BMC Plant Biology

RESEARCH

Open Access

Transcriptomic analysis reveals molecular phenological changes during the flower-to-fruit transition in *Vanilla planifolia* Andrews (Orchidaceae)

Olga Andrea Hernández-Miranda[1,2], Jorge E. Campos[1], Estela Sandoval-Zapotitla[3], Ulises Rosas[3], María Teresa Ortiz-Melo[1] and Victor Manuel Salazar-Rojas[1*]

Abstract

Background The transition from flower to fruit, encompassing flower formation to fruit maturation, has been extensively studied in model plants such as *Arabidopsis thaliana*. However, the Orchidaceae family, including *Vanilla planifolia*, exhibits a unique phenomenon known as post-pollination syndrome (PPS), where pollination initiates ovule development but often leads to premature ovary drop. This phenomenon significantly impacts the yield and stability of *V. planifolia* crops. Understanding the molecular mechanisms underlying PPS is essential for improving crop production. This study explores transcriptomic and histological variations to identify key molecular and phenological changes in the ovary during the flower-to-fruit transition in *V. planifolia*.

Results The flower-to-fruit transition in *Vanilla planifolia* involves dynamic changes in gene expression and phenotypic events, which can be categorized into four distinct stages: (1) Pre-pollination: Ovary differentiation is characterized by the enrichment of nitrogen metabolism and photoperiod-responsive pathways. The upregulation of *VpVRN5-like* and *VpNAC14-like* suggests their roles in photoperiod-induced flowering and ovarian tissue differentiation in response to nitrate availability. (2) Pollination: Key events include nucellar filament branching and the functional enrichment of pathways associated with growth and responses to light intensity. The upregulation of *VpMBS1-like* indicates its involvement in regulating and adapting to high light conditions. (3) Post-pollination: This stage is marked by embryo sac formation and pollen tube elongation, with enrichment in auxin response pathways. The upregulation of *VpIAA6-like* and *VpRALF27-like* suggests their roles in auxin signaling during ovule development. (4) Fertilization: Seed development is associated with the enrichment of abiotic stress response pathways and carbohydrate transport. The upregulation of *VpAAE3-like*, *VpPR1-like*, and *VpSWET12-like* suggests functions in stress responses and sucrose transport, potentially linked to fungal interactions or symbiosis.

Conclusions This study characterizes the molecular and phenological changes occurring during the flower-to-fruit transition in *V. planifolia* by integrating transcriptomic analysis with anatomical data on post-pollination syndrome.

*Correspondence:
Victor Manuel Salazar-Rojas
adnbic@gmail.com

Full list of author information is available at the end of the article

Based on functional predictions, this approach provides valuable insights into the mechanisms governing this transition in plants exhibiting PPS and identifies candidate genes for future experimental validation in *V. planifolia*.

Clinical trial number Not applicable.

Keywords Molecular phenology, Orchids, Post-pollination syndrome, Transcriptome, Vanilla

Background

The flower-to-fruit transition (FFT) is a pivotal process in angiosperm evolution, linking flower development to fruit formation [1]. In model plants like *Arabidopsis thaliana* (L.) Heynh (Brassicaceae) and *Solanum lycopersicum* L. (Solanaceae), FFT is marked by two key events: (1) Pre-pollination, which establishes the final stage of flower development, and (2) Pollination-fertilization, which involves the transfer of pollen to the stigma, the growth of pollen tubes through the gynoecium, and ultimately the union of gametes to form the embryo [1, 2]. While these events are well-understood in model organisms, the dynamics of FFT can differ substantially in other species [3].

In the family Orchidaceae, for example, FFT is characterized by a distinct variant known as the post-pollination syndrome (PPS). This syndrome involves significant structural and physiological changes in the perigonium and gynostemium-ovary, and it is marked by a delay between pollination and fertilization. Notably, the female gametophytes are absent or incipient before pollen contact activates hormonal mechanisms that drive ovule development and pollen tube elongation [3]. In orchids like *Vanilla planifolia* Andrews (Orchidaceae), this extended FFT period makes the plant more susceptible to environmental stresses, which can disrupt the transition and lead to premature ovary drop, significantly affecting yield [3–5].

Despite the importance of PPS in orchid cultivation, particularly for ornamental purposes, most research has focused on post-harvest flower life and pigmentation changes, with little attention given to the molecular changes during the flower-to-fruit transition itself. This gap in knowledge presents an opportunity for a more integrative approach to studying FFT in orchids and other non-model plants [3].

One promising method to explore these complex processes is through genome-wide transcriptomic analyses. Differential transcriptomics offers insights into the developmental stages of plants by examining gene expression patterns across tissues and experimental conditions [6]. This approach, often referred to as "molecular phenology," has been used extensively in model organisms to link gene expression profiles to developmental stages [7]. For example, in *Solanum pimpinellifolium* L. (Solanaceae), spatiotemporal variations in gene expression related to cell division, photosynthesis, and auxin transport were observed during FFT [8]. Similarly, in *Pyrus bretschneideri* Rehd (Rosaceae), changes in mesocarp cell growth were linked to the up-regulation of genes related to carbohydrate metabolism and auxin transport during Pollination stage [9].

Transcriptomic analysis not only uncovers gene expression dynamics but also reveals how environmental signals influence the FFT process [7, 10, 11]. This approach could be particularly valuable for studying non-model plants like *V. planifolia*, where little is known about the molecular mechanisms driving FFT, especially in relation to PPS [3, 4, 12]. While histological studies have provided some understanding of ovarian development in *V. planifolia* [13–15], the functional aspects, such as the role of auxin metabolism-related genes in fruit development, remain poorly understood [4].

This study aims to bridge this knowledge gap by integrating transcriptomic and histological analyses to uncover the molecular phenology of FFT in *V. planifolia* and other orchids with PPS. Our findings may offer a deeper understanding of FFT in plants with PPS and facilitate the identification of candidate genes for functional validation of this process in *V. planifolia.*

Materials and methods

Biological material

The biological material was collected in the Totonacapan region, in the municipality of San Rafael, Veracruz State, Mexico [16]. Plants were sampled from reproductive individuals older than five years, grown in crop fields under an intensive shade mesh system. Environmental conditions in these fields included a temperature of 37 °C, relative humidity of 70%, light intensity of 92.59 μmol/m²/s, and a photoperiod of 14 h of light and 10 h of darkness. Each sample was collected twice, with no visible signs of pathogens, from the middle region of the plant [17].

Flowers and ovaries were collected at four developmental stages of *V. planifolia*: (1) Pre-pollination (Pre-pol: floral bud before anthesis), (2) Pollination (Pol: flower in anthesis with pollen on the stigma), (3) Post-pollination (Post-pol: ovary 25 days after pollination), and (4) Fertilization (Fer: ovary with fertilized ovules, 60 days after pollination) [4]. An herbarium specimen was deposited in the ethnobotanical collection at the *Herbario IZTA, Facultad de Estudios Superiores Iztacala, UNAM* (Boucher number: 3559-IZTA).

Anatomic analysis

The gynostemium was subdivided into apical, medial, and basal regions. Samples were taken from the medial region of the ovary. These samples were fixed in Formaldehyde Alcohol Acetic Acid (10%:50%:5% + 35% water solution, F.A.A.) for 24 h, then washed with water to remove the fixative solution. Prior to dehydration, the samples were softened with 10% ethylenediamine for 24 h. Dehydration was carried out in a gradual series of tert-butyl alcohol solutions (30%, 50%, 60%, 75%, 80%, 90%, 100%). Infiltration and embedding were performed using Paraplast (Sigma-Aldrich). Longitudinal sections of the gynostemium and transverse sections of the ovary were obtained at 5 μm thickness using a rotary microtome [18]. The sections were stained with fast green and safranin. Finally, the samples were mounted using synthetic resin. The resulting preparations were deposited in the collection of the *Laboratorio de Apoyo a la Investigación del Jardín Botánico*, *Instituto de Biología*, *Universidad Nacional Autónoma de México*, and subsequently observed and analyzed using a photomicroscope (Carl Zeiss Axioskop) with bright field, phase contrast, and polarization modes. We analyzed the dermal, fundamental, and vascular tissues of the gynostemium and ovary/fruit tissues. Images were captured using a digital camera (Rising View V x64.4.10.17214.2020).

RNA isolation

The collected tissue was immediately frozen in liquid nitrogen on-site and stored at -70 °C until further processing. While still frozen, the gynostemium was dissected and pulverized in liquid nitrogen to facilitate total RNA isolation using the *Direct-zol RNA Miniprep Plus kit*, following the manufacturers instructions (ZYMO RESEARCH). The integrity of the total RNA was assessed by 1% agarose gel electrophoresis, and the average total RNA concentration was 569 ng/μL, determined using the *NanoDrop One* (Thermo Fisher Scientific, Carlsbad, CA). Total RNA integrity and purity were further evaluated with a *Bioanalyzer 2100* (Agilent Technologies, Santa Clara, CA), and RNA samples with RIN values > 6 were selected for subsequent cDNA synthesis, library construction, and sequencing.

cDNA library synthesis and next-generation sequencing

cDNA libraries were constructed from total RNA using *TruSeq v2 RNA Sample Preparation Kits* (Illumina, Inc., San Diego, CA). A total of eight libraries, representing four developmental stages with two consistent biological replicates (assessed by Pearson correlation and dispersion analysis), were generated, indexed, and pooled. Sequencing was performed at the *Unidad Universitaria de Secuenciación Masiva y Bioinformática* of the *Instituto de Biotecnología*, *Universidad Nacional Autónoma de México* (UUSMB IBT-UNAM) using the Illumina *NextSeq 500* system, generating 76-bp paired-end reads.

De Novo assembly and functional annotation

We performed quality analysis of the raw data using *FastQC version 0.11.9*. The Phred quality score for each nucleotide was calculated, along with sequence length, GC content, duplicated sequences, and adaptors [19]. Based on this information, we trimmed and filtered the data using *Trimmomatic version 0.39* to eliminate adaptors, with a maximum error rate of 2 bp every 30 bases and a minimum Phred score of 15. Sequences with a minimum length of 30 bp were retained [20]. To generate contigs, a de novo assembly of the *V. planifolia* transcriptome was performed using *Trinity version 2.4*. We assessed assembly quality using metrics such as total genes, GC content, median contig length, average contig length, total assembled bases, and the ExN50 statistic, ultimately calculating the E95N50 value [21].

Subsequently, we performed an analysis using *BUSCO v4* software to evaluate the integrity of conserved orthologs, utilizing the *Liliopsida* odb10* database with predefined parameters [22]. Amino acid sequences were identified by searching for open reading frames with *TransDecoder* software [23]. Transcript annotations were carried out using *Trinotate software*, employing three annotation strategies with their respective databases [24]: (1) Sequence homology search via *BLASTp* and *BLASTx* against the *UniProt* database [25]; (2) Protein domain identification using the Markov chain model and the *PFAM* domain database [26]; and (3) GO-term (Gene Ontology) assignment through the *Pannzer* annotation database [27] (pipelines available at https://github.com/Andrea-H-M/Vanilla_TFF/tree/main).

