# Regulation of early-stage tomato fruit growth by the splicing factor RS2Z36

[1*]Stavros Vraggalas, [1*]Remus RE Rosenkranz, [2]Mario Keller, [3]Yolanda Pérez-Pérez, [1]Samia Bachiri, [1]Kerstin J Zehl, [4]Jessica Bold, [4,5]Stefan Simm, [6]Arindam Ghatak, [6]Wolfram Weckwerth, [7]Leila Afjehi-Sadat, [6]Palak Chaturvedi, [3]Pilar S Testillano, [8,9]Michaela Müller-McNicoll, [2,10]Kathi Zarnack, [1+]Sotirios Fragkostefanakis

[1] Institute of Molecular Biosciences, Molecular and Cell Biology of Plants, Goethe University Frankfurt, Frankfurt, Germany

[2] Buchmann Institute of Molecular Life Sciences & Institute of Molecular Biosciences, Computational RNA Biology, Goethe University

[3] Pollen Biotechnology of Crop Plants Group, Margarita Salas Center of Biological Research, CIB-CSIC, Ramiro de Maeztu 9, Madrid 28040, Spain

[4] Institute of Bioinformatics, University Medicine Greifswald, Greifswald, Germany

[5] Institute of Bioanalysis, Coburg University of Applied Sciences, Coburg, Germany

[6] Molecular Systems Biology (MOSYS), Department of Functional and Evolutionary Ecology, University Vienna, Djerassiplatz 1, 1030 Vienna, Austria

[7] Mass Spectrometry unit, Research Support Facilities, Faculty of Life Sciences, University Vienna, Djerassiplatz 1, 1030 Vienna, Austria

[8] Institute of Molecular Biosciences, RNA Regulation in Higher Eukaryotes, Goethe University Frankfurt, Frankfurt, Germany

[9] Max-Planck Institute of Biophysics, Frankfurt, Germany

[10] Theodor Boveri Institute, Biocenter, Julius Maximilians University Würzburg, Am Hubland, 97074 Würzburg, Germany

*Equal contribution

[+]Correspondence: S.F. fragkost@bio.uni-frankfurt.de

Number of Tables: 0

Number of Figures: 6

Word count: 5035

Number of supplementary Figures: 1

Number of supplementary Tables: 8

***Running title: Regulation of RNA splicing during tomato fruit growth***

**Abstract**

Fruit size and shape are determined by cell division and expansion. In tomato, the model plant for fleshy fruit development, both processes are associated with changes in gene expression, including transcription and RNA processing. While several transcription factors have been identified as regulators of genes involved in fruit development, the contribution of splicing regulators remains unknown. Profiling of RNA splicing-related genes revealed stage-preferential expression of several factors, including RS2Z36, a serine/arginine-rich splicing factor highly expressed during the early stages of fruit development. Mutations in RS2Z36 lead to smaller ellipsoid fruits than the wild type, with histological analysis showing elongated pericarp cells in the ovaries along the longitudinal axis. RNA-seq analysis uncovered numerous alternatively spliced genes involved in different biological processes. Complementary proteome analysis identified proteins with altered abundance and peptides corresponding to novel RNA splice isoforms, underscoring the role of RS2Z36 in shaping both transcriptome and proteome diversity. Immunofluorescence screening of cell wall proteins further demonstrated an increased abundance of LM6-recognized arabinan epitopes of rhamnogalacturonan-I pectins and arabinogalactan protein (AGP) epitopes in *rs2z36.1* pericarp cells compared to the wild type. These findings highlight RS2Z36 as a critical regulator of alternative splicing, impacting fruit development through diverse transcriptome and proteome changes.

**Keywords:** alternative splicing, arabinogalactan protein, fruit, growth, ovary, serine/arginine-rich protein, splicing factor, tomato

## Introduction

The growth of fleshy fruits depends on tightly-coordinated cell division and expansion programmes. Tomato (*Solanum lycopersicum*) is a model plant for studying fleshy fruit development. Fruit morphology is genetically determined and the core loci that influence fruit size and shape, contributing to the diversity seen in modern tomatoes through domestication and breeding, have been identified (Paran and Van Der Knaap, 2007).

*LC* and *FAS* in tomato, which encode the orthologs of WUSCHEL (WUS) and CLAVATA3 (CLV3) respectively, regulate locule number by altering the spatial and temporal expression of *SlWUS***,** leading to the formation of multilocular and enlarged fruits compared to wild progenitors (Cong *et al.*, 2008; Muños *et al.*, 2011; Chu *et al.*, 2019). These mutations expand the *SlWUS* expression in young floral buds, thereby increasing floral meristem size and positively influencing tomato fruit size through transcriptional regulation of genes involved in meristem maintenance and floral organ development (Chu *et al.*, 2019). The genes *SUN* and *OVATE* control the elongated shape (Xiao *et al*., 2008). *SUN* codes for a protein belonging to the IQ67 family which is involved in calcium signalling and cellular trafficking through interactions with calmodulins (CaMs) and kinesin-light chain-related proteins (Abel *et al.*, 2013; Clevenger *et al.*, 2015). *sun* mutation caused by a gene duplication mediated by the retrotransposon Rider lead to high expression of *SUN* and consequently to enhanced fruit growth (Xiao *et al.*, 2008; Jiang *et al.*, 2009). *OVATE* codes for a repressor of transcription and an *ovate* allele with a mutation that causes a premature termination codon (PTC) results in elongated fruit shape (Liu *et al.*, 2002; Wang *et al.*, 2007). OVATE has been linked to microtubule organization and cell division patterns through interaction with members of the TONNEAU1 Recruiting Motif (TRM) superfamily (Snouffer *et al.*, 2020).

The growth of tomato fruits, from ovary development to the mature green stage where fruits reach their final volume, is governed by a complex interplay of cellular processes that regulate cell division and expansion with the two processes partially overlapping during the very early stage of fruit development (Renaudin *et al.*, 2017; Penchovsky and Kaloudas, 2023). In tomato, the phase of intensive cell division occurs predominantly in the pericarp and lasts approximately 8–10 days post-anthesis, while in some cases it can be extended to 25 days. Meanwhile, the expansion of the mesocarp cells contributes to a dramatic increase in fruit size, whereby the orientation of cell expansion is important for pericarp growth and anisotropic cell expansion, while expansion along the abaxial-adaxial axis affects pericarp thickness (Musseau *et al.*, 2017). The enlargement of mesocarp cells is driven by vacuolar expansion and cytoplasmic growth, as well as endoreduplication (Cheniclet *et al.*, 2005). SlCCS52A plays a central role in facilitating the transition from mitotic cycles to endocycles, allowing cells to achieve higher ploidy levels contributing to the increased cell size (Mathieu-Rivet *et al.*, 2010). Furthermore, the cytoskeleton provides structural support for cellular enlargement by facilitating vesicle trafficking and cell wall remodelling (Mauxion *et al.*, 2021). Genes like *SUN* and *OVATE* modulate cytoskeletal organization, influencing cell division patterns and the final shape of the fruit (Wang *et al.*, 2019).

Hormonal signalling plays a pivotal role throughout fruit development (Azzi *et al.*, 2015). Auxins are key regulators of fruit set and early growth, and the transcriptional repressor Sl-IAA17 controls fruit size via the regulation of endoreduplication-related cell expansion (Su *et al.*, 2014). Gibberellins (GAs), regulated by genes like *SlGA2ox1*, further stimulate growth during the early developmental stages (Chen *et al.*, 2016). In addition, metabolic processes underpin all phases of growth, supplying the energy and building blocks necessary for cell division and expansion (Azzi *et al.*, 2015). Sugar metabolism, regulated by genes such as *SlLIN5* and *SlHXK1*, ensures the availability of carbohydrates for cell wall synthesis and energy production (Zanor *et al.*, 2009; Li *et al.*, 2023).

Ovary-to-fruit transition and initial fruit development are associated with major transcriptome changes, e.g. the differential expression of approximately 700–2800 genes in pericarp tissue of ovaries during the first 5 days post-anthesis (Pattison *et al.*, 2015; Zhang *et al.*, 2016). The differential expression of hundreds of transcription factors during the distinct stages of fruit development underscores the extent of changes in the transcriptome landscape. Surprisingly, other levels of regulation of gene expression during fruit development such as alternative splicing have received less attention. In tomato, 65% of protein coding genes have been estimated to undergo alternative splicing (Clark *et al.*, 2019). In general, the process of alternative splicing has been linked to fruit ripening and physiochemical quality attributes (Chen *et al.*, 2019; Colanero *et al.*, 2020; Zhou *et al.*, 2022). However, neither the functional relevance of specific alternative splicing events during the early stages of fruit development nor the regulators of alternative splicing have been characterized till now. RNA-seq analysis in early stages of tomato fruit development has shown a high number of alternative splicing events, while several genes are alternatively spliced in a stage-specific manner (Sun and Xiao, 2015; Wang *et al.*, 2016*a*). This also includes several members of the auxin response factor (ARF) gene family which mediate the transcription of auxin responsive genes (Zouine *et al.*, 2014).

Pre-mRNA splicing is mediated by the spliceosome, a large RNA-protein complex composed of small nuclear ribonucleoproteins (snRNPs) and various associated proteins (Reddy *et al.*, 2013; Meyer *et al.*, 2015). The core snRNPs, namely U1, U2, U4/U6, and U5, play distinct roles in recognizing splice sites, facilitating catalysis and ensuring accurate splicing. Each snRNP contains a small nuclear RNA (snRNA) bound by a set of Sm-like (LSM) proteins, which provide structural stability and are essential for snRNP assembly (Perea-Resa *et al.*, 2013).

The selection of different splice sites in the same pre-mRNA leads to the production of multiple mRNA variants from a single gene. Alternative splicing enhances proteome diversity but is also an important mechanism for the regulation of mRNA levels by introducing premature stop codons (PTC), leading to nonsense-mediated mRNA decay (NMD), or by modulating transcript stability and translation efficiency (Reddy *et al.*, 2013). The main types of alternative splicing in plants include intron retention

(IR), cassette exon (CE) or exon skipping, alternative 3' splice site (A3SS), and alternative 5' splice site usage (A5SS) (Staiger and Brown, 2013).

