BMC Plant Biology

RESEARCH Open Access

Transcriptomic landscape of *seedstick* in *Arabidopsis thaliana* funiculus after fertilisation

Maria João Ferreira[1], Jessy Silva[1,2], Hidenori Takeuchi[3,4], Takamasa Suzuki[5], Tetsuya Higashiyama[4,6,7] and Sílvia Coimbra[1*]

Abstract

Background In Angiosperms, the continuation of plant species is intricately dependent on the funiculus multifaceted role in nutrient transport, mechanical support, and dehiscence of seeds. SEEDSTICK (STK) is a MADS-box transcription factor involved in seed size and abscission, and one of the few genes identified as affecting funiculus growth. Given the importance of the funiculus to a correct seed development, allied with previous phenotypic observations of *stk* mutants, we performed a transcriptomic analysis of *stk* funiculi from floral stage 17, using RNA-sequencing, to infer on the deregulated networks of genes.

Results The generated dataset of differentially expressed genes was enriched with cell wall biogenesis, cell cycle, sugar metabolism and transport terms, all in accordance with *stk* phenotype observed in funiculi from floral stage 17. We selected eight differentially expressed genes for transcriptome validation using qPCR and/or promoter reporter lines. Those genes were involved with abscission, seed development or novel functions in *stk* funiculus, such as hormones/secondary metabolites transport.

Conclusion Overall, the analysis performed in this study allowed delving into the STK-network established in Arabidopsis funiculus, fulfilling a literature gap. Simultaneously, our findings reinforced the reliability of the transcriptome, making it a valuable resource for candidate genes selection for functional genetic studies in the funiculus. This will enhance our understanding on the regulatory network controlled by STK, on the role of the funiculus and how seed development may be affected by them.

Keywords *Arabidopsis thaliana*, Abscission, Funiculus, RNA-sequencing, Seed development, SEEDSTICK

*Correspondence:
Sílvia Coimbra
scoimbra@fc.up.pt
[1]LAQV/REQUIMTE, Biology Department, Faculty of Sciences, University of Porto, Rua do Campo Alegre s/n, Porto 4169-007, Portugal
[2]School of Sciences, University of Minho, Campus de Gualtar, Braga 4710-057, Portugal
[3]Institute for Advanced Research, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
[4]Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Furo-cho, Chikusa- ku, Nagoya 464-8601, Japan
[5]Department of Biological Chemistry, College of Biosciences and Biotechnology, Chubu University, Kasugai 487-8501, Aichi, Japan
[6]Division of Biological Science, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan
[7]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan

Background

Seeds are intricate units essential to establish the next sporophytic generation of plants. In Angiosperms, the development of a seed starts with a key process called double fertilisation, where one sperm cell fuses with the egg cell, originating the diploid zygote, while the other fuses with the central cell, giving rise to the triploid endosperm [1]. In the next stages of seed growth and maturation, an influx of nutrients, minerals, sugars and signals from the maternal plant through the vasculature of the funiculus is required [2]. Finally, when the seed is mature, the formation of a seed abscission zone (SAZ) occurs, composed of a separation layer with thin cell walls that will degenerate at the end of the process and an adjacent layer with lignified cells (lignified layer), which produces the tension necessary for the separation of the tissues [3]. This process allows seeds to detach from the plant.

The funiculus is a specialised sporophytic structure that anchors the ovule to the placenta within the gynoecium of flowering plants, being the contact route for communication between them [4]. It arises during ovule primordium and remains intact throughout ovule development until seed dehiscence occurs. Remarkably, there is a lack of information regarding *Arabidopsis thaliana* case-study compared to other plants such as *Brassica napus* [5–7] or *Phaseolus vulgaris* [8], in which topics like anatomy or genetic regulation of funiculus development and underlying biological function have been extensively studied. So far, only one studied focused on clarifying these topics using *A. thaliana.* Khan and colleagues work showed using light microscopy that the funiculus is composed of an outer epidermis, a parenchymatous cortex, and an internal vascular core with xylem and phloem [9].

A comparative transcriptomic study between the zygotic regions and subregions of the seed and the funiculus in *A. thaliana* revealed the accumulation of several exclusive mRNAs in this stalk structure, presenting the funiculus as a transcriptionally distinct region from the other seed tissues [9]. Of the 20 genes identified as funiculus-specific, some were associated with transport and others with carbon metabolism, but over 30% did not have an annotated function, thus, showing that regulatory networks in the funiculus require further investigation.

SEEDSTICK (*STK*) is one of the few genes identified to be involved in funicular patterning and growth [10, 11]. It belongs to the MADS-box transcription factor (TF) family and is considered a master regulator of ovule development. Observations using a reporter line with the entire *STK* genomic region fused to a green fluorescent protein revealed high expression of this gene specifically in the ovules' sporophytic tissues, in which the funiculus is included, from the beginning of ovule development until after fertilisation, when seed is developing [12]. This expression pattern, together with many genetic-based studies performed over the years, supports the pivotal role of STK in orchestrating various processes during plant reproduction. STK redundantly controls ovule identity together with SHATTERPROOF1 (SHP1) and SHP2 [11] and participates in transmitting tract development by interacting with NO TRANSMITTING TRACT (NTT) [13] and CESTA (CES) [14]. STK also determines fruit size by regulating cytokinin homeostasis and controlling *FRUITFULL* expression [15]. Furthermore, STK is involved in seed development and size, affecting processes such as flavonoids biosynthesis [12], cell wall organisation and remodelling properties [16–19] and cell cycle [20]. Observations at floral stage 17 [21] showed that *stk* mutants develop thicker and longer funiculi compared to wild-type (WT) and present defects in the SAZ, resulting in a decrease in the ability to dehiscence [11]. Balanzà and colleagues [22] proposed a model with the STK network controlling the formation of the SAZ. According to this study, STK interacts with SEUSS corepressor, forming a complex that will repress *HECATE 3*, enabling the correct deposition of lignin in the vasculature of funiculus cells. Nonetheless, only TFs have been reported as regulators for the establishment of the lignification pattern. However, in similar abscission processes occurring in the floral organ abscission zone (reviewed by Niederhuth and colleagues [23]) or silique dehiscence zone (reviewed by Yu and colleagues [24]), several intervenients besides TFs have been reported: hormones (like ethylene, auxin, jasmonic acid), small peptides and their receptors (INFLORESCENCE DEFICIENT IN ABSCISSION is perceived by HAESA or HAESA-LIKE 1), as well as cell wall modifying enzymes (cellulases, hemicellulases and pectinases).

Despite the significant achievements on STK role during seed shattering, many questions remain unanswered: Which downstream targets participate in seed abscission? Is STK controlling other processes in the funiculus? Is seed development being affected by them? Only one transcriptome on *stk* single mutant is currently available [12], however those data derive from inflorescences and fruits until five days after pollination, which do not include floral stage 17 [21]. Moreover, considering the different tissues contained in that transcriptome, it would be difficult to study, at the molecular level, processes happening in a specific structure such as the funiculus.

In this study, we explore the transcriptomic data from *stk* and WT funiculi at floral stage 17, in order to investigate at the molecular level, the processes affected by *STK* absence in the funiculus. We provide information about enriched biological processes and molecular functions underpin the mutant funiculus phenotype, and we validate with qPCR and/or promoter reporter lines the expression of candidate genes, encoding proteins such

as pectinases, laccases, or sugar transporters. Overall, we believe that by leveraging our results, novel players in the STK-mediated network in the funiculus can be found, and, importantly, whose functions may be affecting seed development and shattering.

Materials and methods

Plant material and growth conditions

Wild-type (WT) *Arabidopsis thaliana* (L.) Heynh. Columbia-0 seeds and the *stk-2* mutant line [11], referred to as *stk*, were sown directly on soil (COMPO SANA®, Germany) and kept for 48 h at 4 °C in the dark to induce stratification. Afterwards, all plants were grown under long-day conditions (16 h light at 22 °C and 8 h darkness at 18 °C) with 50–60% relative humidity and light intensity at 180 µmol m^{-2} s^{-1}.

Sample collection for RNA-sequencing

A procedure for funiculi collection from dissected siliques was developed to address the question of which genes were differentially expressed in the funiculus in the absence of STK. Under a SZ61 stereo microscope (Olympus), WT and *stk* siliques from stage 17 (st 17) flowers [21] were first dissected, seeds were removed, and only the septum with the attached funiculi was placed into a 50 µL droplet of in vitro ovule culture medium [25] in a plastic dish (Φ 3.5 cm) (Supplementary Fig. 1A–C). Each funiculus was separated using a surgical knife and transferred with a tungsten needle to another plastic dish containing 10 µL of medium (Supplementary Fig. 1D, E). Afterwards, each funiculus was transferred to a 1.5 mL tube containing 40 µL of RNAlater™ Stabilization Solution (Ambion).

RNA-sequencing library preparation

mRNA was extracted from two biological replicates of 100 funiculi from st 17 of both WT and *stk*, using the Dynabeads™ mRNA DIRECT™ Micro Kit (Invitrogen), according to the manufacturer's instructions. Sequencing libraries were prepared following the manufacturer's instructions in the Low Sample protocol, using the TruSeq® RNA Sample Preparation v2 (Illumina Inc.) and sequenced with a NextSeq 500 sequencer (Illumina). Sequencing was performed as reported in Kadokura and colleagues work [26].

RNA-sequencing data analysis

Raw reads were cleaned and trimmed using Trimmomatic v.0.38 [27]. The resulting reads were aligned against the Arabidopsis genome (Araport11) using STAR v.2.7.0 software with default parameters [28]. The total number of reads as well as the uniquely mapped reads to the *A. thaliana* genome are reported in Supplementary Table 1. Afterwards, featureCounts v.1.6.3 [29] was used to generate the count matrix and to calculate gene expression values as raw read counts. Count read values were analysed using the DESeq2 package v.1.38.3 [30] from R software v1.30.1. A pre-filtering step was applied to remove lowly expressed genes. Specifically, genes with a total count of one or less across all samples were excluded from further analysis. Regularized logarithm (rlog) transformation of the count matrix was used for both principal component plot (PCA) analysis and heatmap clustering. PCA was generated using the *plotPCA* of DESeq2 package and data were visualised with ggplot2 package v.3.4.1 [31] in R. Heatmap showing clusters of samples was obtained with pheatmap package v.1.0.12 in R and using the top 1000 genes selected based on the mean of the rlog normalised counts. The total number of genes expressed in *stk*, WT or both samples was plotted using a Venn diagram, and *plotMA* function together with ggplot2 were used to show the $\log_2$FoldChange ($\log_2$FC) of all genes. To identify differentially expressed genes (DEGs), the raw count data were analysed using the *DESeq* function. A cut-off of p-value, p-adjust value and $\log_2$FC was applied to select differentially expressed genes: p-value and p-adjust value < 0.05 and a $-1 < \log_2 FC > 1$.

Gene set enrichment analysis

To determine enriched gene ontology (GO) terms in the *stk* funiculus transcriptome compared to the WT, genes were ranked as: $-\text{Log}_{10}(p\text{-value})$ multiplied by the sign of the Log_2FC [32]. A Parametric Analysis of Gene Set Enrichment [33] with Hochberg multi-test adjustment method was performed using the ranked gene list on agriGO v.2.0 ([34]; http://systemsbiology.cau.edu.cn/agriGOv2/, accessed in July 2020), from which biological process and molecular function GO terms with false discovery rate (FDR) < 0.05 were obtained and analysed.