Differential expression analysis

A pseudoalignment-based mapping of reads to the assembled transcriptome was performed using *Salmon v0.8.0* and *Trinity 2.4* to quantify transcript expression in TPM (transcripts per million) [28]. Differential expression analysis was conducted by comparing pairs of developmental stages (Post-pol vs. Fer, Pol vs. Fer, Pol vs. Post-pol, Pre-pol vs. Fer, Pre-pol vs. Post-pol, and Pre-pol vs. Pol) using the *DESeq2* package in R 4.2.3 [29]. Significance in differential expression was determined by a Log_2 fold change ≥ 2 and a p-value ≤ 0.00001 (pipelines available at https://github.com/Andrea-H-M/Vanilla_TFF/tree/main).

Heatmap cluster analysis

To observe the global profile of differentially expressed genes (DEGs) during FFT, we used the normalized TPM (Log_2) to perform a hierarchical cluster analysis heatmap using bidirectional clustering, generating a dendrogram

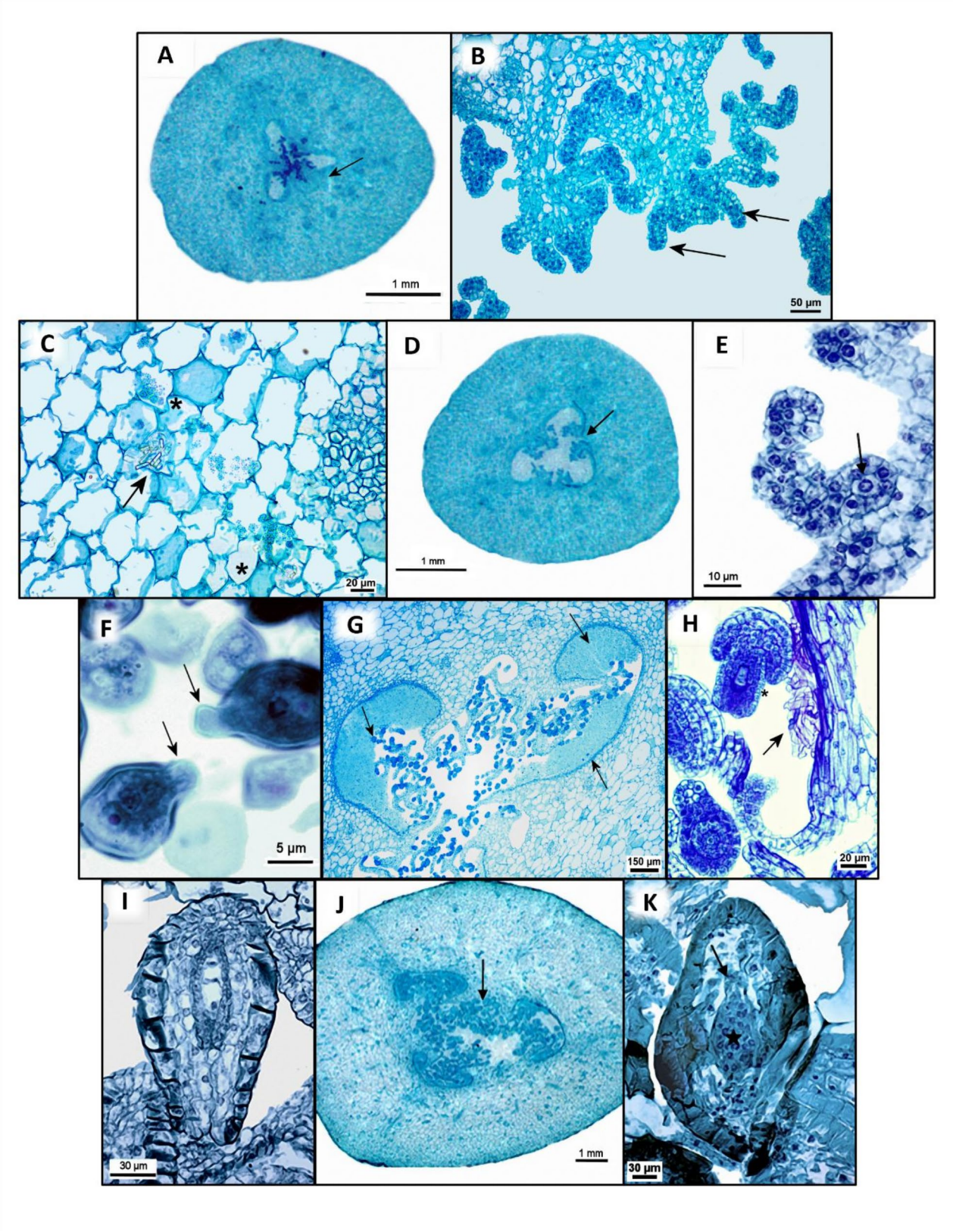

Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Tissue-level flower-to-fruit transition in *Vanilla planifolia*. (**A**) Transverse section of the ovary, central cavity with nucellar filaments (arrow). (**B**) Histological section of bifurcated nucellar filaments (arrows). (**C**) Ground tissue with raphides (arrow) and starch grains (asterisks). (**D**) Transverse section of the ovary enlarged central cavity. (**E**) Nucellar filament with archesporial cell (arrow). (**F**) Pollen grain germination and emerging pollen tube (arrows). (**G**) Transverse section of the ovary, central cavity showing seed initiation wrapped in transmission tissue (arrows). (**H**) Pollen tubes (arrow) near ovules (star), visible at higher magnification. (**I**) Embryo sac (arrow). (**J**) Transverse section of the ovary, central cavity with seeds (arrow). (**K**) Embryo (star) with suspensor (arrow)

based on Euclidian distances, according to the Ward method in the *SAS-JMP* software. Additionally, we used a distance graph and the cubical clustering criteria to estimate the number of groups through the Ward minimum variance method, evaluated through Monte Carlo methods using the *SAS-JMP* software [30].

In Silico identification of functional signatures

To identify functional signatures strongly associated with specific developmental stages, a paired t-test (Students t-test for dependent samples) was performed in R 4.2.3 on the DEGs identified in this study. Genes with significantly higher expression at a specific stage compared to others ($p \leq 0.00001$) were selected for further analysis (pipelines available at https://github.com/Andrea-H-M/Vanilla_TFF/tree/main) [29].

Functional enrichment analysis

Functional enrichment analysis was conducted on the identified DEGs and genes associated with functional signatures at each developmental stage. To achieve this, we performed a *BLASTx* search against the *TAIR 2.9.0* database to identify orthologs in *Arabidopsis thaliana*. The resulting ortholog identifiers were then analyzed using the Planteome 1.1 platform for functional enrichment, employing Chi-square statistics and Gene Ontology (GO) categories to infer biological functions [31]. Data with $p \leq 0.05$ were visualized as bubble plots in R 4.2.3 and further analyzed using *Revigo 1.8.1* to generate enrichment networks, which were visualized in *Cytoscape 3.10.1* [32].

Results

The molecular phenology of the flower-to-fruit transition (TFF) in *Vanilla planifolia* was characterized by integrating key anatomical and transcriptional changes observed across four developmental stages. In plants with post-pollination syndrome (PPS), the absence of distinct macroscopic changes between stages highlights the need for an integrative developmental framework to better understand this process [7].

Anatomical description of the Flower-to-Fruit transition in *Vanilla planifolia*

Significant anatomical changes were observed throughout the development of *V. planifolia*, allowing the classification of four distinct developmental stages:

Stage 1: Pre-Pollination

The ovary is tricarpellary and unilocular (Fig. 1A). Both the outer and inner epidermis are monostratified, with stomata present in the outer epidermis. The inner epidermis delineates the ground tissue and central cavity of the ovary, consisting of a single layer of round, isodiametric cells with thin walls. The ground tissue is composed of parenchyma cells of varying sizes, some of which contain starch grains and raphides (Fig. 1C). The vascular tissue comprises five to six collateral vascular bundles per carpel, arranged in a circular pattern (Fig. 1A). At this stage, transmission tissue begins to develop in the inner epidermis, and placental ridges form, giving rise to placental projections composed of a central row of cells surrounded by an epidermal layer, which is visible in longitudinal sections (Fig. 1B).

Stage 2: pollination

Placental projections branch at this stage (Fig. 1E), with each branch containing a central row of cells surrounded by a single-layered epidermis. The distal cells of the central row have dense, enlarged nuclei compared to the branch epidermal cells (Fig. 1E), corresponding to archesporial cells or megaspore mother cells. Pollen grains are located outside the anther theca and on the stigma, with some observed germinating (Fig. 1F).

Stage 3: Post-Pollination

The funiculus curves, forming an anatropous ovule. The inner integument begins to develop in some projections, consisting of one or two cell layers, while the outer integument comprises two layers, classifying the ovule as bitegmic. Transmission tissue proliferation covers the inner epidermis within the central cavity (Fig. 1G), where pollen tubes are also present (Fig. 1H). In some sections, pollen tubes appear faint but can be distinguished within the transmission tissue (Fig. 1H). A fertilized embryonic sac containing the zygote and two endosperm nuclei is observed in longitudinal sections (Fig. 1I).

Stage 4: fertilization

At this stage, transmission tissue is reduced in the central cavity, where seeds and trichomes are abundant (Fig. 1J). The trichomes contain dense, globular structures. The inner integument surrounds the pro-embryo but remains separated from it. At the micropylar end, the suspensor is visible, while the reduced nucellus is observed at the

chalazal end (Fig. 1K). Although nucellar reduction is not always uniform across sections, different stages can be inferred based on cellular organization and tissue differentiation.

A robust and well annotated transcriptome was generated

Eight libraries, representing four developmental stages, were sequenced with two consistent biological replicates, and the data were assessed using dispersion analysis and Pearson correlation (Fig. S1 and Table S1). This sequencing yielded a total of 147,343,104 reads (Table 1). After trimming and filtering, 66,009,079 paired reads and 332,927 unpaired reads were obtained (Table 1). The paired reads were used for de novo assembly, resulting in 47,843 genes and 97,782 contigs, with an average contig length of 417 bp (Table 1). To assess assembly quality and replicate consistency, we calculated the ExN50 statistic and plotted the expressed percentage (Ex) against the ExN50 contig length. The assembly saturation point was reached at 95% of the total expression, with an ExN50 of 1,555 bp and a total of 14,537 contigs (Table 1; Fig. 2A).

