Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Transcriptome profiling reveals AGAMOUS-LIKE genes associated with the reprogramming of mature endosperm cell fate in hardy kiwifruit

Joanna Malec

joanna.malec@doctoral.uj.edu.pl

Jagiellonian University

Zuzanna Dutkiewicz

Universität Innsbruck

Jan Kanty Trościński

Jagiellonian University

Marzena Popielarska-Konieczna

Jagiellonian University

Małgorzata Czernicka

University of Agriculture in Krakow

Research Article

Keywords: Actinidia arguta, RNA sequencing, Dedifferentiation, Recalcitrant tissue, Non-model plant, MADS-box transcription factors

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11366878/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Transcriptome profiling reveals AGAMOUS-LIKE genes associated with the reprogramming of mature endosperm cell fate in hardy kiwifruit

Joanna Malec[1,2*], Zuzanna Dutkiewicz[3], Jan Kanty Trościński[1,2], Marzena Popielarska-Konieczna[1†], Małgorzata Czernicka [4†],

[1]Department of Plant Cytology and Embryology, Institute of Botany, Faculty of Biology, Jagiellonian University in Kraków, Kraków, Poland

[2]Doctoral School of Exact and Natural Sciences, Jagiellonian University in Kraków, Kraków, Poland

[3]Department of Microbiology, The University of Innsbruck, Innsbruck, Austria

[4]Department of Plant Biology and Biotechnology, Faculty of Biotechnology and Horticulture, University of Agriculture in Kraków, Kraków, Poland

[*] Correspondence: Joanna Malec, joanna.malec@doctoral.uj.edu.pl

[†] MC and MP-K contributed equally as senior authors

Abstract

Background

Mature endosperm is a highly differentiated storage tissue with limited capacity for developmental reprogramming. Although *in vitro* culture can induce endosperm dedifferentiation and regeneration in hardy kiwifruit (*Actinidia arguta*), the transcriptional changes associated with this transition remain poorly understood. This study investigated early reprogramming of mature endosperm in two cultivars with contrasting culture responses, focusing on *AGAMOUS-LIKE* (*AGL*) transcription factors.

Methods

RNA sequencing was performed on freshly isolated endosperm and explants cultured for 14 days on thidiazuron-containing medium from cultivars 'Bingo' and 'Weiki', with three biological replicates per condition. *De novo* transcriptome assembly was followed by functional annotation, differential expression analysis and GO and KEGG enrichment analyses. Selected *AGL* candidates were further

characterised using sequence alignment, conserved-domain analysis, maximum-likelihood phylogenetic reconstruction and three-dimensional protein modelling.

Results

The assembly yielded 149,380 unigenes. Comparisons between cultured and freshly isolated endosperm identified 29,814 differentially expressed genes in 'Bingo' and 35,893 in 'Weiki', with 21,365 shared between the two comparisons. Culture-associated transcriptional changes involved ribosome biogenesis, protein synthesis, cell wall organisation, transport and carbohydrate and lipid metabolism. *AP2/EREBP*, *WRKY* and *MYB* transcription factors exhibited shared induction patterns alongside cultivar-dependent responses in additional regulatory families. Transcripts annotated as *AGL15*, *AGL61* and *AGL62* were up-regulated in both cultivars, whereas additional *AGL61*-, *AGL62*- and *AGL80*-related transcripts displayed distinct expression patterns. Sequence, domain and phylogenetic analyses supported the assignment of selected candidates to Type I and Type II MADS-box lineages. Structural modelling provided complementary support for selected annotations while identifying divergent candidates requiring cautious interpretation.

Conclusions

This study provides a transcriptomic resource for early reprogramming of mature endosperm in *A. arguta* and identifies *AGL* homologues as candidate regulators of this transition. The concurrent induction of transcripts related to endosperm development and embryogenic competence suggests that mature endosperm reprogramming may recruit elements of both developmental programmes. These findings establish a basis for functional testing of *AGL* candidates and investigation of cultivar-dependent endosperm regeneration.

Keywords

Actinidia arguta, RNA sequencing, Dedifferentiation, Recalcitrant tissue, Non-model plant, MADS-box transcription factors

INTRODUCTION

Plant tissues show a high potential to modify their metabolism and development, and consequently - their fate, as an adaptation to environmental fluctuations [1–4]. The conditions of tissue culture allow this capacity to cell reprogramming to be further enhanced. Unorganized cell tissue, known as callus, can theoretically be produced from any living plant cell through the process of dedifferentiation [2, 5]. The mechanisms underlying callus formation and organization remain a crucial area of research, as confirmed by recent studies on *Arabidopsis* [6]. Research suggests that stem-cell-related genes play a key role in dedifferentiation, meaning that meristems composed of stem cells possess higher plasticity and able to change their fate more easily [7]. One of the defining features of stem-like cells is the decondensation of their chromatin, which allows for the initiation of transcriptional processes [8, 9]. In contrast, highly differentiated and specialized tissue are resistant to reprogramming and rarely revert to a stem-like stage. Nevertheless, the process of cell dedifferentiation can be successfully induced in a culture setting through the application of stressors, such as wounding or *via* exogenous plant hormones, including auxins or cytokines [9].

The endosperm, a storage tissue unique to flowering plants that develops within fertilized ovule and persist in the seed, is highly recalcitrant to reprogramming and fate changes. Consequently reports of endosperm dedifferentiation under tissue culture conditions remain scarce. Furthermore, existing reports focus predominantly on the optimization of culture conditions, alongside morphological and histological observations [10–12]. Remarkably, there is complete lack of molecular data concerning endosperm dedifferentiation. From an evolutionary perspective, the endosperm is hypothesized to have evolved from an “altruistic twin embryo” that forfeited its own developmental potential to nourish the second embryo, which forms a seedling and then a new organism [13]. Molecular evidence suggests that the duplication and diversification of specific genes drove the evolution of the endosperm as a developmental novelty in flowering plants [14]. The primary function of the endosperm is to transfer nutrients from maternal tissues to the embryo and store these reserves to fuel post-germinative seedling growth, rather than to facilitate differentiation into other tissue types. Notably, in most flowering plants, strict molecular mechanisms (such as mitotic repression of the central cell) prevent endosperm development in the absence of double fertilization [15]. Furthermore,

during endosperm development, processes like programmed cell death (PCD) act as barriers against fate alteration. This phenomenon has been well documented in the starchy endosperm of cereal grains prior to cell filling [16], and in Arabidopsis endosperm cells adjacent to the embryo to accommodate embryo expansion [17].

Due to the process of double fertilization, endosperm cells typically possess a higher nuclear DNA content than the embryo. Polyploidization plays a pivotal role both as a speciation mechanism in natural environment and as a powerful tool in breeding programs in agriculture [12]. In this context, the in vitro culture of isolated endosperm could represent a promising approach, as it could significantly reduce the time required to obtain plants with higher ploidy levels. Investigating the relationship between ploidy levels and economically important agronomic traits, such as fruit weight, remains a key focus in modern plant breeding [18]. Consequently, establishing reliable methods for endosperm culture continues to be an area of intense research interest [19–21].

Transcriptomic analyses of plant developmental processes represent one of the most dynamically explored areas of plant research [22–24]. This omics-based studies are instrumental in unravelling complex regulatory mechanisms, including those governing dedifferentiation and callus induction. In Arabidopsis, a core set of genes has been identified as crucial for callus induction, operating independently of the initial tissue type or the specific inducer applied [5, 9, 25]. Most of the key genes responsible for start of callus formation are classified as transcription factors (TFs), such as wound-induced APETALA2/ETHYLANE RESPONSE FACTOR (AP2/ERF) transcription factor WOUND INDUCED DEDIFFERENTIATION 1 (WIND1) and its close homologs, WIND2, WIND3, and WIND4 [25, 26]. Several other members of the AP2/ERF family critical for wound-induced callus formation include PLETHORA3 (PLT3), PLT5, and PLT7 [25]. Furthermore, WRKY transcription factors [27], WUSCHEL related homeobox (WOX) genes [28] and MADS-box genes [9] have also been shown to play key roles in the induction of callus.

TFs function as critical regulatory elements that exert strict control over a diverse array of cellular processes and responses to external stimuli by directly modulating gene expression [29]. A typical plant TF is characterized by a sequence-specific DNA-binding domain that enables precise

recognition of and binding to target DNA sequences. Additionally, it features a transcription-regulation domain responsible for mediating its activity as either an activator or a repressor. To facilitate its functional architecture, a TF also possesses an oligomerization site that promotes the assembly of multiprotein complexes, along with a nuclear localization signal ensuring its proper trafficking to the nucleus [30].

MADS-box proteins represent one of the largest and most diverse groups of TFs. They play a crucial role in regulating various developmental pathways across nearly all eukaryotic organisms, including plants [31, 32]. The acronym "MADS" is derived from the initials of its founding members, MINICHROMOSOME MAINTENANCE 1 (MCM1) from yeast, AGAMOUS (AG) from *Arabidopsis thaliana* (Arabidopsis), DEFICIENS (DEF) from *Antirrhinum majus*, and SERUM RESPONSE FACTOR (SRF) from humans [33]. Each MADS-box protein possesses a highly conserved N-terminal domain of approximately 60 amino acids, known as the MADS- box, which specifically recognizes and binds to CArG box sequences (CC[A/T]6GG). Based on sequence characteristics and domain structure, MADS-box TFs are classified into two distinct lineages: Type I and Type II. [34, 35]. While Type II MADS-box proteins feature a characteristic, fully developed domain architecture (including the K domain), Type I MADS-box proteins typically lack this K domain. [36].

Importantly, it has been hypothesized that the duplication and diversification of Type I MADS genes resulted in the evolution of the endosperm, as the developmental novelty in flowering plants [14]. The newly created Type I MADS-box clades, Mγ and Mα*, exhibit preferential expression in floral organs and seeds compared to vegetative organs. Furthermore, studies have demonstrated that certain Type II MADS-box proteins positively regulate endosperm proliferation and cellularization in Arabidopsis [37]. In the genus Actinidia, especially particularly in *A. chinensis*, the MADS-box gene family has been examined primarily for its involvement in bud dormancy regulation [38] and floral sex differentiation [29, 39]. Nevertheless, evidence from Arabidopsis highlights that several genes from this transcription factor family are essential for proper endosperm and embryo development [40]. AGAMOUS-LIKE (AGL) genes constitute a distinct subfamily within the MADS-box transcription

138 factor (TF) family. In plants, they form a large regulatory network involved in controlling the
139 development of virtually every organ. While their crucial importance in flower development is well
140 documented, AGLs also participate significantly in vegetative organogenesis. For instance, *AGL21* is
141 involved in seed germination and lateral root development [41, 42]. Furthermore, *AGL15* has been
142 demonstrated to be crucial for zygotic embryo development, as well as for the formation of embryonic
143 structures derived from other plant tissues through somatic embryogenesis [43–45]. Similarly,
144 *AGL18*, which shares high structural homology with *AGL15*, plays a critical role in the induction of
145 somatic embryogenesis [46]. Furthermore, several evidence indicates that AGL genes exert a
146 significant impact on the proper seed. For example, a study on Arabidopsis interspecies hybrids
147 showed that *AGL62* and *AGL90* are involved in inducing embryo development arrest while
148 simultaneously triggering endosperm hyperproliferation [47].
149 In the majority of dicots, the endosperm is consumed by the developing embryo during seed
150 maturation [48]. In contrast, seeds of the Actinidia genus are characterized by the presence of a
151 substantial amount of endosperm tissue in their mature state [49]. A reproducible and efficient
152 protocol has been established for callus proliferation and adventitious shoot bud induction from
153 isolated mature endosperm of the hardy kiwifruit, *Actinidia arguta (*Fig. 1) [49, 50, 51]. Commonly
154 known as the kiwiberry, this species produces grape-sized, hairless fruits, that are delicate, highly
155 aromatic and exceptionally rich in vitamin C, carotenoids, and polyphenols [52]. Due to these
156 desirable traits, along with its remarkable cold hardiness, *Actinidia arguta* holds significant
157 commercial potential for fresh fruit production in northern latitudes, making efforts to breed new
158 cultivars highly valuable [53]. However, the hardy kiwifruit is a climbing, woody, autotetraploid fruit
159 crop fruit, and non-model plant, which poses substantial challenges in elucidating its molecular
160 mechanisms and gene regulatory networks [54, 55]. To date, transcriptomic analyses in *A. arguta*
161 have focused primarily on dioecy and fruit ripening [40, 54], anthocyanin synthesis [56], low-
162 temperature resistance [57], and plant immunity [55].

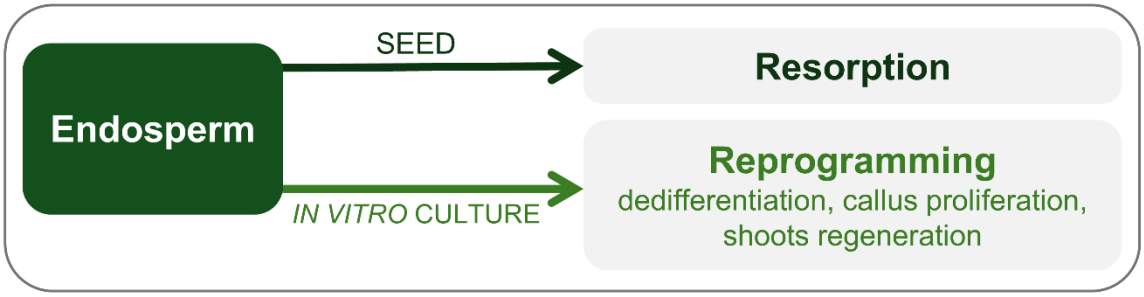


Figure 1 Two fate-pathways of mature *Actinidia arguta* endosperm. In-seed resorption by the growing embryo during seedling germination versus reprogramming following isolation, embryo dissection, and in vitro culture protocols according to [49-51].

Available transcriptomic reports on endosperm development, have predominantly focused on *in planta* processes in Arabidopsis and cereals. These reports are crucial for understanding the regulatory mechanisms driving endosperm formation, from double fertilization through cellularization and proliferation within the ovule [37, 47, 58, 59]. Consequently, it is reasonable to hypothesize that certain genes responsible for these developmental stages also influence the fate of isolated mature endosperm during tissue culture.

The objective of this study is to identify the molecular mechanisms determining cell reprogramming and fate switching of mature endosperm, entering the dedifferentiation pathway after two weeks of isolation and culture. Additionally, since hardy kiwifruit is a non-model plant, the phylogenetic analysis was conducted for selected regulatory genes. Transcriptomic analyses were performed on mature endosperm from two cultivars, 'Bingo' and 'Weiki', of hardy kiwifruit, *Actinidia arguta*. It has been previously shown [51] that the isolated mature endosperm of kiwi berry differed significantly in efficiency response depending on the cultivar. Specifically, the callus induction rate was twofold higher for 'Bingo' compared to 'Weiki'. Furthermore, the indirect shoot bud induction was only detected in this cultivar. We hypothesize that differential expression of genes, including TFs, and genes involved in metabolic and regulatory pathways, drives this endosperm fate switch and underlies these cultivar-specific response. Conversely, transcriptomic data shared by both cultivars may

highlight the core processes and genes critical for endosperm cell fate switching in hardy kiwifruit, independent of the genetic background.

MATERIALS AND METHODS

Plant material

Fruits of hardy kiwifruit (*Actinidia arguta*), cultivars 'Bingo' and 'Weiki' were collected with permission from the Actinidia germplasm collection at Warsaw University of Life Sciences (SGGW) in Poland. The fruits were stored in a refrigerator at 10°C. Then, the fruits were surface sterilized, and the seeds were extracted. The seed coats were then excised, embryos removed, and the endosperm tissue was isolated, according to a protocol previously developed for *A. chinensis* var. *deliciosa* [60].

Isolated endosperm culture

The isolated endosperm explants from both cultivars were put on Murashige and Skoog [61] medium containing micro- and macroelements with vitamins (Duchefa), supplemented with 3% sucrose, 0.6% agar and 0.5 mg/L thidiazuron. This medium composition was selected based on the optimal callus induction efficiency previously reported by Abdullah et al. [51]. The cultures were maintained in 60-mm Petri plates in 25°C in darkness. Following two weeks of culture, explants showing visible bulges – indicative of endosperm proliferation and/or callus induction (Fig. 2) - were collected based on macroscopic observations and transferred to Eppendorf tubes. Freshly isolated endosperms were used as a control. The sample weight was approximately 150 mg (for samples consisting of two-week-old explants) or 200 mg (for freshly isolated endosperm). The tubes containing plant tissue were immediately frozen in liquid nitrogen and stored in -80°C.