Enriched networks among DEGs

To uncover functional links between genes, a list containing up- and down-regulated DEGs was upload to GeneMANIA Cytoscape plugin v3.10.0 [35] software. GeneMANIA constructed individual interaction networks for each data source, including gene co-expression, physical interactions, shared protein domains, and GO biological process annotations. Each network was assigned a weight based on its relevance to the query genes, optimising these weights through a linear regression model to best fit the known associations of the input genes. To enhance the robustness of our analysis, only GO biological processes with a FDR < 0.05 were considered. The individual networks were then combined into a single composite network using the optimised weights. Label propagation algorithms were applied to this composite network to infer connections between genes,

spreading functional labels from known genes to predict the functions of unknown genes.

Sample collection, RNA isolation and cDNA synthesis for validation by qPCR

A total of six samples for RNA extraction were collected from funiculi at st 17, representing two genotypes (WT and *stk*), from three biological samples. One hundred funiculi from one plant were used for each biological replicate. Samples were immediately frozen in liquid nitrogen and stored at −80 °C for follow-up experiments. Total RNA was extracted using the RNeasy Plant Mini Kit (QIAGEN) according to the manufacturer's protocol, with a minor change: the recommended quantities for reagents were reduced by half. RNA concentration was measured using a spectrophotometer (DS-11 Series Spectrophotometer/Fluorometer). Following the manufacturer's instructions, 21 ng of total RNA from funiculi at st 17 was used for cDNA synthesis using the Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher). The cDNA products were diluted to 0.21 ng/µL with nuclease-free water prior to qPCR.

Primer design and qPCR analysis for validation of RNA-sequencing

Primers were designed using Primer3 v.4.1.0 [36–38], following the parameters of Ferreira and colleagues [39]. For *MLP28*, Litholdo and colleagues [40] primers were used, and for *ACTIN 7* (*ACT7*), *YELLOW-LEAF-SPECIFIC GENE 8* (*YLS8*) and *HISTONE 3.3* (*HIS3.3*), Ferreira and colleagues [39] primers were used. Supplementary Table 2 presents all the primers used for qPCR. Primer specificity was confirmed by conventional PCR and electrophoresis on 1% (w/v) agarose gel. qPCR was performed according to Ferreira and colleagues [39]. Briefly, reactions were prepared in a 10 µL final volume containing 5 µL of 2x SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad), 0.125 µL of each specific primer pair at 250 nM, 0.75 µL of nuclease-free water, and 4 µL of diluted cDNA template. The reactions were performed in 96-well plates and run on a CFX96 Real-Time System (Bio-Rad) under the following cycling conditions: initial denaturation at 95 °C for 30 s, followed by 40 cycles at 95 °C for 15 s and 60 °C for 30 s, and an additional data acquisition step of 15 s at the optimal acquisition temperature. All reactions were run in three technical replicates and all assays included non-template controls (NTCs). Using three points of a 5-fold dilution series (1:5, 1:25 and 1:125) from a pool of funiculi cDNA (containing both WT and *stk* cDNA), a standard curve was generated to estimate the PCR efficiency of each primer pair using CFX Maestro software v.2.0 (Bio-Rad). The slope and coefficient of determination (R^2) were obtained from the linear regression line and the amplification efficiency (E) was calculated according to the formula $E = (10^{-1/slope} - 1) \times 100\%$. The R^2 value should be higher than 0.980 and the E value should be between 90% and 110%. After completion of the amplification reaction, melt curves were generated by increasing the temperature from 65 °C to 95 °C, with fluorescence readings acquired at 0.5 °C increments. From the melt curve, the optimal temperature for data acquisition (3 °C below the melting temperature of the specific PCR product) was determined and the specificity of the primers was confirmed (Supplementary Table 3). The sample maximisation method was chosen as the run layout strategy, in which all samples for each defined set were analysed in the same run, thus, different genes were analysed in distinct runs [41]. The quantitative cycle, baseline correction and threshold setting were automatically calculated using CFX Maestro software v.2.0 (Bio-Rad). The qPCR products were verified using 1% (w/v) agarose gels. The expression levels of the target genes were quantified in all samples using three reference genes (*ACT7*, *YLS8* and *HIS3.3*; [39]), whose expression stability between *stk* and WT funiculi at st 17 was verified using the RNA-sequencing (RNA-seq) data (Supplementary Table 4). Relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method [42]. Data were statistically analysed using the CFX Maestro software v.2.0 (Bio-Rad), which compared the relative expression differences using a Student's t-test. Statistical significance was set to a *p*-value threshold of 0.05. The Pearson correlation coefficient (r) was calculated to determine statistical correlation between the RNA-seq and qPCR data.

Construct generation and plant transformation

Genomic regions corresponding to the promoters of three genes (*ADPG1*, *MLP28* and *SCPL41*) were amplified using DreamTaq DNA polymerase (Thermo Fisher) for *ADPG1* and *SCPL41*, and Phusion DNA polymerase (Thermo Fisher) for *MLP28*, using the primer pairs described in Supplementary Table 2. The PCR products were cloned into pENTR™/D-TOPO (Invitrogen) and subcloned into the binary vector pBGWFS7 containing the *β-glucuronidase* (*GUS*) gene [43]. All constructs were confirmed by DNA sequencing. Expression vectors were delivered into *Agrobacterium tumefaciens* GV3101 (pMP90RK) and were used to transform *A. thaliana* (WT and *stk* genotype) by the floral dip method [44]. Transformant lines were obtained using glufosinate ammonium (BASTA®) as a selection agent.

Detection of GUS activity

For histochemical analysis, siliques at st 17 from T3 generation plants were fixed in acetone 90% (v/v) (Thermo Fisher) at -20 ºC overnight (ON). Samples were then washed twice with sodium buffer [0.2 M sodium dihydrogen phosphate (NaH_2PO_4; Panreac), 0.2 M sodium

hydrogen phosphate (Na_2HPO_4; Sigma Aldrich)] for 10 min and incubated in a X-Gluc solution [50 mM Na_2HPO_4, 50 mM NaH_2PO_4, 0.2% (v/v) Triton X-100 (Sigma Aldrich), 2 mM potassium hexacyanoferrate(II) trihydrate ($C_6FeK_3N_6^{+2} \cdot 3H_2O$; Sigma Aldrich), 2 mM potassium hexacyanoferrate(III) ($C_6FeK_3N_6^{+3}$; Sigma Aldrich), 1 mg/mL X-Gluc (5-bromo-4-chloro-3-indolyl β-D-glucuronic cyclohexylammonium salt; Biosynth)] ON at 37 °C. After chemical GUS detection, samples were rinsed with a 90% (v/v) ethanol, followed by 70% (v/v) ethanol for 10 min and incubated in a clearing solution [160 g of chloral hydrate (Sigma Aldrich), 100 mL of deionised water, and 50 mL of glycerol (Sigma Aldrich)], where they were kept at 4 °C ON. The following day, siliques were dissected under a stereomicroscope (model C-DSD230; Nikon) using hypodermic needles (0.4×20 mm; Braun). The funiculi attached to the septum were maintained in a drop of clearing solution and covered with a cover slip. Samples were observed under an upright Axio Imager AZ microscope (Zeiss) equipped with differential interference contrast optics. Images were captured with an Axiocam MRc3 camera (Zeiss) and processed using Zen 2011 Software (blue edition; Zeiss).

Image processing

All images were processed for publication using Fiji [45] and Adobe Photoshop CC 2019.

Results

RNA-sequencing analysis of *stk* funiculi at St 17

To investigate if the transcriptional profiles of *stk* and WT funiculi were similar, a principal component analysis (PCA) and hierarchical clustering were conducted (Fig. 1A). PCA showed that 90% of the total variance could be explained by the first two components (53% by PC1 and 37% by PC2). A cluster of *stk* replicates could be identified, but WT replicates showed a slight discrepancy between each other, with one of the replicates being closer to the *stk* replicates. However, the expression analysis of the top 1000 genes clearly showed two distinct clusters, one for *stk* replicates and the other for WT replicates (Fig. 1B). The heatmap also demonstrated that the expression of these genes between *stk* and WT samples was similar, suggesting a low number of differentially expressed genes (DEGs) between the two genotype samples. Overall, the results showed clusters between replicates, even though the expression profiles of *stk* and WT samples were not completely distinguishable.

To obtain an overview of the funiculi transcriptome, the total number of expressed genes in *stk* and WT samples was examined. Both datasets presented a similar number of genes (21841 for *stk* and 21723 for WT), from which 1293 genes (5.9%) were exclusively expressed in *stk* funiculi, and 1175 genes (5.4%) were only present in WT funiculi. The commonly expressed genes expressed in both datasets numbered 20,548 (Fig. 2A). Differential expression analysis revealed a total of 169 DEGs (Supplementary Table 5), with p-value and p-adjust value < 0.05

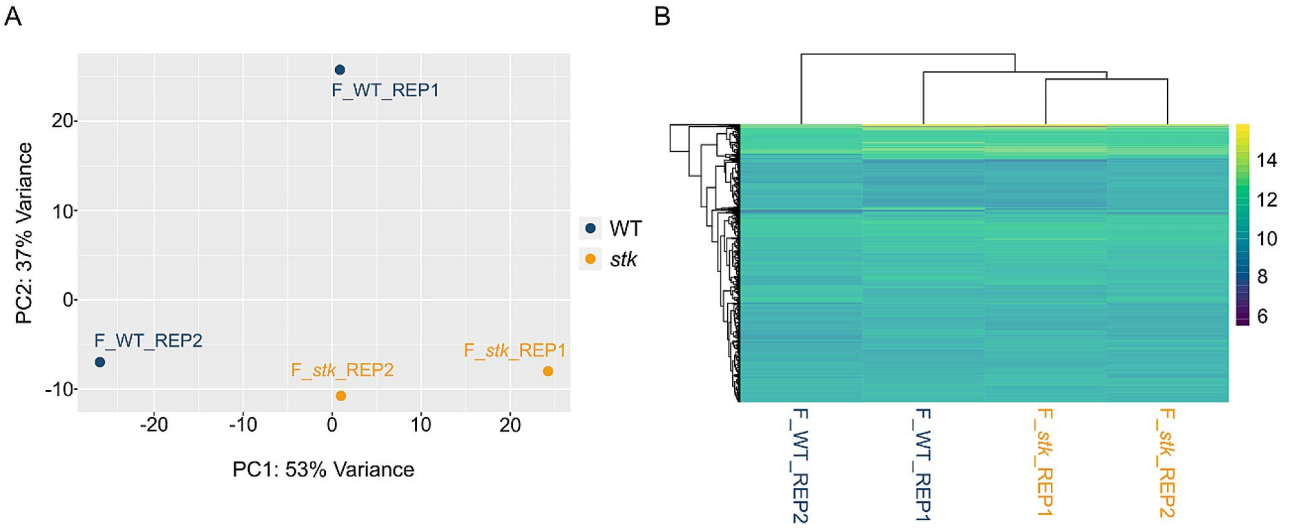

Fig. 1 Assessment of samples quality. (**A**) Principal component analysis (PCA) conducted on rlog normalised gene expression values of *stk* and WT samples. X – and Y – axes show PC1 and PC2, respectively, with the amount of variance contained in each component being 53% and 37%, respectively. Each point in the plot represents a biological replicate of 100 funiculi, with a total of four biological replicates in the plot. Symbols of the same colours are replicates of the same experimental group, where orange represents *stk* funiculi, and blue represents WT funiculi. (**B**) Heat map of the top 1000 genes selected based on the mean of the rlog normalised counts. Individual samples are shown in columns and genes in rows. The upper axis shows the clusters of samples, and the left vertical axis shows the clusters of genes. The colour scale represents the relative expression of genes: purple indicates low relative expression levels and yellow indicates high relative expression levels

