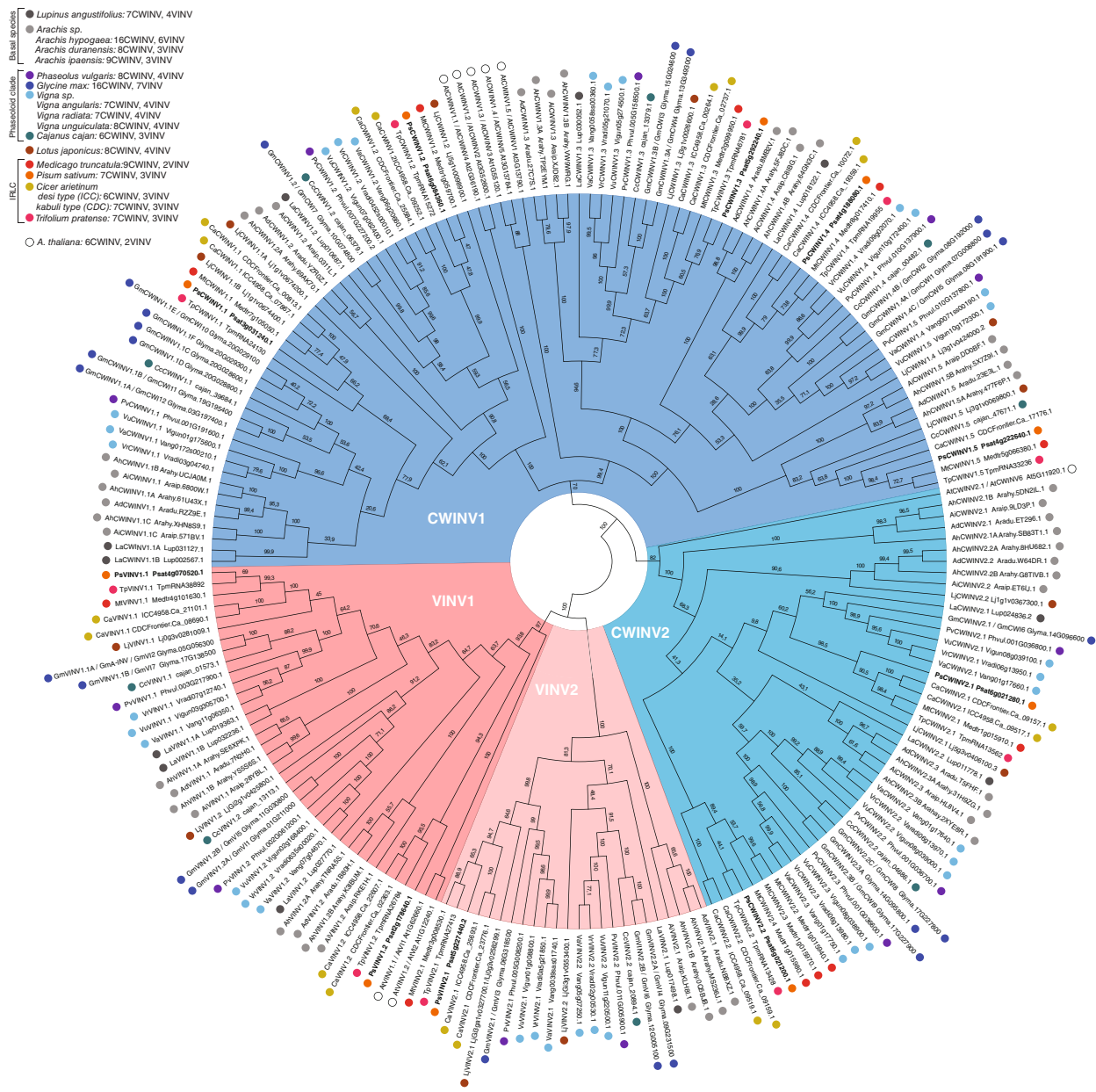


FIGURE 2 Phylogenetic tree of the *Fabaceae* acid invertase gene families. A total of 196 acid invertase accessions were retrieved from 16 legume genomes, along with eight *Arabidopsis* accessions. Acid invertases were named upon phylogenetic grouping into particular cell wall invertase clades (CWINV1 and CWINV2) and vacuolar invertase clades (VINV1 and VINV2)

MtCWINV2.4 only), these aromatic residues being involved in substrate binding from sucrose to fructans (detailed in De Coninck et al., 2005; Le Roy et al., 2007).

3.3 | Invertase expression in *Pisum sativum*, mining a transcriptome atlas

Then, we performed a genome-wide expression approach by screening the transcriptome data available at the Pea Gene Atlas (Alves-Carvalho et al., 2015) to obtain gene expression profiles of the

invertase family in *P. sativum* (Figure 3 and Table S3A). We identified the top three genes that were highly and constitutively expressed in all plant organs (*PsCINV2.1*, *PsCINV3.1* and *PsVINV1.1*), whereas three other genes seemed completely “switched off” in most organs (*PsCWINV1.3*, *PsCINV1.3*, and *PsCINV4.2*). Interestingly, several invertases were specifically expressed only in floral tissues (*PsCWINV1.1*, *PsCWINV1.5*, *PsVINV1.2*, and *PsVINV2.1*) based on this transcriptome data (Figure 3A).

We also screened for gene candidates whose expression was high in seed (Figure 3B). In addition to the top three highly and constitutively expressed genes mentioned before (*PsCINV2.1*, *PsCINV3.1*, and

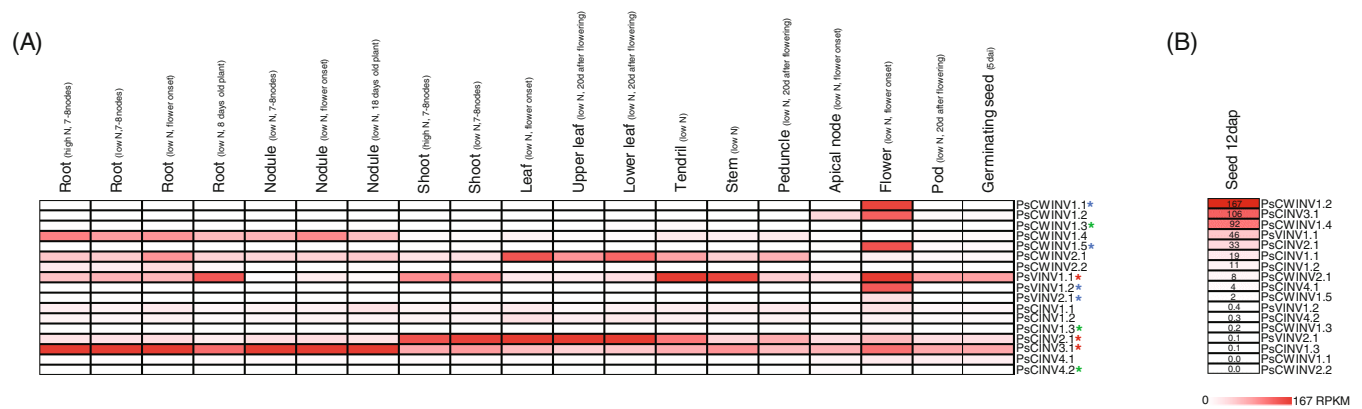


FIGURE 3 Gene expression patterns of the pea invertase gene family. (A) Transcriptomic profiles of the pea invertase gene family (retrieved from the Pea Gene Atlas, Alves-Carvalho et al., 2015). Red stars indicate the top3 most expressed genes (*PsCINV2.1*, *PsCINV3.1*, and *PsVINV1.1*), which also show a constitutive expression conserved in *M. truncatula* (Table S3B). Blue stars indicate four genes expressed in flowers (*PsCWINV1.1*, *PsCWINV1.5*, *PsVINV1.2*, and *PsVINV2.1*). Green stars indicate three genes that seem not expressed in most plant organs (*PsCWINV1.3*, *PsCINV1.3*, and *PsCINV4.2*). (B) Transcriptomic profiles of the pea invertase gene family in seeds 12 days after pollination, which corresponds to the early seed-filling stage. Genes were ordered by gene expression level (normalized RPKM values from high to low), *PsCWINV1.2* being the gene with the highest expression in seed 12 dap. Both heat maps show normalized expression (white: No expression, red: High expression, based on the RPKM values)

PsVINV1.1), we also identified *PsCWINV1.2*, *PsCWINV1.4*, *PsCINV1.1*, and *PsCINV1.2* with significant expression during the early seed-filling stage (12 dap; Figure 3B).