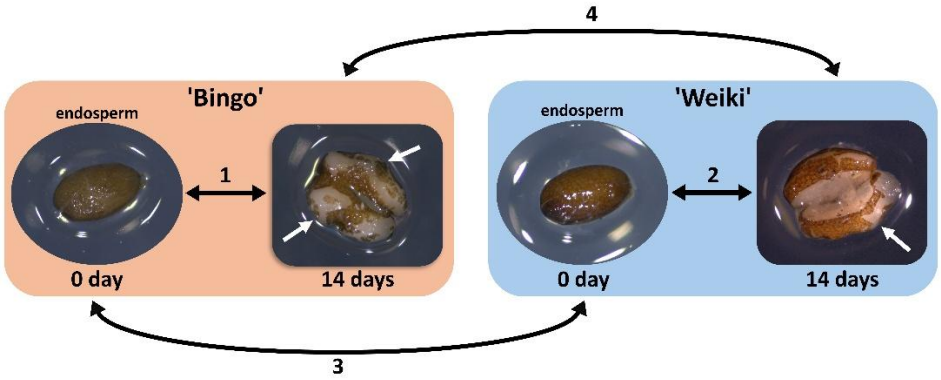


Figure 2 Scheme of sample comparisons used for differential gene expression analysis. The explants on day 0 are fresh, isolated endosperm with the embryo and external, thick seed coat removed. Note the explants after 14 days in culture, with visible bulges in the responding regions (arrows)**.**

RNA extraction

Because of the high concentration of proteins and starch in endosperm tissue, and no working RNA-isolation kits, the protocol for total RNA extraction based on two-phase extraction method developed for wheat (*Triticum aestivum*) by Mortal et al. [62], was used. Briefly, 150-200mg (depends on sample type) of plant tissue was grinded with liquid nitrogen, transferred to the Falcon tubes with the pre-warmed EB and Phenol-Chloroform–Isoamylalcohol (PCI) mixture (25:24:1 V/V). Then, a sample was divided into two Eppendorf tubes and centrifuge. For phase separation, the centrifugation of supernatant with PCI mixture was repeated twice. Then to the aqueous phase 3M NaOAC and 100% ethanol were added, and placed in -80°C for 20 min. After this time, the sample was centrifuged for RNA pellet precipitation. The obtained pellet was dissolved in DEPC water. For RNA purification, first 4M LiCl was added, set in -80°C for 2 hours, then centrifuged and after supernatant removing, 2M LiCl was added to the tubes and centrifuged again. The purified pellet was washed with 70% ethanol and after supernatant discarded was left to completely dry. Finally, the RNA pellet was dissolved in DEPC water and combined to the one Eppendorf tube. Samples were stored in -80°C.

The concentration and quality of isolated RNA were checked spectrophotometrically using NanoDrop 2000c (Thermo Scientific™) by the absorbance at 230, 260, and 280 nm wavelengths measurements. Additionally, the quality and integrity of RNA were determined using 1% (w/v) agarose gel electrophoresis under denaturing conditions. Only high quantity RNA samples (OD260/280 = 2.0-2.2, OD260/230 ≥ 2.0, RIN ≥ 8.0, 28S:18S ≥ 1.5), were employed to construct the sequencing library.

RNA sequencing

RNA-seq Library Construction and Illumina Sequencing

Messenger RNA (mRNA) was purified from total RNA samples using poly-T oligo-attached magnetic beads. Following fragmentation, first-strand cDNA was synthesized using random hexamer primers, and second-strand cDNA synthesis was performed using dTTP for non-directional library preparation. The workflow was completed via end repair, A-tailing, adapter ligation, size selection, PCR amplification, and purification. Constructed libraries were quantified using a Qubit fluorometer (Thermo Fisher Scientific), while their size distribution was assessed via a Bioanalyzer (Agilent Technologies, Inc). Properly quantified libraries were pooled and sequenced on Illumina platforms (Novogene Co., Ltd.) to generate paired-end raw reads. All RNA-seq datasets generated in this study have been deposited in the NCBI SRA database under BioProject PRJNA1542276.

Data Filtering and Transcriptome Assembly

Raw data in FASTQ format were processed through CASAVA base calling for sequencing error rate and base quality score (Phred) evaluation. To obtain high-quality clean reads, raw reads were filtered using Novogene's internal scripts to remove adapter contamination (P5: 5'-AATGATACGGCGACCACCGAGA-3'; P7: 5'-CGTATGCCGTCTTCTGCTTG-3'), reads containing more than 10% uncertain nucleotides (N > 10%), and low-quality reads where over 50% of bases had a Phred score ≤ 5. To reconstruct the Actinidia arguta transcriptome, clean reads were assembled de novo using Trinity software (v2.6.6; parameters: --minkmercov 3, --min_glue 4). The assembly was perform through the sequential Inchworm, Chrysalis, and Butterfly modules using the

default k-mer size of 31. To reduce redundancy, hierarchical clustering of contigs was performed using Corset (v1.09; parameters: -f true, -m 10), filtering out sequences with fewer than 10 mapped reads. Transcriptome completeness was subsequently assessed using BUSCO (v3.0.2; in transcriptome mode: -m tran against the embryophyta_odb9 lineage dataset [63].

Functional Annotation and CDS Prediction

Comprehensive functional annotation of unigenes was performed by querying seven major databases. Diamond (v2.1.6) was used with an E-value threshold of 1 × 10-5 against the NCBI non-redundant protein (nr), KOG/COG, and Swiss-Prot databases (http://www.ebi.ac.uk/uniprot/). The NCBI nucleotide (nt) database was searched using NCBI BLAST (v2.9.0; E-value = 1 × 10-5). Protein domains were predicted via hmmscan from the HMMER 3.1 package (E-value = 0.01) against the Pfam database [64]. Gene Ontology (GO) terms were assigned using Blast2GO (b2g4pipe_v2.5; E-value = 1 × 10-6) and Novogene scripts (http://www.geneontology.org/), while metabolic pathway annotations were completed using the KEGG Automatic Annotation Server (KAAS) and Diamond (v2.1.6; E-value = 1 × 10-5) (http://www.genome.jp/kegg/). Coding protein sequences (CDS) were extracted from BLAST hits or predicted using TransDecoder (v3.0.1), and plant transcription factors were identified and classified using iTAK software via hmmerscan.

Quantification and Differential Expression Analysis

Clean reads were mapped back to the Corset-filtered transcriptome reference, and expression levels were quantified using RSEM (v1.2.28; parameters: --estimate-rspd, -mismatch-rate 0.3). RSEM utilized Bowtie alignments to calculate read counts for each gene, which were then normalized to Fragments Per Kilobase of transcript sequence per Millions base pairs sequenced (FPKM) values. Differential gene expression analysis between groups with biological replicates was performed using the DESeq2 R package (v1.26.0) based on a negative binomial distribution model, adjusting the false discovery rate (FDR) with the Benjamini-Hochberg (BH) method. Genes with an absolute $\log_2$FoldChange greater than 1 ($|\log_2(\text{Fold Change})| > 1$) and an adjusted p-value (padj) below 0.05 were considered significantly differentially expressed genes (DEGs). GO enrichment of DEGs was

analysed using GOseq (v1.32.0) and topGO (v2.32.0) with a corrected p-value < 0.05, while KEGG pathway enrichment was assessed via KOBAS (v3.0; corrected p-value < 0.05). Finally, protein-protein interaction (PPI) networks were predicted by mapping DEG sequences against the STRING database using NCBI BLAST (v2.2.28+; E-value = 1 × 10-10).

Bioinformatic analysis

The transcripts assembled *de novo* were annotated against various databases to assign their likely functions, based on which gene names were tentatively assigned. To confirm the annotations, by identifying sequences homologous to the selected Type I MADS-box AGLs, a BLASTx search was performed against the Non-Redundant (NR) protein database for *Actinidia* species using the default parameters. The two sequences with the highest query coverage were selected for further analysis. Finally, a comparative analysis of the proteins annotated to the studied transcripts was cariied out using the UniProtKB [65] database, including both reviewed (Swiss-Prot) and unreviewed (TrEMBL) data The selected target transcripts were translated using the ExPASy Translate Tool (https://web.expasy.org/). Each of the six resulting reading frames was aligned with the reference homologue sequences using the MUSCLE algorithm to identify the correct open reading frame. A multiple sequence alignment was subsequently performed between translated transcripts and their corresponding homologous proteins to identify conserved domains and regions with the high sequence similarity. Functional domains and motifs within the analysed sequences were further predicted using the InterPro search tool [66].

Phylogenetic tree construction

All AGL protein sequences were aligned using MAFFT v7 [67], and the resulting multiple sequence alignment was subsequently trimmed using trimAl [68] with the gappyout function. An initial phylogenetic tree was generated using FastTree [69] and was subsequently used as the starting tree for inference of the final maximum-likelihood (ML) phylogenetic tree using IQ-TREE [70]. The tree was

inferred using the following command: iqtree -s fasta_file -nt AUTO -m TEST -t fasttree_file.tree -pre output_folder -bb 1000. All analyses were performed on the LEO5 high-performance computing (HPC) cluster at the University of Innsbruck. The cluster comprises 62 compute nodes, each equipped with two 32-core Intel Xeon 8358 (Icelake) processors, providing a total of 64 CPU cores per node. The resulting phylogenetic tree was visualised and clades were annotated using the Interactive Tree of Life (iTOL) web server [71].

Additionally, motif analysis was performed using the MEME suite [72] with the following parameters: "Any number of repetitions" for motif site distribution, a maximum of 3 motifs per sequence, and a maximum motif width of 75 nucleotides. Subsequently, the sequences of the identified motifs were verified using the InterPro database [66].

3D protein models construction

The corresponding sequences of the AGL proteins of interest were modelled using the AlphaFold Server [73]. For each protein, the model with the highest confidence score was selected and downloaded for further analysis. The most similar corresponding protein sequences and their predicted structures were identified and downloaded from the UniProt database [65]. These sequences were subsequently modelled using the AlphaFold Server. For AGL9, the corresponding experimentally resolved structure was instead obtained from the Protein Data Bank (PDB), entry 8CRA [74]. All protein structures were subsequently visualized and analysed using PyMOL Molecular Graphics System (Reference Format: Schrödinger, LLC. (2024). The PyMOL Molecular Graphics System (Version 3.0). Available at [75]).

RESULTS

Comprehensive transcriptome analysis

RNA sequencing of three biological replicates of two cultivars 'Bingo' and 'Weiki' at two time frames: 0 days (freshly isolated endosperm) and two –weeks of cultured endosperm, was performed on the Illumina HiSeq platform, resulting in a total of 12 samples. The highest number of clean reads was obtained for the control (0 days) samples of 'Weiki' – a total of 93,596,927 clean reads and 28,10 Gbp, with an average of 9.37 Gbp per sample. The lowest number of clean reads was generated for 'Bingo' samples after 14 days of culturing (77,972,321 in total and 23.40 Gbp, with an average 7.80 Gbp per sample). The average GC content was approximately 46% across the samples (Fig. 3a).

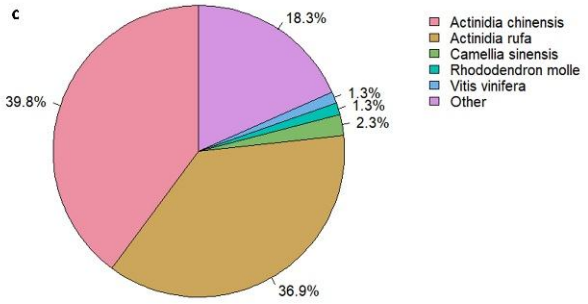
a

Sample	Biological replication	Clean reads	Error rate (%)	Q30 (%)	GC content (%)
Bingo endosperm (0d)	Bingo_0d_1	35,139,443	0.02	94.47	46.92
	Bingo_0d_2	29,620,802	0.02	94.99	46.25
	Bingo_0d_3	22,316,184	0.02	94.71	45.47
	Total	87,076,429			
Bingo 14d	Bingo_14d_1	24,715,659	0.03	94.21	47.08
	Bingo_14d_2	25,920,920	0.02	94.75	46.46
	Bingo_14d_3	27,335,742	0.02	94.25	47.01
	Total	77,972,321			
Weiki endosperm (0d)	Weiki_0d_1	34,186,795	0.02	94.99	45.69
	Weiki_0d_2	29,150,018	0.02	95.03	45.44
	Weiki_0d_3	30,260,114	0.02	94.58	46.34
	Total	93,596,927			
Weiki 14d	Weiki_14d_1	25,127,511	0.02	94.37	47.05
	Weiki_14d_2	28,304,038	0.03	94.16	47.48
	Weiki_14d_3	31,892,682	0.02	94.60	46.21
	Total	85,324,231			

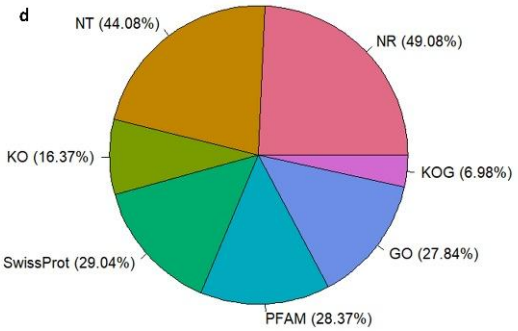
b

Number of transcripts	422,213
Number of unigens	149,380
N50 (for transcripts)	1,643 bp
N50 (for unigens)	1,488 bp
Maximum transcript length	16,469 bp
Minimum transcript length	301 bp
Average transcript length	1,195 bp

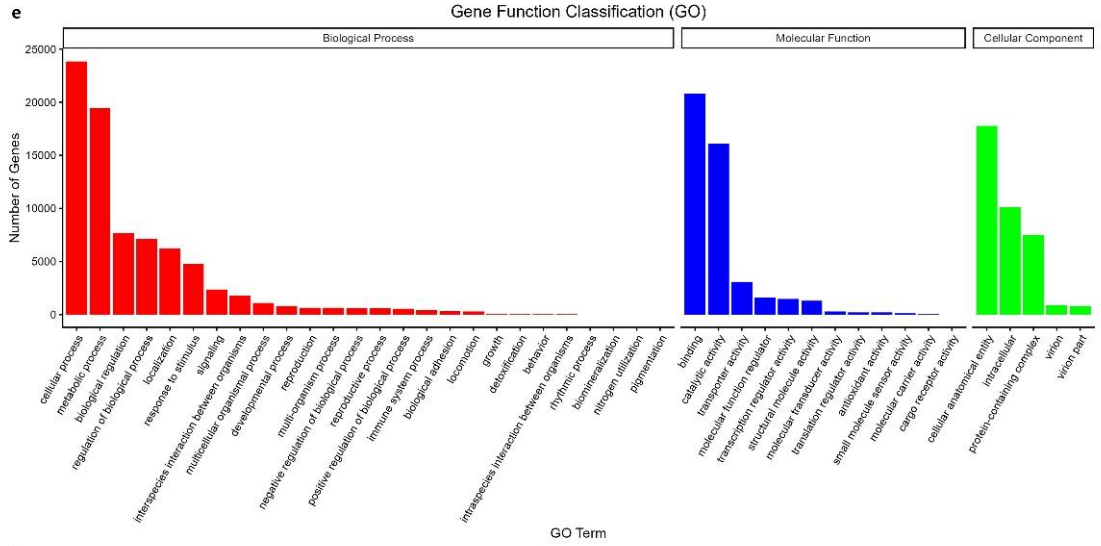
c



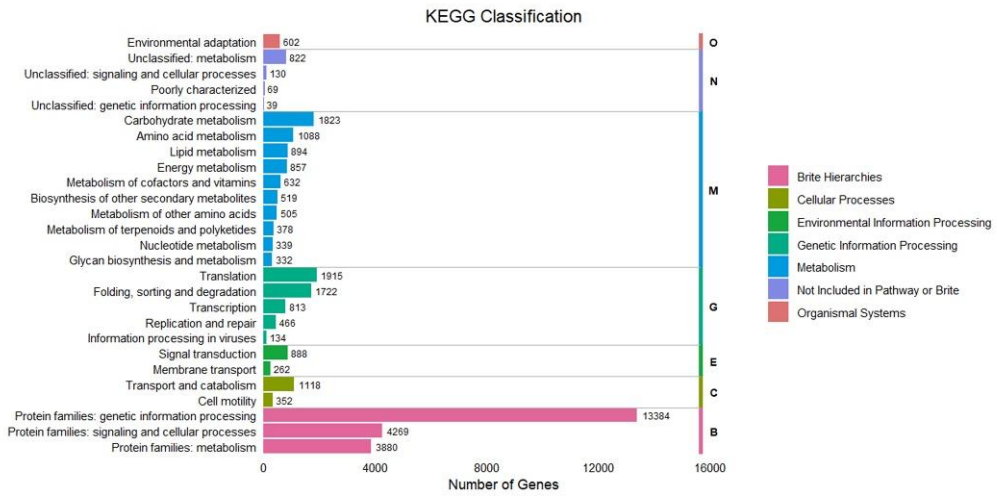
d



e



f



335

Figure 3 Overview of the *Actinidia arguta de novo* transcriptome assembly. (**a**) *Actinidia arguta* transcriptome sequencing data, (**b**) numerical summary of the assembled transcriptome, (**c**) percentage of unigenes annotated to different species, (**d**) percentage of unigenes annotated in various databases, (**e**) number of unigenes annotated in GO, divided into three main categories, (**f**) number of unigenes annotated to KEGG pathways, broken down into main categories.

De novo transcriptome assembly resulted in 422,213 transcripts assigned to 149,380 unigenes, ranging in length from 301 bp to 16,469 bp, with average lengths of 1,195 bp and 1,054 bp for transcripts and unigenes, respectively. The largest number of transcripts was in the range from 500 - 1,000 bp (Fig. 3b).