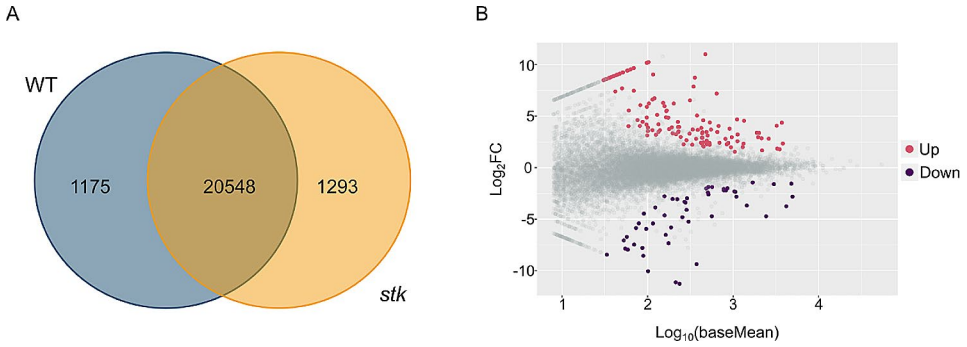

Fig. 2 Global analysis of *stk* vs. WT transcriptome from funiculi at st 17. (**A**) Venn diagram showing the number of genes expressed in each genotype. The number of genes expressed only on *stk* funiculi or WT funiculi are shown, as well as the number of overlapping genes between transcriptomes. WT is represented by a blue circle and *stk* by an orange circle. (**B**) MA plot demonstrates the relationship between the $\log_{10}$ average normalised expression on the x – axis and the significance of the differential expression test expressed as $\log_2$FoldChange ($\log_2$FC) on the y – axis for each gene in the transcriptome. It illustrates the number of DEGs in the *stk* funiculi. Gray dots represent genes that are not significantly differentially expressed, while pink and purple dots represent genes that are significantly up- and downregulated, respectively, with p-adjust value < 0.05

and a $-1 < \log_2$FoldChange ($\log_2$FC) > 1, from which 119 genes were upregulated and 50 were downregulated, as seen in the MA plot (Fig. 2B).

Parametric analysis of gene set enrichment

To better understand the processes being affected in *stk* funiculus compared to WT, a parametric analysis of gene set enrichment (PAGE) was performed. PAGE is a statistically sensitive method that considers gene expression levels to rank an annotated gene cluster. Genes were ranked following the formula: -Log10(p-value) multiplied by the sign of the Log_2FC. Each gene ontology (GO) term has an associated z-score value, which provides information about the significance of the enriched term among the data. Moreover, the z-score can assume a positive or negative value, depending on whether the term includes more up- or downregulated genes, respectively. For a full list of biological and molecular function GO terms, please see Supplementary Table 6. Significant enrichment of 53 and seven GO terms for biological function and molecular function, respectively, was observed among the gene datasets. The most significantly upregulated enriched GO category was cellular cell wall organization or biogenesis (GO:0070882) (Fig. 3A), with more than 100 genes associated. Terms related to dehiscence were active in *stk* transcriptome: dehiscence (GO:0009900), lignin biosynthetic process (GO:0009809) and phenylpropanoid metabolic process (GO:0009698). In addition, many cell wall related terms were found such as hemicellulose metabolic process (GO:0010410), xylan biosynthetic process (GO:0045491) and glucuronoxylan biosynthetic process (GO:0010417). Apart from these, regulation of DNA binding (GO:0051101) was a term that stood out. Concerning the molecular function, only two terms were upregulated, oxidoreductase (GO:0016722) and acyl-CoA hydrolase activity (GO:0047617), with the first one presenting a higher fold enrichment (Fig. 3B).

PAGE analysis identified the GO term glycoside catabolic process (GO:0016139) as the most inhibited (Fig. 4A). The absence of *STK* in the funiculus affected the sucrose metabolism (GO:0005985), starch metabolism (GO:0005982) and the immune effector process (GO:0002252). Furthermore, the term cell cycle (GO:0007049) appeared in the analysis, containing more than 400 genes, as well as other related terms like posttranscriptional regulation of gene expression (GO:0010608) and gene silencing (GO:0016458). The glutathione metabolic process and gynoecium development were also downregulated in the *stk* funiculi transcriptomic profile. The enriched molecular functions for the downregulated genes revealed terms exclusively related to enzymatic activity: transferase activity, sucrose alpha-glucosidase and beta-glucosidase activity (Fig. 4B).

Expression of funiculus-specific genes in the transcriptome

The expression profiles of genes previously identified as funiculus-specific [9] were used to assess the reliability of the transcriptome. All 20 funiculus-specific genes were detected in the RNA-seq data (Table 1), with two of them (*EXPANSIN-LIKE B1* and *GRETCHEN HAGEN 3.3*) being downregulated in the *stk* mutant. These genes represent an expansin-like, involved in cell wall extension, and an IAA-amido synthase, implicated in auxin responses, respectively.

Enriched networks on DEGs list

To further investigate enriched networks among DEGs and how genes correlated with each other's regarding physical interactions, co-expression and shared protein domains, the up- and downregulated DEGs were uploaded separately to the GeneMANIA program. All significant biological functions predicted by this tool with false discovery rate (FDR) < 0.05 were listed in Supplementary Table 7. Regarding the downregulated genes

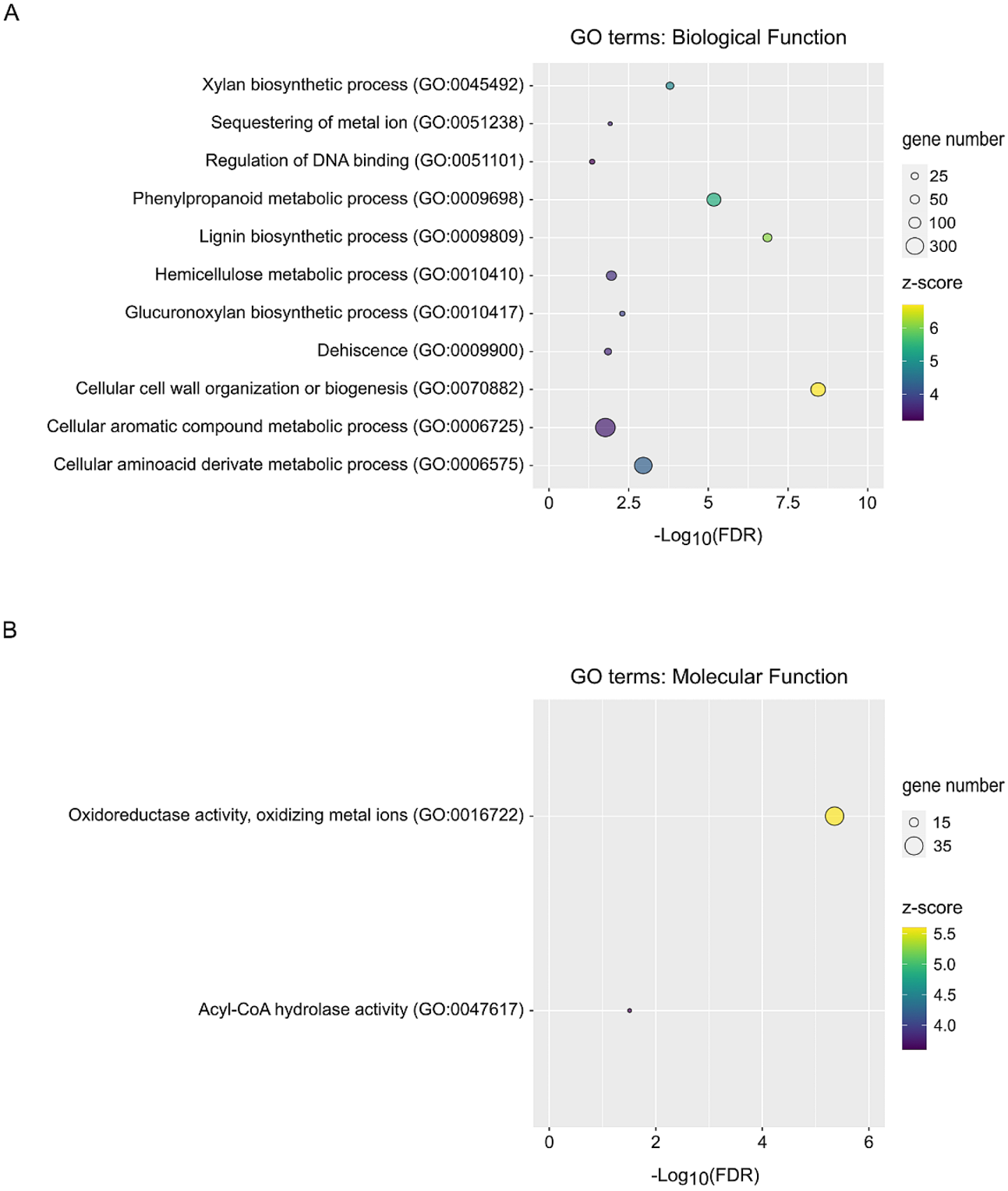

Fig. 3 Parametric GO gene set enrichment analysis of activated pathways in *stk* vs. WT transcriptome from funiculi at st 17. GO enrichment was evaluated at two different levels: (**A**) Biological Processes and (**B**) Molecular Function. Bubble plot diagrams show the significance [-Log_{10}(false discovery rate; FDR)] on the x-axis, the fold enrichment (z-score) in a gradient of colour from purple (underrepresented) to yellow (overrepresented), and the number of genes in each pathway