Regulatory proteins such as serine/arginine-rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs) modulate splice site selection and are considered as core players of alternative splicing (Staiger and Brown, 2013). SR proteins, characterized by RNA recognition motifs (RRMs) and arginine/serine-rich (RS) domains, regulate splicing by binding to cis-regulatory sequences and recruiting spliceosome components (Barta *et al.*, 2010; Howard and Sanford, 2015; Wegener & Müller-McNicoll 2019). Several studies conducted in Arabidopsis have indicated an important role of different SR proteins in plant development, including development of reproductive and vegetative organs, and flowering time (Lopato *et al.*, 1999; Ali *et al.*, 2007; Zhang and Mount, 2009; Stankovic *et al.*, 2016; Yan *et al.*, 2017). In tomato, RS2Z35 and RS2Z36 proteins have been shown to be involved in the regulation of heat stress sensitive alternative splicing and thereby contribute to thermotolerance (Rosenkranz *et al.*, 2024). Additionally, members of the SR gene family show distinct expression patterns during tomato development (Rosenkranz *et al.*, 2021). However, evidence for the role of SR proteins, or other splicing factors, in tomato plant development is lacking.

Here, an initial global expression profiling of RNA splicing-related genes in different tomato fruit tissues and developmental stages revealed that several splicing factors exhibit tissue- and stage-specific expression patterns, reflecting putative specialized roles in regulating pre-mRNA splicing during fruit development and ripening. RS2Z36 is a member of the plant-specific clade of SR splicing factors which is expressed preferentially during the early stages of fruit development. We used two tomato CRISPR/Cas9 lines of *rs2z36* with PTC that cannot produce the full length protein, to examine its role in fruit development. Both mutant lines produced fruits of smaller size and a more ellipsoid shape compared to wild type (WT). We integrated transcriptome and proteome analysis in tomato fruit ovaries that uncovered RS2Z36-dependent changes in RNA and protein abundance of splice variants that are potentially associated with fruit growth and shape. We propose a potential model in which RS2Z36 acts as an important regulator of alternative splicing during the early phase of tomato fruit development, fine-tuning transcript and protein diversity required for fruit size and morphology determination. Our study underscores the critical role of RNA splicing in shaping developmental traits in tomato.

## Materials and Methods

### Plant material

For all experiments in this study, *Solanum lycopersicum* cv. Moneymaker was used as a wild type (WT) or as a background for the generation of mutants. The mutants *rs2z35*, *rs2z36* (here: *rs2z36.1*) and the double *rs2z35 rs2z36* were described in Rosenkranz *et al*. (2024). Generation of *rs2z36.2* is described below. All plants were grown in the greenhouse with a diurnal cycle of 25°C/16 h with 120 µE

illumination and 22°C/8 h without illumination. Plants used for protoplast isolation were grown under sterile conditions in half-strength Murashige and Skoog (MS) medium solidified with GELRITE in Heraeus growth chambers (25°C/16 h with 120 µE illumination and 22°C/8 h without illumination). Mesophyll protoplast and their transformation was performed as described previously (Mishra *et al.*, 2002). For immunoblot analysis, 100.000 protoplasts were transformed with a total of 20 µg plasmid DNA (10 µl pRT-3HA-SR; adjusted to 20 µl with a mock plasmid pRT-Neo expressing the neomycin phosphatase coding gene). After PEG-mediated transformation, protoplasts were incubated for 24 h under light (120 µE). To clone pRT-3HA-SR plasmids, the genomic region of *RS2Z36* was amplified by gene specific oligonucleotides (Fw: ACCGGTACCTCCTCGTTATGATGATCGT; Rv: ACCTCTAGAGCGCAAGTTTCAAGGTGACT) from genomic DNA and cloned into a pRT vector together with an 3HA-Tag (HA: human influenza hemagglutinin) via Acc65I/XbaI sites. The generation of the *RS2Z36* WT-CDS encoding plasmid is described in Rosenkranz *et al.* (2021).

**Generation of transgenic plants**

For the generation of an additional *rs2z36* line, the same procedure was followed as described in Rosenkranz *et al*. (2024). In brief, two sgRNAs targeting RS2Z36 (gRNA1: GACGGACCCGTTCACGTGATC; gRNA2: GTGATGGACGCCGCATAATTG) were fused to the AtU6 promoter (Addgene #46968) using the MoClo Toolkit (Addgene #1000000044), and combined into the final plasmid pICSL002208 (provided by Dr. Nicola Patron from the Earlham Institute in the UK), which carries the *Streptococcus pyogenes Cas9* gene as well as the kanamycin resistance gene *NPTII*. They are under the control of the CaMV35S and NOS promoter, respectively. Plant transformation was performed according to McCormick *et al*. (1986) and generated plants were genotyped as described in Rosenkranz *et al*. (2024). Homozygous and T-DNA-free plants were used for analysis.

**Immunoblot analysis**

Total proteins were extracted from pelleted protoplasts by adding high salt lysis buffer (20 mM Tris-HCl, pH 7.8; 0.5 M NaCl; 25 mM KCl; 5 mM $MgCl_2$; 30 mM EDTA; 0.5% NP40 substitute; 0.2% sarcosyl; 5% sucrose; 5% glycerol; 0.1% ß-mercaptoethanol; supplemented with cOmplete Protease Inhibitor Cocktail, EDTA-free) (Merck) and extracts were treated as described by Port *et al.* (2004). Equal quantities of proteins were separated by sodium dodecylsulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and transferred onto 0.45 µm nitrocellulose membranes (GE Healthcare). For total protein detection, the membrane was stained with Ponceau Red. Detection of HA-tagged proteins was performed using a combination of primary (αHemagglutinin (HA), 1:2000, BioLegend) and secondary antibody (αMouse IgG, HRP-conjugated, 1:10,000, Sigma-Aldrich), followed by detection using the ECL Kit (Perkin-Elmer Life Sciences) and imaging (ECL ChemoStar 6 by Intas Science Imaging).

**Morphological and histological analyses**

For morphological analysis of the fruit shape and weight, mature fruits were collected and measured using a caliper and a precision scale. For morphological analysis of earlier developmental stages, flowers were cut along the proximal-distal axis, scanned, and the dimensions of fruits and ovaries were measured using the software ImageJ. For cell shape analysis, histological sections of ovaries at 0 days post anthesis (dpa) were prepared according to Pérez-Pérez *et al.* (2023). In brief, ovaries were fixed using 4% paraformaldehyde in phosphate-buffered saline (PBS), dehydrated in acetone, and embedded in Historesin (TECHNOVIT® 8100). Semi-thin sections were cut along the proximal-distal axis using an ultramicrotome (LKB ULTROTOME III®), mounted on glass slides and stained with toluidine blue. Pericarp cells were analysed by bright field light microscopy and measured using ImageJ. Fruit- and cell-index was calculated by dividing the maximal length of the fruit/cell by the maximal equatorial diameter.

**Immunofluorescence**

Immunofluorescence assays were performed following the protocol described by Pérez-Pérez *et al.* (2019). Semi-thin sections of ovaries (0 dpa: anthesis) were prepared as described for histological analysis, washed with PBS for 1 min, incubated for 5 min in 5% bovine serum albumin (BSA) in PBS, and then treated with the primary rat monoclonal antibody (LM6, LM19 & LM20, PlantProbes, Leeds, United Kingdom) for 1 h. Sections were then washed three times with PBS for 4 min each, incubated 45 min in the dark with the secondary antibody (anti-rat IgG-conjugated to Alexa 488, Thermo Fisher Scientific, Rockford, IL, United States) diluted 1:25 in 1% BSA, and washed three more times with PBS for 4 min each. For DNA staining, DAPI (1 mg/ml) was added to the sections, incubated for 10 min, and washed twice with $ddH_2O$ for 5 min each. The sections were then mounted with Mowiol and imaged using a confocal microscope (Leica SP5, Leica Microsystems, Wetzlar, Germany). To allow comparison between the signals, the same excitation and emission-capture settings were applied to each preparation corresponding to the same antibody. Controls were prepared following the same procedure, except the primary antibody was omitted.

**RNA isolation, cDNA synthesis and qRT-PCR**

For RNA isolation, leaflets or ovaries were homogenized and proceeded using TRIzol (Invitrogen) following the manufacturer´s instructions. For cDNA synthesis, 1 µg of total RNA was first treated with DNaseI (Roche) following reverse transcription by the RevertAid Reverse Transcriptase (Thermo Fisher Scientific) according to manufactures protocol using oligo$(dT)_{24}$ VN primers. Quantitative real-time PCR (qRT-PCR) was performed using the PowerUp SYBR Green Master Mix (Thermo Fisher Scientific; StepOne Plus Real-Time PCR system, Thermo Fisher Scientific) with gene specific primers (RS2Z36) (Fw: AATAGCACTCGTCTCTATGTG; Rv: GGATCACTAAATTCTACGAAGG). Relative transcript levels were calculated based on the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). For

normalization, EF1α (Fw: TGATCAAGCCTGGTATGGTTGT; Rv: CTGGGTCATCCTTGGAGTT) was used as an internal control. At least 3 biological replicates were performed.

**RNA-Seq and transcriptome analysis**

RNA-seq was performed to identify differentially expressed genes (DEGs) in ovaries of WT and *rs2z36* mutant lines. RNA-Seq and downstream transcriptome analyses were performed as described in Rosenkranz *et al*. (2024). Total RNA was extracted in three biological replicates of 10-15 ovaries (2 dba) per sample using TriZol (Invitrogen), followed by DirectZol column purification (Zymo Research) with DnaseI treatment. RNA quality was assessed using the Qubit IQ Assay Kit (Thermo Fisher Scientific) and RNA libraries were constructed from high-quality RNA (IQ score > 8.3) using the NEBNext Ultra II non-directional RNA Library Kit with polyA selection, followed by sequencing as 150 nt paired-end reads on an Illumina NovaSeq 6000 machine, performed by Admera Health, LLC (New Jersey, USA). Initial data processing was carried out using the Galaxy web platform (https://usegalaxy.eu) (Jalili *et al.*, 2021). Reads were aligned to the *S. lycopersicum* cv. Heinz 1706 genome (SL4.0 build) and annotated using ITAG4.0 (Fernandez-Pozo *et al.*, 2015) with STAR (Galaxy version 2.7.8a) (Dobin *et al.*, 2013), allowing soft clipping and up to 5% mismatches. Quantification of gene expression was performed using htseq-count (Galaxy version 0.9.1) (Anders *et al.*, 2015). Using the R environment (version 4.1.1) (R Core Team, 2020), removal of low abundance transcripts (sum reads <10 across all samples) was followed by differential expression analysis using DESeq2 (version 1.32.0) (Love *et al.*, 2014). Principal component analysis (PCA) was applied to identify batch effects or outlier samples, revealing WT replicate 3 as an outlier, which was excluded from further analysis. To control for high dispersion or low counts, $\log_2$-transformed fold change ($\log_2$FC) shrinkage was applied with the apeglm package (version 1.14.0) (Zhu *et al.*, 2019). Genes were considered significantly differentially expressed if their expression levels in the mutant vs. WT showed $|\log_2 FC| > 1$ and an FDR-adjusted P-value (padj) <0.05.