Since the sequencing was performed *de novo* and a reference genome was unavailable in existing databases, unigenes were annotated based on sequence similarity to known species. The highest number of unigenes was assigned to two related species: *Actinidia chinensis* (29,213; 39.8%) and *Actinidia rufa* (27,088; 36.9%) (Fig. 3c).

Functional analysis

In order to obtain information on the structure and function of the mapped unigenes, their annotation was performed using the following databases: NCBI, SwissProt/UniProt, PFAM, KOG, GO, and KEGG, employing BLASTx and BLASTp tools with an E-value threshold of less than 1.0×10^{-10} (Fig. 3d).

The unigenes were analysed within the Gene Ontology (GO) framework and classified into three main categories: biological processes (BP; 79,543), molecular functions (MF; 45,239), and cellular components (CC; 37,019) (Fig. 3e). In total, 42,377 unigenes were successfully annotated (Supplementary Material 1: Tab. S1).

Within the BP category, the most highly represented GO terms were: *cellular process* (23,846), *metabolic process* (19,427), *biological regulation* (7,693), *regulation of biological process* (7,128),

localization (6,246), *response to stimulus* (4,749), *signaling* (2,361), *interspecies interaction between organisms* (1,789), and *multicellular organismal process* (1,112).

Among the unigenes annotated in the MF category, the most abundant terms included: *binding* (20,822), *catalytic activity* (16,066), *transporter activity* (3,086), *molecular function regulator* (1,604), *transcription regulator activity* (1,468), and *structural molecule activity* (1,307).

In the CC category, the most frequently represented terms were: *cellular anatomical entity* (17,770), *intracellular* (10,100), and *protein-containing complex* (7,480).

To further investigate the biological functions of the transcripts, functional annotation was conducted using the KEGG database, which enabled the assignment of 24,149 unigenes to 29 pathways, categorized into the following major groups: Brite Hierarchies, Cellular Processes, Environmental Information Processing, Genetic Information Processing, Metabolism, Organismal Systems, and Not Included in Pathway or Brite (Fig. 3f; Supplementary Material 1: Tab. S2).

The most highly represented pathways included: *protein families: genetic information processing*, *protein families: signaling and cellular processes*, *protein families: metabolism*, *translation*, *carbohydrate metabolism*, *folding, sorting and degradation*, *transport and catabolism*, and *amino acid metabolism*.

Differential gene expression analysis

A differential gene expression analysis was performed to investigate which genes may be involved in altering the developmental pathway of mature endosperm.

Based on the *de novo*-assembled transcriptome used as a reference, gene expression was analysed across the following sample comparisons: Bingo 14d vs. 0d (1), Weik 14d vs. 0d (2), Bingo vs. Weiki 0d (3), and Bingo vs. Weiki 14d (4) (Fig.2). Differentially expressed genes (DEGs) were identified based on the log2 Fold-Change (log_2FC) values and statistical significance determined by the adjusted p-value (padj), using a threshold of $|log_2FC| > 1$ and $padj < 0.05$. The largest number of DEGs (35,893) was observed when comparing Weiki 14d vs. 0d, while the fewest were identified when comparing the control samples, Bingo vs. Weiki 0d.

The analysis revealed 35,893 DEGs for the Weiki14d vs. 0d comparison, representing the highest number across all tested conditions. In contrast, the Bingo 14d vs. 0d comparison yielded 29,814 DEGs, a value approximately 1.2 times lower. Considerably fewer DEGs were identified in comparisons between the two cultivars at the same time points: 1,482 for Bingo vs. Weiki 0d and 8,985 for Bingo vs. Weiki 14d.

Differences in the number of statistically significant DEGs may contribute to the observed variation between the studied cultivars in their capacity for endosperm proliferation and, consequently, organogenic callus formation. The distribution of DEGs across the analysed comparisons is presented as volcano plots (Fig. 4). Across all four comparisons, the majority of DEGs were up-regulated. The highest number of up-regulated genes (19,548) was detected in the Weiki 14d vs. 0d comparison (Fig. 4b), whereas the lowest number (881) was observed in Bingo vs. Weiki 0d (Fig. 4c). However, this may also reflect differences not only between the cultivars themselves but also in the duration of the culture.

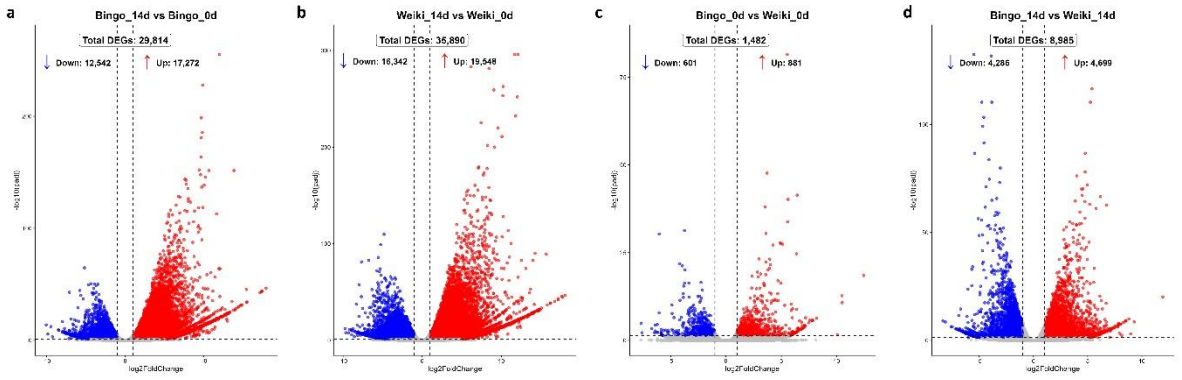


Figure 4 Differentially expressed genes presented in the volcano plots for (**a**) Bingo 14d vs Bingo 0d, (**b**) Weiki 14 d vs Weiki 0d, (**c**) Bingo vs Weiki 0d and (**d**) Bingo vs Weiki 14d; padj <0.05 with the numbers of total, up- and downregulated DEGs (padj ≤ 0.05; log2FoldChange ≥ 1.0).

Venn diagram analysis further demonstrated substantial overlap between Bingo14d vs. 0d and Weiki 14d vs. 0d, with 21,365 common DEGs, most of which were up-regulated (13,520). Interestingly, the uniquely expressed genes within these comparisons were predominantly down-regulated (Fig. 5a). By contrast, Bingo vs. Weiki 0d and Bingo vs. Weiki 14d shared only 555 common genes, of which 360

were up-regulated. On day 14 of culture, Bingo vs. Weiki 14d comparison revealed 8,430 uniquely expressed genes, including 4,339 up-regulated genes (Fig. 5b), which is 9.1 and 8.3 times higher, respectively, than the values obtained for the Bingo vs. Weiki 0d comparison.

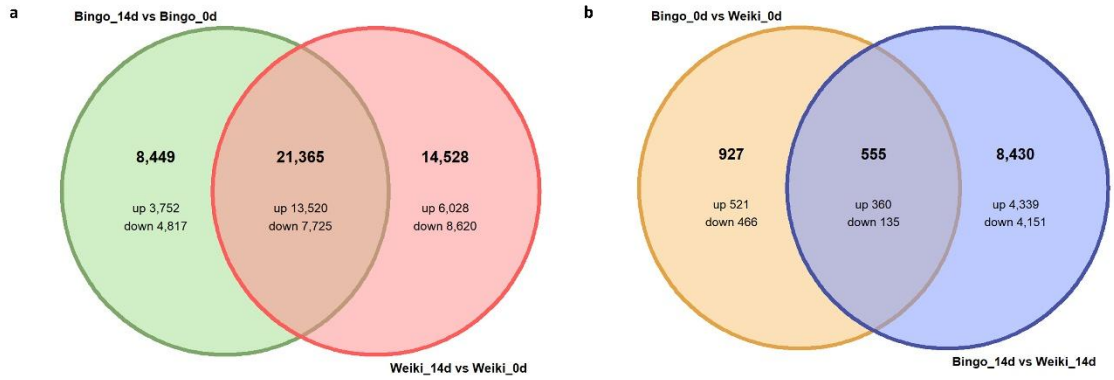


Figure 5 Differentially expressed genes presented in Venn diagrams for (**a**) Bingo 14d vs Bingo 0d vs Weiki 14 d vs Weiki 0d and (**b**) Bingo vs Weiki 0d vs Bingo vs Weiki 14d comparisons.

To gain further insights, GO term enrichment analysis was performed, using Venn diagrams to identify terms that were common or unique to the studied comparisons (Supplementary Material 2). Additionally, a list of transcripts showing differential expression that were common or unique to the compared experimental combinations was prepared (Supplementary Material 3).

The first Venn diagram (Fig. 5a) compares the Bingo14d vs.0d and Weiki14d vs. 0d contrasts, identifying DEGs that are shared between both cultivars after 14 days of culture, as well as those that are cultivar-specific. Both cultivars exhibited a common functional pattern, including increased protein synthesis, as indicated by terms related to “ribosome biogenesis” and ribosomal structure, as well as processes associated with cell wall organization and formation. In addition, shared responses involved enhanced “transmembrane transport” and “microtubule-based movement”, together with the activation of carbohydrate and lipid metabolism. A wide range of enzymatic activities was also observed, including hydrolases, transferases, and oxidoreductases. At the same time, some differences between the cultivars were evident. The 'Weiki' cultivar showed specific enrichment in pathways related to sulfur compound metabolism and modified amino acids. In contrast, the 'Bingo' cultivar was

associated with proteins localized in the “cytosol” and with “motor activity”, suggesting differences in intracellular transport and movement between the two cultivars.

The second Venn diagram (Fig. 5b) compares the Bingo vs. Weiki 0d and Bingo vs. Weiki 14d contrasts, highlighting DEGs that are consistently dependent on the cultivar at both time points, as well as those specific to the 'Bingo' cultivar at particular stages of culture. Only one GO term, “lipid droplet”, was shared between these comparisons, suggesting that lipid storage processes remain relatively stable in this cultivar regardless of culture duration. In contrast, after 14 days of culture, significant changes were observed in the expression of genes involved in lipid and carbohydrate metabolism in the 'Bingo' cultivar. Although both up- and down-regulated DEGs were identified within these pathways, the majority of DEGs, including those encoding hydrolases and oxidoreductases, were down-regulated, suggesting an overall reduction in metabolic activity and energy-related processes. Additionally, the enrichment of the term “transcription regulator activity” points to the activation of regulatory pathways controlling gene expression. These changes are accompanied by structural adjustments within the cell, particularly in the cytosol and cell wall, as indicated by terms such as “cell wall” and “external encapsulating structure”.

Additionally, KEGG pathway enrichment analysis was conducted to examine the metabolic pathways associated with the identified DEGs (Supplementary Material 4). For Bingo vs. Weiki 0d and Bingo vs. Weiki 14d, 16 statistically significant common pathways were found. Among the unique pathways, only one pathway was found - *carotenoid biosynthesis*, for the Bingo vs. Weiki 14d comparison. In contrast, Bingo 14d vs. 0d and Weiki 14d vs. 0d shared enrichment in only three pathways: *oxidative phosphorylation*, *ribosome*, and the *citrate cycle (TCA cycle)*.

Analysis of TRANSCRIPTION FACTORs differential expression

We identified differential expression in 15 families of transcription factors (TFs) across the studied sample combinations (Supplementary Material 5). Among these, we focused our analysis on seven key families: AP2/EREBP, NAC, WRKY, HD-ZIP, MYB, bZIP/ABF, and MADS-box. These families revealed pronounced variations in transcriptional patterns, which were dependent on both the

duration of in vitro culture (0 vs. 14 days) and the genotype. Data visualization via heatmaps (hierarchical clustering, average linkage, Euclidean distance) demonstrated that the primary axis of sample differentiation was culture duration (Supplementary Material 6). This was reflected by the consistent grouping of the Weiki 14d vs. 0d and Bingo 14d vs. 0d comparisons, whereas inter-cultivar comparisons (Bingo vs. Weiki at 0d and 14d) exhibited more dispersed, genotype-specific expression patterns.

Within the AP2/EREBP family, some of the most dynamic expression changes were observed, with a clear segregation of transcripts into two major clusters corresponding to up-regulated and down-regulated genes. In the Bingo 14d vs. 0d and Weiki 14d vs. 0d comparisons, a strong induction of numerous transcripts ($log_2FC > 5$) was evident. This coherent cluster included *18900(Ath_ERF012)* ('Bingo' $log_2FC = 11.35$; 'Weiki' $log_2FC = 10.17$) and *117494(Ath_ERF003)* ('Bingo' $log_2FC = 11.19$; 'Weiki' $log_2FC = 9.37$). Notably, at the initial time point (0d), distinct groups of genes with reduced expression were identified in 'Bingo' relative to 'Weiki', indicating that the two cultivars possess different baseline transcriptional states prior to culture exposure. A marked down-regulation was recorded for *119363(Ath_ERF1A)* ($log_2FC = -5.74$), which persisted through day 14 of the culture ($log_2FC = -3.38$). Furthermore, a significant up-regulation of two genes, *73090(Ath_ERF109)* and *2480(Ath_ERF098)*, was observed in the Bingo vs. Weiki 14d combination ($log_2FC = 9.29$ and 6.31, respectively), while their expression was simultaneously strongly reduced in the Weiki 14d vs. 0d comparison ($log_2FC = -8.98$ and -5.58, respectively). The heatmap structure further suggests the co-regulation of AP2/EREBP gene subsets whose activation correlates with culture progression, consistent with their established roles in stress responses and developmental reprogramming.

The NAC family exhibited a relatively uniform expression profile, characterized by a predominant trend of strong induction at day 14. The heatmap revealed a major cluster of up-regulated transcripts, including *11744(Ath_NAC043)* and *11312(Ath_NAC090)*. In the Weiki 14d vs. 0d combination, specific induction was noted for *12074(Ath_NAC043)* and *121580(Gmax_NAC6)*. In a smaller, distinct cluster for the Bingo 14d vs. 0d combination, the up-regulation of several transcripts was observed, with the highest values reached by *18220(Ath_NAC056)* ($log_2FC = 7.06$) and *36795(Ath_NAC056)* ($log_2FC = 6.97$). Interestingly, a strong up-regulation of two transcripts was

recorded: *525(Ath_NAC090)* ($log_2FC = 6.03$) for the Bingo vs. Weiki 0d combination, and *18831(Ath_NAC083)* ($log_2FC = 6.50$) for Bingo vs. Weiki 14d. This may indicate a significant role for these genes in the processes occurring in fresh 'Bingo' endosperm, potentially determining its further development in in vitro cultures. In contrast to AP2/EREBP, the number of strongly downregulated NAC genes was limited, suggesting they primarily function as activators of repair and morphogenetic processes.

In the WRKY family, complex and contrasting expression patterns were detected. Heatmaps revealed a large cluster of strongly induced transcripts - led by *Ath_WRKY9*, *Ath_WRKY14*, and *Ath_WRKY44* in Bingo 14d vs. 0d, and *Ath_WRKY71* in Weiki 14d vs. 0d - alongside genes exhibiting reduced expression after 14 days, particularly in 'Weiki'. The segregation of samples according to genotype highlights a significant cultivar-specific component in WRKY regulation, reflecting their role in integrating environmental and stress-related signals.

The HD-ZIP family displayed a more structured clustering pattern, with a clear distinction between activated genes (e.g., *Ath_ATHB-52* and *Ath_HAT7* in 'Weiki' 14d vs. 0d) and repressed genes (e.g., in Bingo vs. Weiki 14d, $log_2FC = -5.86$). The clarity of these clusters suggests coordinated regulation of HD-ZIP genes in controlling tissue organization and the formation of new structures.

The MYB family exhibited the highest degree of heterogeneity. In the 14 vs. 0 comparisons, a coherent group of strongly up-regulated transcripts included *Amaj_MYB306*, *Achi_MYB123*, and *Ath_RAX3*. A cultivar-specific feature of 'Bingo' was the activation of an additional subset of genes (e.g., *12642(Ath_MYB93)* and *14027(Ath_MYB35)*, as well as *115728(Ath_MYB105)* and *8676(Ath_MYB108)*), which were simultaneously down-regulated in the Weiki 14d vs. 0d combination, indicating a distinct trajectory of metabolic reprogramming in this cultivar.