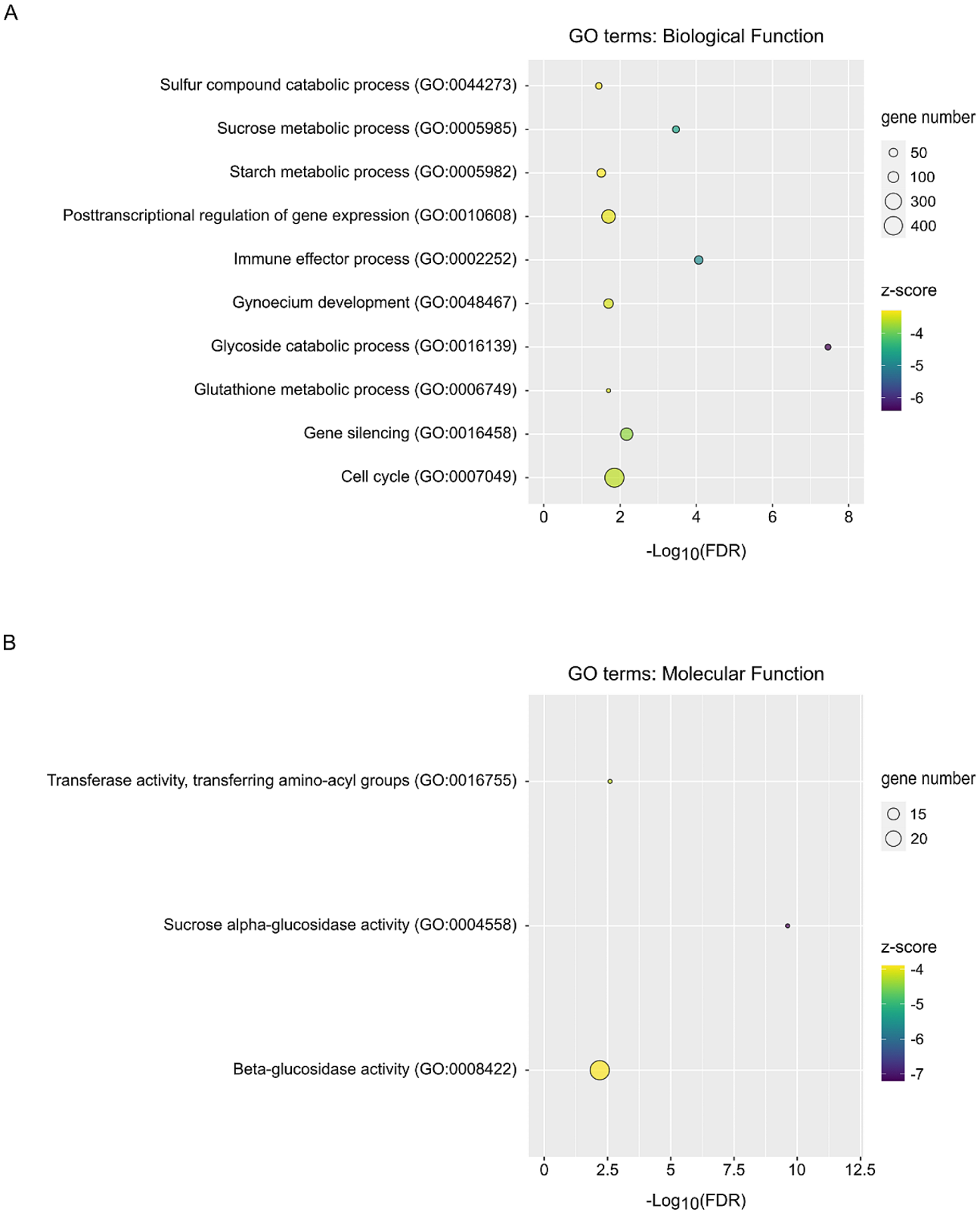

Fig. 4 Parametric GO gene set enrichment analysis of inhibited pathways in *stk* vs. WT transcriptome from funiculi at st 17. GO enrichment was evaluated at two different levels: (**A**) Biological Processes and (**B**) Molecular Function. Bubble plot diagrams show the significance [$-\text{Log}_{10}$(false discovery rate; FDR)] on the x-axis, the fold enrichment (z-score) in a gradient of colour from purple (overrepresented) to yellow (underrepresented), and the number of genes in each pathway

Table 1 Expression of funiculus-specific genes [9] on *stk* vs. WT transcriptome from funiculi at St 17

Locus	Annotation	baseMean	log-$_2$FC	*p*-value	*p*-adjust value
AT4G17030	**EXPANSIN-like B1**	**4177.40**	**−3.75**	**1.05E-09**	**1.22E-06**
AT2G23170	**GRETCHEN HAGEN 3.3**	**4791.69**	**−1.56**	**1.45E-04**	**2.04E-02**
AT2G47160	REQUIRES HIGH BORON 1	1278.61	−1.92	2.53E-03	1.71E-01
AT2G22990	SINAPOYL-GLUCOSE 1	8570.09	−1.56	2.91E-03	1.86E-01
AT2G28305	LONELY GUY 1	2643.74	−1.70	1.51E-02	5.23E-01
AT2G21220	Auxin-responsive protein	996.61	1.69	2.68E-02	7.05E-01
AT5G46240	K + ATPase 1	426.75	−1.65	6.09E-02	9.53E-01
AT2G22980	SERINE CARBOXYPEPTIDASE-like 13	271.23	−2.25	6.15E-02	9.54E-01
AT5G67420	LOB domain protein 37	931.28	0.71	1.69E-01	1.00E+00
AT5G37730		616.69	−0.32	5.91E-01	1.00E+00
AT3G43190	SUCROSE SYNTHASE 4	3314.72	−0.62	2.31E-01	1.00E+00
AT4G01440	Nodulin MtN21 family protein	1294.70	0.12	8.52E-01	1.00E+00
AT1G73630	Calcium-binding protein	896.12	0.30	6.95E-01	1.00E+00
AT1G68520	Zinc finger (B-box type) family protein	197.29	0.19	8.12E-01	1.00E+00
AT1G80760	NOD26-like intrinsic protein 6;1	923.37	0.52	5.50E-01	1.00E+00
AT1G64130	Contains domain Bet v1-like	259.08	0.49	5.05E-01	1.00E+00
AT1G79430	ALTERED PHLOEM DEVELOPMENT	553.35	−0.38	5.33E-01	1.00E+00
AT1G78490	CYTOCHROME P450, FAMILY 708, SUBFAMILY A, POLYPEPTIDE 3	376.31	0.52	5.03E-01	1.00E+00
AT2G15310	ADP-RIBOSYLATION FACTOR B1A	587.33	−0.99	1.96E-01	1.00E+00
AT2G35960	NDR1/HIN1-like 12	593.43	−0.40	5.74E-01	1.00E+00

Two genes on this list (*EXPANSIN-like B1* and *GRETCHEN HAGEN 3.3*) were downregulated in the transcriptome (Bold cells). The baseMean, log$_2$FoldChange (log$_2$FC), p-value and p-adjust value for each gene can also be seen in the table

in the *stk* transcriptome, the program detected a network related to carbohydrate transmembrane transporter activity (Fig. 5A). *SUGARS WILL EVENTUALLY BE EXPORTED TRANSPORTER 4* (*SWEET4*) and *SWEET7* were the two DEGs identified as involved in this process, being co-expressed and physically interacting with other members of SWEET family. Another network that stood out was associated with floral development (Fig. 5A), containing four DEGs: *CELL WALL INVERTASE 4* (*CWINV4*), *ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE 1* (*ADPG1*), *CYTOCHROME P450, FAMILY 90, SUBFAMILY D, POLYPEPTIDE 1* (*CYP90D1*) and *ARGONAUTE 9* (*AGO9*). However, no co-expression was detected between these genes. These two selected enriched networks share not only a common gene, *SWEET13*, but are also linked by co-expression. For instance, *CWINV4* is co-expressed with *SWEET4*, while *ADPG1* is co-expressed with *SWEET7*. For the upregulated DEGs, the most enriched network with smallest FDR identified by the program was associated with plant type-cell wall biogenesis (Fig. 5B). Nine DEGs were identified such as *NAC DOMAIN CONTAINING PROTEIN 43* (*NAC043*), *NAC012*, *LACCASE 4* (*LAC4*) and *KNOTTED-LIKE HOMEOBOX OF ARABIDOPSIS THALIANA 7* (*KNAT7*). In addition, other DEGs and genes that the program identified as involved with cell wall biogenesis can be visualised. For instance, *WRKY12* is co-expressed with both *NAC043* and *NAC012*, *LAC4* is co-expressed with *LAC2*, which is also a DEG in the present transcriptome, and is co-localised with *KNAT7*.

RNA-sequencing data validation through qPCR and promoter expression analysis

To validate the RNA-seq data, we selected a group of up- and downregulated genes from the transcriptome (Table 2), which we considered as potential candidates to be further studied. First, we analysed the GO terms and enriched networks, and we chose genes known to be involved in abscission and seed development. *LAC2* and *LAC4*, members of laccase family, are involved in lignin biosynthesis [46, 47], *SWEET4* and *SWEET7*, belonging to sweet family, are related to sugar transport [48, 49] and *ADPG1*, with pectinase activity, is implicated in abscission [50]. A comprehensive literature search was conducted on all deregulated genes from the transcriptome to identify other potential candidates that were not recognised on the enriched networks analysis. Two distinct gene families were subsequently selected for further investigation. *MAJOR LATEX PROTEIN-like 28* (*MLP28*) and *MLP168* are part of the major latex protein family, which is involved with hormone signalling [51], and *SERINE CARBOXYPEPTIDASE*-like *41* (*SCPL41*), an enzyme with membrane lipid metabolism functions [52]. All genes selected were expressed in pistil (Supplementary

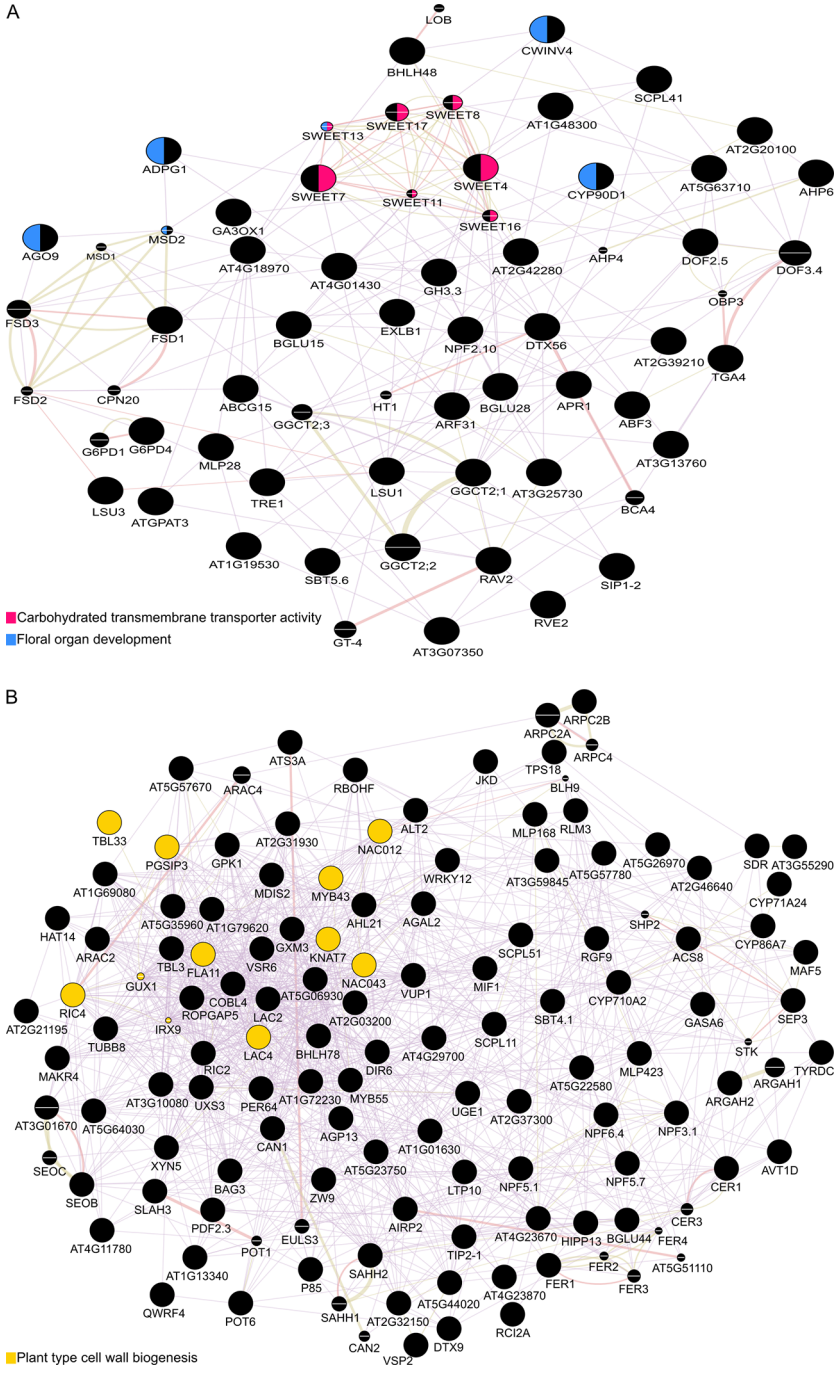

Fig. 5 Expression networks of *stk* vs. WT transcriptome from funiculi at st 17 derived from GeneMANIA. (**A**) Networks known from the downregulated differentially expressed genes (DEGs). (**B**) Networks identified from the upregulated DEGs. Genes added as relevant interactors by the GeneMANIA program are shown with a white dash in the middle of the circle, and the size is proportional to the number of interactions they have. The purple edges denote co-expression network interactions, the orange edges represent physical interactions between genes and the green edges indicate shared domain. Different colours in each circle show different functions, as indicated

Fig. 2) and silique (Supplementary Fig. 3) according to the online database ePlant [53] and Arabidopsis eFP Browser [54], respectively.