**Alternative splicing analysis**

Differential alternative splicing (DAS) events in ovaries (2 dba) was analysed using MAJIQ and VOILA software packages (version 2.3) (Vaquero-Garcia *et al.*, 2016) as described in Rosenkranz *et al*. (2024). First, using *majiq build* on genome-aligned RNA-Seq reads along with the corresponding ITAG4.0 GFF3 annotation file, a splice graph was generated for each gene based on local splicing variations (LSVs). Subsequently, changes in the splice junction usage and intron retention (delta percent selected index, ΔPSI) were quantified and compared between the *rs2z36* lines and the WT using *majiq deltapsi*. *Voila modulize* was employed to reconstruct binary splicing events (e.g., intron retention) from the identified LSVs. This step also calculated the probability of the |ΔPSI| of an LSV surpassing the predefined threshold of 2% (--changing-between-group-dpsi 0.02). In total, 14 distinct binary events were reconstructed, stemming either from single LSVs (e.g. intron retention) or paired LSVs (e.g. cassette exons), whereby *voila modulize* reported two junctions for each LSV. Significance of

differential alternative splicing events between *rs2z36.1* and WT was determined based on the following criteria: (i) at least one LSV junction pair had $|P(|\Delta PSI| \geq 0.02) \geq 0.9$ and $|\Delta PSI| \geq 0.05$, (ii) all involved junctions had $|\Delta PSI| \geq 0.025$, (iii) junction pairs of the same LSV exhibited inverse regulation, and (iv) the lower $|\Delta PSI|$ value in a junction pair was at least 50% of the higher $|\Delta PSI|$ value. Only binary events were analysed (IR, CE, A5SS and A3SS) while complex events were excluded.

**Proteome analysis and identification of novel protein isoforms**

High-throughput shotgun proteomics was done according to (Chaturvedi et al., 2013) with following modifications: Tomato ovaries from WT and *rs2z36.1* (three biological replicate each) were extracted for proteome analysis. Each replicate consists of ~17 ovaries (2 days before anthesis) which were freeze-dried in liquid $N_2$ and ground using mortar and pestle. The proteins were extracted, pre-fractionated (40 µg of total protein were loaded onto the gel (1D SDS-PAGE), trypsin digested and desalted (using a C18 spec plate) according to a previously described method (Chaturvedi et al., 2013; Ghatak et al., 2016). One µg of purified peptides was loaded onto a C18 reverse-phase analytical column (Thermo Scientific, EASY-Spray 50 cm, 2 µm particle size). Separation was achieved using a two-and-a-half-hour gradient method, starting with a 4–35% buffer B (v/v) gradient [79.9% ACN, 0.1% formic acid (FA), 20% ultra-high purity water (MilliQ)] over 90 minutes. Buffer A (v/v) consisted of 0.1% FA in high-purity water (MilliQ). The flow rate was set to 300 nL/min. Mass spectra were acquired in positive ion mode using a top-20 data-dependent acquisition (DDA) method. A full MS scan was performed at 70,000 resolution (m/z 200) with a scan range of 380–1800 m/z, followed by an MS/MS scan at 17,500 resolution (m/z 200). For MS/MS fragmentation, higher energy collisional dissociation (HCD) was used with a normalized collision energy (NCE) of 27%. Dynamic exclusion was set to 20 seconds.

Raw data were searched with the SEQUEST algorithm present in Proteome Discoverer version 1.3 (Thermo Scientific, Germany) described previously (Chaturvedi et al 2013; Chaturvedi et al., 2015; Ghatak et al., 2020). The identified proteins were quantitated based on total ion count and normalised using the normalised spectral abundance factor (NSAF) strategy (Paoletti et al., 2006). The MS/MS spectra of the identified proteins and their meta-information from databases have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD060312.

## Splice variant identification via long-read sequencing

Splicing libraries were sequenced at Heinrich Heine University Düsseldorf. These CCS PacBio Isoseq reads were processed with Iso-Seq tools. After sequencing the barcodes, polyA and concatemers were removed from the ccs reads with lima–isoseq (v1.11.0) and isoseq3 refine (v3.3.0). Afterwards the reads were clustered and polished with isoseq3 cluster –split-bam 0 and isoseq3 polish. The processed reads were mapped with pbmm2 align –preset ISOSEQ –sort (v1.3.0) against the reference genome Solanum lycopersicum cv. Heinz version ITAG4.0. For each sample sqanti3 generates a gtf file which was analysed by Eventpointer (v 2.4.0, with R v 3.6.3) to get a qualitative overview of splicing events. Eventpointer identifies alternative splicing events as “Cassette Exon”, “Alternative 3’ Splice Site”, “Alternative Last Exon”, “Retained Intron”, “Alternative 5’ Splice Site”, “Alternative First Exon”, “Mutually Exclusive Exons” and “Complex Event”. Existing databases were used to generate a splice variant atlas including Isoseq-Tofu, Mikado Transcriptome, NCBI-SRA_2017-2018, and NCBI-SRA_until-2016 (JBrowser from SGN). We merged all with our data using gffcompare (-S –no-merge; v 0.11.6). The tracking file was analyzed to identify shared transcripts within the datasets. Fasta sequences from splice variants were extracted and translated in all three possible forward frames of peptide sequences using the emboss tool transeq -trim -frame F (v. EMBOSS:6.6.0.0). For each translated frame of each transcript the open reading frames (ORF) were taken as input for the proteome discoverer to identify potential peptides. Later the abundant peptides were sorted if they are based on the canonical form or coming from a splice variant only.

## Co-expression analysis

Normalized expression values (RPKM) for splicing related genes were obtained by Tomato Expression Atlas (https://tea.solgenomics.net/; Supplemental Table 1). The genes were categorized based on their annotation and orthology analysis as previously shown (Supplemental Table 2; Rosenkranz *et al.*, 2021). Clusters for co-expression trends were generated by Mfuzz (Kumar and Futschik, 2007) and visualized as a heatmap with heatmapper (www.heatmapper.ca/).

## Results

### Expression of SR proteins during tomato fruit growth and development

Previous studies have shown that many multi-exon genes expressed during tomato fruit growth and development undergo alternative splicing (Sun and Xiao, 2015; Wang *et al.*, 2016*b*). To identify the molecular mechanisms that regulate RNA splicing during this developmental process we examined the expression profile of genes coding for splicing regulators using the Tomato Expression Atlas database (https://tea.solgenomics.net/), which provides information on different tissues of fruits from anthesis to ripe stage (Shinozaki *et al.*, 2018). The set of genes included orthologues of Arabidopsis splicing-related genes including spliceosomal components such as snRNPs as well as associated factors such as hnRNPs and SR proteins (Rosenkranz *et al.*, 2021). The samples included different fruit tissues and for each tissue a series of developmental stages as shown in the right panel of Fig. 1A. A co-expression analysis resulted in six clusters each comprising 21 to 57 genes (Fig. 1A-B). Genes in cluster 1 are expressed at higher levels at anthesis and then gradually decrease during fruit development, while those of cluster 5 exhibit a strong transient increase at 5 days from anthesis. Genes of cluster 2 show a gradual increase from anthesis which peaks at mature green (MG) fruits and then drops during fruit ripening, while those of cluster 6 show a similar induction but are maintained at high expression levels through fruit ripening (Fig. 1A-B). Genes in cluster 3 and 4 exhibit a steady expression, with only those of cluster 4 show a strong increase specifically in seed embryos (Fig. 1A-B).

The stage-specific expression patterns of genes encoding RNA splicing factors and associated proteins suggest that distinct splicing complexes might be selectively assembled to regulate splicing in a developmental stage-specific manner (Fig. 1C). For instance, U1 snRNP-coding genes are predominantly found in clusters 4, 5, and 6, whereas U4/U6 snRNP-coding genes are mainly grouped in clusters 1 and 2 (Fig. 1D). Among the gene families encoding core alternative splicing regulators, SR genes exhibit broad distribution across all clusters, indicating their involvement throughout various stages (Fig. 1D). In contrast, hnRNP genes are largely concentrated in clusters 2 and 3, suggesting a more restricted stage-specific role (Fig. 1D). These patterns highlight the dynamic and coordinated regulation of splicing machinery components during different stages of development.

### RS2Z36 is involved in fruit growth

We focused on cluster 5, as it contains genes showing a peak in expression in the narrow period of 5-10 dpa, coinciding with the end of cell division and the beginning of cell expansion (Azzi *et al.*, 2015). Among the genes of cluster 5, Solyc09g005980 codes for RS2Z36, a protein that belongs to the plant specific clade of SR proteins and was previously shown to be important for thermotolerance, while a mutant line of this gene does not have any phenotype in vegetative tissues under control conditions (Rosenkranz *et al.*, 2021, 2024). Interestingly, RS2Z35, the paralogue of RS2Z36, belongs to cluster 6, thereby having a different expression profile, suggesting that the two genes might have distinct roles in the regulation of alternative splicing during fruit development.

To examine whether RS2Z36 plays a role in tomato growth and development, we utilized CRISPR/Cas9 mutants with two guide RNAs (gRNAs) targeting the RRM coding region in exons 2 and 4 (Fig. 2A). The *rs2z36.1* line has an insertion of one nucleotide, generating a premature termination codon (PTC) in the RRM coding region (Fig. 2A). Both mutant lines carry a second deletion of one nucleotide in the 3' half of the RRM coding region, resulting in a PTC in *rs2z36.2* as well (Fig. 2A). We have previously shown that *rs2z36.1* has an important role in heat stress sensitive alternative splicing and thermotolerance in leaves and seedlings (Rosenkranz *et al.*, 2024). Expression analysis by qRT-PCR revealed that the transcript levels of *RS2Z36* were reduced in the mutant lines compared to the WT (Fig. 2B). This indicates that the generation of a PTC in both mutants likely results in the degradation of the *RS2Z36* RNA by NMD.

To examine whether a protein could be synthesized in the *rs2z36 mutant* lines, the genomic region of the gene from the start to the stop codon from wild type and both mutant lines were cloned under the control of CaMV35S promoter and in frame with three consecutive human influenza hemagglutinin (HA) peptide coding sequences on their N-termini. Following plasmid transfection of tomato protoplasts, immunoblot analysis with an anti-HA antibody showed that only the WT gRS2Z36 protein could be detected but not the mutants, further supporting the assumption that the mRNAs of both mutants are targeted for NMD and are not translated to a protein (Fig. 2C).