We mapped contigs against the *Liliopsida odb10** database, identifying 85% of the contigs as complete orthologs, 8% as fragmented, and 7% as missing sequences (Fig. 2B). Additionally, we searched for open reading frames and performed functional annotation of the amino acid sequences using the *UniProt* (40,233 genes), *PFAM* (46,231 genes), and *Pannzer* (31,804 genes) databases for Gene Ontology (GO)-term assignments (Fig. 2C). We identified 30,753 genes associated with biological processes, 31,804 with molecular functions, and 31,921 with cellular components (Fig. S2-S5). Most of the assembled genes were found to be conserved orthologs, as annotated by the selected databases.

Table 1 De *Novo* assembly of FFT in *Vanilla planifolia*

Sequencing and processing		Transcriptome assembly statistics		Contig statistics	
Total number of sequenced reads	147,343,104	Total genes	47,843	Median contig length (bp)	417
Paired reads	66,009,079	Total transcripts	97,782	Average contig length (bp)	838.94
Non-paired reads	332,927	GC percentage (%)	43.91	Total assembled bases	40,137,513
		E95N50 (bp)	1,555		
		Contigs E95	14,537		

Differential gene expression shows patterns of activity during development

A total of 1,090 differentially expressed genes (DEGs) were identified across six pairwise comparisons of the four developmental stages analyzed: (1) Pre-pol vs. Pol, (2) Pre-pol vs. Post-pol, (3) Pre-pol vs. Fer, (4) Pol vs. Post-pol, (5) Pol vs. Fer, and (6) Post-pol vs. Fer (Fig. 3, Fig. S6, Tables S2–S8). Differential expression analysis was performed using a threshold of Log_2 Fold Change ≥ 2, p-value ≤ 0.00001, with a false discovery rate (FDR) correction for multiple comparisons. The total of 1,090 DEGs represents the unique number of differentially expressed genes, considering their presence across multiple comparisons but counted only once.

Expression profiles during the flower-to-fruit transition are associated with post-pollination syndrome

To assess expression patterns, hierarchical clustering and functional enrichment analyses were conducted, identifying four major clusters of genes associated with specific functions in ovary development (Fig. 4, Fig. S7). These patterns reflect the differential regulation of key processes at each stage, highlighting the activation of genes involved in biosynthesis, cellular organization, primary metabolism, and responses to biotic and abiotic stress.

To characterize expression patterns associated with the flower-to-fruit transition (FFT), we analyzed 1,090 differentially expressed genes (DEGs) across developmental stage comparisons using hierarchical clustering and functional enrichment analysis. The hierarchical clustering analysis and heatmap visualization revealed four distinct clusters, each enriched with gene ontology (GO) terms related to various biological processes (Fig. 4 and Fig. S7).

Cluster 1 (C1) Genes involved in biosynthesis and membrane-cell wall organization, including processes related to macromolecule metabolism, cell wall formation, wounding response, membrane organization, supramolecular fiber organization, and phospholipid biosynthesis (Table S9).

Cluster 2 (C2) Genes involved in stress responses and primary metabolism, such as hydric stress, fungal infection, and positive regulation of cellular and developmental processes (Table S10).

Cluster 3 (C3) Genes associated with cellular structure and tissue function, including radiation response, tissue development, cytokinesis, and reproductive processes (Table S11).

Cluster 4 (C4) Genes related to macronutrient metabolism and morphogenesis, particularly growth and development, lipid metabolism, nitrogen compound biosynthesis,

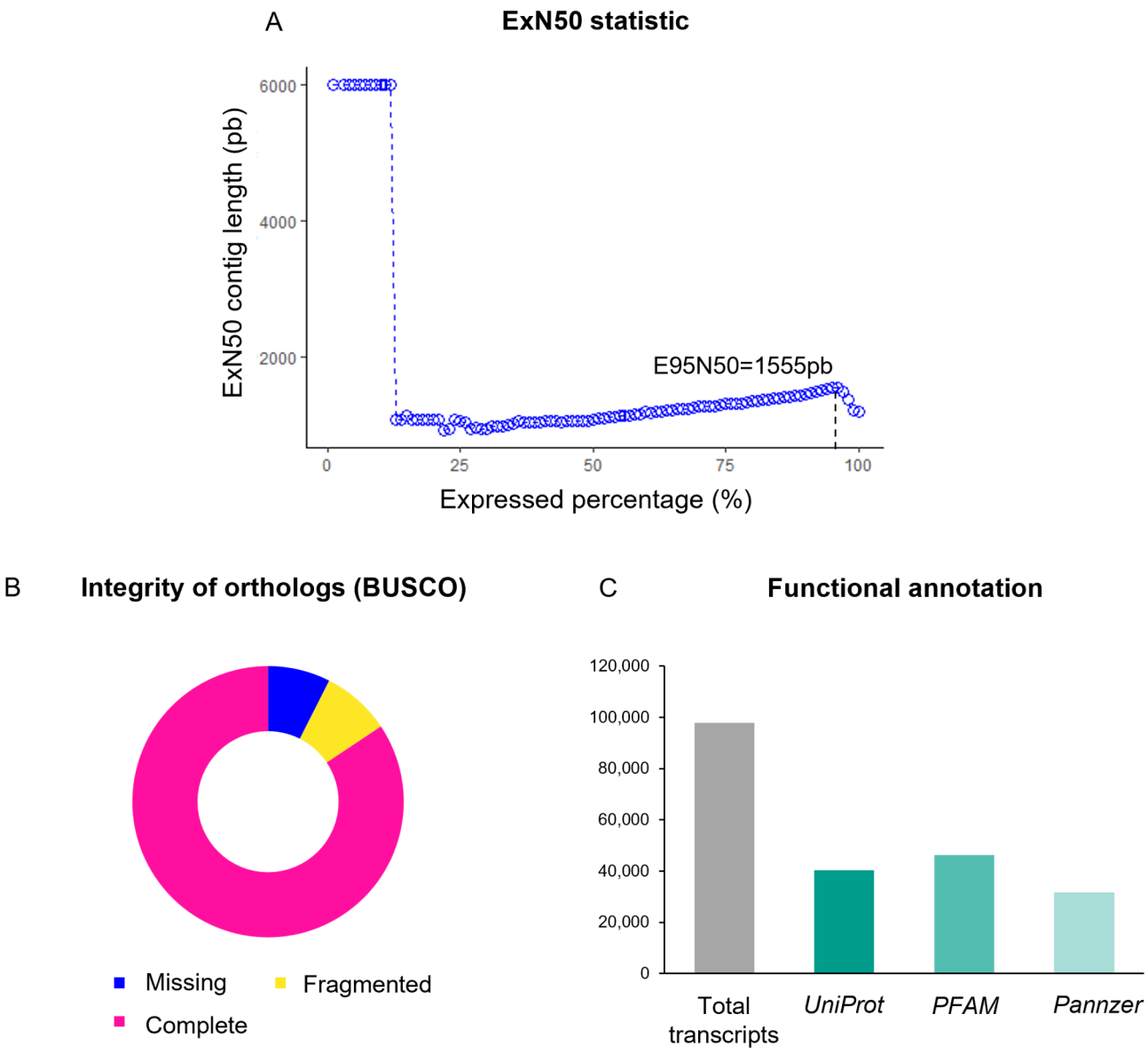

Fig. 2 Metrics of the assembly and functional annotation of the transcriptome. (**A**) Expression percentage vs. ExN50 *contig* length. (**B**) The integrity of conserved orthologs mapped to the *Liliopsida odb10** database. (**C**) *Contigs* with functional annotation in databases *UniProt*, *PFAM*, *and Pannzer*

phosphate regulation, and carbohydrate responses (Table S12).

We also identified two distinct transcriptional profiles during FFT in *V. planifolia*: **Profile 1 (P1)**: Upregulated patterns in clusters C1, C3, and most C2 genes during the Pre-pollination and Pollination stages. **Profile 2 (P2)**: Upregulated patterns in clusters C4, C2, and some C1 genes during the Post-pollination stage. Additionally, an induction pattern was observed in clusters C4, C3, and some C1 genes during the Fertilization stage. These expression profiles reflect biological functions typical of post-pollination syndrome (PPS) and highlight genes induced at each developmental stage.

These expression profiles reflect biological functions characteristic of post-pollination syndrome (PPS) and highlight stage-specific gene induction. Overall, upregulated DEGs were more prevalent than downregulated DEGs, suggesting differential recruitment of transcriptional machinery at each stage. This pattern underscores the unique molecular signature of the flower-to-fruit transition in *V. planifolia*.

In Silico identification of functional signatures in the molecular phenology of the flower-to-fruit transition

To identify functional signatures most strongly associated with specific developmental stages, a paired t-test (Students t-test for dependent samples) was applied

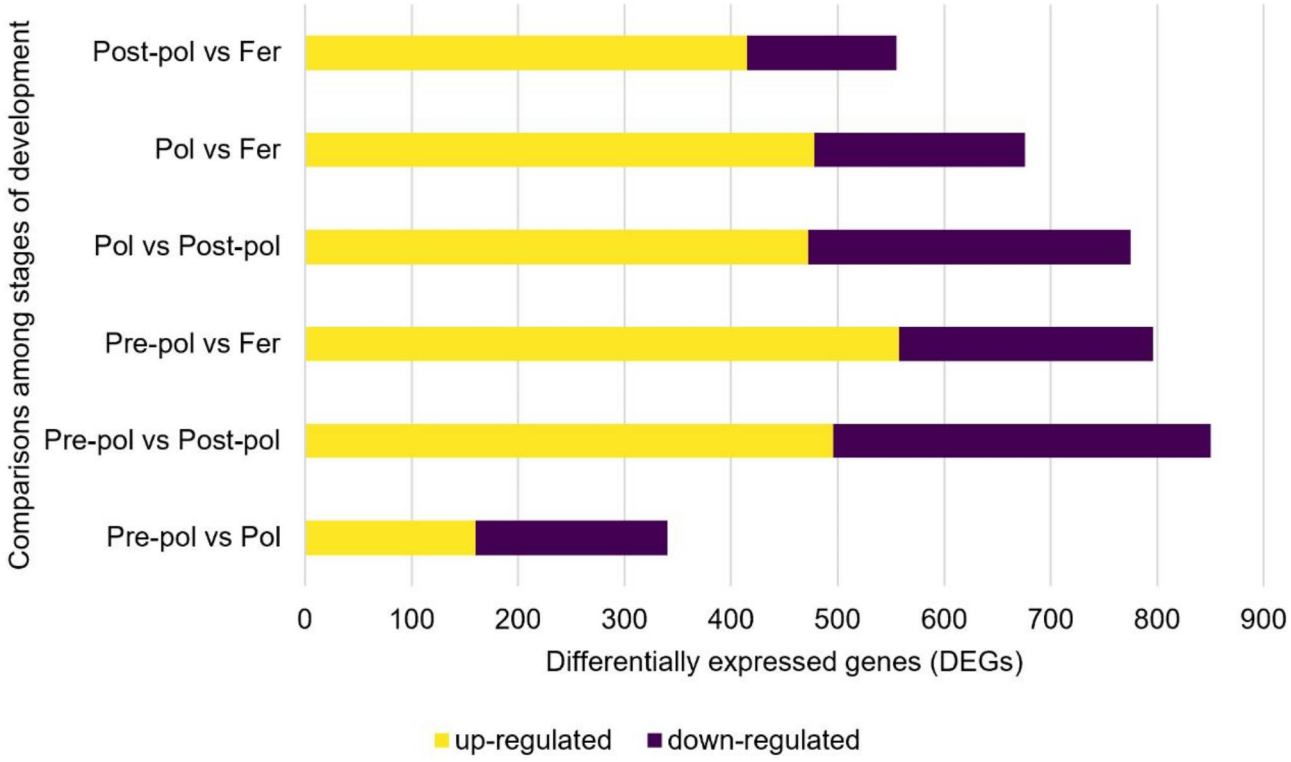

Fig. 3 Differentially expressed genes (DEGs) between six pairwise comparisons of developmental stages. Up-regulated genes are shown in yellow, and down-regulated genes are shown in purple. Pre-pol: Pre-pollination stage, Pol: Pollination stage, Post-pol: Post-pollination stage and Fer: Fertilization stage

to differentially expressed genes (DEGs). This analysis identified genes with significantly higher expression ($p \leq 0.00001$) at a specific stage compared to others, resulting in the identification of 30 genes for Pre-pol, 13 for Pol, 41 for Post-pol, and 43 for Fer (Tables S13–S16). Functional enrichment analysis was then performed on these functional signatures.