3.4 | Invertase gene expression decreases sharply from embryogenesis to seed-filling stages, consistent with higher sucrose and lower monosaccharide contents

We explored by RT-qPCR the invertase gene expression pattern in seeds harvested at 7 dap (embryogenesis stage) and 14 dap (seed-filling stage). Only a single cell wall invertase (*PsCWINV1.2*) showed a substantially high expression profile in seed (Figure 4A). Regarding cytosolic invertases, *PsCINV3.1*, *PsCINV2.1*, *PsCINV1.1*, and *PsCINV1.2* showed a high expression profile (ranked by order of expression; Table S4). Among vacuolar invertases, only *PsVINV1.1* showed a substantially high expression profile at the embryonic stage (Figure 4A). Altogether, our expression data verified the gene candidates identified from the gene Atlas (Figure 3B).

All INV were highly expressed during the embryogenesis stage. Indeed, cytosolic *PsCINV1.1*, *PsCINV1.2*, *PsCINV3.1*, *PsCINV4.1*, *PsCINV4.2*, cell wall *PsCWINV1.2*, *PsCWINV1.5*, and vacuolar invertases *PsVINV1.1*, *PsVINV1.2*, *PsVINV2.1* were all significantly downregulated during the seed-filling stage (Figure 4A). Among them, *PsVINV1.1* showed the highest expression difference (fold change >500, Table S4) from embryogenesis to seed-filling stages.

In parallel, we also tested the expression pattern of putative invertase inhibitors, differentially regulated from embryogenesis to seed-filling stages. Here, *PsC/VIF2* (*Psat0s2823g0080*) showed the highest expression level compared with the three other putative invertase inhibitors (Figure 4B). Also, *PsC/VIF2* showed an upregulation tendency from

embryogenesis to seed-filling stage, while *PsCIF1* (*Psat0s4829g0040*) and the two other putative inhibitors (*Psat1g101000* and *Psat1g136960*) were all significantly repressed at the filling stage (Figure 4B). Altogether, our result suggests that *PsC/VIF2* is the major inhibitor present in seed organs, likely regulating *PsCWINV* and *PsVINV* at the post-translational level; therefore impacting on soluble sugar levels from embryogenesis to seed-filling stages.

Our result suggests that invertase candidates may drive the sink strength towards seeds by cleaving sucrose and providing energy-rich metabolites (glucose and fructose) to sustain the embryo development. Their downregulation at later seed developmental stages probably allows a high sucrose status to fill the seed with starch reserve. To verify this hypothesis, we quantified soluble sugar content in seeds harvested at the embryonic stage (7 dap) and at the filling stages (14 dap). According to previous observations made in other legume species, pea seed contained significantly higher levels of sucrose (+161%) and lower levels of monosaccharides (glucose –80% and fructose –85%) at the seed-filling stage than at the embryogenesis stage (Figure 4C). This elevated sucrose content at the seed-filling stage is consistent with the concurrent repression of several invertases (Figure 4A).

3.5 | Water deficit impacts *PsCINV1.1*, *PsCWINV1.2*, *PsVINV1.1*, vacuolar transporter expression, and seed sugar contents

A water deficit (WD) corresponding to 40% of the water holding capacity was applied, starting from the pollination of the first flowering node, while well-watered (WW) plants were maintained to 80% of the water holding capacity. The imposed water stress was verified by monitoring the volumetric soil water content, which significantly decreased by 50% due to water deficit (Figure S4A). At the

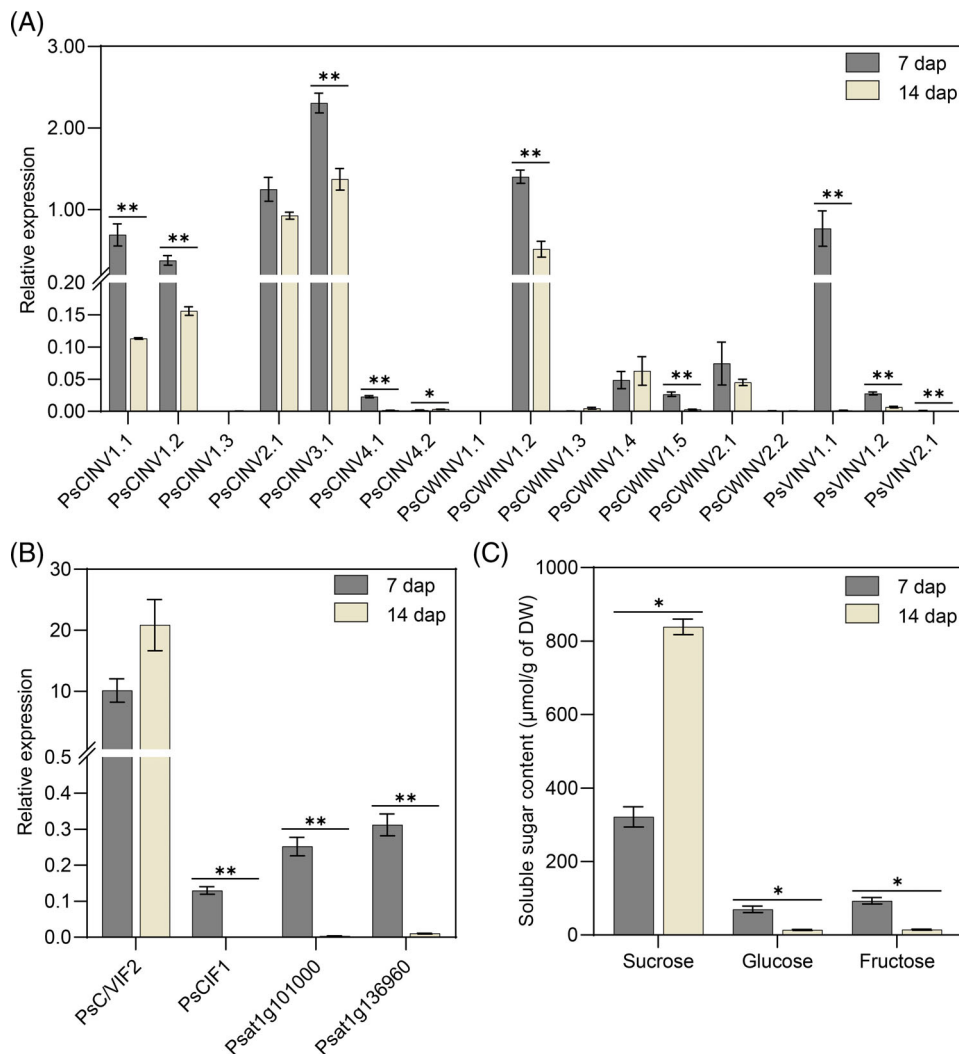


FIGURE 4 (A) Expression profile of the pea invertase family in seeds harvested at 7 dap (embryonic stage) or 14 dap (seed-filling stage). (B) Expression profile of invertase inhibitors in seeds harvested at 7 or 14 dap. (C) Quantification of sucrose, glucose and fructose in seeds harvested at 7 or 14 dap. All measurements were performed on well-watered plants. Experiments were carried out using seeds from five biological replicates for gene expression analysis ($n = 5$, A and B) while sugar quantification was performed with seeds from four replicates of three pooled plants to obtain sufficient material at 7 dap and seeds of four individual plants at 14 dap ($n = 4$, C). Asterisks represent significant differences from 7 dap seeds determined by Kruskal-Wallis test ($*p < 0.05$; $**p < 0.01$)

plant level, stomatal conductance was reduced by 91% and 83% after 7 and 14 days of water deficit, respectively (Figure S4B). Thereafter, we collected seeds subjected to 1 week (7 dap) of water deficit at the embryogenesis stage and seeds subjected to 2 weeks of water deficit (14 dap) at the seed-filling stage.