The mutant plants were phenotypically similar to the WT plants at the vegetative stage under control conditions, but showed noticeable differences in fruit morphology (Fig. 2D). The ripe fruits of both mutant lines had lower weight (Fig. 2E), similar length but reduced equatorial diameter (Fig. 2F-G), resulting in more elliptical fruits (Fig. 2H) when compared to WT. To investigate whether this phenotype is specific for RS2Z36, we also examined a CRISPR mutant line for RS2Z35 as well as a double knockout line (*rs2z25 rs2z36.1*) which have been previously described (Rosenkranz *et al.*, 2024). While the fruits of *rs2z35* plants did not appear to have an altered shape (Fig. 2E-H), the fruits of the double mutant were reduced in their length and diameter and showed a fruit index similar to those of the single *rs2z36* mutants (Fig. 2E-H). We did not observe any significant effect on the timing of the ripening process (not shown). Overall, these results suggest an involvement of RS2Z36 in fruit morphology.

Knowing that RS2Z36 is mainly expressed during the first days after anthesis (Fig. 1B), we examined the size and shape index of the ovaries from flowers at stages during the period 4 days before to 4 days after anthesis. The ovaries of the *rs2z36* lines had a more elliptical shape compared to WT at all time points (Fig. 3A-B). To examine whether this phenotype is associated with alterations in cellular morphology, we performed longitudinal sections and examined the shape index of pericarp cells from ovaries of flowers at anthesis, based on the maximum length of the equatorial diameter of cells from the equatorial region of the pericarp (Fig. 3C-D). In both mutant lines, the pericarp cells were more elongated compared to the cells in the WT ovaries, suggesting that the elliptical ovary and fruit shapes are due to cell elongation resulting from RS2Z36 mutations (Fig. 3C-D).

**RS2Z36-dependent alternative splicing in tomato ovaries**

A transcriptome analysis was performed on RNA isolated from ovaries at 2 dba stage from wild type and the *rs2z36.1* line. MAJIQ was used to analyse alternative splicing events (delta percent selected index [|ΔPSI|] ≥ 0.05 and probability changing [$P_{(|\Delta PSI| \geq 0.02)}$] ≥ 0.9), as previously described (Rosenkranz *et al.*, 2024; Supplemental Table 3). Only binary alternative splicing events were analysed, namely IR, CE, A5SS and A3SS while complex events were excluded. In total, we found 240 differential alternative splicing (DAS) events occurring in 230 genes (Fig. 4A). The majority of the events belong to A3SS (93 DAS events), followed by the other three main alternative splicing event types, namely IR, A5SS and CE with 56, 50 and 42 events, respectively (Fig. 4A).

For CE, A5SS and A3SS, the ΔPSI values were equally distributed in both directions, indicating that the mutation of RS2Z36 results both in enhancement and suppression of alternative splicing in a gene specific context (Fig. 4B). However, for IR, there was a significant shift towards intron splicing, suggesting that for these events, RS2Z36 functions primarily as intron splicing silencer (Fig. 4B).

A comparison of the RS2Z36-dependent DAS genes in ovaries with the RS2Z36-dependent DAS genes in leaves from plants grown under control conditions (Rosenkranz et al. 2024) revealed only 29 common DAS genes in both organs, suggesting that RS2Z36 has organ and tissue specific roles in tomato plant development (Fig. 4C). In addition, the iCLIP dataset generated on heat stressed leaves (Rosenkranz *et al.*, 2024) was utilized to investigate whether the RNAs of ovary DAS genes are likely to be directly bound by RS2Z36. 41% of the RS2Z36-dependent DAS genes in ovaries are listed as iCLIP targets, suggesting a possible direct regulation of these alternative splicing events by RS2Z36 (Fig. 4C).

Analysis of changes in gene expression revealed a total of 182 differentially expressed genes (DEGs), of which 137 were upregulated and 45 downregulated in the *rs2z36.1* compared to WT ovaries ($\log_2$FC >1 or <-1, padj < 0.01) (Fig. 4D; Supplemental Table 4). Only one gene was detected both as DAS and DEG, namely Solyc07g017530 annotated as “conserved oligomeric Golgi complex subunit 3”.

Gene ontology analysis of DAS genes or DEGs did not yield an enrichment of a particular biological process for the regulated genes (not shown). Based on the MapMan annotation, several genes coding for proteins related to RNA transcription, signalling, transport, and protein degradation and modification were among differentially expressed and spliced genes (Fig. 4E; Supplemental Table 5). We did not find any known tomato fruit growth related gene to be affected by *RS2Z36* mutations. However, we identified DAS genes or DEGs coding for ethylene, gibberellin and auxin related genes, cell wall proteins as well as cyclins, which could potentially be associated with fruit growth (Fig. 4E, Supplemental Tables 3-5).

**Proteome analysis for the identification of protein isoforms**

To identify whether differential expression or altered splicing in the mutants results in changes in the levels of specific proteins, we performed a proteome analysis of protein extracts from ovaries of WT and *rs2z36.1* plants. The identified peptides were filtered to retain only those matching annotated

proteins in the tomato genome, the products of canonical/constitutive splicing (Fig. 5A). We further performed differential abundance analysis in the *rs2z36.1* ovaries relative to the WT. From 1,256 annotated proteins, we identified 165 proteins with significantly altered levels (differentially abundant proteins; DAP) between the WT and *rs2z36.1* ovaries, with 96 being more abundant in WT and 69 in the mutant (Fig. 5B; Supplemental Table 6). The proteins that were more abundant in WT ovaries included, among others, several ribosomal proteins, four cell wall related proteins (endo-1,4-beta-xylanase, expansin precursor 5, expansin and pectinesterase), as well as SCL29 (Solyc01g005820), a member of the SR protein family (Supplemental Table 6). Two profilin and six histone proteins, among others, were more abundant in *rs2z36.1* ovaries compared to WT.

Furthermore, the peptides that could not be mapped to the annotated tomato proteome were screened against a library of putative tomato proteins isoforms derived from novel RNA splice variants based on an in-house Iso-Seq database (ThermoTom unpublished) and other database sources including Isoseq-Tofu, Mikado Transcriptome, NCBI-SRA_2017-2018, and NCBI-SRA_until-2016. The RNA splice variant library was generated from RNA of tomato roots, leaves, flower buds, immature and mature green fruits, exposed to 20, 35, 40, 45 °C for one hour (unpublished). Using the long-read library to screen for possible open reading frames (ORFs) that are not corresponding to the annotated proteome allowed us to identify novel splice variants with high accuracy, capturing full-length transcripts and providing a more comprehensive view of the splicing landscape. Considering that it is currently not known which RNA variants are translated, all possible ORFs that are generated by these RNAs were considered. This analysis yielded peptides that are putatively encoded by RNA splice variants of 449 genes found in our library (Fig. 5C, Supplemental Table 7), with 54 peptides being more abundant in WT and 50 in the *rs2z36.1* ovaries, which are further referred to as DAP isoforms (Fig. 5B; Supplemental Table 7). With the exception of two peptides that correspond to two splice variants of a single gene, all others are predicted to derive from different genes.

To our surprise, only 1 of the peptides identified as DAP isoform derived from an RS2Z36-dependent DAS gene. Solyc12g044860 codes for an ATP-dependent RNA helicase harboring a DEXDc and a HELICc domain (Fig. 5D). Using a combination of short and long read mapping, we were able to assign two additional protein isoforms. A5SS in intron 4 and retention of introns 6-8 result in a shorter protein isoform that carries only the HELICc domain. This variant is predicted to derive from a short ORF in exon 8 and the flanking introns, meaning that this mRNA has very long 5´ and 3´ untranslated regions (UTRs; Fig. 5D). Additionally, we were able to map another peptide to a splice variant identified by the Iso-seq library but not the DAS analysis by short reads, which spans exon 5 to exon 7 complex splicing event and in this case results in an RNA with a premature termination codon and a long 3´-UTR, which yields a truncated helicase missing the HELICc domain (Fig. 5D).

The absence of an overlap of DAS and DAP isoforms, which could also be due to the absence of complex splicing events in our DAS inventory, did not allow us to directly connect changes in alternative splicing

with specific protein abundance; however, changes in proteome diversity and abundance indicate a potential indirect link. While Gene Ontology analysis did not yield any enrichment of a particular biological process for either the DAP proteins and DAP isoforms, in both lists we were able to find proteins that might be associated with pathways regulating fruit growth and morphology such as cell wall proteins, proteins involved in cell division and cycle, and hormone metabolism (Supplemental Table 8).

## Immunostaining of cell wall components

The transcriptome and proteome analyses revealed several genes as DAS, DEG or DAP that might be associated with the phenotype of elongated pericarp cells and several of these genes were related to the cell wall. To investigate this further, we used three antibodies, namely LM6, LM19, and LM20, to detect epitopes of specific cell wall components in WT and mutant ovary sections using immunofluorescence. We observed specific labelling patterns over cell walls and controls without primary antibody did not show any labelling in all cases. Labelling with LM19 and LM20, which detect partially and highly methylesterified homogalacturonan (HG), respectively, showed no noticeable differences of localization and signal intensity between WT and mutant samples (Fig. 6). Both antibodies displayed consistent labeling patterns across the sections examined, suggesting no alteration in HG distribution or abundance in the mutant. In contrast, LM6, which recognizes (1→5)-α-L-arabinan epitopes in the arabinan components of rhamnogalacturonan-I (RG-I) and arabinogalactan protein (AGPs), showed a higher immunofluorescence intensity in the mutant ovary sections compared to WT, as well as wider distribution of the signal that homogeneously covered the cell walls in the mutant, while the WT showed a patchy pattern of labelling along cell wall structure (Fig. 6). The increased signal and modified pattern of LM6 labelling suggests enhanced arabinan deposition or changes in its structure in the mutant.

## Discussion

Despite the well-accepted importance of alternative splicing as a mechanism for controlling gene expression, its role in developmental processes like fruit development remains poorly understood. This study shows that the SR splicing factor RS2Z36 is crucial for the early establishment of fruit shape and growth. Our findings demonstrate that RS2Z36 regulates mesocarp cell morphology, as two mutant lines with disrupted RS2Z36 function exhibit elongated mesocarp cells and more ellipsoid fruits (Fig. 2-3). This phenotype is evident even in ovaries, consistent with the elevated expression of RS2Z36 during anthesis and up to five days post-anthesis (Fig. 2-3).

Interestingly, in vegetative organs, RS2Z36 is induced under heat stress, while under control conditions it exhibits low expression compared to reproductive tissues (Rosenkranz *et al.*, 2021). This expression pattern may explain the absence of vegetative phenotypes under control conditions. Furthermore, alternative splicing of RS2Z36 in vegetative organs leads to reduced levels of protein-coding transcripts, potentially further reducing its functional impact in these tissues (Rosenkranz *et al.*, 2021).