Pre-pollination stage

The functional signature of Pre-pollination is characterized by significant overall functional enrichment ($p \leq 0.05$) in processes related to mRNA metabolism, regulation of nitrogenous compound metabolism, and carbohydrate biosynthesis (Fig. 5A).

To explore the relationships among enriched functional categories, a network analysis was conducted, revealing two global networks. Network 1 connects the GO term "short-day photoperiod response" with categories related to the development of carpels, meristems, and reproductive shoots, as well as responses to biotic (viruses, other organisms) and abiotic (heat, vernalization, salinity, and stress) stimuli (Fig. 5B). In contrast, Network 2 is primarily associated with the regulation of defense responses to bacteria, floral development, viral processes, and gene expression via genomic imprinting. It also encompasses the regulation of nitrogenous compound metabolism, protein modification and organization, carbohydrate metabolic processes, and mRNA-related functions (Fig. 5C and Table S17).

Based on the major functional categories identified in the functional enrichment networks (Fig. 5), along with information from the *Planteome*, *UniProt*, and *TAIR* databases [25, 31, 33] and a review of functional studies in the literature, two genes were identified as Pre-pollination stage-specific functional signatures: (1) *VpNAC14-like* (*Vanilla planifolia NAC014-like*), which is associated with nitrogen metabolism and abiotic stress tolerance [34, 35], and (2) *VpVNR5-like* (*Vanilla planifolia VERNALIZATION 5-like*), which plays a role in flowering induction under short-day conditions [36]. These genes showed significantly higher differential expression during the Pre-pollination stage, as confirmed by heatmap analysis (Fig. 5D).

Pollination stage

The functional signature of Pollination stage is characterized by significant overall functional enrichment ($p \leq 0.05$) in processes related to growth, biosynthesis, and responses to acid stress (Fig. 6A and Table S18). To explore the relationships among enriched functional categories, a network analysis was performed, revealing two global networks. The first network is associated with biotic stress (defense against viruses and symbionts) and abiotic stress responses (light intensity, salt stress,

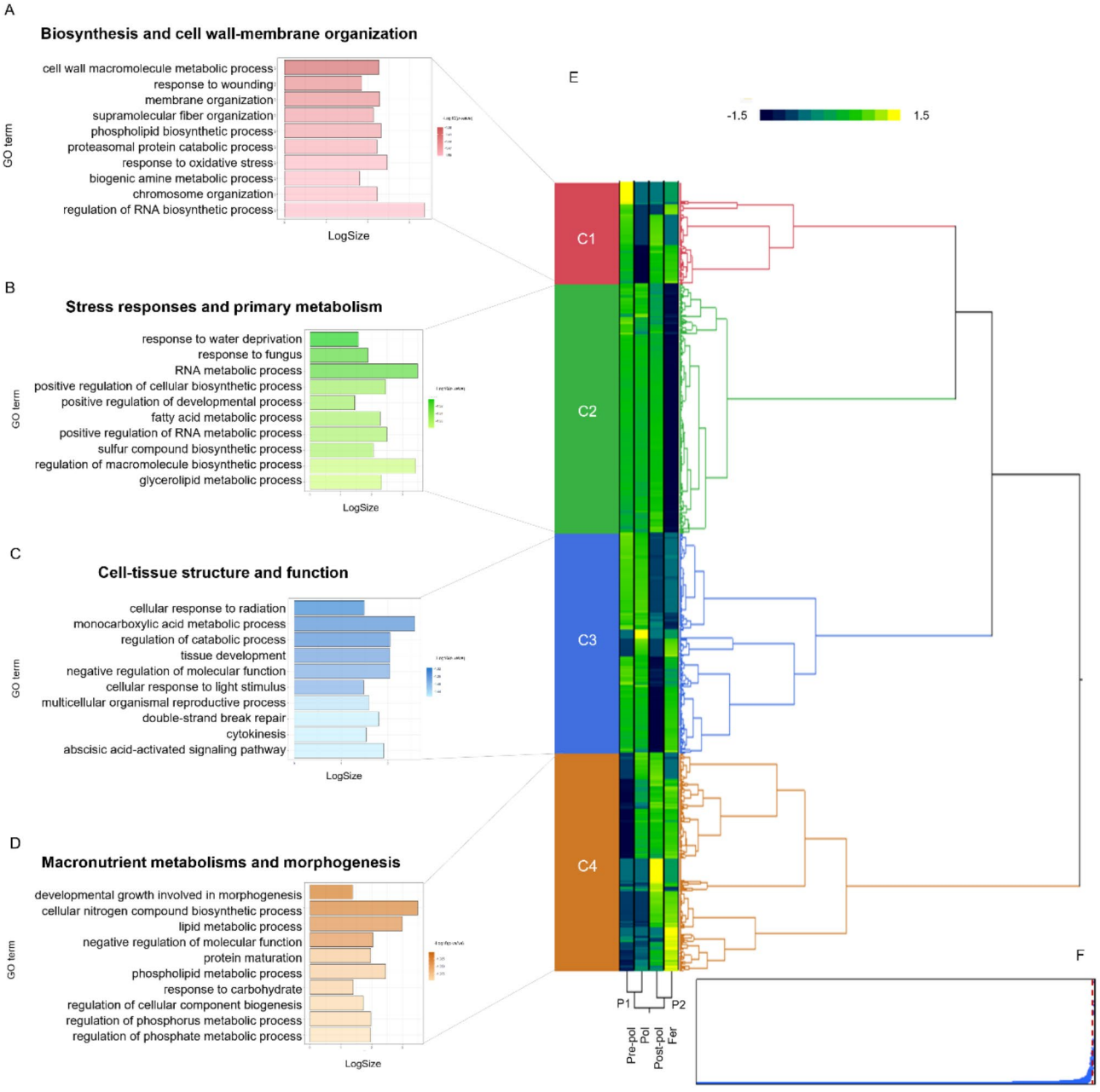

Fig. 4 Hierarchical clustering analysis and functional enrichment of the 1,090 DEGs during FFT. (**A**) Functional enrichment of cluster 1 (C1). (**B**) Cluster 2 (C2). (**C**) Cluster 3 (C3). (**D**) Cluster 4 (C4). The color intensity of the graphs represents the enrichment -Log_{10}*p*-value. (**E**) Profile 1 (P1), Profile 2 (P2). Clustering analysis of DEGs during FFT. (**F**) Distance plot of the cluster number identified. Pre-pol: Pre-pollination stage, Pol: Pollination stage., Post-pol: Post-pollination stage and Fer: Fertilization stage

oxidative stress, and temperature adaptation) (Fig. 6B). The second network involves the epigenetic regulation of gene expression and includes two key processes: (1) regulation of development, biomolecule metabolism, cellular communication, and stress responses, and (2) regulation of cellular component biogenesis, protein-DNA complex organization, and morphogenesis, including embryonic and organ development (Fig. 6C).

Based on the major functional categories identified in the functional enrichment networks (Fig. 6) and information from the *Planteome*, *UniProt* and *TAIR* databases [25, 31, 33], as well as a review of functional studies in the literature, *VpMBS1-like* (*Vanilla planifolia METHYLENE BLUE SENSITIVITY 1-like*), which is associated with plant protection against light stress [37], was identified as a Pollination stage-specific functional signature. This gene showed significantly higher differential expression