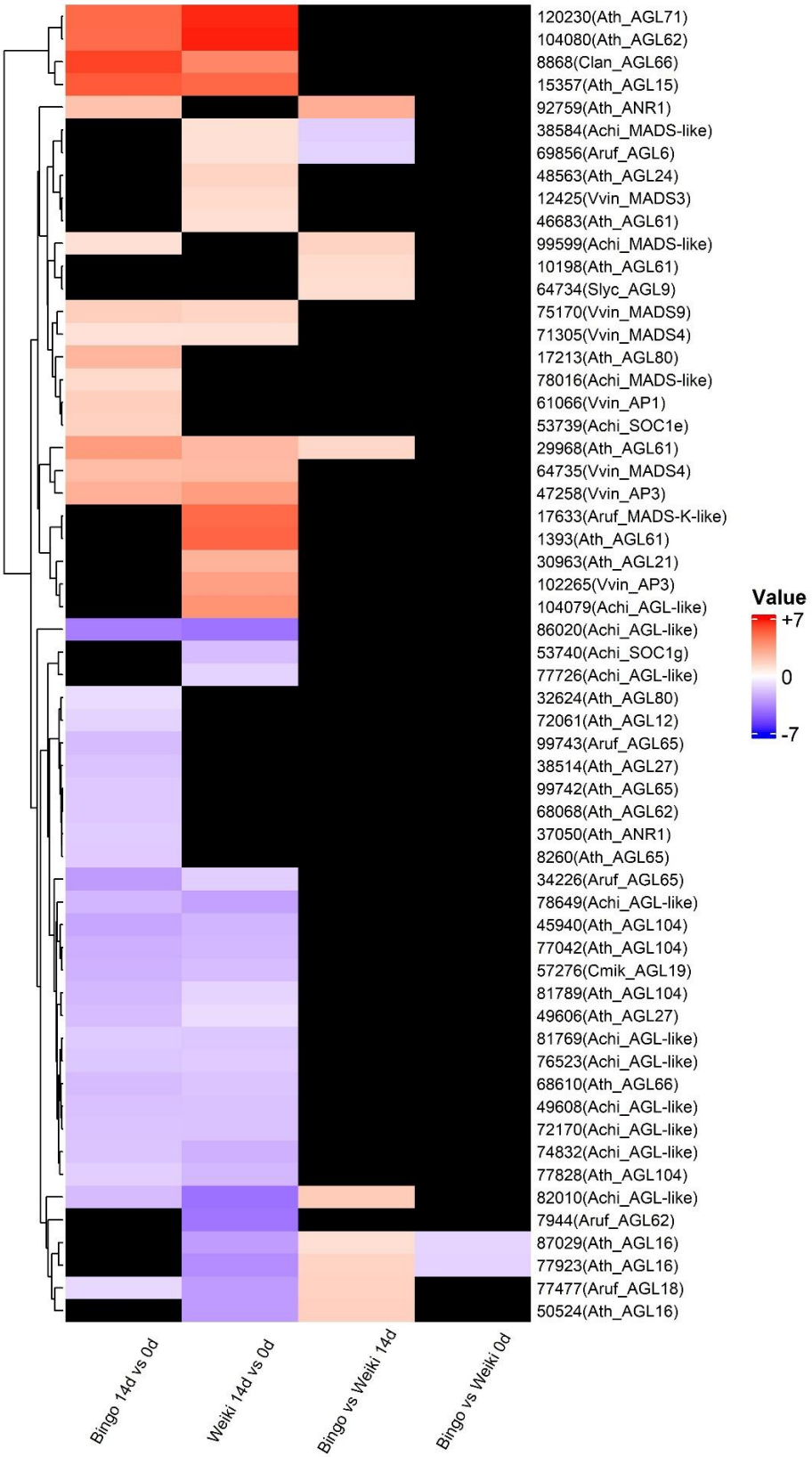
Despite a smaller number of transcripts, the bZIP/ABF family exhibited functionally coherent expression patterns, particularly regarding ABA-mediated responses. Significant upregulation at day 14 was observed for *Ath_BZIP11*, *Ath_FD*, and *Ath_BZIP44*. Additionally, increased expression of two transcripts was recorded for the Bingo 14d vs. 0d combination (*15810(Ath_BZIP34)*) and the Bingo vs. Weiki 14d combination *(47715(Ath_ABI5)*). The only transcript showing up-regulation in the Bingo vs. Weiki 0d combination was *19580(Ath_BZIP44)* ($log_2FC = 4.58$), which also showed

increased expression in the Weiki 14d vs. 0d combination ($log_2FC = 6.32$). This suggests that this family may act as a key sensor of osmotic stress during callus formation.

Hierarchical clustering of differentially expressed MADS-box genes revealed significant transcriptional changes in endosperm cultured in vitro that depend on both time and genotype (Fig. 6). The most significant expression differences were observed in temporal comparisons, suggesting that the transition from the initial state (0 days) to 14 days of culture primarily shapes the transcriptional profile of MADS-box genes. In both cultivars, a distinct group of genes showed strong downregulation over time, displaying relatively high expression at day 0 but markedly reduced expression after 14 days. Conversely, another cluster comprised genes that were strongly induced during culture, characterized by low or undetectable expression at 0 days and substantial up-regulation at 14 days. These temporal changes were more pronounced in the 'Bingo' cultivar, which exhibited a greater number of strongly induced genes and more extensive repression of genes initially expressed in freshly isolated endosperm. Between-cultivar comparisons at 0 days showed relatively minor differences, with only two transcripts - both annotated as *87029(Ath_AGL16)* and *77923(Ath_AGL16)* - showing significant down-regulation. However, by day 14, divergence became evident. Most DEGsshowed upregulation in 'Bingo', including: *77477(Aruf_AGL18)*, *50524(Ath_AGL16)*, *87029(Ath_AGL16)*, *77923(Ath_AGL16)*, *82010(Achi_AGL-like)*, *29968(Ath_AGL61)*, *10198(Ath_AGL61)*, *64734(Ath_AGL9)*, *99599(Achi_MADS-like)*, and *92759(Ath_ANR1)*. Two transcripts were down-regulated: *38584(Achi_MADS-like)* and *69856(Aruf_AGL6)*. Comparison of 14-day cultured endosperm with fresh endosperm revealed cultivar-independent down-regulation of transcripts such as *77477(Aruf_AGL18)*, *77828(Ath_AGL104)*, *68610(Ath_AGL66)*, *49606(Ath_AGL27)*, and *57276(Cmik_AGL19)*. In contrast, increased expression was observed for *64735(Vvin_MADS4)*, *47258(Vvin_AP3)*, *29968(Ath_AGL61)*, *75170(Vvin_MADS9)*, *15357(Ath_AGL15)*, and *104080(Ath_AGL62)*. Genotype-specific DEGs were also identified. In 'Bingo' (14d vs. 0d), unique down-regulation occurred in *37050(Ath_ANR1)*, *8260(Ath_AGL65)*, and *32624(Ath_AGL80)*, while induction was noted for *61066(Vvin_AP1)* and *53739(Achi_SOC1e)*. In 'Weiki', unique up-regulated DEGs included

102265(Vvin_AP3), *30963(Ath_AGL21)*, *1393(Ath_AGL61)*, and *48563(Ath_AGL24)*, while *7944(Aruf_AGL62)* and *53740(Achi_SOC1g)* were down-regulated.

Overall, culture duration exerted a dominant influence on transcriptional profiles, while significant genotype-specific effects remained evident. The induction of genes from the MYB, WRKY, and AP2/EREBP families suggests that they constitute a core regulatory module supporting stress adaptation, hormonal signaling, and the transcriptional reprogramming required to initiate dedifferentiation. Meanwhile, the divergent signatures of MADS-box, NAC, and bZIP/ABF genes may reflect enhanced cellular plasticity and a more robust capacity for forming regenerative centers. These distinct transcription factor profiles ultimately contribute to the variations in developmental potential observed between the two cultivars during long-term in vitro culture.



552

553 **Figure 6** Heatmap of MADS-box genes differentially expressed in endosperm tissue in 'Bingo' and

554 'Weiki' cultivars after 0 and 14 days of culture.

MADS-box family genes – differential gene expression analysis and functional annotation

Among the transcripts showing differential expression, we identified those annotated as genes belonging to the MADS-box family. Many of these annotations corresponded to genes previously described in the model species *Arabidopsis thaliana*, as well as in species belonging to the genus *Actinidia, including Actinidia chinensis*, *Actinidia rufa* and others.

Gene names were constructed according to the following format. For example: *15357(Ath_AGL15)* – where 15357 represents the transcript identifier, 'Ath' is the abbreviation for the species *Arabidopsis thaliana*, and AGL15 is the transcript annotation corresponding to the gene as recorded in databases. Based on information from the UniProt database, we classified these genes into five groups according to their function: regulation of flowering, flower development, endosperm development, root growth and other functions (Fig. 7; Supplementary Material 7). The resulting distribution reflects the diverse regulatory roles of these transcription factors throughout plant development.

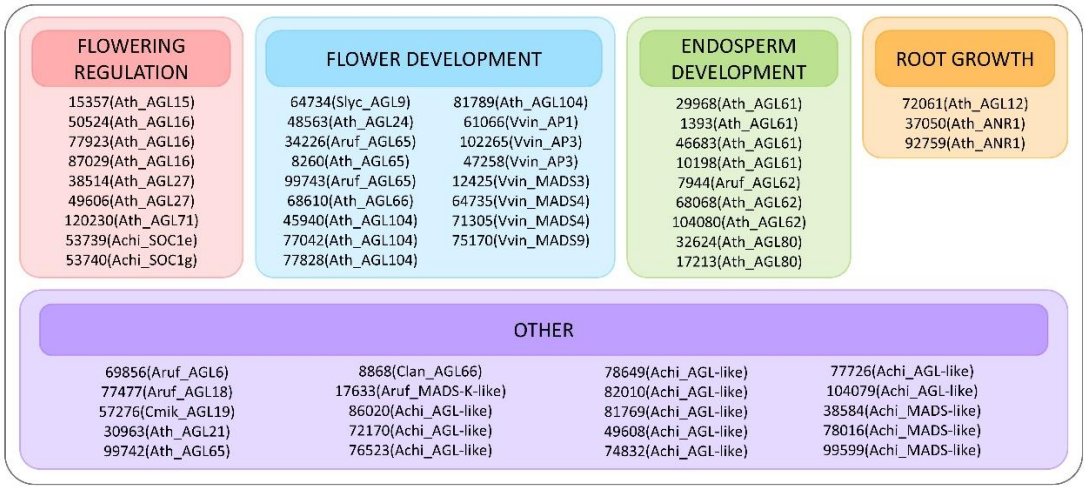


Figure 7 The classification of annotated *Actinidia arguta* transcripts based on their role in plant development.

The first group comprises genes associated with the regulation of flowering. Nine transcripts were assigned to six different genes, four of which were identified in *Arabidopsis thaliana* and two in *Actinidia chinensis*. The encoded MADS-box proteins function as transcription factors that govern the

transition from the vegetative to the reproductive phase and the timing of floral induction. Annotations indicate their involvement in the regulation of RNA polymerase II-dependent gene expression, as well as in embryo development and the establishment of seed dormancy, highlighting the broad regulatory scope of this class.

The second category consists of genes involved in flower development. Seventeen transcripts were annotated to eleven different genes. Five of these genes were assigned to *Arabidopsis thaliana*, another five to *Vitis vinifera*, and only one to *Actinidia rufa*. These gene products serve as key regulators of floral organ identity and inflorescence architecture. Functionally, they belong to the classical Type II MADS-box factors, which control tissue differentiation programs and reproductive organ morphogenesis through the formation of multimeric protein complexes.

Another group includes genes related to endosperm development. A total of nine transcripts were annotated to three genes within this group. Four transcripts were assigned to *AGL61*, three to *AGL62*, and two to *AGL80*. Nearly all genes were identified in *Arabidopsis thaliana*, with the exception of one gene – *AGL62* – which was identified in *Actinidia rufa*. Functional annotations point to their roles in the differentiation of the central cell within the embryo sac, the regulation of endosperm cell divisions, and the control of early seed development. Many of these genes belong to the M-type (Type I) MADS-box group, which is characteristically associated with gametophytic and seed-specific regulatory functions.

The fourth category is composed of genes participating in root growth. The smallest group consisted of three transcripts assigned to two different genes of *Arabidopsis thaliana*. Their products regulate meristematic cell differentiation, cell elongation, and the developmental transition of the plant, suggesting a functional integration of root system development with the overall plant ontogeny.

The final group, designated as others, encompasses MADS-box genes with less clearly defined functions or potentially pleiotropic regulatory effects. In total, 20 transcripts were classified and assigned to genes from five different plant species: *Arabidopsis thaliana*, *Actinidia rufa*, *Actinidia chinensis*, *Cinnamomum micranthum*, and *Camellia lanceoleosa*. While annotations indicate their

involvement in general transcriptional regulation and transcription factor activity, there is a lack of definitive data assigning them to a specific developmental process.

We searched the identified DEGs for genes potentially involved in endosperm development and dedifferentiation processes under controlled *in vitro* conditions. In this study, we focused on genes belonging to the third group ("Endosperm development"), particularly Type I MADS-box genes. We also considered two genes belonging to different groups: "Flower development" and "Flowering regulation", that, according to the literature, may affect endosperm development: *64734(Sl_AGL9)* and *15357(At_AGL15)*.

Analysis of the differential expression of *AGL* family genes in the samples indicated diverse transcriptional patterns related to both the 'Bingo' and 'Weiki' cultivars and the time frame (Tab. 1; detail information in Supplementary Material 8).

Comparing 'Bingo' 14d with 0d revealed a clear induction in the expression of several genes, including *15357(Ath_AGL15)* ($log_2FC = 5.59$), *29968*(*Ath_AGL61)* ($log_2FC = 3.58$), *104080(Ath_AGL62)* ($log_2FC = 5.11$), and *17213(Ath_AGL80)* ($log_2FC = 2.65$). Conversely, the same comparison shows repression of genes such as *7944(Aruf_AGL62)* ($log_2FC = -4.16$) and *68068(Ath_AGL62)* ($log_2FC = -1.65$).

Analysis of Weiki 14d vs 0d shows a similar pattern. The strongest induction was observed in the *104080(Ath_AGL62)* gene ($log_2FC = 6.78$). Additionally, clear overexpression of *15357(Ath_AGL15)* ($log_2FC = 5.24$) and *AGL61* ($log_2FC = 1.18 – 5.32$ depending on paralogues) was observed in 'Weiki'.

A comparison between the cultivars after 14 days (Bingo vs Weiki 14d) reveals more subtle differences. Notably, *64734(Slyc_AGL9)* ($log_2FC = 1.22$), *29968*(*Ath_AGL61)* ($log_2FC = 1.46$), and *10198(Ath_AGL61)* ($log_2FC = 1.33$) showed slightly higher expression levels in 'Bingo' than in 'Weiki'.

Differential expression analysis revealed the presence of genes unique to both the 'Bingo' and 'Weiki' cultivars. In the case of 'Weiki', these included *1393(Ath_AGL61)* and *46683(Ath_AGL61)*, which were upregulated, as well as *7944(Aruf_AGL62)*, which was downregulated. The 'Bingo' cultivar showed decreased expression of *68068(Ath_AGL62)* and *32624(Ath_AGL80)*, and increased expression of *17213(Ath_AGL80)*.

In summary, these results demonstrate that *AGL* family genes undergo dynamic regulation depending on the cultivar and time frame. The induction of *AGL15* and *AGL61* paralogues in both cultivars after 14 days suggests that they may be important regulators of key developmental or adaptive processes. Differences in gene types and their expression patterns observed between the studied cultivars may be associated with differential responses of the endosperm to culture conditions and variations in their capacity for dedifferentiation.

635 **Table 1** Selected AGAMOUS-like genes potentially involved in endosperm dedifferentiation and
636 proliferation identified for Bingo 14d vs 0d, Weiki 14d vs 0d and Bingo vs Weiki 14d combinations
637 with different expression level

Gene ID	*De novo* transcription annotation (description / species)	**Fold change (log2)** | $p_{adj} < 0.05$		
		Bingo 14d vs 0d	Weiki 14d vs 0d	Bingo vs Weiki 14d
UP-REGULATED TRANSCRIPTS				
Cluster-5580.64734	Agamous-like MADS-box protein AGL9 homolog / *Solanum lycopersicum*			**1.22** | 6.5496e-05
Cluster-5580.15357	Agamous-like MADS-box protein AGL15 / *Arabidopsis thaliana*	**5.59** | 4.8476e-10	**5.24** | 5.9565e-09	
Cluster-5580.29968	Agamous-like MADS-box protein AGL61 / *Arabidopsis thaliana*	**3.58** | 4.229e-24	**2.53** | 6.0156e-13	**1.46** | 9.6926e-06
Cluster-5580.1393	Agamous-like MADS-box protein AGL61 / *Arabidopsis thaliana*		**5.33** | 0.01221	
Cluster-5580.46683	Agamous-like MADS-box protein AGL61 / *Arabidopsis thaliana*		**1.18** | 0.013932	
Cluster-5580.10198	Agamous-like MADS-box protein AGL61 / *Arabidopsis thaliana*			**1.34** | 0.033306
Cluster-5580.104080	Agamous-like MADS-box protein AGL62 / *Arabidopsis thaliana*	**5.11** | 3.8482e-14	**6.78** | 8.1466e-21	
Cluster-5580.17213	Agamous-like MADS-box protein AGL80 / *Arabidopsis thaliana*	**2.65** | 0.006041		
DOWN-REGULATED TRANSCRIPTS				
Cluster-5580.7944	AGAMOUS-like 62 / *Actinidia rufa* Agamous-like MADS-box protein AGL29 / *Arabidopsis thaliana*		**-4.17** | 0.022319	
Cluster-5580.68068	Agamous-like MADS-box protein AGL62 / *Arabidopsis thaliana* Agamous-like MADS-box protein AGL62 / *Camellia sinensis*	**-1.65** | 0.043549		
Cluster-5580.32624	Agamous-like MADS-box protein AGL80 / *Arabidopsis thaliana*	**-1.06** | 0.020385		

A comparative analysis of homologous protein sequences corresponding to *AGL* gene transcripts from *Actinidia arguta* revealed a high level of similarity with sequences derived from closely related species, particularly *Actinidia rufa* and *A. chinensis* (Supplementary Material 9). Most alignments were characterized by very low E-values (in the range of 10^{-30} or lower), indicating statistically significant sequence similarity. In a few cases, the percentage identity (Per. Ident.) exceeded 80–90%, especially for transcripts associated with *AGL15, AGL62*, and *AGL80*, suggesting a high degree of sequence conservation. Other alignments showed moderate identity levels (30–60%) and slightly higher E-values, which may reflect the presence of functionally distinct paralogs or evolutionary divergence among homologs.