A comparative analysis between reproductive and vegetative tissues revealed a higher number of RS2Z36-dependent DAS events and genes in ovaries compared to leaves (Fig. 4; Rosenkranz *et al.*, 2024). However, the double *rs2z35 rs2z36* mutant displayed reduced growth rates in vegetative tissues under control conditions, suggesting that RS2Z35 and RS2Z36 have redundant functions in vegetative organs but not in fruit morphology (Rosenkranz *et al.*, 2024). This functional divergence aligns with co-expression analysis, which places RS2Z35 and RS2Z36 in distinct expression clusters (Fig. 1). These observations suggest that different RNA splicing complexes, composed of specific SR family members, may form at different developmental stages, contributing to tissue-specific functions. The possibility of differential complex formation is further supported by studies demonstrating specificity in interactions among SR protein pairs, as well as between SR proteins and spliceosome subunits (Golovkin and Reddy, 1998; Tanabe *et al.*, 2009; Yan *et al.*, 2017). However, detailed investigations into interactions with specific SR family members or spliceosome subunits have not yet been conducted.

Similar to leaves, the most affected alternative splicing type in ovaries upon *rs2z36* mutation is A3SS, suggesting a conserved function, although largely on different RNAs (Fig. 4). In addition, RS2Z36 mutations caused a misregulation of splicing as reported previously for leaves, meaning that for some introns canonical splicing is enhanced while for others alternative splicing is favoured (Fig. 4). Only in the case of IR, ΔPSI values were shifted predominantly to negative values, suggesting that for this type of splicing event, RS2Z36 functions as intron splicing repressor.

Alternative splicing has two major effects: it can affect transcript and protein diversity by generating mRNA variants that are translated to protein isoforms, or transcript and protein abundance by shifting RNAs for non-sense mRNA decay (NMD). While some principles for NMD targets have been proposed, two observations might speak against a generalized model: (i) almost no overlap between DAS genes and DEGs; (ii) identification of peptides from proteome analysis that likely derive from transcript variants with long 3´-UTRs, a typical feature of NMD targets. It is therefore assumed that alternative splicing of a subset of genes, e.g. transcription factors and other RNA transcription related genes, links splicing to changes in gene expression (Fig. 4) as we have previously shown for HSFA2 (Hu *et al.*, 2020). Interestingly, several DEGs code for proteins that are putatively involved in RNA processing, suggesting that there is a level of cross-regulation among the two processes, alternative splicing and transcription.

41% of the DAS genes appear to be direct targets of RS2Z36 based on iCLIP data obtained from heat stress leaves in a previous study (Rosenkranz *et al.*, 2024). The rest of the genes could be potentially regulated by other splicing factors that are misregulated in the *rs2z36* lines, as for example SR32 and SR33 are identified as RS2Z36-dependent DAS in ovaries, while SCL29 was identified as a differentially abundant protein in the proteome analysis. These findings suggest a complex network of

interactions where RS2Z36 may influence other splicing regulators, contributing to its more global effects on RNA processing.

A manual search in our transcriptome and proteome datasets for genes that have been shown to be involved in fruit development did not yield to the identification of any RS2Z36 target that could confidently explain the phenotypes of the mutant lines. Particularly, we did not find any of the known regulators to be DEG, DAS or DAP, suggesting that RS2Z36 regulates fruit growth and morphology through different targets. Several candidates could be listed as putative regulators, including DAS genes involved in hormonal signalling (auxin response factors Solyc12006350 and Solyc08g008380; ethylene responsive transcription factor-like Solyc09g059510, AP2-like ethylene responsive transcription factor TOE3), genes related to cytoskeleton (actin-related protein 8 Solyc12g049440), and calmodulin-related genes (Soly10g081170, Solyc02g079040, Solyc12g099340). Further, SlAP2a, a gene previously shown to be involved in fruit ripening and fruit size, but not shape, is an RS2Z36-dependent DAS gene (Chung *et al.*, 2010). While histological analysis point to an anisotropic expansion of mesocarp cells and not to cell division as a cause for the elongated fruit phenotype, three cyclin genes were identified as RS2Z36-dependent DAS (Solyc01g044558, Solyc10g078330, Solyc01g107030). Another interesting candidate among DAS genes is Solyc07g008670, annotated as LONGIFOLIA 1 (LNG1). LNG1 regulates leaf morphology by promoting longitudinal polar cell elongation by associating with microtubules (Lee *et al.*, 2006). These candidates warrant further investigation to elucidate their roles in mediating the observed phenotypes in *rs2z36* mutants.

Cell wall-related modifications are also likely to be related to fruit size and morphology. We identified a pectin acetyltransferase (Solyc10g086480) and a pectin lyase-like protein (Solyc09g075460) as DAS, and two cellulose synthases, a glycine-rich cell wall protein (Solyc12g098225), a hydroxyproline-rich glycoprotein (Solyc03g043740), two chitinases Solyc04g072000, Solyc10g017980), as well as pectin methylesterase inhibitors (Solyc12g009270, Solyc06g034370) and pectin acetyltransferase (Solyc08g005800) in the list of RS2Z36-dependent DEGs. We could confirm the altered abundance of LM6-recognized epitopes corresponding to (1,5)-α-L-arabinosyl residues found in the arabinan components of rhamnogalacturonan-I pectic polymers and AGPs in the cell walls of *rs2z36.1* ovaries compared to WT (Fig. 6), while two other antibodies recognizing partially and highly methylesterified HG did not yield any difference between the lines. AGPs are involved in cell expansion (Yang *et al.*, 2007) while inhibition of proline hydroxylation and thereby glycosylation of hydroxyproline-rich glycoproteins including AGPs results in stunted growth (Fragkostefanakis *et al.*, 2014, 2018). Changes in pectin metabolism during the early stages of tomato fruit development have been reported (Fragkostefanakis *et al.*, 2012; Terao *et al.*, 2013; Kutyrieva-Nowak *et al.*, 2024), however, how they contribute exactly to cell differentiation, patterning, division and growth is not understood.

In conclusion, this study highlights the role of RS2Z36 in shaping early fruit morphology by regulating mesocarp cell elongation through alternative splicing and accompanied changes in gene expression and protein abundance and diversity. Our findings suggest tissue- and stage-specific functions of RS2Z36 and hint to potentially similar functions of other SR genes in developmental programs with high relevance for crops. Further investigation of candidate genes may reveal key regulators underpinning fruit growth and morphology establishment and contribute with additional layers of knowledge to the current understanding of fleshy fruit development in tomato.

## Figure Legends

**Figure 1. Expression analysis of tomato RNA splicing related genes in tomato tissues during development and ripening.** (A) Heat map and clusters of RNA splicing related genes in different tomato tissues and developmental stages. Data have been acquired from the Tomato Expression Atlas database (Shinozaki *et al.*, 2018). The panel on the right indicates the different stages of each tissue that were included in the analysis. (B) Mean expression of genes from each cluster in pericarp tissue during fruit development. The x-axis refers to the samples shown in the panel above for pericarp. (C) Overview of pre-mRNA splicing regulation. The scheme is adopted from (Meyer *et al.*, 2015). (D) Composition of gene clusters.

**Figure 2. Mutations of *RS2Z36* gene affect fruit size and shape.** (A) CRISPR/Cas9-mediated mutations in two independent tomato lines. Boxes indicate exons, and colours coding regions for protein domains. UTR: untranslated region, PTC: premature termination codon, RRM: RNA-recognition motif, ZnK: zinc knuckle, RS: serine/arginine-rich region, SP: serine/proline-rich region. Arrows indicate the gRNA targeting region and the blue nucleotides the PAM sequence. (B) Expression of *RS2Z36* transcript levels in leaves of WT and *rs2z36* mutants based on qRT-PCR. Bars indicate the mean of three biological replicates ±SEM. Asterisks indicate statistically significant differences based on a t-test (* $p < 0.05$, ** $p < 0.01$). (C) Immunodetection of WT and mutated RS2Z36 in tomato protoplasts. Protoplasts were transformed with DNA carrying the coding sequence of the WT RS2Z36 (CDS), or the gDNA from WT or the *rs2z36* mutants. All constructs were N-terminally tagged with a 3xHA peptide. Proteins were detected with an αHA antibody. The lower blot indicates the large subunit of RuBisCO after Ponceau staining. (D) Representative images of WT and *rs2z36* mutant fruits at ripening stage. (E) Weight, (F) length, (G) equatorial dimeter, and (H) fruit index of ripe WT, *rs2z36.1*, *rs2z36.2*, *rs2z35* and the double *rs2z35 rs2z36.1*. Box plots show the median from 15-20 fruits. Statistically significant differences are based on a ANOVA tests with post-hoc Duncan's multiple range test ($p < 0.05$).

**Figure 3. The shape of ovaries from *rs2z36* mutant lines is more elliptical compared to wild type.** (A) Representative flowers of WT and *rs2z36* lines. An illustration of the ovary shape is shown on the top right. (B) Fruit index at 4 and 2 days before anthesis (dba), at anthesis (0) and at 4 days post anthesis (dpa). Box plots are based on 15-20 fruits. (C) Representative images of the equatorial region from

longitudinal sections of fruits of WT and mutants at anthesis after staining with Toluidine blue. (D) Cell index from the same region as shown in C, from 4-6 independent fruits (>40 cells from each fruit from 3 sections). Statistically significant differences between the WT and the mutants are indicated by two ($p < 0.01$) or three asterisks ($p < 0.001$) based on t-tests.

**Figure 4. Regulation of alternative splicing and gene expression by RS2Z36 in tomato ovaries**. (A) Number of DAS events and genes in four main alternative splicing categories. CE: cassette exon, IR: intron retention, A5SS and A3SS: alternative 5´ and 3´ splice site selection, respectively. (B) Violin plot of the absolute ΔPSI values of DAS events. Red line is the median and the grey lines above and below are the quartiles. (C) Pie chart showing the overlap of DAS genes identified in ovaries with DAS genes and iCLIP targets reported in leaves by Rosenkranz *et al*. (2024). (D) Number of up- and downregulated DEGs and comparison to DAS genes. (E) List of top biological processes that DAS genes and DEGs are involved in. A full list is provided in Supplemental Table 5.