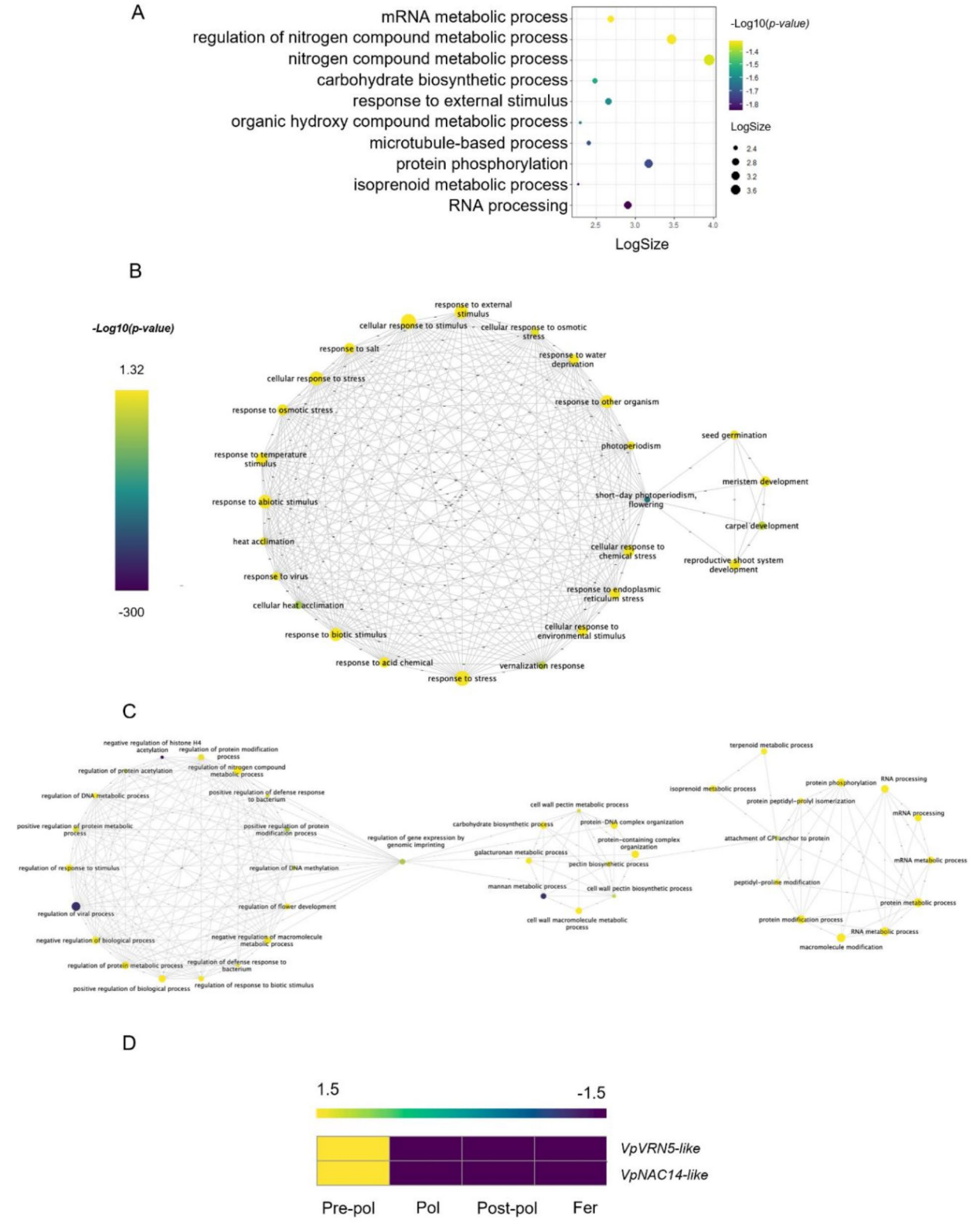

Fig. 5 Functional signatures in *Vanilla planifolia* during the Pre-pollination stage. (**A**) Functional enrichment dispersion plot displaying the ten most enriched GO terms ($p \leq 0.05$), with the vertical axis showing GO terms and the horizontal axis representing the Log_{10}-transformed number of annotations. (**B**-**C**) Functional enrichment maps, where purple indicates low statistical significance and yellow denotes significant data. (**D**) Heatmap of differentially expressed genes (Log_2 fold change ≥ 2, $p \leq 0.00001$), highlighting upregulated genes (normalized TPM) in yellow

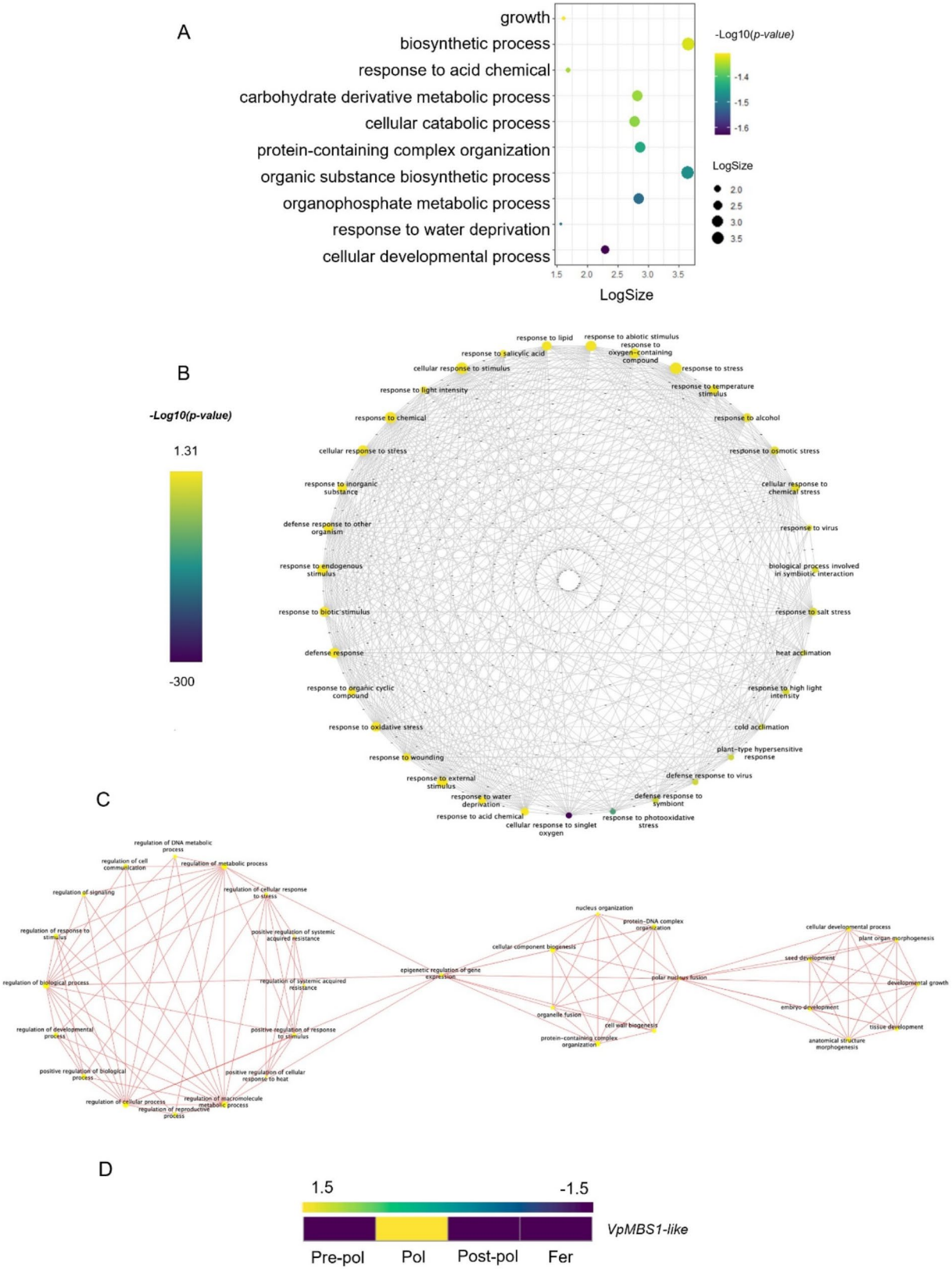

Fig. 6 Functional signatures in *Vanilla planifolia* during the Pollination stage. (**A**) Scatter plot of functional enrichment, showing the top ten enriched GO terms (p-value ≤ 0.05). The vertical axis represents the GO terms, while the horizontal axis displays the Log_{10}-transformed number of annotations. (**B**-**C**) Functional enrichment maps, where data with low statistical significance are shown in purple and statistically significant data in yellow. (**D**) Heatmap of differentially expressed genes (Log_2 fold change ≥ 2, p-value ≤ 0.00001), highlighting upregulated genes (normalized TPM) in yellow

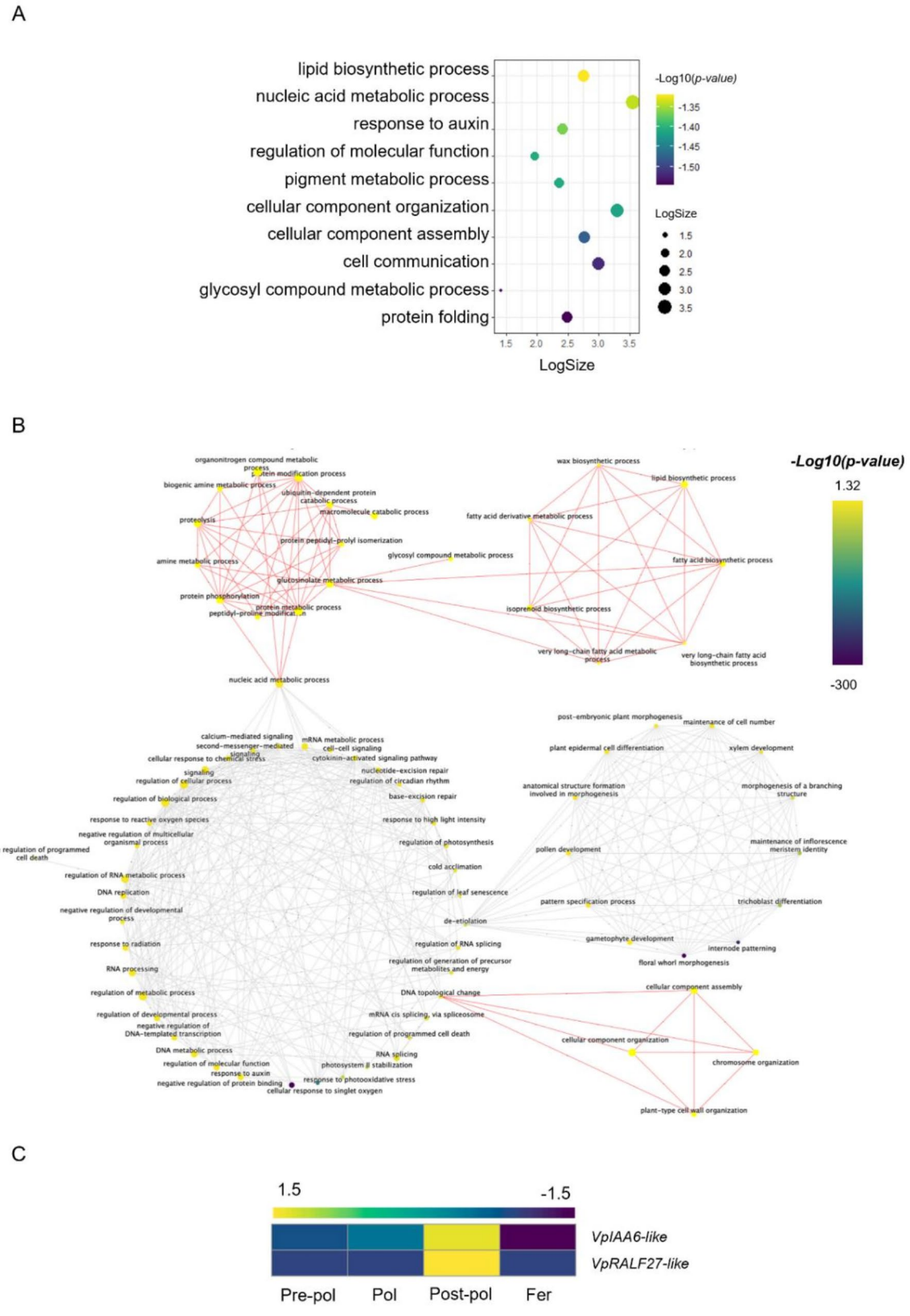

Fig. 7 (See legend on next page.)

(See figure on previous page.)