We explored by RT-qPCR the gene expression of the pea invertase gene family, focusing on gene candidates significantly regulated by water stress. Out of the 17 identified invertases, *PscCINV1.1*, *PscCWINV1.2*, and *PsvINV1.1* showed significant deregulation of their gene expression in response to water deficit at the embryonic stage (Figure 5A). Only water stress applied to the embryonic stage impacts the gene expression of *PscCINV1.1* and *PscCWINV1.2* (Figure 5A), no effect was observed when applied at later stages of seed-filling (Figure S5). *PsvINV1.1* expression significantly increased in response to WD at both developmental stages (Figures 5A and S5), however invertase expression level remained low at the seed-filling stage (Figure S5).

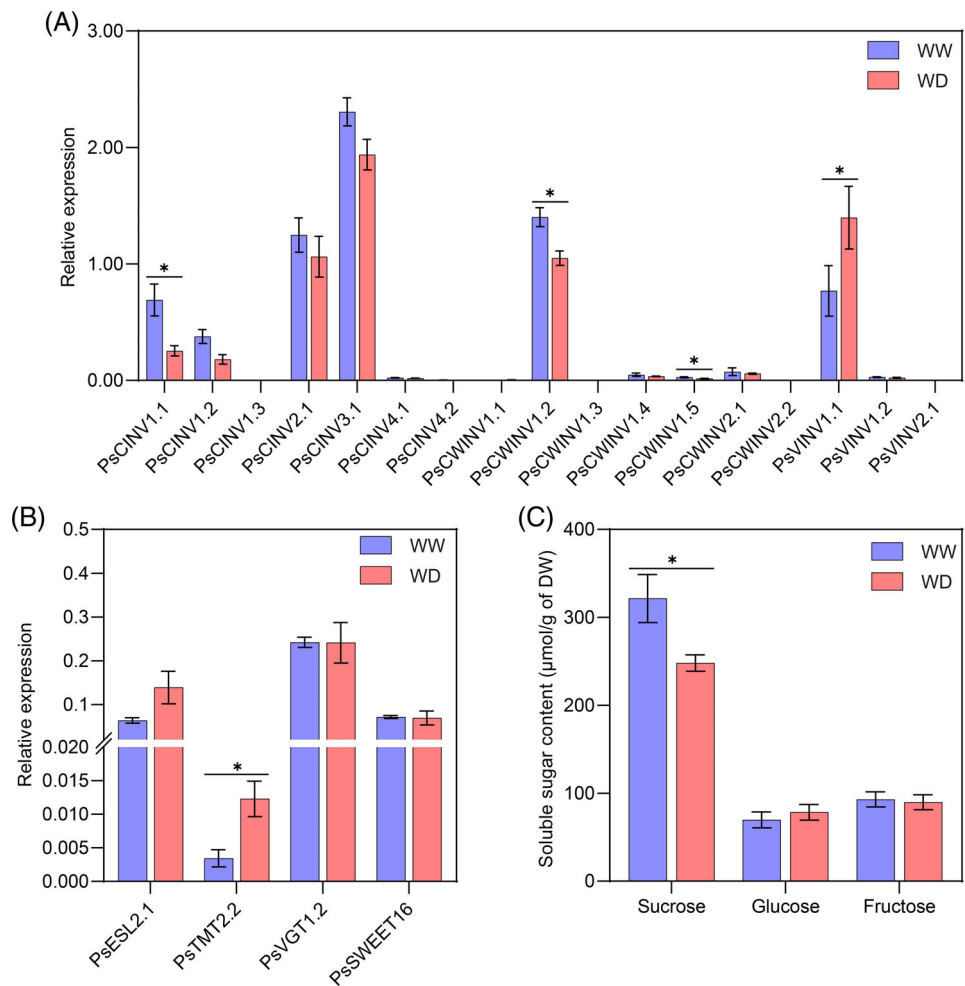
Also, *PscCWINV1.5* was significantly downregulated by water stress during the embryonic stage but showed a low expression level (Figure 5A). At the seed-filling stage, two invertases showed differential regulation of their gene expression, a decrease for *PscCINV4.2* and an increase for *PsvINV1.1* in response to WD (Figure S5) with an

expression level close to qPCR detection threshold (Table S4). Therefore, we preferred to withdraw these candidates with an extremely low expression level from further discussion.

In parallel to gene expression changes of relevant invertases, we also quantified soluble sugar contents of seeds subjected to WD. Overall, soluble sugar content decreased in seeds subjected to water stress during embryogenesis (Figure 5C). At early stage of seed development (7 dap), we observed a significant decrease in sucrose (−23%) in seeds subjected to WD when compared with WW control (Figure 5C). These results suggest that seed sink strength is significantly decreased in water-deficient conditions.

Finally, we also tested the expression pattern of relevant sugar transporters located on the tonoplast. We observed a significant induction upon water stress for the Tonoplast Monosaccharide Transporter (TMT) *PstTMT2.2*. An up-regulation tendency was observed for Early responsive to dehydration Six-Like (ESL) *PseSL2.1* upon water stress, but no regulation for Vacuolar Glucose Transporter (VGT) *PsvGT1.2* and Sugar Will Eventually be Exported Transporters (SWEET) *PswSWEET16* (Figure 5B). These results suggest a readjustment of intracellular sugar reserves, likely to maintain cell turgescence in response to water stress.

FIGURE 5 (A) Expression profile of the pea invertase family in seeds harvested at 7 dap in well-watered (WW, blue) and water-deficient (WD, red) conditions. (B) Expression profile of vacuolar sugar transporters in seeds harvested at 7 dap in WW and WD conditions. (C) Quantification of sucrose, glucose and fructose in seeds harvested at 7 dap in WW and WD conditions. Experiments were carried out using seeds from five biological replicates for gene expression analysis ($n = 5$, A and B) while sugar quantification was performed with seeds from four replicates of three pooled plants ($n = 4$, C). Asterisks represent significant differences from WW control plants determined by Kruskal–Wallis test ($*p < 0.05$; $**p < 0.01$)



4 | DISCUSSION

4.1 | A new and simple nomenclature: How to correctly name plant invertases?

Our comprehensive phylogenetic analysis (Figures 1 and 2) sets a milestone for the scientific community, as we propose a new and simple nomenclature to correctly name plant invertases. We raise the attention of the scientific community to consistently annotate invertase genes according to their phylogenetic position rather than naming accessions following the chronological order of their discovery. We believe that this new nomenclature will enable consistent mapping and annotation of invertases across plant genomes.