**Figure 5. Changes in the proteome of ovaries due to *rs2z36* mutation.** (A) Proteome analysis workflow. (B) Number of differentially abundant annotated proteins (DAP annotated) and peptides to be the results of alternative splicing (DAP isoforms), expressed at higher levels either in the WT or *rs2z36.1* ovaries. (C) Venn diagram of DAS genes and genes corresponding to differentially abundant annotated proteins and isoforms. (D) Gene structure, read coverage, RNA splice variants, and protein isoforms with the indicated domains of *Solyc12g044860* gene. Grey boxes in the splice variants indicate the coding regions. Red boxes indicate the coding regions for the identified unannotated peptides. Violet colour indicates RNA variants. The asterisk indicates a splice variant identified in the Iso-seq library but is not an RS2Z36-dependent DAS event.

**Figure 6. Immunolocalization of AGP epitopes in ovaries at anthesis.** Longitudinal sections of ovaries from WT and *rs2z36.1* mutant line were immunolabeled with LM6 (top) recognizing (1→5)-α-L-arabinan epitopes in rhamnogalacturonan-I (RG-I) and arabinogalactan protein (AGPs), and LM19 (middle) and LM20 (bottom) detecting partially and highly methylesterified homogalacturonan (HG) epitopes. In blue, DAPI staining indicates nuclei. On the left, micrographs of bright field controls are shown. Bars are 75 µm. In the left panel, a magnification of the area indicated by the red box is shown. A negative control is provided in Supplemental Figure 1.

**Supplementary data**

Supplemental Figure 1. Immunostaining of WT and *rs2z36.1* ovary sections in the absence of a primary antibody (negative control).

Supplemental Table 1. Expression (RPKM) of splicing related genes in different tomato fruit tissues across different developmental stages.

Supplemental Table 2. Categories of splicing related genes.

Supplemental Table 3. List of all identified differentially splice events in wild type and *rs2z36.1* ovaries.

Supplemental Table 4. List of differentially expressed genes in wild type and rs2z36.1 ovaries.

Supplemental Table 5. Mapman categories of DAS genes and DEGs.

Supplemental Table 6. Differentially abundant proteins (annotated) in wild type and *rs2z36.1* mutant ovaries.

Supplemental Table 7. Identified peptides corresponding to novel protein isoforms derived from RNA splice variants in wild type and *rs2z36.1* mutant ovaries.

Supplemental Table 8. MapMan categories of annotated DAP and DAP isoforms in wild type and *rs2z36.1* mutant ovaries.

**Acknowledgements**

We thank Holger Schranz for the cultivation of plants, Daniela Bublak for the technical support and Enrico Schleiff for the valuable discussions. SV acknowledges the support with a Short Term Scientific Mission by the COST Action EPICATCH (CA19125). SF, SS and PC acknowledge network opportunities by RECROP COST Action (CA22157). This work has been supported with a grant by DFG (FR 3776/5-1) to SF.

**Author contributions**

SF, SV, RRER: conceptualization; SV, RRER, MK, YPP, SB, KJZ, JB, SS, PC, AG, LAS: methodology; SF, RRER, MK, JB, SS, PC, KZ, AG: formal analysis; SV, RRER, MK, YPP, SB, KZ, JB, SS, PC: investigation; SF, SS, PC, PST, KJZ, WW, LAS: resources; SF, SV, RRER, MK, SS, PC, KJZ: data curation; SF, SV, RRER: writing - original draft; SV, RRER, MK, YPP, SB, KZ, JB, SS, PC, PST, MMM, KZ, SF: writing - review & editing; SV, RRER, YPP, SF: visualization; SF, SS, PST, MMM, KZ: supervision; SF, PC: funding acquisition

**Conflict of interest**

No conflict of interest declared

**Funding**

This study was supported with a grant from DFG (FR3776/5-1) to SF.

**Data availability**

The RNA-Seq data underlying this article are available in the NCBI Gene Expression Omnibus (GEO) Database at www.ncbi.nlm.nih.gov/geo/ with accession number GSE286889. Proteome data is available in PRIDE database (https://www.ebi.ac.uk/pride/) with the accession number PXD060312.

## References

**Abel S, Bürstenbinder K, Müller J**. 2013. The emerging function of IQD proteins as scaffolds in cellular signaling and trafficking.

**Ali GS, Palusa SG, Golovkin M, Prasad J, Manley JL, Reddy ASN**. 2007. Regulation of plant developmental processes by a novel splicing factor. PLoS ONE **2**.

**Anders S, Pyl PT, Huber W**. 2015. HTSeq-A Python framework to work with high-throughput sequencing data. Bioinformatics doi: 10.1093/bioinformatics/btu638.

**Azzi L, Deluche C, Gévaudant F, Frangne N, Delmas F, Hernould M, Chevalier C**. 2015. Fruit growth-related genes in tomato. Journal of Experimental Botany **66**, 1075–7086.

**Barta A, Kalyna M, Reddy ASN**. 2010. Implementing a Rational and Consistent Nomenclature for Serine/Arginine-Rich Protein Splicing Factors (SR Proteins) in Plants. The Plant Cell **22**, 2926–2929.

**Chaturvedi P, Ischebeck T, Egelhofer V, Lichtscheidl I, Weckwerth W.** 2013. Cell-specific analysis of the tomato pollen proteome from pollen mother cell to mature pollen provides evidence for developmental priming. J Proteome Res **12**, 4892-4903.

**Chaturvedi P, Doerfler H, Jegadeesan S, Ghatak A, Pressman E, Castillejo MA, Wienkoop S, Egelhofer V, Firon N, Weckwerth W**. 2015. Heat-Treatment-Responsive Proteins in Different Developmental Stages of Tomato Pollen Detected by Targeted Mass Accuracy Precursor Alignment (tMAPA). J Proteome Res 14, 4463-4471.

**Chen L, Li W, Li Y, Feng X, Du K, Wang G, Zhao L**. 2019. Identified trans-splicing of YELLOW-FRUITED TOMATO 2 encoding the PHYTOENE SYNTHASE 1 protein alters fruit color by map-based cloning, functional complementation and RACE. Plant Molecular Biology **100**.

**Chen S, Wang X, Zhang L, *et al.*** 2016. Identification and characterization of tomato gibberellin 2-oxidases (GA2oxs) and effects of fruit-specific SlGA2ox1 overexpression on fruit and seed growth and development. Horticulture Research **3**.

**Cheniclet C, Rong WY, Causse M, Frangne N, Bolling L, Carde JP, Renaudin JP**. 2005. Cell expansion and endoreduplication show a large genetic variability in pericarp and contribute strongly to tomato fruit growth. Plant Physiology **139**.

**Chu YH, Jang JC, Huang Z, van der Knaap E**. 2019. Tomato locule number and fruit size controlled by natural alleles of lc and fas. Plant Direct **3**.

**Chung MY, Vrebalov J, Alba R, Lee J, McQuinn R, Chung JD, Klein P, Giovannoni J**. 2010. A tomato (Solanum lycopersicum) APETALA2/ERF gene, SlAP2a, is a negative regulator of fruit ripening. Plant Journal **64**.

**Clark S, Yu F, Gu L, Min XJ**. 2019. Expanding alternative splicing identification by integrating multiple sources of transcription data in tomato. Frontiers in Plant Science **10**.

**Clevenger JP, Van Houten J, Blackwood M, Rodríguez GR, Jikumaru Y, Kamiya Y, Kusano M, Saito K, Visa S, van der Knaap E**. 2015. Network analyses reveal shifts in transcript profiles and metabolites that accompany the expression of SUN and an elongated tomato fruit. Plant Physiology **168**.

**Colanero S, Tagliani A, Perata P, Gonzali S**. 2020. Alternative Splicing in the Anthocyanin Fruit Gene Encoding an R2R3 MYB Transcription Factor Affects Anthocyanin Biosynthesis in Tomato Fruits. Plant Communications **1**.

**Cong B, Barrero LS, Tanksley SD**. 2008. Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nature Genetics **40**.

**Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR**. 2013. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics **29**.

**Fernandez-Pozo N, Menda N, Edwards JD, *et al.*** 2015. The Sol Genomics Network (SGN)-from genotype to phenotype to breeding. Nucleic Acids Research doi: 10.1093/nar/gku1195.

**Fragkostefanakis S, Dandachi F, Kalaitzis P**. 2012. Expression of arabinogalactan proteins during tomato fruit ripening and in response to mechanical wounding, hypoxia and anoxia. Plant Physiology and Biochemistry **52**, 112–118.

**Fragkostefanakis S, Kaloudas D, Kalaitzis P**. 2018. Pyridine 2,4-dicarboxylic acid suppresses tomato seedling growth. Frontiers in Chemistry **6**, 1–9.

**Fragkostefanakis S, Sedeek KEM, Raad M, Zaki MS, Kalaitzis P**. 2014. Virus induced gene silencing of three putative prolyl 4-hydroxylases enhances plant growth in tomato (Solanum lycopersicum). Plant Molecular Biology **85**, 459–471.

**Ghatak A, Chaturvedi P, Nagler M, Roustan V, Lyon D, Bachmann G, Postl W, Schrofl A, Desai N, Varshney RK, Weckwerth W**. 2016. Comprehensive tissue-specific proteome analysis of drought stress responses in Pennisetum glaucum (L.) R. Br. (Pearl millet). J Proteomics **143**, 122-135.

**Ghatak A, Chaturvedi P, Bachmann G, Valledor L, Ramsak Z, Bazargani MM, Bajaj P, Jegadeesan S, Li W, Sun X, Gruden K, Varshney RK, Weckwerth W**. 2020. Physiological and Proteomic Signatures Reveal Mechanisms of Superior Drought Resilience in Pearl Millet Compared to Wheat. Front Plant Sci 11: 600278.

**Golovkin M, Reddy ASN**. 1998. The plant U1 small nuclear ribonucleoprotein particle 70k protein interacts with two novel serine/arginine-rich proteins. Plant Cell **10**.

**Howard JM, Sanford JR**. 2015. The RNAissance family: SR proteins as multifaceted regulators of gene expression.

**Hu Y, Mesihovic A, Jiménez-Gómez JM, Röth S, Gebhardt P, Bublak D, Bovy A, Scharf K-DD, Schleiff E, Fragkostefanakis S**. 2020. Natural variation in HsfA2 pre-mRNA splicing is associated with changes in thermotolerance during tomato domestication. New Phytologist **225**, 1297–1310.

**Jalili V, Afgan E, Gu Q, Clements D, Blankenberg D, Goecks J, Taylor J, Nekrutenko A**. 2021. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. Nucleic Acids Research **48**.

**Jiang N, Gao D, Xiao H, Van Der Knaap E**. 2009. Genome organization of the tomato sun locus and characterization of the unusual retrotransposon Rider. Plant Journal **60**.