The relative expression levels of eight selected genes, *MLP168*, *LAC2*, *LAC4*, *SWEET4*, *SCPL41*, *MLP28*, *ADPG1* and *SWEET7* were quantified by qPCR in *stk* and WT funiculi at st 17 [21]. The transcript levels were normalised to three reference genes (*ACT7*, *YLS8 and HIS3.3*), previously tested according to Ferreira and colleagues [39] guidelines to infer about their appropriateness for funiculi samples, and are presented relative to WT funiculi transcript levels (Fig. 6). The analysis

Table 2 List of differentially expressing genes candidates from the *stk* vs. WT funiculi transcriptome

Locus	Gene symbol	Gene name	*stk* vs. WT			Group
			baseMean	log_2FC	p-adjustvalue	
AT1G35310	MLP168	MAJOR LATEX PROTEIN-like 168	58.92	9.44	1.21E-03	Up
AT2G29130	LAC2	LACCASE 2	356.87	8.73	1.20E-06	
AT2G38080	LAC4	LACCASE 4	1341.28	4.55	2.31E-03	
AT3G28007	SWEET4	SUGARS WILL EVENTUALLY BE EXPORTED TRANSPORTER 4	781.16	−2.21	7.95E-04	Down
AT5G42230	SCPL41	SERINE CABOXYPEPTIDASE-like 41	1035.41	−2.31	1.64E-05	
AT1G70830	MLP28	MAJOR LATEX PROTEIN-like 28	4941.77	−2.81	1.06E-09	
AT3G57510	ADPG1	ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE 1	123.03	−3.86	4.92E-02	
AT4G10850	SWEET7	SUGARS WILL EVENTUALLY BE EXPORTED TRANSPORTER 7	69.68	−7.47	5.67E-03	

Gene locus, gene symbol and name, baseMean, log_2FoldChange (log_2FC) and the p-adjust value are detailed for each gene. The list is organised in descending order according to the fold changes

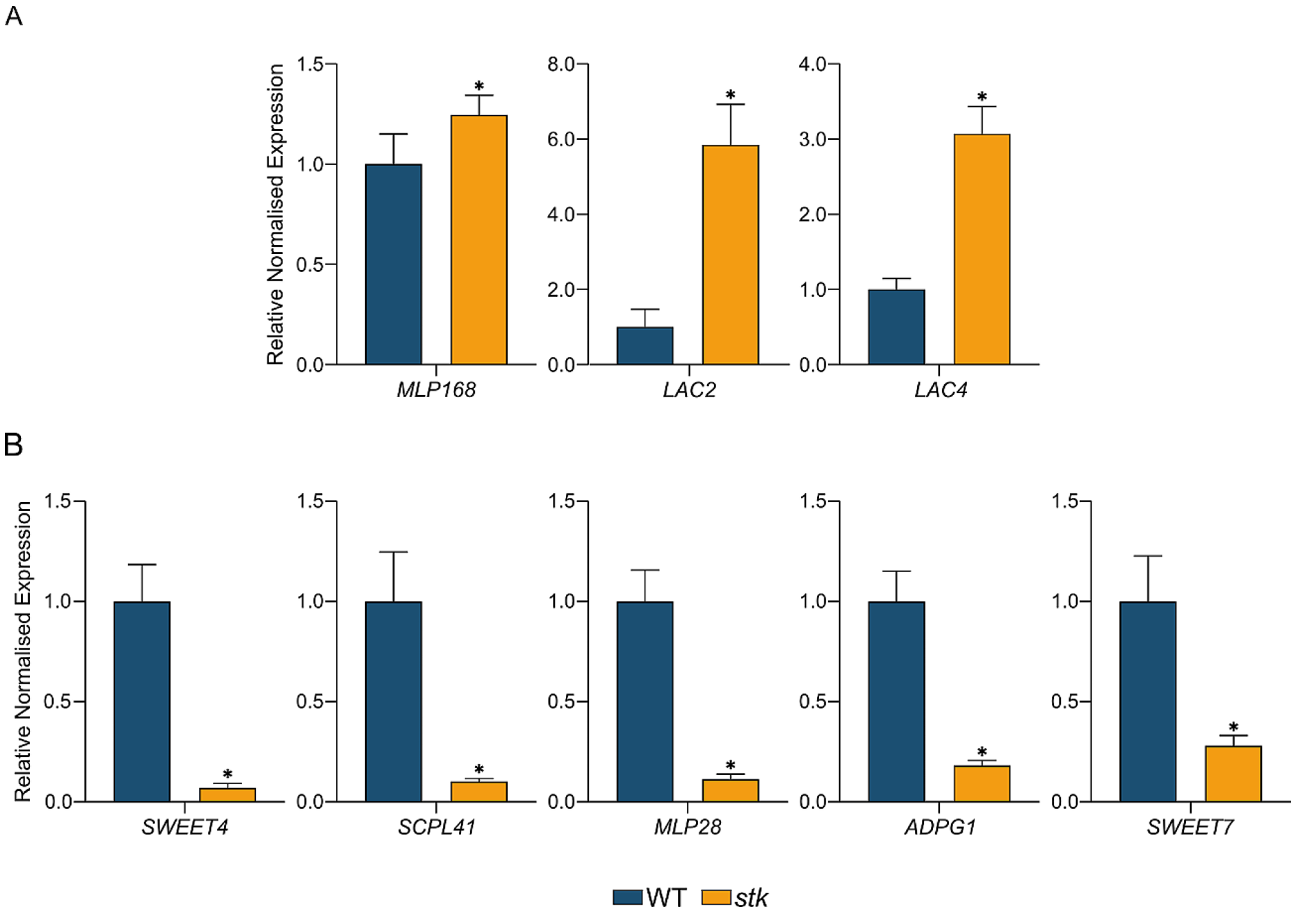

Fig. 6 Relative normalised expression levels of *MLP168*, *LAC2*, *LAC4*, *SWEET4*, *SCPL41*, *MLP28*, *ADPG1* and *SWEET7* in WT and *stk* funiculi at st 17 [21]. (**A**) and (**B**) Relative normalised expression of upregulated and downregulated genes, respectively, in *stk* funiculi according to the RNA-seq data. The transcript levels were normalised to *ACT7*, *YLS8 and HIS3.3* reference genes [39]. The data correspond to the ratio of the expression compared to the WT. Error bars represent the standard error of the mean of three independent biological replicates, each with three technical replicates. Statistical analysis was performed using a Student's t-test. Asterisks indicate significant differences from WT (*, $p < 0.05$)

revealed that *MLP168*, *LAC2* and *LAC4* expression was higher on *stk* funiculi compared to WT (Fig. 6A). In contrast, *SWEET4*, *SCPL41*, *MLP28*, *ADPG1* and *SWEET7* were downregulated on *stk* samples (Fig. 6B). These results were consistent with the transcriptome dataset (Table 2), with a correlation coefficient of 0.809 (Supplementary Fig. 4), indicative of good correlation between RNA-seq and qPCR expression values for candidate genes.

A promoter-driven GUS assay was also performed to further confirm the RNA-seq transcriptome. For instance, we analysed the promoter activity of *MLP28*

(p*MLP28*), *SCPL41* (p*SCPL41*) and *ADPG1* (p*ADPG1*). The promoter regions of these genes were fused to the *GUS* reporter gene, and funiculi from both *stk* and WT transgenic plants were observed at st 17 of flower development [21], corresponding to the same type of sample used to obtain the transcriptome (Fig. 7). The selection of downregulated genes facilitated the observation of differences in the promoters' expression between *stk* and WT funiculi. GUS activity driven by p*MLP28* was observed in the entire WT funiculus and septum, whereas in the *stk* background, the detection of the promoter ceased in these two tissues (Fig. 7A and D). In the case of p*SCPL41*, GUS expression was strong in the WT funiculus, however, in the *stk* background, there was no expression of *GUS* in the funiculus (Fig. 7B and E). For p*ADPG1*, the GUS signal was restricted to the SAZ established in the WT funiculus and no GUS staining was visible in the *stk* funiculus (Fig. 7C and F). Overall, these results confirmed the transcriptome data and qPCR, which showed a decrease in transcript levels of *MLP28*, *SCPL41* and *ADPG1* in the *stk* funiculus.

Discussion

Seeds are essential for plants, as they harbour the new offspring, and for humans, as a major source of calories. The processes involved in seed development are complex and require an accurate coordination of several regulatory networks established in different reproductive tissues [2]. The funiculus connects the maternal plant to the ovule and lately to the seed, contributing to nutrient transport, mechanical support, and seed positioning. When one of these functions is altered, as observed in *stk* mutants [11], seed development is disturbed. This influences seed size and viability as well as seed dispersal, ultimately compromising the reproductive success of the plant. Deciphering the genetic networks controlled by STK in the funiculus is, therefore, of major importance in the frame of elucidating how seed development is being disrupted.

Here we analysed the funiculus transcriptional profile of *stk* and WT st 17 of flower development to infer the processes affected by STK absence and gene regulatory networks enriched among the DEGs. We show that cell wall biogenesis, lignification and sugar transport are enriched in the transcriptome and which DEGs are known to be involved in these processes. We then focus on selecting candidate genes for future studies, by validating their expression using qPCR and/or promoter reporter lines.