Simply, we named invertases belonging to respective phylogenetic clades (e.g. “CINV1” for cytosolic invertase clade 1; see Figure 1). We then added a “.number” extension when multiple paralogous genes were present within the same clade (e.g. *PsCINV1.1*, *PsCINV1.2*, and *PsCINV1.3*). We tried our best to correctly identify orthologous genes between different species by employing identical extensions (e.g. *MtCINV1.1* and *PsCINV1.1*) when such accessions clearly appeared within the same branch of a particular clade. Finally, we also added a capital letter extension to clearly identify all representative accessions (e.g. *GmCINV1.1A* and *GmCINV1.1B*) when multiple paralogs or genes

occurring from genome duplication appeared within the same branch (e.g. in polyploid species like *G. max* and *A. hypogaea*).

We renamed the Arabidopsis invertase gene family accordingly (Figures 1 and 2). Previously, CINV accessions in Arabidopsis were named with letters (A–H; Ji, Van den Ende, et al., 2005a; Xiang et al., 2011), while numbers were used to annotate acid invertases (Sherson et al., 2003; Xiang & Van den Ende, 2013; Table S1). Such previous annotations were not consistent with the phylogenetic relations among INV genes.

We also renamed the soybean invertase gene family accordingly (Figures 1 and 2). Previously, 32 invertase genes were identified in *G. max* (Su et al., 2018), and here we identified four additional GmCWINV accessions (Table S1), now making a total of 36 invertases in the soybean genome. Interestingly, all newly identified accessions consist of tandem repeats in this complex tetraploid genome.

4.2 | Invertase gene family in Fabaceae and Arabidopsis genomes

Acid invertases exhibit little homology to neutral invertases. CINV belong to the glycoside hydrolase family 100 enzymes (GH100), while CWINV and VINV both belong to the GH32 family (Van den Ende

et al., 2009; Wan et al., 2018). Therefore, we built two separate trees for phylogenetic analysis of these distinct families (Figures 1 and 2).

Fabaceae genomes comprised four clades of cytosolic invertase (CINV1, CINV2, CINV3, and CINV4), two cell wall invertase clades (CWINV1 and CWINV2) and two vacuolar invertase clades (VINV1 and VINV2). In contrast, the Arabidopsis genome only seems to comprise accessions in the VINV clade 1 (AtVINV1.1 and AtVINV1.2), but no accession in the VINV2 clade (Figure 2). In parallel, Arabidopsis genome is lacking accessions from the cytosolic CINV4 clade (Figure 1). Therefore, VINV2 and CINV4 clades may be legume-specific clades (Table 1). Further phylogenetic analysis with multispecies comparison across the plant kingdom would be needed to verify this hypothesis.

Regarding cell wall invertase, the CWINV2 clade comprised two gene accessions for IRLC species and *Lotus japonicus*, while Phaseoloid species comprised three CWINV2 paralogs. Interestingly, all CWINV2 accessions cluster with the Arabidopsis accession AtCWINV6 (here renamed AtCWINV2.1), which did not show an invertase activity but was rather characterized as a fructan exohydrolases (FEH; De Coninck et al., 2005). Here, we show that accessions belonging to the CWINV2 clade comprise a FEH signature in their protein sequence (Figure S3). They lack the aspartate (D-239) necessary for binding and hydrolyzing sucrose (Le Roy et al., 2007). Moreover, the cysteine-containing domain “WECPD” essential for the catalytic activity of cell wall invertases is replaced by “VIECPD” (Sturm, 1999, Figure S3). In our study, *PsCWINV2.1* was slightly expressed during the embryogenesis stage of seed development (Figure 4A). The putative role of such FEH enzymes in plants that do not synthesize fructans (Incoll & Bonnett, 1993) remains an open question.

4.3 | Expression profile of the pea invertase gene family

Three genes were highly expressed in all plant tissues (*PsCINV2.1*, *PsCINV3.1*, and *PsVINV1.1*). Interestingly, the orthologous genes in Medicago (*MtCINV2.1*, *MtCINV3.1*, and *MtVINV1.1*) also showed the same constitutive expression pattern (Table S3B), indicating that these orthologous pairs may have an important and conserved function among legumes. Expression of CINVs, notably the expression of *PsCINV2.1* and *PsCINV3.1*, dominated both during embryogenesis and seed-filling stages in our conditions (Figures 4A, 5A, and S5) and overall showed a high expression in all plant organs (Figure 3). Unfortunately, the role of such CINVs during seed development has not yet been deciphered. Wan et al. (2018) showed that CINV had higher evolutionary constraints than acid invertases because CINV evolved under stronger purifying selection, and they postulated an important role for controlling cytosolic sugar homeostasis. Here, we could also hypothesize that *PsCINV2.1* and *PsCINV3.1*, which are highly expressed with a conserved constitutive pattern among species, may confirm this important role in controlling cytosolic sugar homeostasis for cellular function (Wan et al., 2018). Another hypothesis for this high and constitutive expression of CINV could arise from earlier work

on so-called “futile cycles,” where enzymes would be involved in the continuous degradation and re-synthesis of sucrose (Nguyen-Quoc & Foyer, 2001).

In contrast, three other genes (*PsCWINV1.3*, *PsCINV1.3*, and *PsCINV4.2*) were not expressed in any of the plant tissues available for the reference Caméor cultivar (Alves-Carvalho et al., 2015; Figure 3), and were barely detected in the seed expression data generated in this study using the Safran variety (Figure 4A). Consistent with the expression data, *PsCWINV1.3* protein from the reference genome cultivar lacks the β -fructosidase motif (NDPD/NG; Figure S3), indicating that this invertase may be non-functional (Wan et al., 2018). Noteworthy, *MtCWINV1.3*, the gene ortholog in *M. truncatula*, also lacks this motif (Figure S3). Conversely, the same *PsCWINV1.3* accession from a different pea cultivar (Alaska variety) showed a high expression in roots induced by wounding (Zhang et al., 1996) and its protein harbored a functional β -fructosidase motif. It could be possible that the gene expression pattern of such tissue-specific invertase (as opposed to constitutively expressed ones), as well as the mutation of functional motifs might be the result of varietal selection, thereby regulating the enzymatic activity and the nutritional quality of the cultivar. Consistent with this, the GIF1 gene, encoding a CWINV in rice, showed a restricted expression pattern during grain-filling compared with the wild rice allele, as well as mutations in the GIF1 regulatory sequence selected through domestication (Wang et al., 2008).

4.4 | *PsCWINV1.2* is a major candidate providing sugar nutrient and signal to the developing embryo

Among all INV, cell wall invertases are particularly involved in sugar partitioning (Ruan, 2014), regulating both sucrose influx from the seed coat and generating glucose signal to the developing embryo.

Firstly, CWINVs regulate the sink strength by cleaving sucrose molecules carried by sugar transport systems. Notably, sucrose influx towards seed is mediated through *PsSUT1.1*, the main sucrose transporter loading the phloem (Doidy et al., 2019), and through SWEETs facilitating the release of sucrose into the seed apoplast (Hageman & Van Volkenburgh, 2021). Thus, sugar flux is partly controlled through the catalytic activity of CWINV localized to the seed coat, determining the magnitude of the transmembrane sucrose gradient. In a pea mutant overexpressing *SUT1* sucrose transporter and *AAP1* amino acid transporter, differential expression of *PsCWINV1.2* was observed (Grant et al., 2021). Grant et al. (2021) also reported co-upregulated expression of *PsCWINV1.2* with a suite of sucrose transporting clade III SWEETs in pea seed coat. Thus, they provided an insight into how *PsCWINV1.2* expression in seed coats could be upregulated by a source-driven influx of sugars and amino acids. This suggests a mechanism integrating both source and sink activities.