**Kumar L, Futschik ME**. 2007. Mfuzz: A software package for soft clustering of microarray data. Bioinformation **2**.

**Kutyrieva-Nowak N, Leszczuk A, Ezzat L, *et al.*** 2024. The modified activity of prolyl 4 hydroxylases reveals the effect of arabinogalactan proteins on changes in the cell wall during the tomato ripening process. Frontiers in Plant Science **15**, 1–21.

**Lee YK, Kim GT, Kim IJ, Park J, Kwak SS, Choi G, Chung W Il**. 2006. LONGIFOLIA1 and LONGIFOLIA2, two homologous genes, regulate longitudinal cell elongation in Arabidopsis. Development **133**.

**Li J, Liu Y, Zhang J, Cao L, Xie Q, Chen G, Chen X, Hu Z**. 2023. Suppression of a hexokinase gene SlHXK1 in tomato affects fruit setting and seed quality. Plant Physiology and Biochemistry **205**.

**Liu J, Van Eck J, Cong B, Tanksley SD**. 2002. A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proceedings of the National Academy of Sciences of the United States of America **99**.

**Livak KJ, Schmittgen TD**. 2001. Analysis of relative gene expression data using real-time quantitative PCR and. Methods **25**, 402–408.

**Lopato S, Kalyna M, Dorner S, Kobayashi R, Krainer AR, Barta A**. 1999. atSRp30, one of two SF2/ASF-like proteins from Arabidopsis thaliana, regulates splicing of specific plant genes. Genes and Development **13**, 987–1001.

**Love MI, Huber W, Anders S**. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology doi: 10.1186/s13059-014-0550-8.

**Mathieu-Rivet E, Gévaudant F, Sicard A, *et al.*** 2010. Functional analysis of the anaphase promoting complex activator CCS52A highlights the crucial role of endo-reduplication for fruit growth in tomato. Plant Journal **62**.

**Mauxion JP, Chevalier C, Gonzalez N**. 2021. Complex cellular and molecular events determining fruit size.

**McCormick S, Niedermeyer J, Fry J, Barnason A, Horsch R, Fraley R**. 1986. Leaf disc transformation of cultivated tomato (L. esculentum) using Agrobacterium tumefaciens. Plant Cell Reports **5**, 81–84.

**Meyer K, Koester T, Staiger D**. 2015. Pre-mRNA splicing in plants: In vivo functions of RNA-binding proteins implicated in the splicing process. Biomolecules **5**, 1717–1740.

**Mishra SK, Tripp J, Winkelhaus S, Tschiersch B, Theres K, Nover L, Scharf K-D**. 2002. In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of thermotolerance in tomato. Genes & development **16**, 1555–1567.

**Muños S, Ranc N, Botton E, *et al.*** 2011. Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. Plant Physiology **156**.

**Musseau C, Just D, Jorly J, Gévaudant F, Moing A, Chevalier C, Lemaire-Chamley M, Rothan C, Fernandez L**. 2017. Identification of two new mechanisms that regulate fruit growth by cell expansion in tomato. Frontiers in Plant Science **8**.

**Paran I, Van Der Knaap E**. 2007. Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper.

**Pattison RJ, Csukasi F, Zheng Y, Fei Z, van der Knaap E, Catalá C**. 2015. Comprehensive tissue-specific transcriptome analysis reveals distinct regulatory programs during early tomato fruit development. Plant Physiology **168**, 1684–1701.

**Penchovsky R, Kaloudas D**. 2023. Molecular factors affecting tomato fruit size. Plant Gene **33**, 100395.

**Perea-Resa C, Hernández-Verdeja T, López-Cobollo R, Castellano M del M, Salinas J**. 2013. LSM proteins provide accurate splicing and decay of selected transcripts to ensure normal Arabidopsis development. Plant Cell **24**.

**Pérez-Pérez Y, Carneros E, Berenguer E, Solís MT, Bárány I, Pintos B, Gómez-Garay A, Risueño MC, Testillano PS**. 2019. Pectin de-methylesterification and AGP increase promote cell wall remodeling and are required during somatic embryogenesis of quercus suber. Frontiers in Plant Science **9**.

**Pérez-Pérez Y, Solís MT, Albacete A, Testillano PS**. 2023. Opposite Auxin Dynamics Determine the Gametophytic and Embryogenic Fates of the Microspore. International Journal of Molecular Sciences **24**.

**Port M, Tripp J, Zielinski D, Weber C, Heerklotz D, Winkelhaus S, Bublak D, Scharf K-D**. 2004. Role of Hsp17.4-CII as coregulator and cytoplasmic retention factor of tomato heat stress transcription factor HsfA2. Plant physiology **135**, 1457–1470.

**R Core Team**. 2020. R: A language and environment for statistical computing R Foundation for

Statistical. URL: https://www.r-project.org/. *in press*.

**Reddy ASN, Marquez Y, Kalyna M, Barta A**. 2013. Complexity of the Alternative Splicing Landscape in Plants. The Plant Cell doi: 10.1105/tpc.113.117523.

**Renaudin JP, Deluche C, Cheniclet C, Chevalier C, Frangne N**. 2017. Cell layer-specific patterns of cell division and cell expansion during fruit set and fruit growth in tomato pericarp. Journal of Experimental Botany **68**.

**Rosenkranz RREE, Bachiri S, Vraggalas S, Keller M, Simm S, Schleiff E, Fragkostefanakis S**. 2021. Identification and Regulation of Tomato Serine/Arginine-Rich Proteins Under High Temperatures. Frontiers in Plant Science **12**, 1–18.

**Rosenkranz RRE, Vraggalas S, Keller M, *et al.*** 2024. A plant-specific clade of serine/arginine-rich proteins regulates RNA splicing homeostasis and thermotolerance in tomato. Nucleic Acids Research, 1–15.

**Shinozaki Y, Nicolas P, Fernandez-Pozo N, *et al.*** 2018. High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening. Nature Communications **9**.

**Snouffer A, Kraus C, van der Knaap E**. 2020. The shape of things to come: ovate family proteins regulate plant organ shape.

**Staiger D, Brown JWS**. 2013. Alternative Splicing at the Intersection of Biological Timing, Development, and Stress Responses. The Plant Cell doi: 10.1105/tpc.113.113803.

**Stankovic N, Schloesser M, Joris M, Sauvage E, Hanikenne M, Motte P**. 2016. Dynamic distribution and interaction of the arabidopsis SRSF1 subfamily splicing factors. Plant Physiology **170**.

**Su L, Bassa C, Audran C, Mila I, Cheniclet C, Chevalier C, Bouzayen M, Roustan JP, Chervin C**. 2014. The auxin Sl-IAA17 transcriptional repressor controls fruit size via the regulation of endoreduplication-related cell expansion. Plant and Cell Physiology **55**.

**Sun Y, Xiao H**. 2015. Identification of alternative splicing events by RNA sequencing in early growth tomato fruits. BMC Genomics doi: 10.1186/s12864-015-2128-6.

**Tanabe N, Kimura A, Yoshimura K, Shigeoka S**. 2009. Plant-specific SR-related protein atSR45a interacts with spliceosomal proteins in plant nucleus. Plant Molecular Biology **70**, 241–252.

**Terao A, Hyodo H, Satoh S, Iwai H**. 2013. Changes in the distribution of cell wall polysaccharides in early fruit pericarp and ovule, from fruit set to early fruit development, in tomato (Solanum lycopersicum). Journal of Plant Research **126**, 719–728.

**Vaquero-Garcia J, Barrera A, Gazzara MR, Gonzalez-Vallinas J, Lahens NF, Hogenesch JB, Lynch KW, Barash Y**. 2016. A new view of transcriptome complexity and regulation through the lens of local splicing variations. eLife **5**.

**Wang S, Chang Y, Guo J, Chen JG**. 2007. Arabidopsis Ovate Family Protein 1 is a transcriptional repressor that suppresses cell elongation. Plant Journal **50**.

**Wang Y, Clevenger JP, Illa-Berenguer E, Meulia T, Van Der Knaap E, Sun L**. 2019. A Comparison of sun, ovate, fs8.1 and Auxin Application on Tomato Fruit Shape and Gene Expression. Plant and Cell Physiology **60**.

**Wang K, Jiao Z, Xu M, Wang Y, Li R, Cui X, Gu L, Zhang S**. 2016*a*. Landscape and Fruit Developmental Regulation of Alternative Splicing in Tomato by Genome-Wide Analysis. Horticultural Plant Journal **2**, 338–350.

**Wang K, Jiao Z, Xu M, Wang Y, Li R, Cui X, Gu L, Zhang S**. 2016*b*. Landscape and Fruit Developmental Regulation of Alternative Splicing in Tomato by Genome-Wide Analysis. Horticultural Plant Journal doi: 10.1016/j.hpj.2017.01.007.

**Wegener M, Müller-McNicoll M.** 2019. View from an mRNP: The Roles of SR Proteins in Assembly, Maturation and Turnover. Adv Exp Med Biol.;1203:83-112. doi: 10.1007/978-3-030-31434-7_3. PMID: 31811631.

**Xiao H, Jiang N, Schaffner E, Stockinger EJ, Van Der Knaap E**. 2008. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science **319**.

**Yan Q, Xia X, Sun Z, Fang Y**. 2017. Depletion of Arabidopsis SC35 and SC35-like serine/arginine-rich proteins affects the transcription and splicing of a subset of genes. PLoS Genetics **13**, 1–29.

**Yang J, Sardar HS, McGovern KR, Zhang Y, Showalter AM**. 2007. A lysine-rich arabinogalactan protein in Arabidopsis is essential for plant growth and development, including cell division and expansion. Plant Journal **49**.

**Zanor MI, Osorio S, Nunes-Nesi A, *et al.*** 2009. RNA interference of LIN5 in tomato confirms its role in controlling brix content, uncovers the influence of sugars on the levels of fruit hormones, and demonstrates the importance of sucrose cleavage for normal fruit development and fertility. Plant Physiology **150**.