Further comparative analysis of the identified homologous proteins focused on their molecular and functional properties. The lengths of the aligned amino acid sequences varied widely – from 172 to 471 residues – with most proteins falling within the 200–300 amino acid range, which is typical for members of the MADS-box transcription factor family. Protein molecular masses correlated with sequence lengths and ranged from 19.8 to 48.5 kDa.

All analysed proteins were assigned to the MADS-box transcription factor family. Domain analyses performed with InterPro, Pfam, and SMART confirmed the presence of conserved characteristic motifs, including the MADS-box, K-box, and MEF2-like domains. The structural models available in the AlphaFold database revealed a high degree of similarity among the proteins and preservation of key interaction interfaces. Protein-protein interaction data obtained from the STRING database further suggested that these AGL proteins are involved in complex transcriptional regulatory networks, although current experimental evidence is largely based on the *Arabidopsis thaliana* model.

Subcellular localization predictions indicated that all identified proteins are predominantly nuclear, consistent with their regulatory functions; however, AGL15 and AGL61 may undergo dynamic nucleocytoplasmic shuttling. Expression profiling demonstrated distinct spatiotemporal patterns: transcripts of AGL61, AGL62, and AGL80 were specifically detected in the central cell of the embryo sac and in early endosperm, whereas AGL9 and AGL15 showed broader expression across floral

664 organs, embryos, and young seedlings. Notably, AGL15 expression was inducible by the auxin 2,4-D,
665 pointing to a hormonally regulated mechanism of transcriptional control.
666
667 ## *AGL* transcripts – protein alignment and domain architecture
668 A multiple sequence alignment of the translated amino acid sequences of the selected transcripts and
669 their homologues was performed using MAFFT (Supplementary Material 10). The alignment focused
670 on regions of the highest conservation and sequence homology, with the highly divergent N- and C-
671 terminal flanking regions truncated to enhance clarity and the reliability of the comparative analysis.
672 The resulting alignment was of high quality, with the analysed transcripts showing homology to
673 established reference sequences. This high degree of conservation supports the accurate identification
674 and annotation of these transcripts as members of the MADS-box gene family. A highly conserved
675 MADS domain (comprising approximately 55–60 amino acids) was identified at the N-terminus, the
676 defining feature of this class of transcription factors.
677 Further structural analysis revealed distinct domain architectures among the sequences. In several
678 transcripts, the MADS domain was followed by an intervening domain (I) and a well-conserved
679 keratin-like domain (K), which is essential for protein dimerisation and protein–protein interactions.
680 The presence of this M-I-K domain organisation confirms that these genes belong to Type II
681 (MIKC-type) of the MADS-box family. Conversely, a subset of proteins lacked the K domain and
682 featured significantly shorter, more divergent C-terminal regions. These proteins possessed only the
683 MADS domain, a configuration typical of Type I (M-type) genes.
684 The alignment revealed notable differences in the length and conservation of the C-terminal regions
685 between the two types. MIKC-type proteins exhibited longer structured regions corresponding to the
686 K domain, whereas M-type proteins were characterised by higher sequence variability and the
687 absence of this motif.

Phylogenetic tree of AGLs potentially involved in endosperm development

Phylogenetic analysis enabled the classification of *Actinidia arguta* transcripts belonging to the MADS-box gene family (Fig. 8). Based on the Maximum Likelihood (ML) reconstruction, the phylogenetic tree revealed several distinct clades corresponding to homologs of AGL6, AGL15, AGL61, AGL62, AGL80, and SEPALLATA (SEP1/SEP3/AGL9). The inferred phylogenetic relationships showed strong general support, as indicated by high bootstrap values (≥80) in most clades.

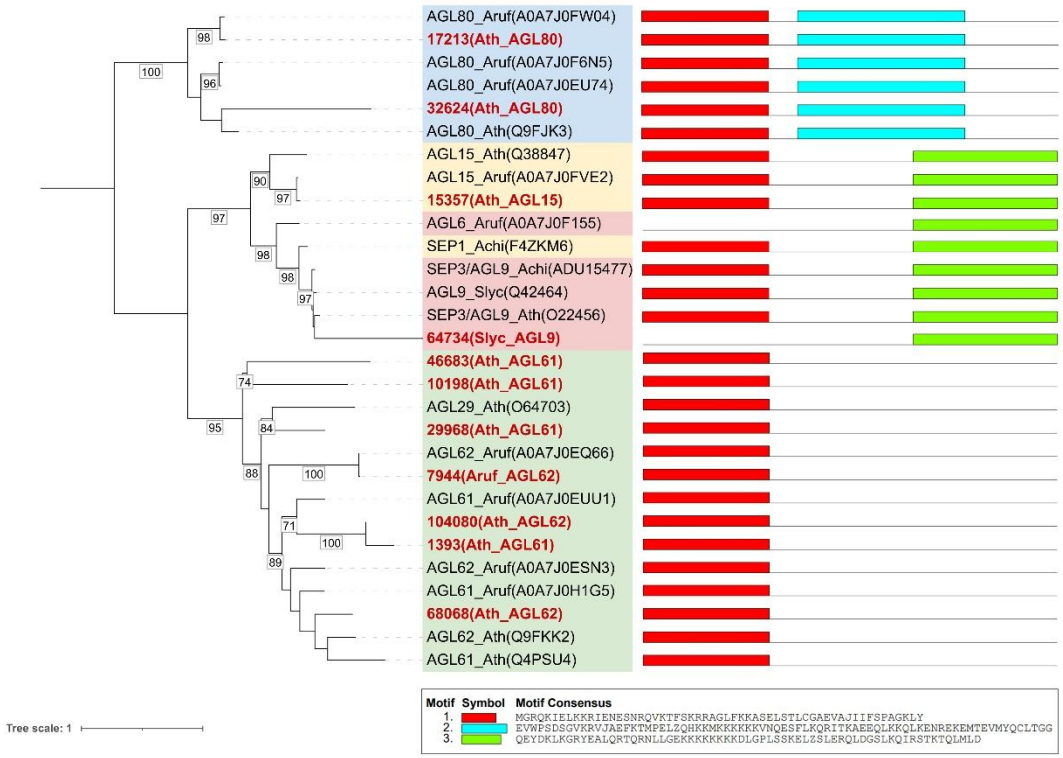


Figure 8 Phylogenetic analysis of 29 sequences, including the analysed *Actinidia arguta* transcripts (highlighted in red) and their homologues, alongside their predicted conserved motifs identified using MEME. The phylogenetic tree illustrates the relationships among the analysed sequences, with bootstrap support values indicated at the corresponding nodes. Three motifs were identified: motif 1 (red), the shortest motif, present in most sequences; motif 2 (blue), the longest motif, detected only in AGL80-related sequences; and motif 3 (green), detected in AGL9- and AGL15-related sequences.

The largest cluster comprised members of the AGL29/AGL61/AGL62/ subfamily. *A. arguta* transcripts assigned to AGL62, such as transcript 7944(Aruf_AGL62), clustered closely with AGL62 orthologs from related *Actinidia* species and *Arabidopsis thaliana*, supporting their orthologous relationships. Similarly, transcripts assigned to AGL61 formed well-supported clusters with the corresponding reference sequences, indicating a high degree of evolutionary conservation within this lineage. Notably, the close phylogenetic proximity and partial intermixing of AGL61 and AGL62 sequences within this cluster likely reflect their shared ancestry and close evolutionary relationship within the Type I MADS-box gene family.

A distinct and well-supported clade was identified for AGL80 homologs. The close clustering of *A. arguta* transcripts 32624(Ath_AGL80) and 17213(Ath_AGL80) with AGL80 sequences from other taxa suggests a high degree of evolutionary conservation within this subfamily. Another region of the phylogenetic tree comprised regulators associated with floral organ development, including AGL6, AGL15 and member of the SEPALLATA group, such as SEP1, SEP3, and AGL9. In particular, transcript 64734(Slyc_AGL9) clustered within the SEP/AGL9 clade together with well-characterized homologous, supporting its classification within this subfamily.

Bootstrap analysis based on 1000 replicates provided strong support for the major clades and helped assess the consistency of the phylogenetic reconstruction. The deeper nodes anchoring the principal lineages, particularly the AGL80 clade (100% bootstrap support) and the SEP/AGL6/AGL15 group (97% bootstrap support), were strongly supported. In contrast, several nodes within the AGL61/AGL62 cluster displayed lower to moderate bootstrap support, indicating that the relationships among some sequences within this group were less consistently resolved. Nevertheless, the overall agreement between the bootstrap consensus and the maximum likelihood (ML) tree topology further supports the stability of the major relationships inferred.

Furthermore, domain analysis revealed three distinct structural features among the analysed proteins . A highly conserved MADS domain (indicated in red) was formally identified via InterPro in all sequences, confirming their classification as members of the MADS-box transcription factor family. The K-box domain (indicated in green) was detected exclusively in proteins belonging to the

AGL6/SEP and AGL15 lineages, supporting their classification as canonical MIKC-type or Type II MADS-box proteins. In contrast, the K-box domain was absent in the AGL61/AGL62 and AGL80 subfamilies, consistent with their classification as Type I MADS-box proteins, which generally lack the characteristic K domain. A third domain (highlighted in light blue) was detected specifically in AGL80-related sequences. InterPro annotated part of this region as a coiled-coil structural motif, however, its precise structural and functional significance remains unclear and requires further investigations.

Three-dimensional structural modelling of AGL proteins

To further validate the automated annotation and homologs identified via NCBI BLASTx searches, three-dimensional structural models were generated for selected AGL protein sequences. The 64734(Slyc_AGL9) transcript was structurally verified using the experimentally determined AGL9 structure (PDB entry 8CRA, chain 2) as a reference (see Fig. 9). The predicted model exhibited high structural similarity to the experimental template, with clear alignment across the main helical regions (Fig. 9a). Furthermore, key residues involved in complex interactions within the 8CRA structure (M150, L154, L157, L164 and L171) were highly conserved in the predicted protein and mapped to positions M180, L184, L187, L194 and L201, respectively. Structural alignment also confirmed that the predicted protein (salmon) covers a fragment corresponding to chain 2 (green) of the 8CRA template, while the rest of chain 2 (light blue) and its interacting partner (chain 1, dark blue) were not included in the model (Fig. 9b). Collectively, these structural insights strongly support the classification of 64734(Slyc_AGL9) as an AGL9 homologue.

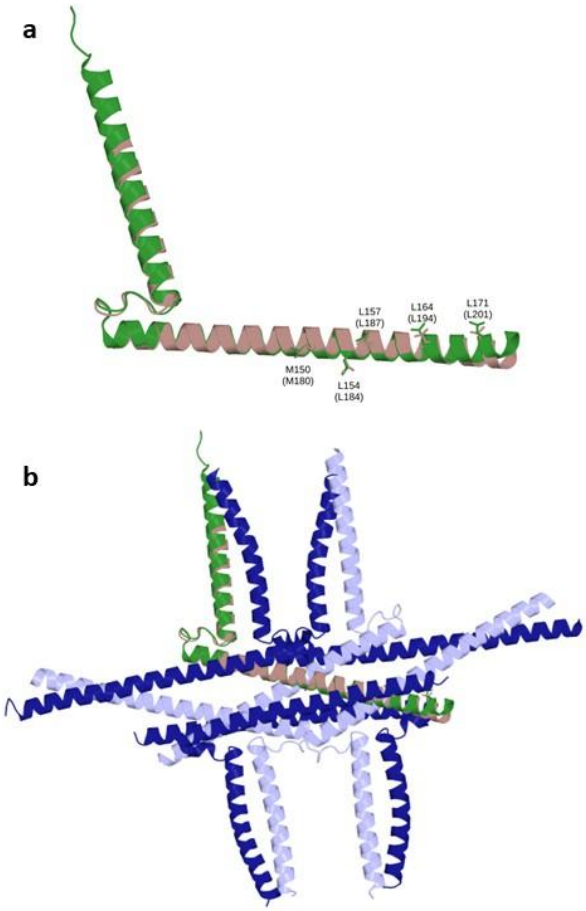


Figure 9 Three-dimensional structural analysis of AGL9. (**a**) Structural comparison of the experimentally resolved AGL9 structure (PDB entry 8CRA, chain 2; green) and the predicted structure of 64734(Slyc_AGL9) (salmon). Residues previously identified as involved in binding in the 8CRA structure are indicated, with the corresponding residues in the predicted structure given in parentheses. (**b**) Structural comparison showing the position of the predicted AGL9 structure (salmon) relative to the complete 8CRA structure. The matching fragment of chain 2 is shown in green, the remaining portion of chain 2 in light blue, and chain 1 in dark blue.

The remaining AGL transcripts were structurally modelled using AlphaFold (Supplementary Material 11). Overall, the predicted structures showed a high degree of correspondence with the available reference structures, particularly within the central regions containing conserved domains, such as the MADS-box. In contrast, greater structural variation was observed in the N- and C-terminal regions.

Several cases, however, showed limited structural correspondence. For the two AGL62 transcripts, 68068(Ath_AGL62) and 104080(Ath_AGL62), discrepancies in the sequence alignment were reflected in pronounced structural differences, which prevented reliable superimposition with the corresponding reference structure. Similarly, the 32624(Ath_AGL80) transcript showed structural correspondence only within a short region. The limited structural similarity, together with the pronounced divergence observed in the phylogenetic analysis, indicates that the structural features of this protein should be interpreted with caution. Nevertheless, the combined phylogenetic and structural analyses provided valuable insights into its evolutionary relationships and supported the overall characterisation of the studied AGL proteins.

DISCUSSION

The present study investigates the molecular features associated with the cell fate switch of mature endosperm in two cultivars of hardy kiwifruit (*Actinidia arguta*), 'Bingo' and 'Weiki'. Transcriptomic data revealed distinct gene expression profiles for each sample type (freshly isolated tissue versus explants after two weeks of culture) and cultivar. Particular attention was given to genes specifically expressed in 'Bingo', owing to its more effective endosperm dedifferentiation and callus proliferation induction compared to 'Weiki'. To our knowledge, this is the first report concerning the transcriptional profiles of mature endosperm undergoing reprogramming.

Previously, we demonstrated that defined culture conditions (including the isolation of the endosperm, the removal of the seed coat and embryo, and the use of Murashige and Skoog medium supplemented with TDZ as exogenous PGR) effectively unblock the restriction on endosperm cell fate switching in *A*. *arguta*, leading to callus proliferation and adventitious shoot bud regeneration [50, 51]. In the present study, two-week-old explants selected for transcriptomic analysis exhibited macroscopically visible changes compared to freshly isolated endosperm. As demonstrated by Sala-Cholewa et al. [49], a distinct enlargement of explants was detectable as early as seven days of culture. Furthermore, histological analysis confirmed that dedifferentiation and initial cell divisions occurred within one to two weeks of culture. Based on these observations, we hypothesized that the transcriptional profiles of

791 the studied explants would be significantly altered compared to the freshly isolated mature endosperm
792 used as the starting explants.

793 There are limited reports concerning the changes in isolated endosperm tissue under culture
794 conditions that alter the fate of this specific storage tissue. Most studies focus on the optimization of
795 culture protocols, including the type of basal medium, plant growth regulators (PGRs), the
796 developmental stage of the isolated endosperm (mature or immature) and the efficiency of the explant
797 as indicated by callus proliferation and organogenesis [10–12]. Some biochemical and histological
798 studies on the development of isolated endosperm have been conducted on *Passiflora* [76, 77] and
799 *Actinidia* [49]. To the best of the present author's knowledge, no reports have been published on
800 molecular studies exploring the reprogramming of mature endosperm through dedifferentiation and
801 proliferation pathways.

802 ***Activation of genes involved in cellular processes***

803 Dedifferentiation induces profound structural remodeling within the plant cell. Storage organelles are
804 broken down to mobilize nutrient reserves, while the nucleus and nucleoli enlarge, signaling enhanced
805 ribosomal RNA synthesis [78]. As stated in the previous histological study, storage reserves – visible
806 as protein bodies and lipid droplets – are mobilized during the first weeks of culture in the mature
807 endosperm tissue of *A*. *arguta* [49]. This aligns with scientific literature reporting that seeds of the
808 *Actinidia* genus are generally rich in proteins and lipids. However, the vast majority of fats are located
809 within the endosperm [79–81]. In the present work unigenes associated with the term "lipid droplet"
810 – which are responsible for the synthesis and structural integrity of protein-coated lipid droplets –
811 were predominantly expressed in the cultivar 'Bingo', both in freshly isolated endosperm and in
812 cultured explants. While these subcellular structures are best known as the primary storage
813 compartments in seeds, they have also been heavily implicated in plant stress response [82].