Fig. 7 Functional signatures in *Vanilla planifolia* during the Post-pollination stage. (**A**) Dispersion plot of functional enrichment (the ten enriched GO terms with a *p*-value ≤ 0.05 are displayed, with the vertical axis representing the GO terms and the horizontal axis showing the Log_{10} of the number of annotations). (**B**) Functional enrichment map (data with low statistical significance are shown in purple, and data with significant statistical significance are shown in yellow). (**C**) Heatmap of differential expression (Log_2 fold change ≥ 2 and *p*-value ≤ 0.00001) showing upregulated genes (normalized TPM) in yellow

during the pollination stage, as confirmed by heatmap analysis (Fig. 6D).

Post-pollination stage

The functional signature of Post-pollination stage is characterized by significant overall functional enrichment ($p \leq 0.05$) in processes related to lipid biosynthesis, nucleic acid metabolism, and auxin responses (Fig. 7A, Table S19). To explore the relationships among the enriched functional categories, a network analysis was performed, revealing a global network encompassing annotations related to nucleic acid metabolism, protein modification and phosphorylation, fatty acid biosynthesis, proteolysis, and nitrogenous organic compound metabolism (Fig. 7B). This network also includes processes such as RNA signaling, regulation of biological processes, cell-cell signaling, cell wall organization, chromosomal and cellular organization, and responses to auxins and cytokinins, linking it to post-embryonic morphogenesis as well as gametophyte and pollen development (Fig. 7B). Based on the major functional categories identified in the functional enrichment networks (Fig. 7), along with information from the *Planteome*, *UniProt*, *TAIR* databases [25, 31, 33] and a review of functional studies in the literature, two genes were identified as Post-pollination stage-specific functional signatures: (1) *VpRALF27-like* (*Vanilla planifolia RALFL27-like*) and (2) *VpIAA6-like* (*Vanilla planifolia INDOLE-3-ACETIC ACID 6-like*), both implicated in transcriptional regulation and auxin response [38–40]. These genes showed significantly higher differential expression during the Post-pollination stage, as confirmed by heatmap analysis (Fig. 7D).

Fertilization stage

The functional signature of fertilization is characterized by significant overall functional enrichment ($p \leq 0.05$) in processes related to the response to biotic stimuli and interspecies interactions (Fig. 8A). A network analysis was conducted to explore the relationships among the enriched functional categories, revealing two global networks. The first network is associated with formaldehyde catabolism and encompasses metabolic processes regulating proteins, small molecules, organophosphates, and phosphate compounds. It also integrates responses to endogenous stimuli, including those triggered by vitamin B1, jasmonic acid signaling, and lipid and fatty acid responses. Additionally, this network links to reactions to exogenous biotic stimuli (fungi and symbionts) and abiotic stresses (cold, salinity, nutrient availability, and nitrogen) (Fig. 8B). The second network, in contrast, is more focused on the biogenesis of oleic bodies in seeds and cell walls, plastid organization during megagametogenesis, seed sac development, germination, and seed longevity (Fig. 8C and Table S20). Based on the major functional categories identified in the networks (Fig. 8) and information from databases like *Planteome*, *UniProt* and *TAIR* databases [25, 31, 33], as well as functional studies from the literature, three genes were identified as specific functional signatures of the Post-pollination stage: (1) *VpAAE3-like* (*Vanilla planifolia ACYL-ACTIVATING ENZYME 3-like*), (2) *VpPR1-like* (*Vanilla planifolia PATHOGENESIS-RELATED GENE 1-like*), associated with stress response, and (3) *VpSWT12-like* (*Vanilla planifolia SWEET12-like*), involved in sugar transport [41–44]. Heatmap analysis confirmed significantly higher differential expression of these genes during the fertilization stage (Fig. 8D).

Discussion

The flower-to-fruit transition in *Vanilla planifolia* involves molecular phenological changes imperceptible at the macroscopic level

In model plants such as *Arabidopsis thaliana*, the FFT is characterized by functional modifications in genetic networks that regulate development, along with anatomical changes linked to pre-anthesis ovule differentiation, leading to fruit formation shortly after pollination [1]. In contrast, *V. planifolia* follows a distinct developmental pattern in which ovules do not differentiate before anthesis. Instead, pollination triggers their initiation, culminating in fruit formation approximately 45 days later [3].

Despite this major developmental shift, anatomical changes remain imperceptible at the macroscopic level, hindering the visual assessment of early fruit development [4]. However, transcriptomic analyses reveal dynamic gene expression shifts associated with ovule initiation and early fruit development, even in the absence of visible morphological modifications. This underscores the importance of a molecular phenology approach to characterize the FFT in *V. planifolia*, as it enables the identification of key genetic changes associated with specific developmental stages (Figs. 4, 5, 6, 7 and 8) and provides novel insights into the molecular regulation of this unique transition [7].

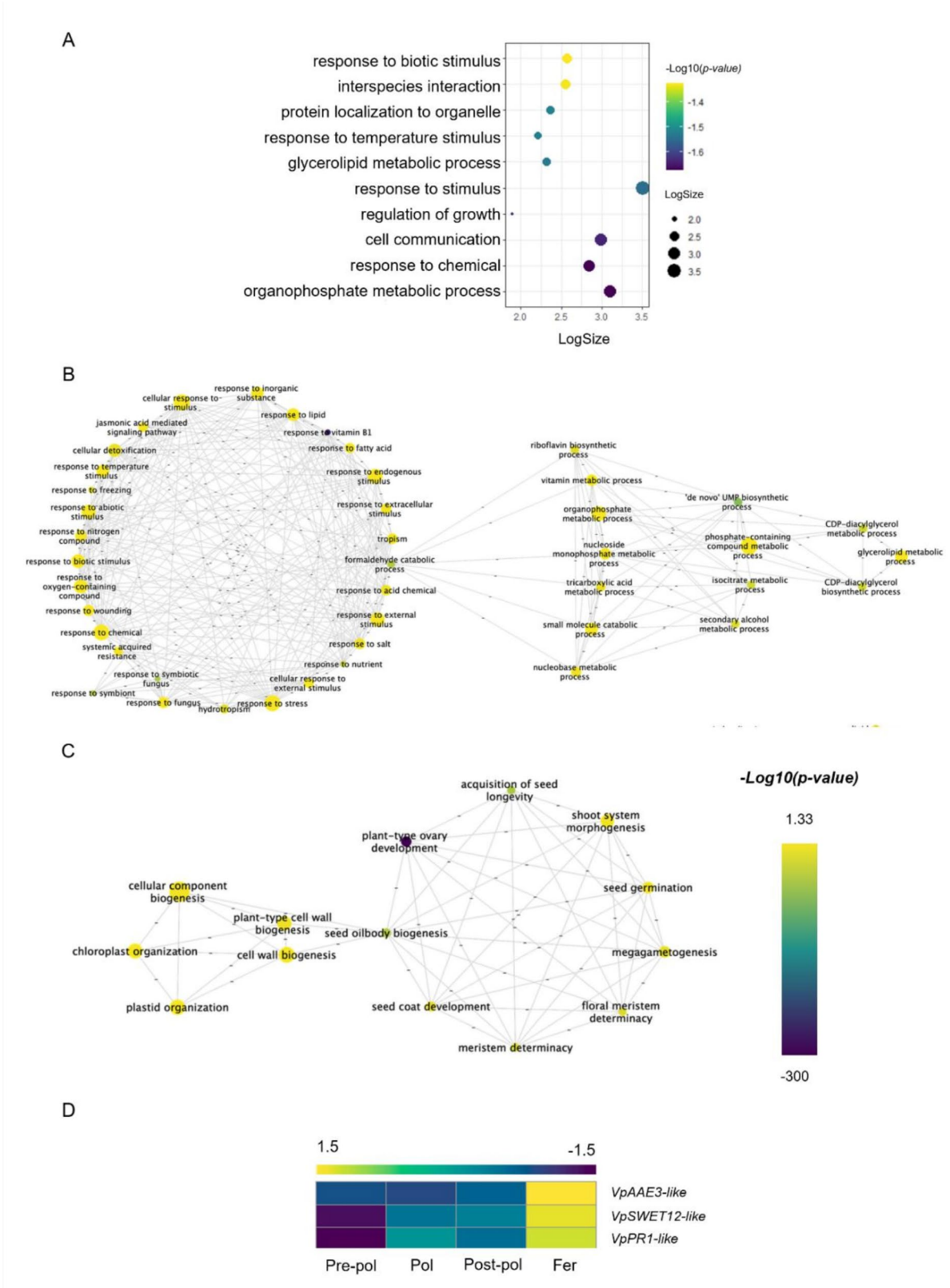

Fig. 8 (See legend on next page.)

(See figure on previous page.)

Fig. 8 Functional signatures in *Vanilla planifolia* during the Fertilization stage. (**A**) Dispersion plot of functional enrichment (The ten enriched GO terms with a *p*-value ≤ 0.05 are shown, with the vertical axis representing the GO terms and the horizontal axis displaying the Log_{10} of the number of annotations). (**B**-**C**) Functional enrichment map (Data with low statistical significance are shown in purple, while data with high statistical significance are shown in yellow). (**D**) Heatmap of differential expression (Log_2 fold change ≥ 2 and *p*-value ≤ 0.00001), highlighting up-regulated genes (normalized TPM) in yellow

The molecular phenology of flower-to-fruit transition in *Vanilla planifolia* comprises four developmental stages

To characterize the molecular phenology of the flower-to-fruit transition (FFT) in *V. planifolia*, we correlated anatomical observations with transcriptomic data, identifying four distinct developmental stages: Pre-pol, Pol, Post-pol, and Fer. Transcriptomic analysis revealed two primary gene expression profiles (P1 and P2), corresponding to the early and late phases of the transition:

P1 (Pre-pol and Pol) Enriched in genes involved in membrane and cell wall biosynthesis, as well as cellular and tissue organization (Fig. 4E).

P2 (Post-pol and Fer) Characterized by the activation of genes associated with macronutrient metabolism and morphogenesis (Fig. 4E).

This molecular classification aligns with the post-pollination syndrome observed in *V. planifolia*, wherein pollination functions as a regulatory switch that activates distinct gene expression programs before and after pollination [3]. Stage-specific functional signatures were identified at key developmental transitions (Figs. 5, 6, 7 and 8), with significantly upregulated gene expression ($p \leq 0.00001$) at distinct phases of the transition. These findings integrate transcriptomic and anatomical data, providing deeper insights into the molecular mechanisms underlying post-pollination syndrome and the FFT process in *V. planifolia*.

Stage 1: Pre-pollination involves ovarian differentiation, nitrogen metabolism, stress and short-day response

During the Pre-pollination stage, the differentiation of the ovary inner wall takes place (Fig. 1B). At the transcriptional level, enrichment networks integrate environmental signals such as photoperiod, heat acclimatization, and nitrogen metabolism (Fig. 5). These networks are associated with specific gene signatures, including *VpNAC14-like* (linked to nitrogen metabolism and abiotic stress tolerance) and *VpVRN5-like* (a flowering inducer) (Fig. 5) [34, 35].