Secondly, invertases also regulate developmental processes by generating signaling metabolites (also see Section 4.5 below). Indeed, the ratio of sucrose to hexose is likely a signal switch for the high mitotic divisions required during embryogenesis, and the phase of cell expansion and starch storage during nutritional filling. In *Vicia faba*,

the large-seeded genotype delays this signal, thereby maintaining cell division, compared with the small-seeded genotype which terminates its invertase peak earlier (Weber et al., 1996). This switch is likely governed by *VfCWINV1*, which is highly expressed in the seed coat of fava bean (Weber et al., 1996). Thus, *VfCWINV1* influences mitotic activity, therefore the final pool of cells for cell expansion during the seed-filling period. In this study, we identified *PscCWINV1.2*, the orthologous gene of *VfCWINV1* in *V. faba* (Weber et al., 1995), as the cell wall invertase gene with the highest expression during embryogenesis, both according to the transcriptomic atlas (Figure 3B) and the results of our RT-qPCR analysis (Figure 4A). Here, we hypothesize that *Fabaceae* orthologs *VfCWINV1* and *PscCWINV1.2* cleave sucrose in the seed coat, providing monosaccharides towards the endosperm and the developing cotyledons (Weber et al., 1995).

In conclusion, *PscCWINV1.2* may be the major invertase driving sucrose cleavage from the seed coat to the apoplast, thereby feeding sugar nutrients and generating glucose signal to the developing embryo. In maize, a single cell wall invertase (*ZmCWINV2*) also seems predominantly expressed during early seed development (Shen et al., 2022). Based on this cereal grain model, the authors conclude that a spatial-temporal expression of *CWINV*, *SUT*, *STP*, and *SWEET* govern sugar transport and utilization from maternal to filial tissues during the critical stage of early grain set (Shen et al., 2022).

4.5 | Invertases are potential markers of the transition between embryogenesis and seed-filling stages

Transcriptional regulation of invertases clearly marks the transition between seed developmental stages (Figure 4A). Indeed, the transition between embryogenesis and seed-filling was marked by a sharp transcriptional decrease in *INV* gene expression (Table S4). In parallel to the gene repression (Figure 4A), this developmental transition is also marked by a decrease in enzymatic activities (Borisjuk et al., 2002; Wang et al., 2018; Weber et al., 2005) and consequently higher sucrose and lower monosaccharide contents (Figure 4C). Noteworthy, both *CWINVs* and *VINVs* are also regulated through post-translational protein-protein interaction with small inhibitors (Rausch & Greiner, 2004). Here, the high expression of *PscVIF2* (Figure 4B) confirmed the importance of post-translational regulation during seed development. The gene ortholog *GmC/VIF2* was also predominantly expressed in developing soybean seeds (Tang et al., 2017). This suggests that invertases are also highly regulated at the post-translational level, putatively through the interaction with the dominant inhibitor *C/VIF2*. We also show that both invertases and their inhibitors are highly regulated through transcriptional regulation (Figure 4A, B), as their expression pattern marked the transition between early embryonic and later seed-filling stages. In conclusion, changes in the invertase gene expression, enzymatic activity, their inhibitors and sugar profile (sucrose to hexose ratio) mark the seed developmental phases, specifically the transition between embryogenesis and the later stages of nutrient storage.

Interestingly, the transition between the embryogenesis and the seed-filling stages also corresponds to the final stage of seed abortion (FSSA, i.e. the stage at which seeds no longer abort; Duthion & Pigeaire, 1991). Here, we postulate that *INV* genes might take part in the FSSA transition, notably through the provision of glucose, an energy-rich and signaling metabolite. In other words, by facilitating sucrose unloading and by generating glucose signals, invertases might suppress apoptosis and stimulate cell division for fruit and seed set (Ruan et al., 2012). Consistent with this discussion, it has been shown that reducing vacuolar *GhVINV* in the maternal seed coat caused programmed cell death in the endosperm and embryos, leading to seed abortion in cotton (Wang & Ruan, 2016). Here, seeds harvested at embryogenesis (7 dap), which showed a high invertase expression, had not yet reached the FSSA, while seeds at the filling stage (14 dap) with a significant reduction of gene expression (mainly *PscCINV1.1*, *PscCINV1.2*, *PscCINV3.1*, *PscCWINV1.2*, and *PscVINV1.1*; Figure 4A) had passed the FSSA.

Among all invertases, *PscVINV1.1* showed the highest fold change repression (\log_2 FC = -9 , which equals a 507-fold repression, Table S4) from embryogenesis to seed-filling stages. Therefore, *PscVINV1.1* may be used as a robust gene marker of the transition between the embryogenesis and the seed-filling stages. The main cell wall invertase expressed in seed, *PscCWINV1.2*, also showed a marked downregulation from embryogenesis to seed-filling stages (\log_2 FC = -1.4 , which equals a threefold repression, Table S4). The temporal downregulation of *CWINV* activity results from progressive mechanical crushing of seed coat cells expressing the enzyme by the developing embryo (Tegeger et al., 1999; Weber et al., 1995). Both *CWINV* and *VINV* are highly active in the maternal tissue of young seeds of legumes (Weber et al., 2005) and their expression later decreases during the filling stage (Figure 4A). According to the invertase control hypothesis of seed development (Weber et al., 1997; 2005), *CWINV* rather regulate the unloading of assimilates, elevate hexose levels, and control the nature and amount of sugars supplied to the embryo, thus regulating the cell division phase (also see Section 4.4 above). In parallel, *VINV* adjust sink size to the available resources and thus can act as stress sensors (also see Section 4.6 below).

In conclusion, invertases seem to represent key gene markers of the transition between embryogenesis and seed-filling stages and could be interesting candidates for evaluating both embryonic sink strength and FSSA, notably in response to abiotic stress.

4.6 | Invertase gene expression in response to drought in seed crops

The crucial role of *INV* enzymes in seed development and drought response has been highlighted in previous studies (Ruan et al., 2010). Applying a water deficit during seed development decreases seed weight and seed number by impacting carbon fluxes towards reproductive sink organs (Du et al., 2020; Guilioni et al., 2003; Henriot et al., 2019; Ney et al., 1993). Our results confirmed the

transcriptional regulation of three major candidate genes (*PsCINV1.1*, *PsCWINV1.2*, and *PsVINV1.1*; Figure 5A) in seeds of plants exposed to water deficit. Noteworthy, *PsCINV1.1*, *PsCWINV1.2*, and *PsVINV1.1* invertases were highly expressed and differentially regulated by water stress at the embryonic stages (Figure 5A). At later stage of seed-filling, *PsCINV1.1* and *PsCWINV1.2* lost this expression sensitivity to water stress and *PsVINV1.1* only showed a weak expression (Figure S5). Overall gene expression change is thought to be higher (induced or repressed) at the early set stages of seed development than at the later ones (Figure 4A). The embryogenesis stage not only establishes cell number through embryonic division, thus determining yield potential, but also is the most sensitive stage to stresses (Shen et al., 2022). We hypothesize that once the seed passed the FSSA, little difference occurs for sugar transporters and invertase expression.

Here, we identified three major invertases regulated by developmental stage and water deficit (Figure 5A). Not only *PsCINV1.1*, *PsCWINV1.2*, and *PsVINV1.1* showed a significant decrease in expression from embryogenesis to seed-filling stages (Figure 4A), but they were also significantly deregulated by water stress (Figure 5A). This suggests the combined influence of the major invertase type (CWINV, CINV, and VINV) controlling carbon provision towards developing seeds, sugar homeostasis and cellular osmotic pressure during water stress.