We started by performing a collection assay to obtain RNA-seq data from *stk* and WT funiculi of st 17 of flower development. A total of 169 genes (p-adjust value < 0.05) were found to be deregulated in the transcriptome (Supplementary Table 5), indicating that the transcriptome profiles of *stk* and WT were not completely distinct, as supported by PCA. Commonly, RNA-seq studies comparing different tissues or developing stages tend to show a more evident separation between different biological

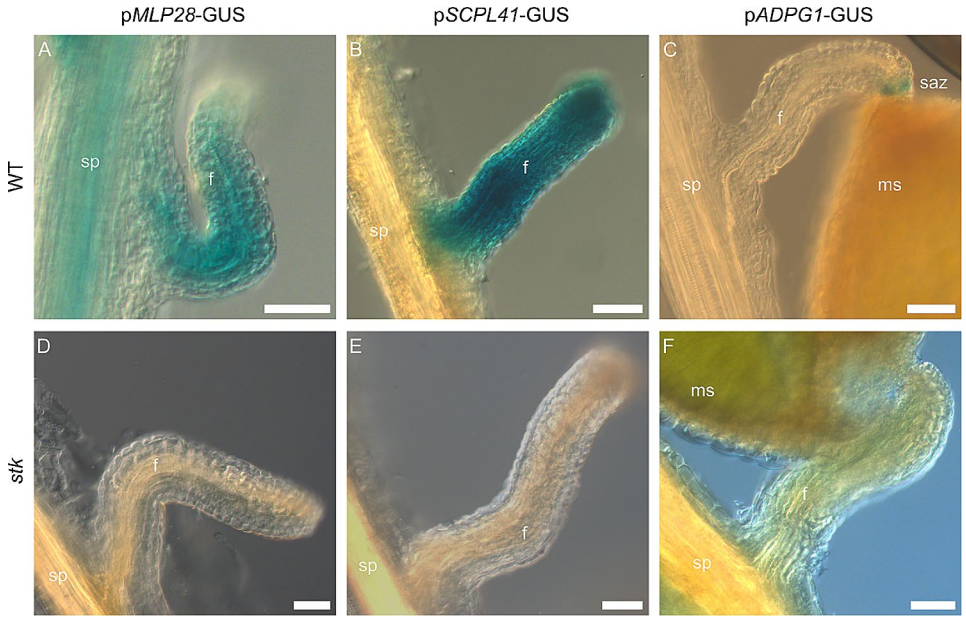

Fig. 7 Histochemical localisation of GUS activity in WT (**A**–**C**) and *stk* (**D**–**F**) funiculi at st 17 [21], expressing p*MLP28*-GUS, p*SCPL41*-GUS or p*ADPG1*-GUS constructs. (**A** and **D**) p*MLP28* expression was detected in WT funiculus but not in the *stk* background. (**B** and **E**) p*SCPL41* was strongly expressed in WT funiculus, while the GUS activity ceased in *stk* funiculus. (**C** and **F**) p*ADPG1* signal was restricted to the SAZ of WT funiculus, but it was not detected on the *stk* funiculus. Three independent transgenic lines were observed for each situation. Abbreviations: f – funiculus; ms – mature seed; saz – seed abscission zone; sp – septum. Scale bar = 50 µm

samples and, thus, a high number of DEGs [55–57]. Moreover, the transcriptomic data generated hereby were from low-input mRNA, which could explain the reduced coverage of low expressed transcripts and, consequently, the detection of low expressed genes as DEGs. Generally, the number of DEGs is slightly lower when compared to libraries obtained from higher amounts of input RNA [58]. It is worth noting that all genes classified as funiculus-specific by Khan and colleagues [9] were present in the transcriptome, providing a high level of confidence for further analysis.

The most actively enriched function in *stk* funiculi was related to cell wall biogenesis (GO:0070882), consistent with previous phenotypic observations [11, 22]. After fertilisation, the WT funiculus vasculature proliferates in a synchronised manner with the cortex and epidermis [9], with growth being more accentuated in *stk* mutants. Therefore, terms related to the production of cell wall components such as polysaccharides [xylan (GO:0045491), glucuronoxylan (GO:0010417) and hemicellulose (GO:0010410)] were found to be enriched in PAGE analysis. The proliferation of cells is characterised by activation of the transcriptional machinery and, in the case of *stk*, transcription needs to be intensified to allow the overgrowth of the funiculus. This may explain the DNA binding (GO:0051101) GO term enrichment. The *stk* mutant has defects in seed dehiscence [11] and the reason was later revealed by Balanzà and colleagues [22], who concluded that an irregular accumulation of lignin in *stk* SAZ caused this phenotype. This control of lignin production by STK was reflected in the enrichment of GO terms for dehiscence (GO:0009900), lignification (GO:0009809) and phenylpropanoid metabolism (GO:0009698), which is a lignin precursor [59]. The analysis of molecular functions revealed the induction of genes related to oxidoreductase activity (GO:0016722). In plants, enzymes such as catalases or peroxidases present this type of activity, and are responsible for regulating the cellular level of hydrogen peroxide, a reduced oxygen species [60]. In the present transcriptome, *PEROXIDASE 64* (*PRX64*) was upregulated in *stk*, and many studies using loss-of-function mutants or reporter lines observation [61–63] have concluded that PRX64 is involved in lignification. The link between this peroxidase and lignification in SAZ has not yet been described, making it an interesting gene to be further analysed. The most inhibitory process in *stk* transcriptome was the glycoside catabolic process (GO:0016139). Glycosides are a wide range group of plant secondary metabolites with multifaceted roles in regulating growth, defense and reproduction in plants [64–66]. Glycoside hydrolases are a family of enzymes responsible for the removal of specific sugar moieties from glycosides. The mutant analysis of a specific enzyme, BRANCHING ENZYME 1, revealed an impairment in carbohydrate metabolism [67], which is another GO term revealed in this transcriptome. Sugars are essential for embryogenesis, and the fact that STK is controlling both glycosylation and sugar catabolism (GO:0005985) might related to the small seed phenotype that *stk* presents. Also related to *stk* seed size are the terms starch metabolism (GO:0005982) and glutathione metabolism (GO:0006749). Starch is a source of sugar production in the funiculus, that can possibly be transported to the endosperm to nourish the embryo [9]. On the other hand, glutathione is a cell redox that helps in preventing an accumulation of reactive oxygen species within the cell [68]. A previous study showed that glutathione is expressed and accumulates in the funiculus as well as it is crucial for proper embryo development and seed maturation [69]. Collectively, these results suggest that STK controls the metabolism and influx of molecules from the funiculus to the seed, which are important for the correct development of this unit.

A closer look at the DEGs list, using GeneMANIA, allowed the identification of different networks and which specific DEGs were part of them, as well as which genes were being co-expressed, shared protein domains or had a known physical interaction with the DEGs. Not all DEGs were associated with a biological process, which may be a result of an uncharacterised gene or the program itself, since GeneMANIA shows improved outcomes if the input gene list is functionally related. Nonetheless, this analysis contributed for the criteria of selecting genes known to be involved in abscission or seed development. One of the enriched networks for downregulated DEGs was associated with carbohydrate transport activity, a process that allows the delivery of sucrose with maternal origin to the seed coat via the funiculus phloem, important for seed development [70–72]. Two genes from the transcriptome were identified in this network, *SWEET4* and *SWEET7*, both with glucose transport activity, indicating that STK promotes the sugar efflux required for seed filling (Eom et al., 2015), particularly in the form of glucose instead of sucrose. This finding, combined with other previously reported data [16, 17, 19, 20], may explain the small seed phenotype that *stk* exhibits [11]. The other network that stood out included terms associated with floral organ development. CWINV4 is an enzyme involved in sugar signalling [73] that hydrolyses sucrose to generate glucose and fructose. It was first found to be involved in nectar production in Arabidopsis [74] and, afterwards, a study reported that CWINV4 together with CWINV2 (mentioned in the paper as CWIN4 and CWIN2) are positive regulators of ovule initiation, through sugar signalling [75]. Remarkably, transcriptomic data obtained from the double mutant transgenic line p*STK*-amiRNA*CWIN24* showed several members of the SWEET family as downregulated

DEGs, including *SWEET4* and *SWEET7*, as well as TFs, such as STK [75]. Therefore, it would be valuable to investigate whether STK involvement in sugar efflux is due to the control of CWINV, which would produce sucrose that could be transported to the seed via SWEET transporters. So far, we had confirmed by qPCR that *SWEET4* and *SWEET7* are downregulated in *stk* funiculi at st 17. Another gene, part of the floral development network, was *ADPG1* that belongs to the polygalacturonase family, which is responsible for disassembling the pectin part of plant cell walls [76]. A genetic approach demonstrated that ADPG1 together with ADPG2 are important for silique dehiscence, whereas the combined action of ADPG1, ADPG2 and QUARTET 2 contributes to anther dehiscence [50]. Observations based on transcriptional β-glucuronidase activity were in agreement with the phenotype mentioned above: *ADPG1* was mainly expressed in silique dehiscence zones and in anthers, immediately before anthesis, but also in SAZ. Using the reporter line generated in the present study, we simultaneously confirmed that *ADPG1* promoter is specifically expressed in the SAZ and that GUS activity ceases in the *stk* background (in accordance with RNA-seq and qPCR results). Notably, the authors showed that *ADPG1* expression in the SAZ is disrupted in the *hecate 3* mutant, a TF controlled by STK and involved in SAZ formation [22]. Despite the absence of an "obvious defect" in the *adpg1* mutants mentioned by the authors, they did not reveal the type of analysis that was performed. Hence, we believe that these findings point to the involvement of ADPG1 with STK during seed abscission. Among the upregulated DEGs, the most enriched network recognised by GeneMANIA was the plant cell wall biogenesis, in which a laccase family member was identified, *LAC4*. Laccases have been associated with lignin biosynthesis [77], functioning as activators of lignin production [47, 78]. Studies have shown that disruption of LAC4, along with other laccases, leads to a decrease in lignin content of several organs, such as stems, siliques and roots [47, 78]. Interestingly, only one study has reported the activity of a specific laccase, LAC2, as a negative regulator of lignin deposition in root vascular tissues under abiotic stress [46]. *LAC2* was identified as being co-expressed with *LAC4*, and qPCR indicated that there was more gene expression in *stk* with respect to WT. Since *stk* SAZ has an overproduction of lignin [22], these findings suggest that LAC2 and LAC4 participate in the STK-network controlling lignin biosynthesis in the funiculus.

Three additional candidate genes were selected for validation: *MLP168* and *MLP28*, members of the pollen allergen Bet v 1 family [79] and the pathogenesis-related protein class 10 as well. Their protein conformation exhibits a hydrophobic cavity capable of binding to ligands, such as cytokinins, brassinolides and secondary metabolites, which MLPs transport to other organs via the plant vasculature, triggering, in some cases, downstream transduction signals in those organs [51]. Extensive studies on MLPs functions, mostly under biotic and abiotic stresses, have been conducted in several plant species, like *A. thaliana*, *Brassica rapa*, and *Vitis vinifera* [51]. Regarding MLP168, it was reported an interaction of this protein with the ABA INSENSITIVE 5 protein in a yeast two-hybrid assay [80]. We validated by qPCR that *MLP168* was upregulated in *stk* funiculi. As for MLP28, observations of an amiRNA-*MLP28* line showed defects on vegetative growth with alterations in leaf morphology and shoot apex, dwarfness and eventual premature plant death [40]. The present transcriptomic data revealed that *MLP28* was downregulated in *stk* funiculi, which was confirmed by qPCR and promoter-driven GUS activity assays. Due to MLPs nature of binding to plant hormones already implicated in several processes of reproduction [81], it seems likely their involvement with ligand transportation from the funiculus to the seed. The last gene selected was *SCPL41*, a serine carboxypeptidase-like protein containing a catalytic triad responsible for cleaving the C-terminal peptide bond in proteins or peptides [82]. The analysis of *scpl41* mutants has revealed an increase in lipid content in seedlings [52]. Since *stk* funiculus cells undergo an augmentation in size that must be accompanied by overproduction of cell components, such as lipids, it is reasonable to think that SCPL41 control of lipid content in the funiculus may be regulated by STK. The qPCR results showed a subexpression of *SCPL41* in *stk* funiculi from st 17 flowers and the promoter line showed expression of GUS in the WT funiculus, which was absent in the *stk* background.