This aligns with findings in *A. thaliana* and *Oryza sativa* L. (Poaceae), where *NAC* genes have been implicated in the response to nitrate deficiency [34, 35]. Furthermore, this gene family has been suggested to play a role in the transition from the vegetative to the reproductive phase, as well as in the development of the ovary and fruit [45].

This aligns with findings in *A. thaliana* and *Oryza sativa*, where *NAC* genes are implicated in responses to nitrate deficiency and in regulating the transition from vegetative to reproductive phases, as well as ovary and fruit development [34, 35, 45]. In plants exhibiting SPP, floral development occurs without ovule differentiation before pollination [3]. Thus, it is hypothesized that *VpNAC14-like* may act as a key regulator of ovary tissue differentiation in response to nitrate availability in *V. planifolia*. However, functional validation is required.

In *A. thaliana*, *AtVRN5* (*Arabidopsis thaliana VERNALIZATION 5*) promotes flowering via epigenetic repression of *AtMAF1* (*Arabidopsis thaliana MOTHER OF FT AND TFL1)* and *AtMAF5* (*Arabidopsis thaliana MOTHER OF FT AND TFL5)* under short-day conditions [36]. Although photoperiod is suggested to play a partial role in *V. planifolia* flowering, other factors such as humidity and nutrient availability may also contribute in tropical climates [17, 46]. Given its function in *A. thaliana*, *VpVRN5-like* may similarly regulate photoperiod-mediated flowering induction in *V. planifolia*, though further experiments are necessary to confirm this.

Stage 2: pollination involves the formation of the megaspore mother cell, the induction of growth genes, and the response to light intensity

At this stage, nucellar filament branching and megaspore mother cell formation occur, while placental proliferation remains minimal (Fig. 1E). These observations align with the low number of differentially expressed genes (DEGs) between Pre-pol and Pol (Fig. 3), which is characteristic of the Pre-pollination stage, where ovary development remains unchanged before pollination [3].

At the transcriptional level, enrichment analysis revealed networks integrating responses to exogenous biotic signals and abiotic factors, including light intensity, oxidative stress, and temperature acclimation (Fig. 6). Among these, a network associated with *VpMBS1-like* showed a functional signature linked to light intensity, photoreception and stress adaptation (Fig. D6) [37].

Notably, the *AtMBS1* (*Arabidopsis thaliana METHYLENE BLUE SENSITIVITY 1*) gene in *A. thaliana* regulates the response to O_2-induced oxidative stress through the β-cyclocitral pathway, facilitating adaptation to high irradiance [37]. In vanilla cultivation, it is well established that plants are exposed to high light intensities during spring, coinciding with the pollination period [17, 46]. Therefore, it is suggested that *VpMBS1-like* genes may

play a role in protecting the ovary from light-induced stress; however, experimental validation is required.

Stage 3: Post-pollination involves the formation of the embryonic sac, elongation of pollen tubes, and induction of genes associated with response to auxins

During Post-pollination stage, the embryonic sac forms, and pollen tubes elongate (Fig. 1I and H). Furthermore, transcriptomic enrichment analysis revealed associations between nucleic acid metabolism, post-translational processes, and RNA signaling with cell-cell communication, auxin responses, and gametophyte and pollen development (Fig. 7).

Stage-specific functional signatures include *VpIAA6-like* and *VpRALF27-like*, associated with transcriptional regulation and auxin response [38–40]. This aligns with findings in *Solanum lycopersicum* and *O. sativa*, where Aux/IAA (Auxin/Indole-3-Acetic Acid) transcription factors regulate reproductive organ morphogenesis [38, 39]. Similarly, *VpRALF27-like* has been linked to pollen tube elongation in *Brassica rapa* [40]. While its function in *V. planifolia* remains unvalidated, these findings suggest that *VpIAA6-like* and *VpRALF27-like* may play similar roles in ovary development and pollen tube elongation.

Stage 4: fertilization is characterized by complete seed development, up-regulation of genes associated with abiotic interactions, and transport of sugars

The Fertilization stage is characterized by the formation of fully developed seeds (Fig. 1K). Enrichment networks integrate functional categories related to biotic responses (interactions with fungi and symbionts) and abiotic stress factors (cold, salinity, and nutrient availability) in connection with seed development (Fig. 8). These networks are associated with *VpAAE3* (abiotic stress response), *VpPR1* (fungal response), and *VpSWT12* (sugar transport) (Fig. 8) [41–44]. This finding aligns with previous reports in *S. lycopersicum*, where the *SlAAE3* (*Solanum lycopersicum ACYL-ACTIVATING ENZYME 3*) gene has been implicated in responses to aluminum stress [43]. Therefore, it is suggested that *VpAAE3-like* may also play a role in metal stress responses in *V. planifolia*. Similarly, studies in *Allium sativum* L. (Amaryllidaceae) have proposed that *AsPR1* (*Allium sativum PATHOGENESIS-RELATED GENE 1*) contributes to *Fusarium* infection responses [41]. Furthermore, *SWT* (*SWEET*) genes are known to mediate sugar translocation for seed filling and modulate defense responses against pathogens in *A. thaliana* and *Zea mays* L. (Poaceae), highlighting their role in symbiosis establishment with fungi [42, 44, 47]. In orchids such as *V. planifolia*, germination is tightly linked to symbiotic mycorrhizal interactions [48]. These findings suggest that *VpSWT12-like* may contribute not only to seed filling but also to facilitating symbiotic relationships or fungal responses, potentially acting in coordination with *VpPR1-like*. Further studies are needed to elucidate these molecular mechanisms during the FFT in *V. planifolia*.

Conclusions

This study investigates the flower-to-fruit transition (FFT) in *Vanilla planifolia*, integrating anatomical data with transcriptomic profiles across four distinct stages of ovary development. This approach allowed us to identify functional signatures and explore the molecular phenology associated with this critical developmental process. Our findings contribute to the understanding of the flower-to-fruit transition in plants exhibiting post-pollination syndrome (PPS), with particular emphasis on *V. planifolia*.

While we identify several candidate genes associated with this transition, it is important to note that their functions have not been experimentally validated in this study. Thus, further functional validation is necessary to confirm their roles in the FFT process. Nevertheless, these candidate genes provide valuable insights into potential biomarkers for future studies focused on agricultural improvement and plant breeding.

This work also highlights the value of investigating non-model plants, such as *V. planifolia*, and species exhibiting PPS. These plants offer unique opportunities for exploring plant reproductive biology and the molecular underpinnings of fruit development. Our study paves the way for future research aimed at elucidating the functional roles of these genes, with implications for crop enhancement and the broader understanding of plant developmental processes.

Abbreviations

FFT	Flower-to-fruit transition
PPS	Post-pollination syndrome
DE	Differential Expression
FDR	False Discovery Rate
A. thaliana	*Arabidopsis thaliana*
S. lycopersicum	*Solanum lycopersicum*
P. bretschneideri	*Pyrus bretschneideri*
V. planifolia	*Vanilla planifolia*
O. sativa	Oryza sativa
UNAM	Universidad Nacional Autonoma de México
Pre-pol	Pre- Pollination
Pol	Pollination
Post-pol	Post- Pollination
Fer	Fertilization
N_2	Liquid Nitrogen
$_{Total}$RNA	Total Ribonucleic Acid
RNA	Ribonucleic Acid
cDNA	Complementary Desoxirribonucleic Acid
UUSMB IBT-UNAM	University Mass Sequencing and Bioinformatics facility of the Instituto de Biotecnología de la Universidad Nacional Autónoma de México
bp	Base pair
GO	Gene Ontology
TPM	Transcripts Per Million
DEGs	Differential Expressed Genes

FAA	Formaldehyde Alcohol Acetic Acid, 10%:50%:5% + 35% water solution
C1	Cluster 1
C2	Cluster 2
C3	Cluster 3
C4	Cluster 4
P1	Profile 1
P2	Profile 2
mRNA	Messenger Ribonucleic Acid
VpMBS1-like gene	*Vanilla planifolia METHYLENE BLUE SENSITIVITY 1-like*
VpIAA6-like gene	*Vanilla planifolia INDOLE-3-ACETIC ACID 6-like*
VpPR1-like gene	*Vanilla planifolia PATHOGENESIS-RELATED GENE 1-like*
VpVNR5-like gene	*Vanilla planifolia VERNALIZATION 5-like*
VpNAC14-like gene	*Vanilla planifolia NAC014-like*
VpRALF27-like gene	*Vanilla planifolia RALFL27-like*
VpAAE3-like gene	*Vanilla planifolia ACYL-ACTIVATING ENZYME 3-like*
VpSWT12-like gene	*Vanilla planifolia SWEET12-like*
AtVRN5 gene	*Arabidopsis thaliana VERNALIZATION 5*
AtMAF1 gene	*Arabidopsis thaliana MOTHER OF FT AND TFL1*
AtMAF5 gene	*Arabidopsis thaliana MOTHER OF FT AND TFL5*
SlAAE3 gene	*Solanum lycopersicum ACYL-ACTIVATING ENZYME 3*
SWT genes	*SWEET*

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06476-z.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Supplementary Material 8

Supplementary Material 9

Supplementary Material 10

Supplementary Material 11

Supplementary Material 12

Supplementary Material 13

Supplementary Material 14

Supplementary Material 15

Supplementary Material 16

Supplementary Material 17

Supplementary Material 18

Supplementary Material 19

Supplementary Material 20

Supplementary Material 21

Supplementary Material 22

Supplementary Material 23

Supplementary Material 24

Supplementary Material 25

Supplementary Material 26

Supplementary Material 27

Acknowledgements

This research is part of O.A.H.M. requirements to obtain a PhD Degree in Biological Sciences in the field of experimental biology knowledge, at Posgrado en Ciencias Biológicas (PCB), of the Universidad Nacional Autónoma de México (UNAM). O.A.H.M. thanks to Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) for the scholarship (CVU, 1003013) awarded to support her PhD studies. We thank the vanilla producers of Totonacapan, especially Mr. Miguel Acosta, for his significant contribution to this work. We would like to thank Ana Sidney Betanzos Ávalos and Ma. Concepción Guzmán Ramos for their invaluable contribution to the processing and anatomical analysis of the gynostemium-ovary during its transition from flower to fruit in Vanilla planifolia.

Author contributions

O.A.H.M. and V.M.S.R.: Conceptualization. O.A.H.M., V.M.S.R., E.S., U.Y.R.L. and J.E.C.C.: methodology. O.A.H.M and V.M.S.R.: software. O.A.H.M., V.M.S.R., and U.Y.R.L.: formal analysis. O.A.H.M and V.M.S.R.: investigation. V.M.S.R, J.E.C.C., and E.S: resources. O.A.H.M. and M.T.O.M.: data curation. O.A.H.M., V.M.S.R., E.S. J.E.C.C., and U.Y.R.L.: writing—original draft preparation. O.A.H.M., V.M.S.R., E.S., J.E.C.C., U.Y.R.L., and M.T.O.M.: writing—review and editing. O.A.H.M., V.M.S.R., E.S., U.Y.R.L. and M.T.O.M.: visualization. V.M.S.R.: supervision. V.M.S.R. and J.E.C.C.: project administration. V.M.S.R.: funding acquisition. All authors reviewed the manuscript.