814 The processes of structural adjustments prepare the cell for further metabolic and mitotic action. In
815 the present study, majority of expressed genes common to both cultivar after culturing were
816 categorized into the classes of the "transcription regulator activity", "ribosome biogenesis", "cell wall

organization", "transmembrane transport", "carbohydrate metabolism", and "catalytic activity" It is noteworthy that genes involved in "protein processing in the endoplasmic reticulum", "MAPK signaling pathway", "regulation of transcription", "transcription regulator activity" and "carbohydrate metabolic process" showed higher activity in cultured 'Bingo' explants. This may be related to shoot regeneration, which has been reported for this cultivar after several weeks of culture [49–51]. The shoot regeneration process from callus is influenced by a number of factors, such as increasing the accessibility of gene loci associated with pluripotency and shoot fate determination [83]. The process of dedifferentiation in plant cells is characterized by a series of complex mechanisms that enable the reversion of a differentiated cell to a stem-like cell [78, 84, 85]. The key mechanism involves the modulation of gene expression followed by the reorganization of chromosome composition. The condensed areas of heterochromatin undergo relaxation, resulting in the formation of transcriptionally active euchromatin. This reactivation of previously silenced genes is contingent upon epigenetic modifications, including DNA methylation and histone acetylation. The present study revealed a significant number of unigenes annotated to "genetic information processing" categories for both cultivars, including "transcription", "translation", "folding, sorting, and degradation".

A majority of the differentiated plant cells are arrested in the G0 phase of the cell cycle. Dedifferentiation enables these cells to restart the cell cycle and re-enter the phases (G1, S, G2 and M), which include mitosis and cytokinesis [78, 85]. In this study, the presence of unigenes annotated to the category "microtubule-based movement" suggests that processes involved in organelle movement and phragmoplast formation are actively induced [86]. This transition requires an increased metabolic rate to provide the energy and molecular building blocks required for DNA replication and mitosis. Endosperm is the tissue characterized by a high content of accumulated storage reserves, the endosperm must activate "catalytic activity" molecular function, along with "carbohydrate metabolism", "amino acid metabolism", and "lipid metabolism" (KEGG pathways). Furthermore, the identification of unigenes classified under "ribosomes" and "ribosomes biogenesis" points to the robust activation of the cellular machinery involved in protein synthesis. Another process correlated with dedifferentiation – involving the reprogramming and rebuilding of cellular components as well

as the removal of superfluous or damaged macromolecules and organelles – is autophagy [87]. In the present study, the up-regulation of unigenes related to the "phagosome" was observed for both cultivars after two weeks of culture.

In the present study, we observed the activity of unigenes related to the "phenylpropanoid biosynthesis" and "flavonoid biosynthesis" categories in cultured endosperm tissue. The phenylpropanoid biosynthesis pathway is associated with the biosynthesis of various secondary metabolites, including lignin and flavonoids [88]. Flavonoids exhibit diverse biological functions, such as protection against ultraviolet radiation, signaling, auxin transport, and participation in stress responses [89]. Furthermore, research indicates that flavonoids, including quercetin, are present in high concentrations in vegetative tissues of *A. arguta* [90]. It is possible that the activation of these genes initiates following the reprogramming of endosperm tissue towards callus formation.

Range of transcription factors engaged in the endosperm dedifferentiation

Transcription factors (TFs) serve as the primary regulators of gene expression, orchestrating multiple functions related to modulation of plant development and stress response. They dictate the activity of other downstream TFs, as well as genes encoding metabolic enzymes and structural proteins. Furthermore, cross-regulation represents a key feature of these gene regulatory networks. TFs control all steps of dedifferentiation [78], particularly chromatin remodeling mediated by epigenetic processes, including DNA methylation and histone modifications (such as acetylation and methylation) [91]. They also govern transcriptional and translational events [92] alongside post-transcriptional mechanisms involving microRNAs [93] and post-translational regulation [94]. Despite their importance to plant biology and basic research, the identification and characterization of TFs driving dedifferentiated endosperm has remained entirely unexplored.

In this study, the activity of a number of transcription factors (TFs) was detected in the analyzed explant samples, indicating the complex regulatory mechanisms are induced during the reprogramming of mature endosperm (Fig. 9). Notably, a shared expression pattern emerged, characterized by the strong upregulation of specific transcripts across all cultured explants. These core

transcripts belong to AP2/ERF-BP, WRKY and MYB families, which are known to mediate signaling cascades and transcriptional regulatory networks involved in plant growth, development, and responses to biotic and abiotic stress factors [95–98]. The similar activation of these genes suggests they form a core regulatory module essential for initiation of mature endosperm reprogramming. In contrast, several other TFs, including members of the NAC, HD-ZIP, bZIP-ABF and MADS-box families, exhibited significant up- or down-regulation that varied depending on the cultivar. Interestingly, many TFs from these families have previously been implicated in regulating endosperm and seed development *in planta*, highlighting their functional versatility.

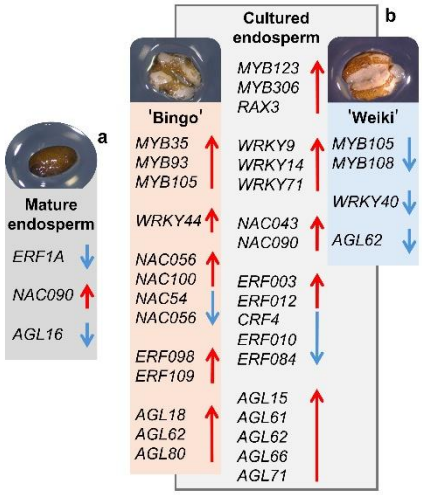


Figure 10 Transcription factor (TF) genes expressed in (**a**) freshly isolated mature endosperm and (**b**) endosperm explants cultured for two weeks. Genes expressed in both cultivars are framed, whereas genes uniquely expressed in 'Bingo' or 'Weiki' are displayed in separate squares.

The AP2/EREBP family represents a large, plant-specific multigene family of TFs whose members contain highly conserved AP2/ERF DNA-binding domains, originally described in the homeotic gene APETALA2 (AP2) and in ethylene-responsive element binding proteins (EREBPs) [99]. These genes are mainly activated in response to various environmental stressors, such as drought [100] or cold [95]. Furthermore, AP2/EREBPs TFs have been implicated in various hormone-related signal transduction pathways, including cytokinin signaling, which regulates plant growth and development [96, 101]. The exogenous application of cytokinins is critical for initiating differentiation in the

mature endosperm of *A*. *arguta*, thereby directing explants toward either organogenic or non-organogenic pathways [49]. In the present study, we observed a significant upregulation of ERF003 and ERF012 in freshly isolated endosperm. However, the expression profiles of other genes within this family varied between the two cultivars at the onset of the culture. Following a two-week culture period, *ERF098* and *ERF109* were found to be expressed at higher levels specifically in the 'Bingo' explants.

WRKY TFs play key roles in modulating the balance between plant immunity and growth during environmental stress responses, thereby optimizing plant adaptation [97]. Regarding endosperm development, *WRKY10* has been reported to participate in endosperm cellularization in Arabidopsis [102], while *WRKY71* is engaged in maintaining the proper function of the aleurone layer in rice [103]. In the present study, the cultured explant exhibited a distinct upregulation of *WRKY9*, *WRKY14*, and *WRKY44* in 'Bingo', whereas *WRKY71* was specifically upregulated in 'Weiki'.

MYB TFs serve as central regulators of plant development and stress responses, including the control of phenylpropanoid metabolism [98]. In the current work, a distinct up-regulation of genes associated with the “phenylpropanoid biosynthesis” pathway was noticed in the cultured explants. Previous reports have also highlighted the specific roles of certain MYB TFs in proper endosperm development. For instance, at the onset of seed formation in Arabidopsis, storage compounds such as oils and storage proteins initially accumulate in both the endosperm and the embryo. However, the endosperm-specific expression of *MYB118* subsequently represses the synthesis of these storage compounds, thereby establishing the embryo as the primary deposition site for nutrient reserves in the mature seed [104]. In present work *MYB93*, *MYB35*, *MYB105* and *MYB108* were significantly upregulated in 'Bingo' and simultaneously down regulated in 'Weiki' after two weeks of culture. Conversely, a strong upregulation of *MYB306*, *MYB123*, and *RAX3* was noticed in both cultivars following the culture period.

NAC TFs play pivotal roles in plant development and stress responses [105]. A defining feature of NAC TFs is their ability to regulate gene expression through diverse mechanisms, including binding to specific DNA motifs within target gene promoters, interacting with other regulatory proteins, and

modulating chromatin structure. Additionally, NAC genes are implicated in post-translational regulatory pathways [94]. In Arabidopsis, certain members of the NAC family induce programmed cell death to eliminate specific endosperm cells, allowing the embryo to grow and expand without any obstacles [106]. In the present experiment, *NAC043* and *NAC090* exhibited significant up-regulation, while *NAC083* demonstrated notable up-regulation after two weeks of culture, particularly in the 'Bingo' samples.

HD-ZIPs TFs are crucial mediators of plant responses to various biotic and abiotic stressors [107]. They also exert an indirect influence on the regulation of secondary metabolite-driven processes. Following the two-week culture period, a distinct upregulation of ATHB-52 and HAT7 was observed specifically in the 'Weiki' cultivar.

Plant bZIPs and ABF TFs often collaborate with other TF families, such as WRKY and NAC, to form complex, non-redundant regulatory networks. Their activity is tightly controlled through post-translational modifications (including phosphorylation and ubiquitination), enabling plants to adapt swiftly to environmental changes [108]. The present study revealed a significant upregulation of *BZIP11*, *FD*, and *BZIP44* in the cultured explants, with a particularly high expression of *BZIP34* and *AB1*5 observed in the 'Bingo' cultivar.

MADS-box and AGLs are potential key regulators in the 'awakening' of mature endosperm

In the present study, the expression of AGL genes, classified as Type I and II MADS-box genes, was identified in endosperm explants following two weeks of culture. Particularly genes AGL61 and AGL62 (from Type I clad Mα*), AGL80 (from Type I clad Mγ), and AGL9, AGL15 (from Type II) were overexpressed. The upregulation of *AGL15* and *AGL61* was evident in both the 'Bingo' and 'Weiki' cultivars, suggesting that these genes play a crucial role in the reprogramming of mature endosperm.

AGL15 is highly expressed in various embryogenic tissues, including Arabidopsis and maize zygotic embryos, dandelion apomictic embryos, alfalfa somatic embryos, and canola microspore-derived

941 embryos [43]. Furthermore, both *AGL15* and *AGL18* are essential for the induction of somatic
942 embryogenesis in Arabidopsis [46].

943 AGL61 forms a heterodimer with AGL80, and their expression is considered exclusive to the central
944 cell of the embryo sac and early endosperm [109, 110]. However, recent findings have revealed that
945 during somatic embryogenesis in *Dimocarpus longan*, the AGL61/80 regulatory network controls the
946 expression of the *BGAL9* gene, which encodes a β-galactosidase enzyme that plays a pivotal role in
947 cell wall modification [111].

948 Inactivation of AGL62 and AGL90 is known to cause dysfunction in endosperm development [47].
949 *AGL62* is responsible for activating small invertase inhibitors in the syncytial endosperm and
950 controlling the rate of embryo growth in Arabidopsis by regulating the expression of InvINH1 [112].
951 The InvINH1 promoter contains an enhancer region enriched with CarG motifs. Although the
952 consensus motif bound by Type I MADS-box proteins has not been studied as extensively as that of
953 Type II, the latter are well documented to bind a consensus sequence known as the CarG-box
954 (CC(A/T)6GG) [113]. Type II MADS-box proteins are often recruited by multiple CarG motifs within
955 target promoters, where they form higher-order heterotetrameric complexes that regulate floral organ
956 identity. Similar motifs to the CarG-box have also been identified as the binding sites for the Type I
957 MADS-box gene *PHE1/AGL37*.

958 LEAFY COTYLEDON2, FUSCA3, and ABA INSENSITIVE3, which encode B3 domain
959 transcription factors that serve as master regulators of embryogenesis, were identified and verified as
960 direct target genes of AGL15 [44]. Downstream genes targeted by these B3-domain factors are often
961 co-regulated by AGL15; among these, *INDOLEACETIC ACID-INDUCED PROTEIN30* (*IAA30*) has
962 been shown to participate actively in promoting somatic embryo development [44].

963 ***Phylogenetic placement and structural modelling support the functional annotation of the studied***
964 ***AGLs***

965 The phylogenetic reconstruction, combined with protein motif analysis, provided valuable insights
966 into the evolutionary diversification and structural organization of the analysed MADS-box family. A

particularly noteworthy finding was the close phylogenetic relationship between the AGL61 and AGL62 lineages, accompanied by their highly conserved domain architecture. In contrast to the classical Type II (MIKC-type) MADS-box proteins, such as SEPALLATA and AGL6, which possess the characteristic K-box domain required for higher-order protein complex formation, members of the AGL61, AGL62, and AGL80 subfamilies lacked this domain, consistent with their classification as Type I MADS-box proteins [114, 115]. Despite their simpler protein architecture, functional studies in model species have demonstrated that Type I MADS-box proteins operate through highly coordinated interaction networks. Although AGL61 and AGL62 do not interact directly, members of these Mα-type lineages share interaction partners from the Mγ-type lineage, most notably AGL80, highlighting the conserved interaction patterns characteristic of Type I MADS-box proteins [116]. In Arabidopsis thaliana, the AGL61-AGL80 heterodimer is indispensable for central cell specification, whereas AGL62 plays a crucial role in regulating early endosperm development [109, 117]. The close phylogenetic proximity of the newly identified *Actinidia arguta* homologs to the AGL61, AGL62, and AGL80 clades, together with their conserved MADS only motif composition, suggest that these genes have retained conserved functions associated with female gametophyte and early seed development. The conservation of both their evolutionary relationships and structural features is consistent with the conserved role of Type I MADS-box genes and their interaction networks in regulating reproductive development across angiosperms [118].

To complement these phylogenetic and motif-based findings, 3D structural modeling provided structural context regarding both domain organization and transcript annotation. For the candidate 64734(Slyc_AGL9) transcript, the predicted model aligned strictly with chain 2 (the SEP3-specific K-domain monomer) of the experimental 8CRA heterotetrameric structure [74], without incorporating chain 1 (AG). This structural selectivity reflects the canonical interface features of class E MADS-box factors, which assemble into higher-order floral quartets alongside class C proteins [119] while maintaining lineage-specific residues within their K-domains [120].

Beyond AGL9, AlphaFold modeling performed without an experimental structural template for the remaining Type I MADS-box candidates (including AGL61, AGL62, and AGL80) further

corroborated their predicted protein architecture. Across these models, high structural conservation was consistently restricted to the central DNA-binding MADS cores, whereas greater conformational flexibility was observed in the terminal extensions – a spatial arrangement consistent with the absence of a structured K-domain in Type I factors. Interestingly, 3D structural superimposition identified several sequences that warrant cautious interpretation. The AGL62 transcripts 68068(Ath_AGL62) and 104080(Ath_AGL62) showed difficulties in structural alignment, whereas 32624(Ath_AGL80) displayed limited structural overlap. These findings corresponded with the relatively long branch lengths observed for these sequences in the phylogenetic analysis. However, such differences may reflect local sequence variation, transcript assembly-related limitations, or uncertainty associated with structure prediction, rather than representing genuine functional divergence. Further analyses would help determine whether the observed sequence variation is associated with substantial changes in the predicted protein structures.

Conclusion

In the present study, we did not focus on the common genes typically associated with dedifferentiation and described in model plant. Due to our work with endosperm tissue, our research centered on *AGL* genes, which play a crucial role in proper endosperm development under *in vivo* conditions. This raises the question: are specific *AGL* genes primarily responsible for dedifferentiation and callus induction in mature *A*. *arguta* endosperm? Several of the genes identified in the present study, such as *AGL15*, coincide with those involved in callus proliferation in Arabidopsis tissue [5]. However, it is noteworthy that others may be specific to endosperm dedifferentiation, such as *AGL61*. By comparing two cultivars with different response efficiencies, this study provides an overview of *AGL* gene regulatory patterns during the dedifferentiation of mature endosperm. The results presented demonstrate that the upregulation of certain *AGL* genes belonging to Types I and II of the MADS-box family is essential to shift the fate of mature endosperm cells, consistent with their role in proper *in planta* endosperm development. Notably, comparative and phylogenetic analyses of *AGL* transcripts in the non-model plant *A*. *arguta* provide a valuable framework for generating future mutants. Such an

approach could precisely establishing their role in endosperm development in this fruit crop species, both *in vivo* and under experimental conditions.