Conclusion

Overall, the presented work provides the first RNA-seq data from *stk* and WT funiculi at st 17. The RNA-seq analysis complemented a previous phenotyping study of *stk* funiculi by showing GO terms and enriched networks related to cell wall biogenesis, sugar transport and abscission. The results demonstrated that STK is mainly a repressor of cell wall and lignin biosynthesis-related genes, but simultaneously an activator of glucose transporters and seed dehiscence genes acting at the separation layer of the SAZ (Fig. 8). The validation of the transcriptome reinforced the reliability of the data obtained and confirmed that genes never associated with funiculus functions such as *MLP*s or *SCPL*s are, in fact, deregulated in *stk* funiculus (Fig. 8). These results provide new insights into the regulatory molecular network controlled by STK in the funiculus and highlight the existence of novel downstream targets of STK, whose functions require further exploration in the context of funiculus and seed development processes.

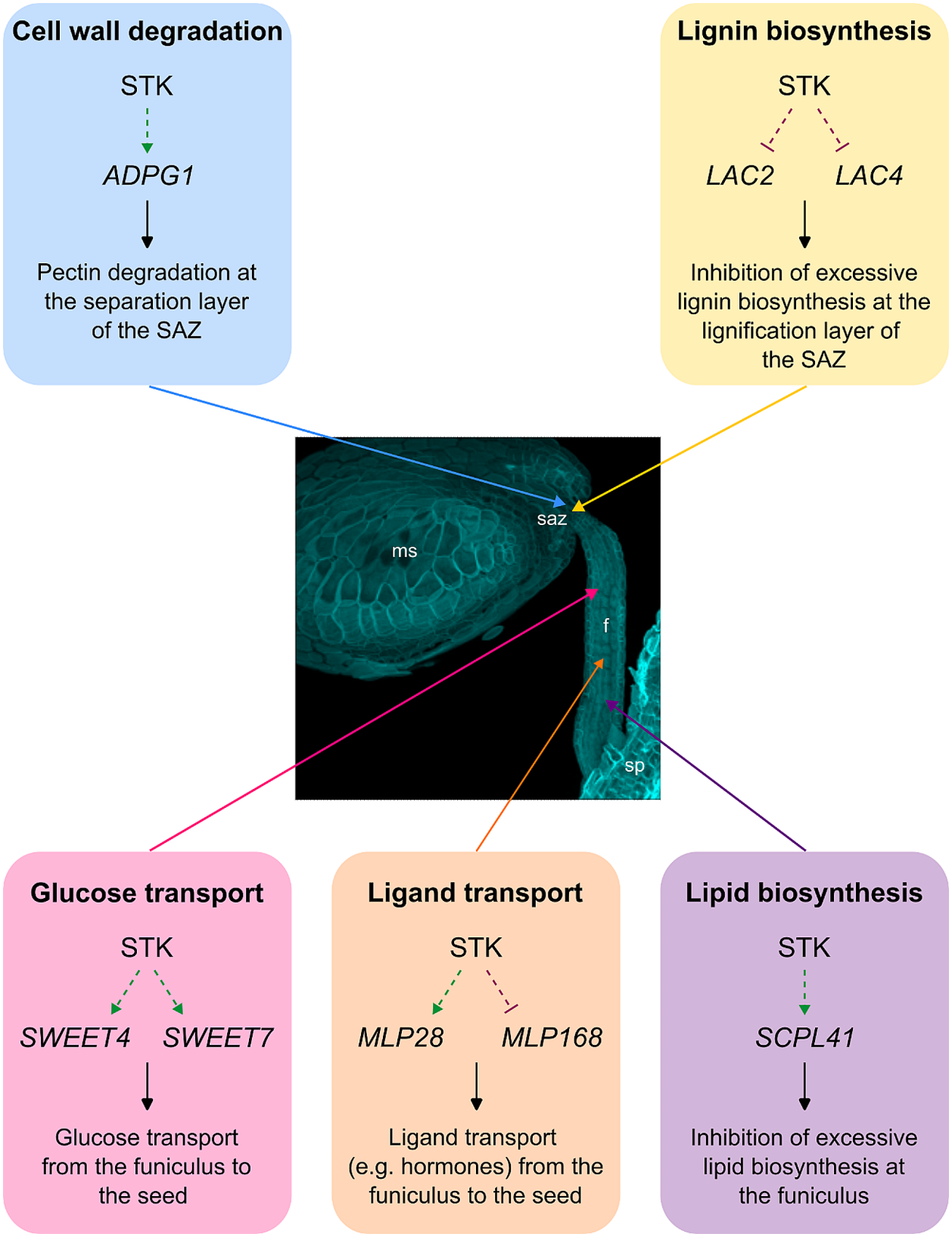

Fig. 8 (See legend on next page.)

(See figure on previous page.)

Fig. 8 A simplified model proposing the involvement of candidate genes from *stk* vs. WT transcriptome from funiculi at st 17 in the STK-mediated network established in the funiculus. (**Blue box**) STK activates the expression of *ADPG1* at the separation layer of the SAZ. This activation leads to pectin degradation of the cell wall, which is essential for seed abscission. (**Yellow box**) STK inhibits the expression of *LAC2* and *LAC4* at the lignification layer of the SAZ. This repression prevents an overproduction of lignin, facilitating seed abscission. (**Pink box**) STK activates *SWEET4* and *SWEET7* expression. These transporters facilitate glucose movement from the funiculus to the seed, that is important for proper seed development. (**Orange box**) STK also regulates *MLP28* and *MLP168* expression. *MLP28* is activated, while *MLP168* is repressed. This fine-tuned regulation ensures correct transport of ligands, such as hormones, from the funiculus to the seed, further supporting seed development. (**Purple box**) STK activates *SCPL41* expression, which repress the overproduction of lipids in funiculus cells. Green dashed arrow – activates gene expression; Pink dashed line with end bars – represses gene activation; black arrow – promotes action. Abbreviations: f – funiculus; ms – mature seed; saz – seed abscission zone; sp – septum

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-05489-4.

Supplementary Material 1

Acknowledgements
We acknowledge Dr Lucia Colombo for providing *stk* seeds, Dr David Haak and Dr Aureliano Bombarely for assistance with RNA-sequencing analysis, and Ayami Furuta for her technical assistance.

Author contributions
MJF, TH, and SC: conceptualisation; MJF, JS, and TS: formal analysis; MJF, JS, HT, TH, and SC: funding acquisition; MJF, JS, and TS: investigation; MJF, and HT: methodology; MJF: project administration; TH, and SC: resources; HT, TH and SC: supervision; MJF, and JS: validation; MJF: visualisation; MJF: writing—original draft; MJF, JS, HT, TH, and SC: writing—review & editing. All authors read and approved the final manuscript.

Funding
This work was supported by Fundação para a Ciência e Tecnologia and Ministério da Ciência, Tecnologia e Ensino Superior (FCT/MCTES) through national funds [grant agreement UIDB/50006/2020 DOI https://doi.org/10.54499/UIDB/50006/2020 to SC]. This research received financial support as well from Fundação para a Ciência e Tecnologia [PhD grant agreement SFRH/BD/143579/2019 to MJF], from "la Caixa" Foundation (ID 100010434) [grant agreement LCF/BQ/DR20/11790010 to JS], and from Japan Society for the Promotion of Science [Grant-in-Aid for Early-Career Scientists (18K14729, 20K15817) and Grant-in-Aid for Transformative Research Areas (22H05677, 23H04740) to HT; grant agreement 22H04980, 22K21352 to TH].

Data availability
The dataset generated in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE247348 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE247348).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 4 April 2024 / Accepted: 5 August 2024
Published online: 13 August 2024

References

1. Hamamura Y, Saito C, Awai C, Kurihara D, Miyawaki A, Nakagawa T, et al. Live-cell imaging reveals the dynamics of two sperm cells during double fertilization in *Arabidopsis thaliana*. Curr Biol. 2011;21:497–502.
2. Becker MG, Hsu S-W, Harada JJ, Belmonte MF. Genomic dissection of the seed. Front Plant Sci. 2014;5:464.
3. Estornell LH, Agustí J, Merelo P, Talón M, Tadeo FR. Elucidating mechanisms underlying organ abscission. Plant Sci Int J Exp Plant Biol. 2013;199–200:48–60.
4. Schneitz K, Hülskamp M, Pruitt RE. Wild-type ovule development in *Arabidopsis thaliana*: a light microscope study of cleared whole-mount tissue. Plant J. 1995;7:731–49.
5. Chan A, Belmonte MF. Histological and ultrastructural changes in canola (*Brassica napus*) funicular anatomy during the seed lifecycle. Botany. 2013;91:671–9.
6. Chan AC, Khan D, Girard IJ, Becker MG, Millar JL, Sytnik D, et al. Tissue-specific laser microdissection of the *Brassica napus* funiculus improves gene discovery and spatial identification of biological processes. J Exp Bot. 2016;67:3561–71.
7. Tan L, Varnai P, Lamport DTA, Yuan C, Xu J, Qiu F, et al. Plant O-hydroxyproline arabinogalactans are composed of repeating trigalactosyl subunits with short bifurcated side chains. J Biol Chem. 2010;285:24575–83.
8. Mawson BT, Steghaus AK, Yeung EC. Structural development and respiratory activity of the funiculus during bean seed (*Phaseolus vulgaris* L.) maturation. Ann Bot. 1994;74:587–94.
9. Khan D, Millar JL, Girard IJ, Chan A, Kirkbride RC, Pelletier JM, et al. Transcriptome atlas of the Arabidopsis funiculus - a study of maternal seed subregions. Plant J. 2015;82:41–53.
10. Lan J, Wang N, Wang Y, Jiang Y, Yu H, Cao X, et al. Arabidopsis TCP4 transcription factor inhibits high temperature-induced homeotic conversion of ovules. Nat Commun. 2023;14:5673.
11. Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, et al. Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature. 2003;424:85–8.
12. Mizzotti C, Ezquer I, Paolo D, Rueda-Romero P, Guerra RF, Battaglia R, et al. SEEDSTICK is a master regulator of development and metabolism in the Arabidopsis seed coat. PLoS Genet. 2014;10:e1004856.
13. Herrera-Ubaldo H, Lozano-Sotomayor P, Ezquer I, Di Marzo M, Chávez Montes RA, Gómez-Felipe A, et al. New roles of NO TRANSMITTING TRACT and SEEDSTICK during medial domain development in Arabidopsis fruits. Development. 2019;146:dev172395.
14. Di Marzo M, Roig-Villanova I, Zanchetti E, Caselli F, Gregis V, Bardetti P, et al. MADS-box and bHLH transcription factors coordinate transmitting tract development in *Arabidopsis thaliana*. Front Plant Sci. 2020;11:526.
15. Di Marzo M, Herrera-Ubaldo H, Caporali E, Novák O, Strnad M, Balanzà V, et al. SEEDSTICK controls Arabidopsis fruit size by regulating cytokinin levels and FRUITFULL. Cell Rep. 2020;30:2846–e28573.
16. Di Marzo M, Viana VE, Banfi C, Cassina V, Corti R, Herrera-Ubaldo H, et al. Cell wall modifications by α-XYLOSIDASE1 are required for control of seed and fruit size in Arabidopsis. J Exp Bot. 2022;73:1499–515.
17. Di Marzo M, Babolin N, Viana VE, De Oliveira AC, Gugi B, Caporali E, et al. The genetic control of SEEDSTICK and LEUNIG-HOMOLOG in seed and fruit development: new insights into cell wall control. Plants. 2022;11:3146.
18. Ezquer I, Mizzotti C, Nguema-Ona E, Gotté M, Beauzamy L, Viana VE, et al. The developmental regulator SEEDSTICK controls structural and mechanical properties of the Arabidopsis seed coat. Plant Cell. 2016;28:2478–92.
19. Paolo D, Orozco-Arroyo G, Rotasperti L, Masiero S, Colombo L, de Folter S, et al. Genetic interaction of SEEDSTICK, GORDITA and AUXIN RESPONSE FACTOR 2 during seed development. Genes. 2021;12:1189.
20. Paolo D, Rotasperti L, Schnittger A, Masiero S, Colombo L, Mizzotti C. The Arabidopsis MADS-domain transcription factor SEEDSTICK controls seed size via direct activation of *E2Fa*. Plants. 2021;10:192.
21. Smyth DR, Bowman JL, Meyerowitz EM. Early flower development in Arabidopsis. Plant Cell. 1990;2:755–67.