Regarding cytosolic invertase, *PsCINV1.1* expression was downregulated by water deficit during embryogenesis (Figure 5A). Previous research in *Arabidopsis* showed that CINVs contribute to reactive oxygen species homeostasis, notably under osmotic stress (Qi et al., 2007; Xiang et al., 2011). Here, we may also hypothesize that CINV may be involved in plant defense against oxidative stress in response to drought.

Regarding cell wall invertases, we previously postulated that CWINV1.2 in *Fabaceae* may drive sucrose unloading towards sink seed. Consistent with this hypothesis, both *PsCWINV1.2* expression and sucrose content decreased in response to water deficit (Figures 5A and 4C), suggesting that sink strength is adaptatively adjusting to plant water resource. Thus, *PsCWINV1.2* seems to regulate the sink strength towards seed by tuning sucrose cleavage to ensure embryo survival subjected to water deficit. By maintaining a hexose/sucrose ratio in response to water deficit similar to the well-watered control (Figure 5C), we postulate that cell division is maintained in embryo and endosperm (Wang et al., 2018), protecting young seeds under water stress early in their development (Ruan, 2014).

Both cell wall and vacuolar invertases also responded to water stress in *Oryza sativa* and *Zea mays* (Andersen et al., 2002; Ji, Raveendran, et al., 2005b). In response to drought, cell wall invertase *OsCINV2* was downregulated, while vacuolar invertase *OsVINV2* was upregulated in rice peduncle (Ji, Raveendran, et al., 2005b), similar to the expression pattern observed for *PsCWINV1.2* and *PsVINV1.1* in pea seeds at 7 dap (Figure 5A). Vacuolar invertases are well documented in response to water stress, especially a vacuolar invertase *IVR2* from maize (Andersen et al., 2002; Kim et al., 2000; Pelleschi et al., 1999; Qin et al., 2004). In maize, *IVR2* rather seemed repressed by drought in ovaries (Andersen et al., 2002) and in adhering perianth (Qin et al., 2004), whereas *OsVINV2* (the ortholog of *IVR2* in rice) was

upregulated by drought in most organs (Ji, Raveendran, et al., 2005b) and so was *PsVINV1* in pea seeds subjected to water deficit (Figure 5A). Andersen et al. (2002) concluded that *IVR2* regulation might confer a critical survival advantage for few versus many seeds in terminal drought environments by reducing sink numbers at a key point in development, thus securing sufficient carbon supply for maturation of a few remaining seeds. Hence, it has been shown that manipulating acid invertases can enhance drought tolerance and improve water use efficiency (Wan et al., 2018).

By adjusting intracellular carbon partitioning, VINV also take part in the regulation of osmotic pressure in response to plant stress. By cleaving sucrose into monosaccharides inside the vacuole, VINV contribute to carbon and water fluxes for the maintenance of cell turgescence in response to water stress. At early stage of seed development, maintaining cell turgescence is also crucial for embryonic expansion. Consistent with this hypothesis, we showed concurrent transcriptional induction of *PsVINV1.1* with tonoplasmic transporters (Figure 5B), mediating sugar fluxes in and out of the vacuoles (Doidy et al., 2012, 2019). Members of the ESL are characterized as vacuolar monosaccharide transporters; their responsiveness to water stress has long been described to protect plant cells from the detrimental effects of dehydration (Kiyosue et al., 1998; Slawinski et al., 2021). Members of the Tonoplast Monosaccharide/Sugar Transporter (TMT/TST) can transport both sucrose and monosaccharides inside the vacuoles and may also be involved in plant stress adaptation, such as salt or drought acclimation (Wormit et al., 2006).

Vacuolar invertases were also shown to respond to drought in several *Fabaceae* species. In *Medicago*, *MtVINV1.1* (*Medtr4g101630*) was detected as an early drought-responsive gene in roots (Zhang et al., 2014) and was identified as the top GWAS candidate for leaf chlorophyll changes in a dehydration treatment with PEG (Kang et al., 2015). In soybean, *GmA-INV* (here renamed *GmVINV1.1A*, *Glyma.05G056300*; Figure 2) was upregulated in response to water stress at 15 dap (Du et al., 2020), which corresponds to the embryonic heart stage (Goldberg et al., 1989). In pea seed, *PsVINV1.1* was also upregulated by water stress during embryogenesis (Figure 5A). Interestingly, *PsVINV1.1* in *P. sativum* is the orthologous gene of both *GmA-INV/GmVINV1.1A* in soybean and *MtVINV1.1* in *Medicago* and *IVR2* also seem to cluster within the VINV1 clade. Thus, candidates belonging to the VINV1 clade seem interesting indicators of the water status in seed crops.

In summary, we release an exhaustive repertoire of the invertase gene family in legumes and highlight key candidate genes as putative gene markers of seed developmental stage and regulators of sugar partitioning in response to water stress.

ACKNOWLEDGMENT

This work was supported by the University of Poitiers, the Centre National de la Recherche Scientifique (CNRS), the Nouvelle-Aquitaine Region with the 2015-2020 State-Region Planning Contracts (CPER) and the European Regional Development Fund (FEDER) and by a doctoral grant from Grand Poitiers awarded to AM. We thank Véronique Biarnes (Terres INOVIA) for fruitful discussions. We thank Vincent

Lebeurre, Helene Paulhac and Florence Thibault for their technical assistance, and all members of the SEVE team (EBI, University of Poitiers) for their contributions.

AUTHOR CONTRIBUTIONS

The experimental plan was designed by Amélie Morin, Cécile Vriet, Joan Doidy, Laurence Maurousset, Nathalie Pourtau and Rémi Lemoine. Samples were collected by Amélie Morin, Benoit Porcheron, Cécile Vriet, Joan Doidy, Laurence Maurousset and Nathalie Pourtau. Phylogenetic analyses were performed by Fadia Kadi and Joan Doidy. Expression data were performed and analyzed by Amélie Morin and Joan Doidy. The manuscript was written by Amélie Morin and Joan Doidy and reviewed by all authors.

DATA AVAILABILITY STATEMENT

Data availability statement: All accessions and transcriptomic data presented in this work can be found in the supplementary data and were retrieved from public databases detailed in the Materials and Methods section.