Author's contributions

MC and MPK conceived and designed the study and supervised the project. MC, MPK, and JM acquired the funding. JM and JKT performed tissue cultures and sample collection. JM and MC jointly performed RNA extraction and prepared the samples for RNA sequencing (RNA-seq). MC conducted the RNA-seq data analyses and collaborated with JM on the interpretation and evaluation of the results. JM, MC, and ZD contributed to the phylogenetic analyses and interpretation of the results. ZD also participated in the structural analyses of the selected proteins. JM and MPK prepared the figures. JM, MPK, and MC wrote the original draft, as well as reviewed and edited the manuscript. All authors have read and approved the final manuscript.

Funding

This work was partially supported by the "Excellence Initiative – Research University" program at the Jagiellonian University in Kraków (2023 edition, to M.P.-K.) and the "Research Support for Doctoral Students" grant from the Faculty of Biology at the Jagiellonian University in Kraków, Poland (2024 edition, to J.M.).

Ethics approval and consent to participate

Permission to collect fruits of Actinidia arguta 'Bingo' and 'Weiki' from the Actinidia germplasm collection at the Warsaw University of Life Sciences (SGGW), Poland, was obtained from Prof. Piotr Latocha.

Availability of data and materials

All raw RNA-seq datasets generated during the current study have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under BioProject accession number PRJNA1542276.

REFERENCES

1. Grafi G. How cells dedifferentiate: a lesson from plants. Dev Biol. 2004;268:1–6. https://doi.org/10.1016/j.ydbio.2003.12.027.

2. Fehér A. Callus, Dedifferentiation, Totipotency, Somatic Embryogenesis: What These Terms Mean in the Era of Molecular Plant Biology? Front Plant Sci. 2019;10:536. https://doi.org/10.3389/fpls.2019.00536.

3. Sugimoto K, Gordon SP, Meyerowitz EM. Regeneration in plants and animals: dedifferentiation, transdifferentiation, or just differentiation? Trends Cell Biol. 2011;21:212–8. https://doi.org/10.1016/j.tcb.2010.12.004.

4. Ikeuchi M, Favero DS, Sakamoto Y, Iwase A, Coleman D, Rymen B, et al. Molecular Mechanisms of Plant Regeneration. Annu Rev Plant Biol. 2019;70:377–406. https://doi.org/10.1146/annurev-arplant-050718-100434.

5. Ikeuchi M, Sugimoto K, Iwase A. Plant callus: mechanisms of induction and repression. Plant Cell. 2013;25:3159–73. https://doi.org/10.1105/tpc.113.116053.

6. Zhai N, Xu L. Pluripotency acquisition in the middle cell layer of callus is required for organ regeneration. Nat Plants. 2021;7:1453–60. https://doi.org/10.1038/s41477-021-01015-8.

7. Yadav RK, Perales M, Gruel J, Ohno C, Heisler M, Girke T, et al. Plant stem cell maintenance involves direct transcriptional repression of differentiation program. Mol Syst Biol. 2013;9:654. https://doi.org/10.1038/msb.2013.8.

8. Fehér A. Somatic embryogenesis — Stress-induced remodeling of plant cell fate. Biochim Biophys Acta BBA - Gene Regul Mech. 2015;1849:385–402. https://doi.org/10.1016/j.bbagrm.2014.07.005.

9. Fehér A. Dedifferentiation of Plant Cells: A Term Covering Multiple Pathways?

10. Thomas TD, Chaturvedi R. Endosperm culture: a novel method for triploid plant production. Plant Cell Tissue Organ Cult. 2008;93:1–14. https://doi.org/10.1007/s11240-008-9336-6.

11. Hoshino Y, Miyashita T, Thomas TD. In vitro culture of endosperm and its application in plant breeding: Approaches to polyploidy breeding. Sci Hortic. 2011;130:1–8. https://doi.org/10.1016/j.scienta.2011.06.041.

12. Wang X, Cheng Z-M, Zhi S, Fengxiang X. Breeding triploid plants: a review. Czech J Genet Plant Breed. 2016;52:41–54. https://doi.org/10.17221/151/2015-CJGPB.

13. Baroux C, Spillane C, Grossniklaus U. Evolutionary origins of the endosperm in flowering plants. Genome Biol. 2002;3:reviews1026. https://doi.org/10.1186/gb-2002-3-9-reviews1026.

14. Qiu Y, Köhler C. Endosperm Evolution by Duplicated and Neofunctionalized Type I MADS-Box Transcription Factors. Mol Biol Evol. 2022;39:msab355. https://doi.org/10.1093/molbev/msab355.

15. Hands P, Rabiger DS, Koltunow A. Mechanisms of endosperm initiation. Plant Reprod. 2016;29:215–25. https://doi.org/10.1007/s00497-016-0290-x.

16. Sabelli PA. Replicate and die for your own good: Endoreduplication and cell death in the cereal endosperm. J Cereal Sci. 2012;56:9–20. https://doi.org/10.1016/j.jcs.2011.09.006.

17. Doll N, Nowack M. Endosperm cell death : roles and regulation in angiosperms. J Exp Bot. 2024;75:4346–59. https://doi.org/10.1093/jxb/erae052.

18. Wu J-H, Ferguson AR, Murray BG, Jia Y, Datson PM, Zhang J. Induced polyploidy dramatically increases the size and alters the shape of fruit in Actinidia chinensis. Ann Bot. 2012;109:169–79. https://doi.org/10.1093/aob/mcr256.

19. Machado MD, Souza CS, Machado M, Reis AC, de Sousa SM, Matos EM, et al. Novel avenues for passion fruit in vitro regeneration from endosperm culture, and morpho-agronomic and physiological traits of triploid Passiflora cincinnata Mast. emblings. Plant Cell Tissue Organ Cult PCTOC. 2022;150:637–50. https://doi.org/10.1007/s11240-022-02318-0.

20. Tang J, Luo X, Zhu Y, Cai N, Chen L, Chen S, et al. In vitro regeneration of triploid from mature endosperm culture of Passiflora edulis “Mantianxing.” Biosci Biotechnol Biochem. 2024;88:412–9. https://doi.org/10.1093/bbb/zbae009.

21. da Silva NT, de Faria RB, Crepaldi GB, Paim-Pinto DL, Viccini LF, de Matos EM, et al. Karyotype characterization of endosperm-derived plants of Passiflora cristalina Vanderplank and Zappi, a wild amazonian passionfruit species. Plant Cell Tissue Organ Cult PCTOC. 2024;156:102. https://doi.org/10.1007/s11240-024-02714-8.

22. Zhu X-T, Sanz-Jimenez P, Ning X-T, Tahir ul Qamar M, Chen L-L. Direct RNA sequencing in plants: Practical applications and future perspectives. Plant Commun. 2024;5:101064. https://doi.org/10.1016/j.xplc.2024.101064.

23. Panahi B, Hamid R, Mohammad Zadeh Jalaly H. Deciphering plant transcriptomes: Leveraging machine learning for deeper insights. Curr Plant Biol. 2025;41:100432. https://doi.org/10.1016/j.cpb.2024.100432.

24. Rurek M. Diversity of organ-specific plant transcriptomes. Acta Biochim Pol. 2025;72:14609. https://doi.org/10.3389/abp.2025.14609.

25. Iwase A, Kondo Y, Laohavisit A, Takebayashi A, Ikeuchi M, Matsuoka K, et al. WIND transcription factors orchestrate wound-induced callus formation, vascular reconnection and defense response in Arabidopsis. New Phytol. 2021;232:734–52. https://doi.org/10.1111/nph.17594.

26. Iwase A, Mitsuda N, Koyama T, Hiratsu K, Kojima M, Arai T, et al. The AP2/ERF Transcription Factor WIND1 Controls Cell Dedifferentiation in *Arabidopsis*. Curr Biol. 2011;21:508–14. https://doi.org/10.1016/j.cub.2011.02.020.

27. Prát T, Hajný J, Grunewald W, Vasileva M, Molnár G, Tejos R, et al. WRKY23 is a component of the transcriptional network mediating auxin feedback on PIN polarity. PLOS Genet. 2018;14:e1007177. https://doi.org/10.1371/journal.pgen.1007177.

28. Zuo J, Niu Q-W, Frugis G, Chua N-H. The WUSCHEL gene promotes vegetative-to-embryonic transition in Arabidopsis. Plant J. 2002;30:349–59. https://doi.org/10.1046/j.1365-313X.2002.01289.x.

29. Oh S, Kim J, Kim Y, Lee M, Kim D. A MADS-Box Gene-Based InDel Marker Discriminating Sex in Actinidia arguta. Horticulturae. 2023;9:1310. https://doi.org/10.3390/horticulturae9121310.

30. Liu L, White MJ, MacRae TH. Transcription factors and their genes in higher plants. Eur J Biochem. 1999;262:247–57. https://doi.org/10.1046/j.1432-1327.1999.00349.x.

31. Latif A, Azam S, Shahid N, Javed MR, Haider Z, Yasmeen A, et al. Overexpression of the AGL42 gene in cotton delayed leaf senescence through downregulation of NAC transcription factors. Sci Rep. 2022;12:21093. https://doi.org/10.1038/s41598-022-25640-1.

32. Han J-P, Wan J-N, Guan Z-L, Xu H, Wang Q-F, Wan T. The origin, evolution, and diversification of MADS-box transcription factors in green plants. Plant Commun. 2025;6. https://doi.org/10.1016/j.xplc.2025.101462.

33. Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM. The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. Nature. 1990;346:35–9. https://doi.org/10.1038/346035a0.

34. Becker A, Theissen G. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Mol Phylogenet Evol. 2003;29:464–89. https://doi.org/10.1016/s1055-7903(03)00207-0.

35. Lai X, Vega-Léon R, Hugouvieux V, Blanc-Mathieu R, van der Wal F, Lucas J, et al. The intervening domain is required for DNA-binding and functional identity of plant MADS transcription factors. Nat Commun. 2021;12:4760. https://doi.org/10.1038/s41467-021-24978-w.

36. Parenicová L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, et al. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. Plant Cell. 2003;15:1538–51. https://doi.org/10.1105/tpc.011544.

37. Zhang S, Mohanty D, Muzaffar A, Ni M. Two MADS-box proteins, AGL9 and AGL15, recruit the FIS-PRC2 complex to trigger the phase transition from endosperm proliferation to embryo

development in Arabidopsis. Mol Plant. 2024;17:1110–28.
https://doi.org/10.1016/j.molp.2024.05.011.

38. Voogd C, Brian LA, Wu R, Wang T, Allan AC, Varkonyi-Gasic E. A MADS-box gene with similarity to FLC is induced by cold and correlated with epigenetic changes to control budbreak in kiwifruit. New Phytol. 2022;233:2111–26. https://doi.org/10.1111/nph.17916.

39. Ye L-X, Luo M-M, Wang Z, Bai F-X, Luo X, Gao L, et al. Genome-wide analysis of MADS-box gene family in kiwifruit (Actinidia chinensis var. chinensis) and their potential role in floral sex differentiation. Front Genet. 2022;13:1043178. https://doi.org/10.3389/fgene.2022.1043178.

40. Li Y, Huang H, Abid M, Gu H, Fang J, Cheng Z, et al. Characterization and Identification of a Ripening-Related Gene AaPG18 in Actinidia arguta. Int J Mol Sci. 2022;23:2597. https://doi.org/10.3390/ijms23052597.

41. Yu L-H, Miao Z-Q, Qi G-F, Wu J, Cai X-T, Mao J-L, et al. MADS-box transcription factor AGL21 regulates lateral root development and responds to multiple external and physiological signals. Mol Plant. 2014;7:1653–69. https://doi.org/10.1093/mp/ssu088.

42. Yu L-H, Wu J, Miao Z-Q, Zhao P-X, Wang Z, Xiang C-B. Arabidopsis MADS-box transcription factor AGL21 acts as environmental surveillance for seed germination by regulating ABI5. 2016;:082164. https://doi.org/10.1101/082164.

43. Perry SE, Lehti MD, Fernandez DE. The MADS-Domain Protein AGAMOUS-Like 15 Accumulates in Embryonic Tissues with Diverse Origins. Plant Physiol. 1999;120:121–30. https://doi.org/10.1104/pp.120.1.121.

44. Zheng Y, Ren N, Wang H, Stromberg AJ, Perry SE. Global Identification of Targets of the Arabidopsis MADS Domain Protein AGAMOUS-Like15. Plant Cell. 2009;21:2563–77. https://doi.org/10.1105/tpc.109.068890.

45. Zheng Q, Zheng Y, Ji H, Burnie W, Perry SE. Gene Regulation by the AGL15 Transcription Factor Reveals Hormone Interactions in Somatic Embryogenesis1[OPEN]. Plant Physiol. 2016;172:2374–87. https://doi.org/10.1104/pp.16.00564.

46. Paul P, Joshi S, Tian R, Diogo Junior R, Chakrabarti M, Perry SE. The MADS-domain factor AGAMOUS-Like18 promotes somatic embryogenesis. Plant Physiol. 2022;188:1617–31. https://doi.org/10.1093/plphys/kiab553.

47. Walia H, Josefsson C, Dilkes B, Kirkbride R, Harada J, Comai L. Dosage-Dependent Deregulation of an AGAMOUS-LIKE Gene Cluster Contributes to Interspecific Incompatibility. Curr Biol. 2009;19:1128–32. https://doi.org/10.1016/j.cub.2009.05.068.

48. Bente H, Köhler C. Molecular basis and evolutionary drivers of endosperm-based hybridization barriers. Plant Physiol. 2024;195:155–69. https://doi.org/10.1093/plphys/kiae050.

49. Sala-Cholewa K, Kozieradzka-Kiszkurno M, Malec J, Trościński JK, Abdullah M, Czernicka M, et al. The appearance of amyloplasts as one of the features distinguishing the initial stages of the development of isolated endosperm in kiwiberry. Acta Soc Bot Pol. 2023;92:1–9. https://doi.org/10.5586/asbp/176092.

50. Popielarska-Konieczna M, Sala K, Abdullah M, Tuleja M, Kurczyńska E. Extracellular matrix and wall composition are diverse in the organogenic and non-organogenic calli of Actinidia arguta. Plant Cell Rep. 2020;39:779–98. https://doi.org/10.1007/s00299-020-02530-2.

51. Abdullah M, Sliwinska E, Góralski G, Latocha P, Tuleja M, Widyna P, et al. Effect of medium composition, genotype and age of explant on the regeneration of hexaploid plants from endosperm culture of tetraploid kiwiberry (Actinidia arguta). Plant Cell Tissue Organ Cult PCTOC. 2021;147:569–82. https://doi.org/10.1007/s11240-021-02149-5.

52. Latocha P. The Nutritional and Health Benefits of Kiwiberry (Actinidia arguta) – a Review. Plant Foods Hum Nutr. 2017;72:325–34. https://doi.org/10.1007/s11130-017-0637-y.

53. Latocha P, Vereecke D, Debersaques F. Kiwiberry commercial production – what stage are we at? Acta Hortic. 2018;:559–64. https://doi.org/10.17660/ActaHortic.2018.1218.76.

54. Lu X-M, Yu X-F, Li G-Q, Qu M-H, Wang H, Liu C, et al. Genome assembly of autotetraploid Actinidia arguta highlights adaptive evolution and enables dissection of important economic traits. Plant Commun. 2024;5. https://doi.org/10.1016/j.xplc.2024.100856.

55. Cui W, Deng C, Yoon M, Žárský V, Rikkerink E. Systematic Analysis of the EXO70 Gene Family in Kiwifruit Species: Evolutionary Selection and Potential Functions in Plant Immunity. 2025;:2025.10.28.684437. https://doi.org/10.1101/2025.10.28.684437.

56. Li Y, Song Z, Zhan X, Li X, Ye L, Lin M, et al. Chromosome-level genome assembly assisting for dissecting mechanism of anthocyanin regulation in kiwifruit (Actinidia arguta). Mol Hortic. 2025;5:18. https://doi.org/10.1186/s43897-024-00139-7.

57. Wen M, Li Y, Lang H, Wang C, Zhang A, Zhang H, et al. Integrated physiological, transcriptomic and metabolomic analyses shed light on the complex mechanisms underlying low temperature resistance in kiwifruit (*Actinidia arguta*). Plant Physiol Biochem. 2026;237:111507. https://doi.org/10.1016/j.plaphy.2026.111507.

58. 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. https://doi.org/10.1002/1873-3468.14570.

59. Guo X, Chen Z, Liu X, Miao L, Chen Y, Yao Y, et al. Dynamic regulation and embryo-endosperm crosstalk revealed by whole transcriptome profiling of wheat seed development. Cell Rep. 2025;44. https://doi.org/10.1016/j.celrep.2025.116195.

60. Góralski G, Popielarska-Konieczna M, Ślesak H, Siwińska D, Batycka M. Organogenesis in endosperm of Actinidia deliciosa cv. Hayward cultured in vitro. 2005.

61. Murashige T, Skoog F. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiol Plant. 1962;15:473–97. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x.

62. Mortaji Z, Tavakkol Afshari R, Alizadeh H, Yazdi Samadi B. RNA Isolation and Expression from Different Dormant and After-Ripened Wheat (Triticum aestivum) Seed Tissues Rich in Polysaccharides and Proteins. Asian J Plant Sci. 7:201–6. https://doi.org/10.3923/ajps.2008.201.206.

63. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, et al. BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. Mol Biol Evol. 2018;35:543–8. https://doi.org/10.1093/molbev/msx319.

64. Paysan-Lafosse T, Andreeva A, Blum M, Chuguransky SR, Grego T, Pinto BL, et al. The Pfam protein families database: embracing AI/ML. Nucleic Acids Res. 2025;53:D523–34. https://doi.org/10.1093/nar/gkae997.

65. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2025. Nucleic Acids Res. 2025;53:D609–17. https://doi.org/10.1093/nar/gkae1010.

66. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30:1236–40. https://doi.org/10.1093/bioinformatics/btu031.

67. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol. 2013;30:772–80. https://doi.org/10.1093/molbev/mst010.

68. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3. https://doi.org/10.1093/bioinformatics/btp348.

69. Price MN, Dehal PS, Arkin AP. FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. Mol Biol Evol. 2009;26:1641–50. https://doi.org/10.1093/molbev/msp077.

70. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020;37:1530–4. https://doi.org/10.1093/molbev/msaa015.

71. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. 2024;52:W78–82. https://doi.org/10.1093/nar/gkae268.

72. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. Nucleic Acids Res. 2015;43:W39–49. https://doi.org/10.1093/nar/gkv416.

73. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630:493–500. https://doi.org/10.1038/s41586-024-07487-w.

74. Hugouvieux V, Blanc-Mathieu R, Janeau A, Paul M, Lucas J, Xu X, et al. SEPALLATA-driven MADS transcription factor tetramerization is required for inner whorl floral organ development. Plant Cell. 2024;36:3435–50. https://doi.org/10.1093/plcell/koae151.

75. The PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

76. da Silva NT, Silva LAS, Reis AC, Machado M, de Matos EM, Viccini LF, et al. Endosperm culture: a facile and efficient biotechnological tool to generate passion fruit (Passiflora cincinnata Mast.) triploid plants. Plant Cell Tissue Organ Cult PCTOC. 2020;142:613–24. https://doi.org/10.1007/s11240-020-01887-2.

77. Mikovski AI, da Silva NT, Silva LAS, Machado M, de Souza Barbosa LC, Reis AC, et al. From endosperm to triploid plants: a stepwise characterization of the de novo shoot organogenesis and morpho-agronomic aspects of an ornamental passion fruit (Passiflora foetida L.). Plant Cell Tissue Organ Cult PCTOC. 2021;147:239–53. https://doi.org/10.1007/s11240-021-02120-4.

78. Maqbool A, Lu M, Muhammad N, Baz NM, Liu M, Liu Z, et al. Dedifferentiation and redifferentiation: the developing cell fate transition of plants. Hortic Plant J. 2025;11:1961–84. https://doi.org/10.1016/j.hpj.2025.08.005.

79. Deng J, Sun T, Cao W, Fan D, Cheng N, Wang B, et al. Extraction Optimization and Functional Properties of Proteins from Kiwi Fruit(Actinidia chinensis Planch.) Seeds. Int J Food Prop. 2014;17:1612–25. https://doi.org/10.1080/10942912.2013.772197.

80. Deng J, Liu Q, Zhang Q, Zhang C, Liu D, Fan D, et al. Comparative study on composition, physicochemical and antioxidant characteristics of different varieties of kiwifruit seed oil in China. Food Chem. 2018;264:411–8. https://doi.org/10.1016/j.foodchem.2018.05.063.

81. Fontechiari L, Accardo F, Bonzanini F, Bonomini MG, Caligiani A, Lolli V, et al. Kiwifruit waste valorisation: Chemical characterisation of by-products and enzymatic production of seed protein hydrolysates. Food Chem Adv. 2026;10:101249. https://doi.org/10.1016/j.focha.2026.101249.

82. Dabisch J, Scholz P, Mullen RT, de Vries J, Ischebeck T. Lipid droplet proteome plasticity in plant evolution, growth and development, and response to stress. J Exp Bot. 2026;77:3426–44. https://doi.org/10.1093/jxb/erag087.

83. Chen C, Hu Y, Ikeuchi M, Jiao Y, Prasad K, Su YH, et al. Plant regeneration in the new era: from molecular mechanisms to biotechnology applications. Sci China Life Sci. 2024;67:1338–67. https://doi.org/10.1007/s11427-024-2581-2.

84. Jiang F, Feng Z, Liu H, Zhu J. Involvement of Plant Stem Cells or Stem Cell-Like Cells in Dedifferentiation. Front Plant Sci. 2015;6. https://doi.org/10.3389/fpls.2015.01028.

85. Lee LR, Torii KU. Cell cycle in plant development and reprogramming. Development. 2026;153:dev205318. https://doi.org/10.1242/dev.205318.

86. Kolli R. From the archives: Roles of microtubule-associated proteins in organelle movement, tip growth, and phragmoplast architecture. Plant Cell. 2025;37:koae331. https://doi.org/10.1093/plcell/koae331.

1289 87. Rodriguez E, Chevalier J, Olsen J, Ansbøl J, Kapousidou V, Zuo Z, et al. Autophagy mediates
1290 temporary reprogramming and dedifferentiation in plant somatic cells. EMBO J.
1291 2020;39:EMBJ2019103315. https://doi.org/10.15252/embj.2019103315.

1292 88. Rahim MA, Zhang X, Busatto N. Editorial: Phenylpropanoid biosynthesis in plants. Front Plant
1293 Sci. 2023;14. https://doi.org/10.3389/fpls.2023.1230664.

1294 89. Falcone Ferreyra ML, Rius S, Casati P. Flavonoids: biosynthesis, biological functions, and
1295 biotechnological applications. Front Plant Sci. 2012;3. https://doi.org/10.3389/fpls.2012.00222.

1296 90. Tan C, Wang Z, Feng X, Pan B, Irfan M, Liu C. Transcriptomic and metabolomics of flavonoid
1297 compounds in Actinidia arguta var. arguta. J King Saud Univ – Sci. 2021;33.
1298 https://doi.org/10.1016/j.jksus.2021.101605.

1299 91. Xu G, Chen Y, Martins LM, Li E, Wang F, Magana T, et al. Transcription factors instruct DNA
1300 methylation patterns in plant reproductive tissues. Nat Cell Biol. 2025;27:2116–27.
1301 https://doi.org/10.1038/s41556-025-01808-5.

1302 92. Chowdhary AA, Mishra S, Mehrotra S, Upadhyay SK, Bagal D, Srivastava V. Chapter 1 - Plant
1303 transcription factors: an overview of their role in plant life. In: Srivastava V, Mishra S, Mehrotra S,
1304 Upadhyay SK, editors. Plant Transcription Factors. Academic Press; 2023. p. 3–20.
1305 https://doi.org/10.1016/B978-0-323-90613-5.00003-0.

1306 93. Samad AFA, Sajad M, Nazaruddin N, Fauzi IA, Murad AMA, Zainal Z, et al. MicroRNA and
1307 Transcription Factor: Key Players in Plant Regulatory Network. Front Plant Sci. 2017;8.
1308 https://doi.org/10.3389/fpls.2017.00565.

1309 94. Puranik S, Sahu PP, Srivastava PS, Prasad M. NAC proteins: regulation and role in stress
1310 tolerance. Trends Plant Sci. 2012;17:369–81. https://doi.org/10.1016/j.tplants.2012.02.004.

1311 95. Yang S, Zhou J, Li Y, Wu J, Ma C, Chen Y, et al. AP2/EREBP Pathway Plays an Important Role
1312 in Chaling Wild Rice Tolerance to Cold Stress. Int J Mol Sci. 2023;24:14441.
1313 https://doi.org/10.3390/ijms241914441.

96. Gentile D, Serino G, Frugis G. CRF transcription factors in the trade-off between abiotic stress response and plant developmental processes. Front Genet. 2024;15. https://doi.org/10.3389/fgene.2024.1377204.

97. Chen X, Zhang T, Wang H, Zhao W, Guo Z. Transcription factor WRKY complexes in plant signaling pathways. Phytopathol Res. 2025;7:54. https://doi.org/10.1186/s42483-025-00349-x.

98. Zhang D, Zhou H, Zhang Y, Zhao Y, Zhang Y, Feng X, et al. Diverse roles of MYB transcription factors in plants. J Integr Plant Biol. 2025;67:539–62. https://doi.org/10.1111/jipb.13869.

99. Riechmann JL, Meyerowitz EM. The AP2/EREBP family of plant transcription factors. Biol Chem. 1998;379:633–46. https://doi.org/10.1515/bchm.1998.379.6.633.

100. Sharoni AM, Nuruzzaman M, Satoh K, Moumeni A, Attia K, Venuprasad R, et al. Comparative transcriptome analysis of AP2/EREBP gene family under normal and hormone treatments, and under two drought stresses in NILs setup by Aday Selection and IR64. Mol Genet Genomics. 2012;287:1–19. https://doi.org/10.1007/s00438-011-0659-3.

101. Rashotte AM, Mason MG, Hutchison CE, Ferreira FJ, Schaller GE, Kieber JJ. A subset of Arabidopsis AP2 transcription factors mediates cytokinin responses in concert with a two-component pathway. Proc Natl Acad Sci. 2006;103:11081–5. https://doi.org/10.1073/pnas.0602038103.

102. Luo M, Dennis ES, Berger F, Peacock WJ, Chaudhury A. MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in Arabidopsis. Proc Natl Acad Sci. 2005;102:17531–6. https://doi.org/10.1073/pnas.0508418102.

103. Bataller S, Davis JA, Gu L, Baca S, Chen G, Majid A, et al. Disruption of the OsWRKY71 transcription factor gene results in early rice seed germination under normal and cold stress conditions. BMC Plant Biol. 2024;24:1090. https://doi.org/10.1186/s12870-024-05808-9.

104. Barthole G, To A, Marchive C, Brunaud V, Soubigou-Taconnat L, Berger N, et al. MYB118 Represses Endosperm Maturation in Seeds of Arabidopsis. Plant Cell. 2014;26:3519–37. https://doi.org/10.1105/tpc.114.130021.

105. Xiong H, He H, Chang Y, Miao B, Liu Z, Wang Q, et al. Multiple roles of NAC transcription factors in plant development and stress responses. J Integr Plant Biol. 2025;67:510–38. https://doi.org/10.1111/jipb.13854.

106. Doll NM, Van Hautegem T, Schilling N, De Rycke R, De Winter F, Fendrych M, et al. Endosperm cell death promoted by NAC transcription factors facilitates embryo invasion in Arabidopsis. Curr Biol. 2023;33:3785-3795.e6. https://doi.org/10.1016/j.cub.2023.08.003.

107. Sharif R, Xie C, Wang J, Cao Z, Zhang H, Chen P, et al. Genome wide identification, characterization and expression analysis of HD-ZIP gene family in *Cucumis sativus* L. under biotic and various abiotic stresses. Int J Biol Macromol. 2020;158:502–20. https://doi.org/10.1016/j.ijbiomac.2020.04.124.

108. Marco Llorca C, Potschin M, Zentgraf U. bZIPs and WRKYs: two large transcription factor families executing two different functional strategies. Front Plant Sci. 2014;5. https://doi.org/10.3389/fpls.2014.00169.

109. Bemer M, Wolters-Arts M, Grossniklaus U, Angenent GC. The MADS Domain Protein DIANA Acts Together with AGAMOUS-LIKE80 to Specify the Central Cell in Arabidopsis Ovules. Plant Cell. 2008;20:2088–101. https://doi.org/10.1105/tpc.108.058958.

110. Steffen JG, Kang I-H, Portereiko MF, Lloyd A, Drews GN. AGL61 Interacts with AGL80 and Is Required for Central Cell Development in Arabidopsis. Plant Physiol. 2008;148:259–68. https://doi.org/10.1104/pp.108.119404.

111. Chen Y, Ma X, Ma W, Xu L, Lai C, Zhang X, et al. AGL61/80 regulates *BGAL9*, which participates in early somatic embryogenesis and increases longan resistance to heat stress. Fruit Res. 2024;4. https://doi.org/10.48130/frures-0024-0005.

112. Hoffmann T, Shi X, Hsu C-Y, Brown A, Knight Q, Courtney LS, et al. The identification of type I MADS box genes as the upstream activators of an endosperm-specific invertase inhibitor in Arabidopsis. BMC Plant Biol. 2022;22:18. https://doi.org/10.1186/s12870-021-03399-3.

113. Aerts N, De Bruijn S, Van Mourik H, Angenent GC, Van Dijk ADJ. Comparative analysis of binding patterns of MADS-domain proteins in Arabidopsis thaliana. BMC Plant Biol. 2018;18:131. https://doi.org/10.1186/s12870-018-1348-8.

114. Masiero S, Colombo L, Grini PE, Schnittger A, Kater MM. The Emerging Importance of Type I MADS Box Transcription Factors for Plant Reproduction. Plant Cell. 2011;23:865–72. https://doi.org/10.1105/tpc.110.081737.

115. Qiu Y, Li Z, Walther D, Köhler C. Updated Phylogeny and Protein Structure Predictions Revise the Hypothesis on the Origin of MADS-box Transcription Factors in Land Plants. Mol Biol Evol. 2023;40:msad194. https://doi.org/10.1093/molbev/msad194.

116. de Folter S, Immink RGH, Kieffer M, Pařenicová L, Henz SR, Weigel D, et al. Comprehensive Interaction Map of the Arabidopsis MADS Box Transcription Factors. Plant Cell. 2005;17:1424–33. https://doi.org/10.1105/tpc.105.031831.

117. Kang I-H, Steffen JG, Portereiko MF, Lloyd A, Drews GN. The AGL62 MADS Domain Protein Regulates Cellularization during Endosperm Development in Arabidopsis. Plant Cell. 2008;20:635–47. https://doi.org/10.1105/tpc.107.055137.

118. Chen F, Zhang X, Liu X, Zhang L. Evolutionary Analysis of MIKCc-Type MADS-Box Genes in Gymnosperms and Angiosperms. Front Plant Sci. 2017;8. https://doi.org/10.3389/fpls.2017.00895.

119. Theissen G, Saedler H. Plant biology. Floral quartets. Nature. 2001;409:469–71. https://doi.org/10.1038/35054172.

120. Puranik S, Acajjaoui S, Conn S, Costa L, Conn V, Vial A, et al. Structural basis for the oligomerization of the MADS domain transcription factor SEPALLATA3 in Arabidopsis. Plant Cell. 2014;26:3603–15. https://doi.org/10.1105/tpc.114.127910.

Supplementary Materials:

Supplementary Material 1; Format: .xlsx; Title: Gene Ontology and KEGG classification of transcripts.; Description: Functional classification of identified transcripts based on Gene Ontology (GO) and KEGG pathway analyses.

Supplementary Material 2; Format: .xlsx; Title: Commonly and uniquely enriched GO terms.; Description: Comparative analysis of enriched GO terms among differentially expressed genes (DEGs) across all studied combinations.

Supplementary Material 3; Format: .xlsx; Title: Lists of common and unique DEGs.; Description: Detailed list of common and unique DEGs identified across different experimental combinations.

Supplementary Material 4; Format: .xlsx; Title: Enriched KEGG pathways associated with DEGs.; Description: Common and uniquely enriched KEGG pathways associated with identified DEGs across all experimental combinations.

Supplementary Material 5; Format: .xlsx; Title: DEGs belonging to 15 transcription factor families.; Description: Comprehensive list of DEGs categorized across 15 major transcription factor (TF) families in the studied sample combinations.

Supplementary Material 6; Format: .pdf; Title: Expression heatmaps of six major transcription factor families.; Description: Visual representation of differential gene expression patterns for AP2/EREBP, NAC, WRKY, HD-ZIP, MYB, and bZIP/ABF transcription factor families.

Supplementary Material 7; Format: .xlsx; Title: Functional classification of MADS-box DEGs.; Description: List of MADS-box DEGs categorized by biological function, including regulation of flowering, floral development, endosperm development, and root growth.

Supplementary Material 8; Format: .xlsx; Title: Detailed breakdown of AGL family DEGs.; Description: Expression profile and classification of identified AGL gene family DEGs across experimental combinations.

Supplementary Material 9; Format: .pdf; Title: Comparative sequence analysis of AGL homologous proteins.; Description: Comparative analysis of homologous protein sequences derived from identified AGL gene transcripts.

Supplementary Material 10; Format: .pdf; Title: Multiple sequence alignment of AGL proteins.; Description: Multiple sequence alignment of translated amino acid sequences from selected AGL transcripts alongside their closest homologs.

Supplementary Material 11; Format: .pdf; Title: AlphaFold structural models of AGL proteins.; Description: Predicted 3D structural models and descriptions of AGL proteins generated using AlphaFold.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial7.xlsx
- SupplementaryMaterial8.xlsx
- SupplementaryMaterial1.xlsx
- SupplementaryMaterial4.xlsx
- SupplementaryMaterial5.xlsx
- SupplementaryMaterial6.pdf
- SupplementaryMaterial10.pdf
- SupplementaryMaterial2.xlsx
- SupplementaryMaterial9.pdf
- SupplementaryMaterial11.pdf
- SupplementaryMaterial3.xlsx