22. Balanzà V, Roig-Villanova I, Di Marzo M, Masiero S, Colombo L. Seed abscission and fruit dehiscence required for seed dispersal rely on similar genetic networks. Development. 2016;143:3372–81.
23. Niederhuth CE, Cho SK, Seitz K, Walker JC. Letting go is never easy: abscission and receptor-like protein kinases. J Integr Plant Biol. 2013;55:1251–63.
24. Yu Y-K, Li Y-L, Ding L-N, Sarwar R, Zhao F-Y, Tan X-L. Mechanism and regulation of silique dehiscence, which affects oil seed production. Front Plant Sci. 2020;11:580.
25. Gooh K, Ueda M, Aruga K, Park J, Arata H, Higashiyama T, et al. Live-cell imaging and optical manipulation of Arabidopsis early embryogenesis. Dev Cell. 2015;34:242–51.
26. Kadokura S, Sugimoto K, Tarr P, Suzuki T, Matsunaga S. Characterization of somatic embryogenesis initiated from the Arabidopsis shoot apex. Dev Biol. 2018;442:13–27.
27. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
28. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.
29. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923–30.
30. Love MI, Huber W, Anders S. Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
31. Wickham H. Getting started with ggplot2. ggplot2. Use R! Cham: Springer International Publishing; 2016.
32. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc. 2019;14:482–517.
33. Kim S-Y, Volsky DJ. PAGE: Parametric Analysis of Gene Set Enrichment. BMC Bioinformatics. 2005;6:144.
34. Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, et al. agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res. 2017;45:W122–9.
35. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.
36. Kõressaar T, Lepamets M, Kaplinski L, Raime K, Andreson R, Remm M. Primer3_masker: integrating masking of template sequence with primer design software. Bioinformatics. 2018;34:1937–8.
37. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. Bioinformatics. 2007;23:1289–91.
38. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. Nucleic Acids Res. 2012;40:e115–115.
39. Ferreira MJ, Silva J, Pinto SC, Coimbra S. I choose you: selecting accurate reference genes for qPCR expression analysis in reproductive tissues in *Arabidopsis thaliana*. Biomolecules. 2023;13:463.
40. Litholdo CG, Parker BL, Eamens AL, Larsen MR, Cordwell SJ, Waterhouse PM. Proteomic identification of putative MicroRNA394 target genes in *Arabidopsis thaliana* identifies major latex protein family members critical for normal development. Mol Cell Proteom. 2016;15:2033–47.
41. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol. 2007;8:R19.
42. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2 – ΔΔCT method. Methods. 2001;25:402–8.
43. Karimi M, Inzé D, Depicker A. GATEWAY™ vectors for Agrobacterium-mediated plant transformation. Trends Plant Sci. 2002;7:193–5.
44. Clough SJ, Bent AF. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. 1998;16:735–43.
45. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9:676–82.
46. Khandal H, Singh AP, Chattopadhyay D. The MicroRNA397b-LACCASE2 module regulates root lignification under water and phosphate deficiency. Plant Physiol. 2020;182:1387–403.
47. Zhao Q, Nakashima J, Chen F, Yin Y, Fu C, Yun J, et al. Laccase is necessary and nonredundant with peroxidase for lignin polymerization during vascular development in Arabidopsis. Plant Cell. 2013;25:3976–87.
48. Kuanyshev N, Deewan A, Jagtap SS, Liu J, Selvam B, Chen L-Q, et al. Identification and analysis of sugar transporters capable of co-transporting glucose and xylose simultaneously. Biotechnol J. 2021;16:e2100238.
49. Liu X, Zhang Y, Yang C, Tian Z, Li J. AtSWEET4, a hexose facilitator, mediates sugar transport to axial sinks and affects plant development. Sci Rep. 2016;6:24563.
50. Ogawa M, Kay P, Wilson S, Swain SM. ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE 1 (ADPG1), ADPG2, and QUARTET2 are polygalacturonases required for cell separation during reproductive development in Arabidopsis. Plant Cell. 2009;21:216–33.
51. Fujita K, Inui H, Review. Biological functions of major latex-like proteins in plants. Plant Sci. 2021;306:110856.
52. Chen J, Li W, Jia Y. The serine carboxypeptidase-like gene *SCPL41* negatively regulates membrane lipid metabolism in *Arabidopsis thaliana*. Plants. 2020;9:696.
53. Swanson R, Clark T, Preuss D. Expression profiling of Arabidopsis stigma tissue identifies stigma-specific genes. Sex Plant Reprod. 2005;18:163–71.
54. Mizzotti C, Rotasperti L, Moretto M, Tadini L, Resentini F, Galliani BM, et al. Time-course transcriptome analysis of Arabidopsis siliques discloses genes essential for fruit development and maturation. Plant Physiol. 2018;178:1249–68.
55. Hofmann F, Schon MA, Nodine MD. The embryonic transcriptome of *Arabidopsis thaliana*. Plant Reprod. 2019;32:77–91.
56. Susaki D, Suzuki T, Maruyama D, Ueda M, Higashiyama T, Kurihara D. Dynamics of the cell fate specifications during female gametophyte development in Arabidopsis. PLoS Biol. 2021;19:e3001123.
57. Zhang Y, Maruyama D, Toda E, Kinoshita A, Okamoto T, Mitsuda N, et al. Transcriptome analyses uncover reliance of endosperm gene expression on Arabidopsis embryonic development. FEBS Lett. 2023;597:407–17.
58. Bhargava V, Head SR, Ordoukhanian P, Mercola M, Subramaniam S. Technical variations in low-input RNA-seq methodologies. Sci Rep. 2014;4:3678.
59. Yao T, Feng K, Xie M, Barros J, Tschaplinski TJ, Tuskan GA, et al. Phylogenetic occurrence of the phenylpropanoid pathway and lignin biosynthesis in plants. Front Plant Sci. 2021;12:704697.
60. Pandey VP, Awasthi M, Singh S, Tiwari S, Dwivedi UN. A comprehensive review on function and application of plant peroxidases. Biochem Anal Biochem. 2017;6:308.
61. Lee Y, Rubio MC, Alassimone J, Geldner N. A mechanism for localized lignin deposition in the endodermis. Cell. 2013;153:402–12.
62. Morel O, Lion C, Neutelings G, Stefanov J, Baldacci-Cresp F, Simon C, et al. REPRISAL: mapping lignification dynamics using chemistry, data segmentation, and ratiometric analysis. Plant Physiol. 2022;188:816–30.
63. Tokunaga N, Kaneta T, Sato S, Sato Y. Analysis of expression profiles of three peroxidase genes associated with lignification in *Arabidopsis thaliana*. Physiol Plant. 2009;136:237–49.
64. Choi H, Ohyama K, Kim Y-Y, Jin J-Y, Lee SB, Yamaoka Y, et al. The role of Arabidopsis ABCG9 and ABCG31 ATP binding cassette transporters in pollen fitness and the deposition of steryl glycosides on the pollen coat. Plant Cell. 2014;26:310–24.
65. Kytidou K, Artola M, Overkleeft HS, Aerts JMFG. Plant glycosides and glycosidases: a treasure-trove for therapeutics. Front Plant Sci. 2020;11:357.
66. Moreira D, Kaur D, Pereira AM, Held MA, Showalter AM, Coimbra S. Type II arabinogalactans initiated by hydroxyproline-O-galactosyltransferases play important roles in pollen–pistil interactions. Plant J. 2023;114:371–89.
67. Wang X, Xue L, Sun J, Zuo J. The Arabidopsis *BE1* gene, encoding a putative glycoside hydrolase localized in plastids, plays crucial roles during embryogenesis and carbohydrate metabolism. J Integr Plant Biol. 2010;52:273–88.
68. Gill SS, Tuteja N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol Biochem. 2010;48:909–30.
69. Cairns NG, Pasternak M, Wachter A, Cobbett CS, Meyer AJ. Maturation of Arabidopsis seeds is dependent on glutathione biosynthesis within the embryo. Plant Physiol. 2006;141:446–55.
70. Chen L-Q, Lin IW, Qu X-Q, Sosso D, McFarlane HE, Londoño A, et al. A cascade of sequentially expressed sucrose transporters in the seed coat and endosperm provides nutrition for the Arabidopsis embryo. Plant Cell. 2015;27:607–19.
71. Stadler R, Lauterbach C, Sauer N. Cell-to-cell movement of green fluorescent protein reveals post-phloem transport in the outer integument and identifies symplastic domains in Arabidopsis seeds and embryos. Plant Physiol. 2005;139:701–12.
72. Weber H, Borisjuk L, Wobus U. Sugar import and metabolism during seed development. Trends Plant Sci. 1997;2:169–74.
73. Ruan Y-L. Signaling role of sucrose metabolism in development. Mol Plant. 2012;5:763–5.

74. Ruhlmann JM, Kram BW, Carter CJ. CELL WALL INVERTASE 4 is required for nectar production in Arabidopsis. J Exp Bot. 2010;61:395–404.
75. Liao S, Wang L, Li J, Ruan Y-L. Cell wall invertase is essential for ovule development through sugar signaling rather than provision of carbon nutrients. Plant Physiol. 2020;183:1126–44.
76. Caffall KH, Mohnen D. The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. Carbohydr Res. 2009;344:1879–900.
77. Wang J, Feng J, Jia W, Chang S, Li S, Li Y. Lignin engineering through laccase modification: a promising field for energy plant improvement. Biotechnol Biofuels. 2015;8:145.
78. Berthet S, Demont-Caulet N, Pollet B, Bidzinski P, Cézard L, Le Bris P, et al. Disruption of LACCASE 4 and 17 results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. Plant Cell. 2011;23:1124–37.
79. Radauer C, Lackner P, Breiteneder H. The bet v 1 Fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands. BMC Evol Biol. 2008;8:286.
80. Wang Y, Yang L, Chen X, Ye T, Zhong B, Liu R, et al. *Major latex protein-like protein 43* (*MLP43*) functions as a positive regulator during abscisic acid responses and confers drought tolerance in *Arabidopsis thaliana*. J Exp Bot. 2016;67:421–34.
81. Zuniga-Mayo VM, Duran-Medina Y, Marsch-Martínez N, De Folter S. Hormones and flower development in Arabidopsis. In: Riechmann JL, Ferrándiz C, editors. Flower Development: methods and protocols. New York, NY: Springer US; 2023. pp. 111–29.
82. Fraser CM, Thompson MG, Shirley AM, Ralph J, Schoenherr JA, Sinlapadech T, et al. Related Arabidopsis serine carboxypeptidase-like sinapoylglucose acyltransferases display distinct but overlapping substrate specificities. Plant Physiol. 2007;144:1986–99.
83. Edgar R, Domrachev M, Lash AE. Gene expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30:207–10.

Publisher's Note