ORCID

Amélie Morin  <https://orcid.org/0000-0003-3318-8778>

Cécile Vriet  <https://orcid.org/0000-0001-6625-268X>

Nathalie Pourtau  <https://orcid.org/0000-0003-1200-4968>

Joan Doidy  <https://orcid.org/0000-0003-3966-3218>

REFERENCES

- Alves-Carvalho, S., Aubert, G., Carrère, S., Cruaud, C., Brochot, A.-L., Jacquin, F. et al. (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.
- Andersen, M.N., Asch, F., Wu, Y., Jensen, C.R., Næsted, H., Mogensen, V. O. et al. (2002) Soluble invertase expression is an early target of drought stress during the critical, abortion-sensitive phase of young ovary development in maize. *Plant Physiology*, 130, 591–604.
- Benedito, V.A., Torres-Jerez, I., Murray, J.D., Andriankaja, A., Allen, S., Kakar, K. et al. (2008) A gene expression atlas of the model legume *Medicago truncatula*. *The Plant Journal*, 55, 504–513.
- Bertioli, D.J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G. et al. (2019) The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nature Genetics*, 51, 877–884.
- Borisjuk, L., Walenta, S., Rolletschek, H., Mueller-Klieser, W., Wobus, U. & Weber, H. (2002) Spatial analysis of plant metabolism: sucrose imaging within *Vicia faba* cotyledons reveals specific developmental patterns. *The Plant Journal*, 29, 521–530.
- Burstin, J., Kreplak, J., Macas, J. & Lichtenzveig, J. (2020) *Pisum sativum* (Pea). *Trends in Genetics*, 36, 312–313.
- Chang, S., Puryear, J. & Cairney, J. (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter*, 11, 113–116.
- Chen, Z., Gao, K., Su, X., Rao, P. & An, X. (2015) Genome-wide identification of the invertase gene family in *Populus*. *PLoS One*, 10, e0138540.
- Dash, S., Campbell, J.D., Cannon, E.K.S., Cleary, A.M., Huang, W., Kalberer, S.R. et al. (2016) Legume information system (LegumeInfo.Org): a key component of a set of federated data resources for the legume family. *Nucleic Acids Research*, 44, D1181–D1188.
- De Coninck, B., Le Roy, K., Francis, I., Clerens, S., Vergauwen, R., Halliday, A.M. et al. (2005) Arabidopsis AtcwlINV3 and 6 are not invertases but are fructan exohydrolases (FEHs) with different substrate specificities. *Plant Cell and Environment*, 28, 432–443.
- 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.
- Doidy, J., Grace, E., Kühn, C., Simon-Plas, F., Casieri, L. & Wipf, D. (2012) Sugar transporters in plants and in their interactions with fungi. *Trends in Plant Science*, 17, 413–422.
- 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.
- Du, Y., Zhao, Q., Chen, L., Yao, X., Zhang, H., Wu, J. et al. (2020) Effect of drought stress during soybean R2–R6 growth stages on sucrose metabolism in leaf and seed. *International Journal of Molecular Sciences*, 21, 618.
- Duthion, C. & Pigeaire, A. (1991) Seed lengths corresponding to the final stage in seed abortion of three grain legumes. *Crop Science*, 31, 1579–1583.
- Goldberg, R.B., Barker, S.J. & Perez-Grau, L. (1989) Regulation of gene expression during plant embryogenesis. *Cell*, 56, 149–160.
- Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J. et al. (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40, D1178–D1186.
- Grant, J.E., Ninan, A., Cripps-Guazzone, N., Shaw, M., Song, J., Pet ík, I. et al. (2021) Concurrent overexpression of amino acid permease AAP1 (3a) and SUT1 sucrose transporter in pea resulted in increased seed number and changed cytokinin and protein levels. *Functional Plant Biology*, 48, 889–904.
- Guilioni, L., Wéry, J. & Lecoq, J. (2003) High temperature and water deficit may reduce seed number in field pea purely by decreasing plant growth rate. *Functional Plant Biology*, 30, 1151–1164.
- Hageman, A. & Van Volkenburgh, E. (2021) Sink strength maintenance underlies drought tolerance in common bean. *Plants (Basel)*, 10, 489.
- Hall, T. (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.
- Hane, J.K., Ming, Y., Kamphuis, L.G., Nelson, M.N., Garg, G., Atkins, C.A. et al. (2017) A comprehensive draft genome sequence for lupin (*Lupinus angustifolius*), an emerging health food: insights into plant-microbe interactions and legume evolution. *Plant Biotechnology Journal*, 15, 318–330.
- Henriet, C., Aimé, D., Térézol, M., Kilandamoko, A., Rossin, N., Combes-Soia, L. et al. (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.
- Incoll, L.D. & Bonnett, G.D. (1993) The occurrence of fructan in food plants. In: Fuchs, A. (Ed.) *Studies in plant science*. Amsterdam: Elsevier, pp. 309–322.
- Ji, X., Van den Ende, W., Van Laere, A., Cheng, S. & Bennett, J. (2005a) Structure, evolution, and expression of the two invertase gene families of rice. *Journal of Molecular Evolution*, 60, 615–634.
- Ji, X.M., Raveendran, M., Oane, R., Ismail, A., Lafitte, R., Bruskiwich, R. et al. (2005b) Tissue-specific expression and drought responsiveness of cell-wall invertase genes of rice at flowering. *Plant Molecular Biology*, 59, 945–964.
- Juárez-Colunga, S., López-González, C., Morales-Eliás, N.C., Massange-Sánchez, J.A., Trachsel, S. & Tiessen, A. (2018) Genome-wide analysis of the invertase gene family from maize. *Plant Molecular Biology*, 97, 385–406.
- Kang, Y., Sakiroglu, M., Krom, N., Stanton-Geddes, J., Wang, M., Lee, Y.C. et al. (2015) Genome-wide association of drought-related and biomass traits with HapMap SNPs in *Medicago truncatula*. *Plant, Cell & Environment*, 38, 1997–2011.
- Kim, J.-Y., Mahé, A., Brangeon, J. & Prioul, J.-L. (2000) A maize vacuolar invertase, IVR2, is induced by water stress. Organ/tissue specificity and diurnal modulation of expression. *Plant Physiology*, 124, 71–84.