Funding

This work was financially supported by El Fondo Sectorial de Investigación para la Educación (SEP- CONACYT), Basic Scientific Research 2015, Project 255952 awarded to VMSR.

Data availability

Data is provided within the manuscript or supplementary information files. The supplementary materials are available at https://github.com/Andrea-H-M/Vanilla_TFF/tree/main/4SupplementaryMaterials. The datasets analyzed during the current study are available at NCBI BioProject ID: PRJNA808623 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA808623/).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Colonia Los Reyes Ixtacala Tlalnepantla, Estado de México, Avenida de los Barrios Número 1, Mexico C.P. 54090, Mexico
[2]Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México. Cto. de Posgrados, Ciudad Universitaria Del. Coyoacán, Ciudad de México C. P. 04510, Mexico
[3]Jardín Botánico, Instituto de Biología, Universidad Nacional Autónoma de México. Cto. Zona Deportiva, Ciudad Universitaria Del. Coyoacán, Ciudad de México C. P. 04510, Mexico

Received: 10 September 2024 / Accepted: 27 March 2025
Published online: 05 April 2025

References

1. Lord EM, Russell SD. The mechanisms of pollination and fertilization in plants. Annu Rev Cell Dev Biol. 2002;18:81–105.
2. Chevalier É, Loubert HA, Zimmerman EL, Matton DP. Cell-cell communication and signalling pathways within the ovule: from its inception to fertilization. New Phytol. 2011;192:13–28.
3. O'Neill SD. Pollination regulation of flower development. Annu. Rev. Plant physiol. Plant mol. Biol. 1997;48:547–74.

4. Hernández MOA, Reyes MDA, Campos CJE, Sandoval ZE, Herrera CBE, Salazar RVM. Expresión diferencial Del Complejo ARF8-IAA25-like-TIR1 Durante El síndrome post-polinización En Dos Genotipos de Vainilla. RFM. 2020;43:575–82.
5. Ruan YL, Patrick JW, Bouzayen M, Osorio S, Fernie AR. Molecular regulation of seed and fruit set. Trends Plant Sci. 2012;17:656–65.
6. Marguerat S, Bähler J. RNA-seq: from technology to biology. CMLS. 2010;67:569–79.
7. Kudoh H. Molecular phenology in plants: in natura systems biology for the comprehensive Understanding of seasonal responses under natural environments. New Phytol. 2016;210:399–412.
8. Pattison RJ, Csukasi F, Zheng Y, Fei Z, Van der Knaap E, Catalá C. Comprehensive Tissue-Specific transcriptome analysis reveals distinct regulatory programs during early tomato fruit development. Plant Physiol. 2015;168:1684–701.
9. Liu L, Wang Z, Liu J, Liu F, Zhai R, et al. Histological, hormonal and transcriptomic reveal the changes upon gibberellin-induced parthenocarpy in Pear fruit. Hortic Res. 2018;5:1.
10. Fajardo TVM, Quecini V. Comparative transcriptome analyses between cultivated and wild grapes reveal conservation of expressed genes but extensive rewiring of co-expression networks. Plant Mol Biol. 2021;106:1–2.
11. García GBE, Salazar JA, Nicolás AM, Razi M, Rubio M, et al. Molecular bases of fruit quality in Prunus species: an integrated genomic, transcriptomic, and metabolic review with a breeding perspective. Int J Mol Sci. 2021;22:1–38.
12. Poppenberger B. The orphan crop Crassocephalum Crepidioides accumulates the pyrrolizidine alkaloid jacobine in response to nitrogen starvation. Front. Plant Sci. 2021. https://doi.org/10.3389/fpls.2021.702985.
13. Swamy BGL. On the Life-History of *Vanilla planifolia*. Bot Gaz. 1947;108:449–56.
14. Roux P. Etudes morphologiques et anatomiques dans le genre Vanilla. En: G. Bouriquet, ed. Le vanillier et la vanilla dans le monde. Paul Lechevalier, 1954;Paris:44–92.
15. Odoux E, Brillout JM. Anatomy, histochemestry, and biochemistry Og glucovanillin, Oleoresin, and mucilage accumulation sites in green mature vanilla pod (*Vanilla planifolia*; Orchidaceae): A comprehensive and critical re-examination. Fruits. 2009;64:221–41.
16. Salazar RVM, Sandoval ZE, Granados HCV, Cruz RY, Herrera CBE, et al. Descripcion estructural y funcional de Caida prematura de Frutos de *Vanilla planifolia* jacks. Ex Andrew. Agroproductividad. 2016;9:17–20.
17. Havkin FD, Belanger FC. Handbook of Vanilla Science and Technology. In Handbook of Vanilla Science and Technology; 2010.
18. Sandoval ZE, Rojas LA, Guzmán RC, Carmona JL, Ponce RM, et al. Cuadernos 38: técnicas Aplicadas al estudio de La anatomía vegetal. Universidad Nacional Autónoma de México; 2005.
19. Andrews. Feb. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc. Accessed 1 2021.
20. Bolger AM, Lohse M, Usadel B, Trimmomatic. A flexible trimmer for illumina sequence data. J Bioinform. 2014;30:2114–20.
21. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Philip D, et al. De *Novo* transcript sequence Recostruction from RNA-Seq: reference generation and analysis with trinity. Nat Protoc. 2013. https://doi.org/10.1038/nprot.2013.084.De.
22. Seppey M, Manni M, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness. Gene Prediction: Methods Protocols. 2019;2019:227–45.
23. Solano DMT, Adame GJ, Gregorio JJ, Jiménez JV, Vega AL, et al. Functional categorization of de *Novo* transcriptome assembly of *Vanilla planifolia* jacks. Potentially points to a translational regulation during early stages of infection by F.sarium oxysporum F. Sp. vanillae. BMC Genom. 2019;20:1–15.
24. Bryant DM, Johnson K, DiTommaso T, Tickle T, Couger MB, et al. A Tissue-Mapped Axolotl de *Novo* transcriptome enables identification of limb regeneration factors. Cell Rep. 2017;18:762–76.
25. UniProt C, UniProt. A hub for protein information. Nucleic Acids Res. 2015;43:D204–12.
26. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, et al. The Pfam protein families database. Nucleic Acids Res. 2004;32:D138–41.
27. Koskinen P, Törönen P, Nokso KJ, Holm L, PANNZER. High-throughput functional annotation of uncharacterized proteins in an error-prone environment. J Bioinform. 2015;31:1544–52.
28. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14:417–9.
29. Love MI, Huber W, Anders S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:1–21.
30. Statistical Discovery JMP. LLC. Software estadístico para una mejora basada en datos. https://www.jmp.com/es_mx/home.html. Accessed 8 Jan 2024.
31. Planteome Project. Ontology Enrichment Analysis Tool. https://planteome.org/oat/index.html. Accessed 1 Sep 2023.
32. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS ONE. 2011;6:e21800.
33. Arabidopsis Information Resource (TAIR). Gene Search. https://www.arabidopsis.org/search/genes. Accessed 22 Feb 2025.
34. Vidal EA, Álvarez JM, Gutiérrez RA. Nitrate regulation of *AFB3* and *NAC4* gene expression in *Arabidopsis* roots depends on *NRT1.1* nitrate transport function. Plant Signal Behav. 2014;9(6):e28501.
35. Nakashima K, Takasaki H, Mizoi J, Shinozaki K, Yamaguchi SK. NAC transcription factors in plant abiotic stress responses. Biochim Biophys Acta Gene Regul Mech. 2012;1819:97–103.
36. Sharma N, Xin R, Kim DH, Sung S, Lange T, et al. No flowering in short day (NFL) is a bHLH transcription factor that promotes flowering specifically under short-day conditions in Arabidopsis. J Dev. 2016;143:682–90.
37. Shumbe L, d'Alessandro S, Shao N, Chevalier A, Ksas B, et al. METHYLENE BLUE SENSITIVITY 1 (MBS1) is required for acclimation of Arabidopsis to singlet oxygen and acts downstream of β-cyclocitral. Plant Cell Environ. 2017;40:216–26.
38. Wang H, Jones B, Li Z, Frasse P, Delalande C, et al. The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. Plant Cell. 2005;17:2676–92.
39. Jain M, Kaur N, Garg R, Thakur JK, Tyagi AK, et al. Structure and expression analysis of early auxin-responsive Aux/IAA gene family in rice (Oryza sativa). Funct Integr Genom. 2006;6:47–59.
40. Shi F, Zhou X, Liu Z, Feng H. Rapid alkalinization factor (RALF) genes are related to genic male sterility in Chinese cabbage (Brassica Rapa L. Ssp. pekinensis). Sci Hortic. 2017;225:480–9.
41. Anisimova OK, Shchennikova AV, Kochieva EZ, Filyushin MA. Pathogenesis-related genes of PR1, PR2, PR4, and PR5 families are involved in the response to fusarium infection in Garlic (*Allium sativum* L). Int J Mol Sci. 2021;22:6688.
42. Li X, Si W, Qin Q, Wu H, Jiang H. Deciphering evolutionary dynamics of SWEET genes in diverse plant lineages. Sci Rep. 2018;8(1):13440.
43. Jin JF, He QY, Li PF, Lou HQ, Chen WW, et al. Genome-wide identification and gene expression analysis of acyl-activating enzymes superfamily in tomato (Solanum lycopersicum) under aluminum stress. Front Plant Sci. 2021;12:754147.
44. Bezrutczyk M, Hartwig T, Horschman M, Char SN, Yang J, et al. Impaired phloem loading in zmsweet13a, B, C sucrose transporter triple knock-out mutants in Zea Mays. New Phytol. 2018;218:594–603.
45. Liu J, Qiao Y, Li C, Hou B. The NAC transcription factors play core roles in flowering and ripening fundamental to fruit yield and quality. Front Plant Sci. 2023;14:1095967.
46. Fouché JG, Jouve L. *Vanilla planifolia*: history, botany and culture in reunion Island. Agronomie. 1999;19(8):689–703.
47. Gebauer P, Korn M, Engelsdorf T, Sonnewald U, Koch C, et al. Sugar accumulation in leaves of Arabidopsis sweet11/sweet12 double mutants enhances priming of the Salicylic acid-mediated defense response. Front Plant Sci. 2017;8:1378.
48. Camarena-Gutiérrez G. Interacción planta-hongos Micorrízicos arbusculares. Rev Chapingo Ser Cienc Amb. 2012;18(3):409–21.

Publisher's note