- Kiyosue, T., Abe, H., Yamaguchi-Shinozaki, K. & Shinozaki, K. (1998) ERD6, a cDNA clone for an early dehydration-induced gene of *Arabidopsis*, encodes a putative sugar transporter. *Biochimica et Biophysica Acta*, 1370, 187–191.
- 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.
- Kreplak, J., Madoui, M.-A., Cápál, P., Novák, P., Labadie, K., Aubert, G. et al. (2019) A reference genome for pea provides insight into legume genome evolution. *Nature Genetics*, 51, 1411–1422.
- Kumar, S., Stecher, G. & Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33, 1870–1874.
- Le Roy, K., Lammens, W., Verhaest, M., De Coninck, B., Rabijns, A., Van Laere, A. et al. (2007) Unraveling the difference between invertases and fructan exohydrolases: a single amino acid (Asp-239) substitution transforms *Arabidopsis* cell wall invertase1 into a fructan 1-exohydrolase. *Plant Physiology*, 145, 616–625.
- Munier-Jolain, N., Biarnès, V., Chaillet, I. & Lecoœur, J. (2010) *Physiology of the pea crop*. Boca Raton: CRC Press.
- Murayama, S. & Handa, H. (2007) Genes for alkaline/neutral invertase in rice: alkaline/neutral invertases are located in plant mitochondria and also in plastids. *Planta*, 225, 1193–1203.
- Ney, B., Duthion, C. & Fontaine, E. (1993) Timing of reproductive abortions in relation to cell division, water content, and growth of pea seeds. *Crop Science*, 33, 267–270.
- Nguyen-Quoc, B. & Foyer, C.H. (2001) A role for ‘futile cycles’ involving invertase and sucrose synthase in sucrose metabolism of tomato fruit. *Journal of Experimental Botany*, 52, 881–889.
- Pelleschi, S., Guy, S., Kim, J.-Y., Pointe, C., Mahé, A., Barthes, L. et al. (1999) *lvr2*, a candidate gene for a QTL of vacuolar invertase activity in maize leaves. Gene-specific expression under water stress. *Plant Molecular Biology*, 39, 373–380.
- Qi, X., Wu, Z., Li, J., Mo, X., Wu, S., Chu, J. et al. (2007) AtCYT-INV1, a neutral invertase, is involved in osmotic stress-induced inhibition on lateral root growth in *Arabidopsis*. *Plant Molecular Biology*, 64, 575–587.
- Qin, L., Trouverie, J., Chateau-Joubert, S., Simond-Côte, E., Thévenot, C. & Prioul, J.-L. (2004) Involvement of the *lvr2*-invertase in the perianth during maize kernel development under water stress. *Plant Science*, 166, 371–379.
- Rausch, T. & Greiner, S. (2004) Plant protein inhibitors of invertases. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1696, 253–261.
- Rolland, F., Baena-Gonzalez, E. & Sheen, J. (2006) Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annual Review of Plant Biology*, 57, 675–709.
- Ruan, Y.-L. (2014) Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annual Review of Plant Biology*, 65, 33–67.
- Ruan, Y.-L., Jin, Y., Yang, Y.-J., Li, G.-J. & Boyer, J.S. (2010) Sugar input, metabolism, and signaling mediated by invertase: roles in development, yield potential, and response to drought and heat. *Molecular Plant*, 3, 942–955.
- Ruan, Y.-L., Patrick, J.W., Bouzayen, M., Osorio, S. & Fernie, A.R. (2012) Molecular regulation of seed and fruit set. *Trends in Plant Science*, 17, 656–665.
- Shen, S., Ma, S., Chen, X.-M., Yi, F., Li, B.-B., Liang, X.-G. et al. (2022) A transcriptional landscape underlying sugar import for grain set in maize. *The Plant Journal*, 1–15. <https://doi.org/10.1111/tpj.15668>
- Sherson, S.M., Alford, H.L., Forbes, S.M., Wallace, G. & Smith, S.M. (2003) Roles of cell-wall invertases and monosaccharide transporters in the growth and development of *Arabidopsis*. *Journal of Experimental Botany*, 54, 525–531.
- Slawinski, L., Israel, A., Artault, C., Thibault, F., Atanassova, R., Laloi, M. et al. (2021) Responsiveness of early response to dehydration six-like transporter genes to water deficit in *Arabidopsis thaliana* leaves. *Frontiers in Plant Science*, 12, 1–21.
- Smykal, P., Aubert, G., Burstin, J., Coyne, C.J., Ellis, N.T.H., Flavell, A.J. et al. (2012) Pea (*Pisum sativum* L.) in the genomic era. *Agronomy*, 2, 74–115.
- Sosso, D., Luo, D., Li, Q.-B., Sasse, J., Yang, J., Gendrot, G. et al. (2015) Seed filling in domesticated maize and rice depends on SWEET-mediated hexose transport. *Nature Genetics*, 47, 1489–1493.
- Sturm, A. (1999) Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning. *Plant Physiology*, 121, 1–8.
- Su, T., Han, M., Min, J., Chen, P., Mao, Y., Huang, Q. et al. (2018) Genome-wide survey of invertase encoding genes and functional characterization of an extracellular fungal pathogen-responsive invertase in *Glycine max*. *International Journal of Molecular Sciences*, 19, 2395.
- Tang, X., Su, T., Han, M., Wei, L., Wang, W., Yu, Z. et al. (2017) Suppression of extracellular invertase inhibitor gene expression improves seed weight in soybean (*Glycine max*). *Journal of Experimental Botany*, 68, 469–482.
- Tegeder, M., Wang, X.D., Frommer, W.B., Offler, C.E. & Patrick, J.W. (1999) Sucrose transport into developing seeds of *Pisum sativum* L. *Plant Journal*, 18, 151–161.
- Van den Ende, W., Lammens, W., Van Laere, A., Schroeven, L. & Le Roy, K. (2009) Donor and acceptor substrate selectivity among plant glycoside hydrolase family 32 enzymes. *The FEBS Journal*, 276, 5788–5798.
- Wan, H., Wu, L., Yang, Y., Zhou, G. & Ruan, Y.-L. (2018) Evolution of sucrose metabolism: the dichotomy of invertases and beyond. *Trends in Plant Science*, 23, 163–177.
- Wang, E., Wang, J., Zhu, X., Hao, W., Wang, L., Li, Q. et al. (2008) Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nature Genetics*, 40, 1370–1374.
- Wang, L., Patrick, J.W. & Ruan, Y.-L. (2018) Live long and prosper: roles of sugar and sugar polymers in seed vigor. *Molecular Plant*, 11, 1–3.
- Wang, L. & Ruan, Y.-L. (2016) Critical roles of vacuolar invertase in floral organ development and male and female fertilities are revealed through characterization of GhVIN1-RNAi cotton plants. *Plant Physiology*, 171, 405–423.
- Wang, L., Zheng, Y., Ding, S., Zhang, Q., Chen, Y. & Zhang, J. (2017) Molecular cloning, structure, phylogeny and expression analysis of the invertase gene family in sugarcane. *BMC Plant Biology*, 17, 109.
- Weber, H., Borisjuk, L., Heim, U., Buchner, P. & Wobus, U. (1995) Seed coat-associated invertases of fava bean control both unloading and storage functions: cloning of cDNAs and cell type-specific expression. *The Plant Cell*, 7, 1835–1846.
- Weber, H., Borisjuk, L. & Wobus, U. (1996) Controlling seed development and seed size in *Vicia faba*: a role for seed coat-associated invertases and carbohydrate state. *The Plant Journal*, 10, 823–834.
- Weber, H., Borisjuk, L. & Wobus, U. (1997) Sugar import and metabolism during seed development. *Trends in Plant Science*, 2, 169–174.
- Weber, H., Borisjuk, L. & Wobus, U. (2005) Molecular physiology of legume seed development. *Annual Review of Plant Biology*, 56, 253–279.
- Wind, J., Smeeckens, S. & Hanson, J. (2010) Sucrose: metabolite and signaling molecule. *Phytochemistry*, 71, 1610–1614.
- Wormit, A., Trentmann, O., Feifer, I., Lohr, C., Tjaden, J., Meyer, S. et al. (2006) Molecular identification and physiological characterization of a novel monosaccharide transporter from *Arabidopsis* involved in vacuolar sugar transport. *The Plant Cell*, 18, 3476–3490.
- Xiang, L., Le Roy, K., Bolouri-Moghaddam, M.-R., Vanhaecke, M., Lammens, W., Rolland, F. et al. (2011) Exploring the neutral invertase-oxidative stress defence connection in *Arabidopsis thaliana*. *Journal of Experimental Botany*, 62, 3849–3862.

- Xiang, L. & Van den Ende, W. (2013) Trafficking of plant vacuolar invertases: from a membrane-anchored to a soluble status. Understanding sorting information in their complex N-terminal motifs. *Plant and Cell Physiology*, 54, 1263–1277.
- Xu, X., Ren, Y., Wang, C., Zhang, H., Wang, F., Chen, J. et al. (2019) OsVIN2 encodes a vacuolar acid invertase that affects grain size by altering sugar metabolism in rice. *Plant Cell Reports*, 38, 1273–1290.
- Mals, Z., Osorio, S., Nunes-Nesi, A., Carrari, F., Lohse, M., Br, U. et al. (2009) RNA interference of LIN5 in tomato confirms its role in controlling brix content, uncovers the influence of sugars on the levels of fruit hormones, and demonstrates the importance of sucrose cleavage for normal fruit development and fertility. *Plant Physiology*, 150, 1204–1218.
- Zhang, J.Y., Cruz, D.E.C.M.H., Torres-Jerez, I., Kang, Y., Allen, S.N., Huhman, D.V. et al. (2014) Global reprogramming of transcription and metabolism in *Medicago truncatula* during progressive drought and after rewatering. *Plant, Cell & Environment*, 37, 2553–2576.
- Zhang, L., Cohn, N.S. & Mitchell, J.P. (1996) Induction of a pea cell-wall invertase gene by wounding and its localized expression in phloem. *Plant Physiology*, 112, 1111–1117.
- Zuber, H., Poignavent, G., Le Signor, C., Aimé, D., Vieren, E., Tadla, C. et al. (2013) Legume adaptation to sulfur deficiency revealed by comparing nutrient allocation and seed traits in *Medicago truncatula*. *The Plant Journal*, 76, 982–996.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Morin, A., Kadi, F., Porcheron, B., Vriet, C., Maurousset, L., Lemoine, R. et al. (2022) Genome-wide identification of invertases in *Fabaceae*, focusing on transcriptional regulation of *Pisum sativum* invertases in seed subjected to drought. *Physiologia Plantarum*, 174(2), e13673. Available from: <https://doi.org/10.1111/ppl.13673>