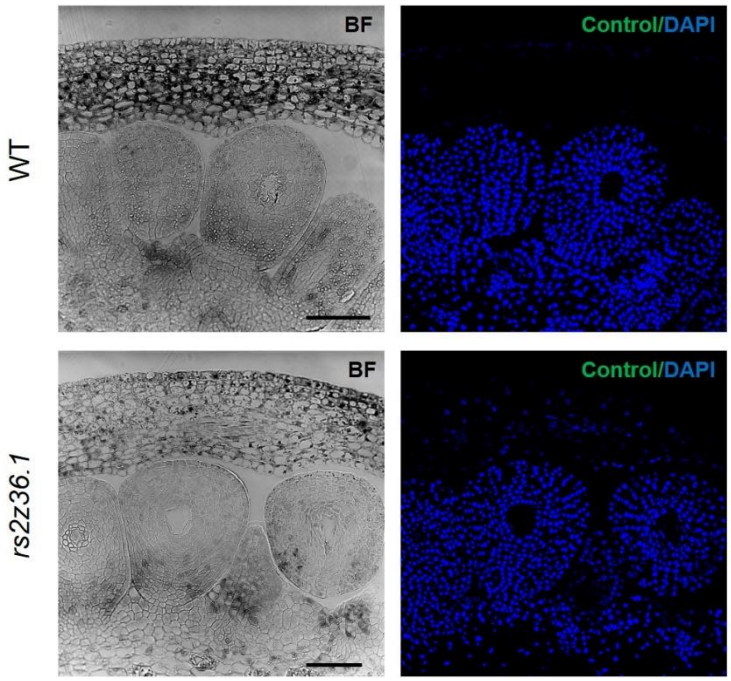
**Zhang XN, Mount SM**. 2009. Two alternatively spliced isoforms of the Arabidopsis SR45 protein have distinct roles during normal plant development. Plant Physiology **150**.

**Zhang S, Xu M, Qiu Z, Wang K, Du Y, Gu L, Cui X**. 2016. Spatiotemporal transcriptome provides insights into early fruit development of tomato (Solanum lycopersicum). Scientific Reports **6**.

**Zhou M, Deng L, Guo S, Yuan G, Li C, Li C**. 2022. Alternative transcription and feedback regulation suggest that SlIDI1 is involved in tomato carotenoid synthesis in a complex way. Horticulture Research **9**.

**Zhu A, Ibrahim JG, Love MI**. 2019. Heavy-Tailed prior distributions for sequence count data: Removing the noise and preserving large differences. Bioinformatics **35**.

**Zouine M, Fu Y, Chateigner-Boutin AL, Mila I, Frasse P, Wang H, Audran C, Roustan JP, Bouzayen M**. 2014. Characterization of the tomato ARF gene family uncovers a multi-levels post-transcriptional regulation including alternative splicing. PLoS ONE **9**.




**Supplemental Figure 1. Negative control of the immunohistochemistry samples.** Bars are 75 µm.

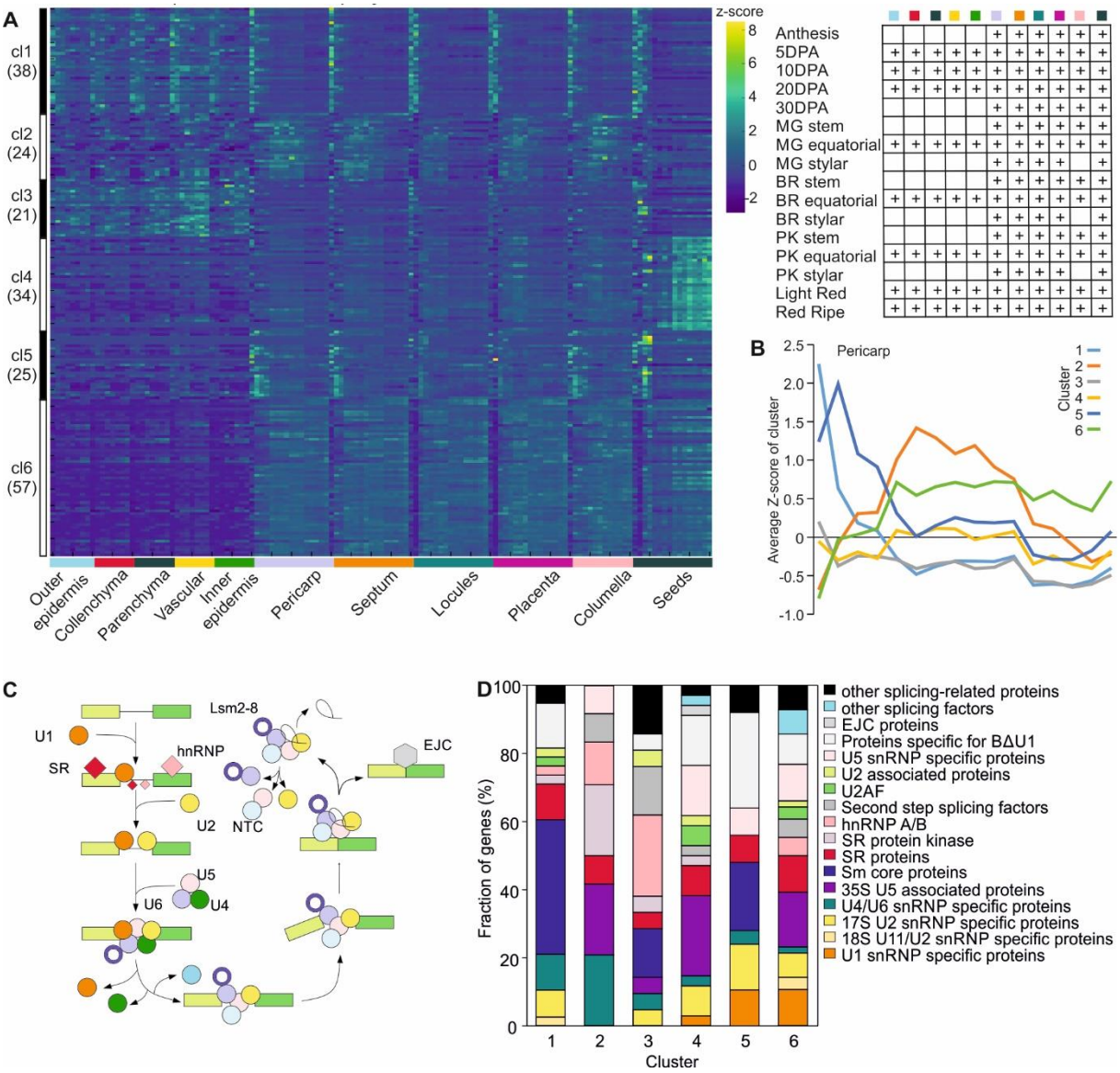



**Figure 1. Expression analysis of tomato RNA splicing related genes in tomato tissues during development and ripening.** (A) Heat map and clusters of RNA splicing related genes in different tomato tissues and developmental stages. Data have been acquired from the Tomato Expression Atlas database (Shinozaki *et al.*, 2018). The panel on the right indicates the different stages of each tissue that were included in the analysis. (B) Mean expression of genes from each cluster in pericarp tissue during fruit development. The x-axis refers to the samples shown in the panel above for pericarp. (C) Overview of pre-mRNA splicing regulation. The scheme is adopted from (Meyer *et al.*, 2015). (D) Composition of gene clusters.

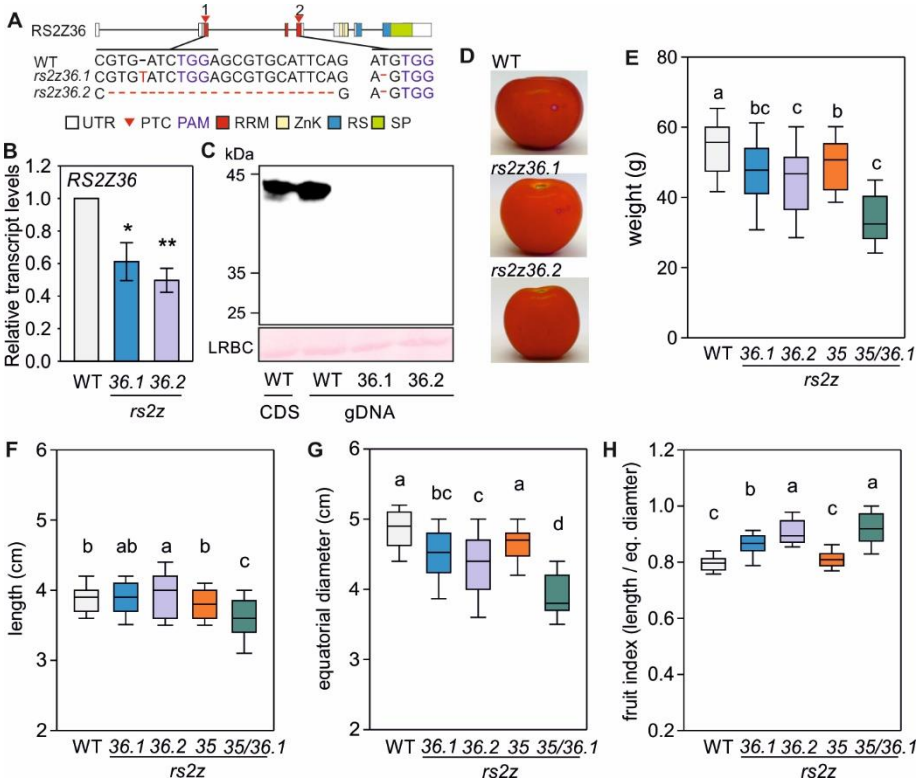



**Figure 2. Mutations of *RS2Z36* gene affect fruit size and shape.** (A) CRISPR/Cas9-mediated mutations in two independent tomato lines. Boxes indicate exons, and colours coding regions for protein domains. UTR: untranslated region, PTC: premature termination codon, RRM: RNA-recognition motif, ZnK: zinc knuckle, RS: serine/arginine-rich region, SP: serine/proline-rich region. Arrows indicate the gRNA targeting region and the blue nucleotides the PAM sequence. (B) Expression of *RS2Z36* transcript levels in leaves of WT and *rs2z36* mutants based on qRT-PCR. Bars indicate the mean of three biological replicates ±SEM. Asterisks indicate statistically significant differences based on a t-test (* $p < 0.05$, ** $p < 0.01$). (C) Immunodetection of WT and mutated RS2Z36 in tomato protoplasts. Protoplasts were transformed with DNA carrying the coding sequence of the WT RS2Z36 (CDS), or the gDNA from WT or the *rs2z36* mutants. All constructs were N-terminally tagged with a 3xHA peptide. Proteins were detected with an αHA antibody. The lower blot indicates the large subunit of RuBisCO after Ponceau staining. (D) Representative images of WT and *rs2z36* mutant fruits at ripening stage. (E) Weight, (F) length, (G) equatorial dimeter, and (H) fruit index of ripe WT, *rs2z36.1*, *rs2z36.2*, *rs2z35* and the double *rs2z35 rs2z36.1*. Box plots show the median from 15-20 fruits. Statistically significant differences are based on a ANOVA tests with post-hoc Duncan's multiple range test ($p < 0.05$).

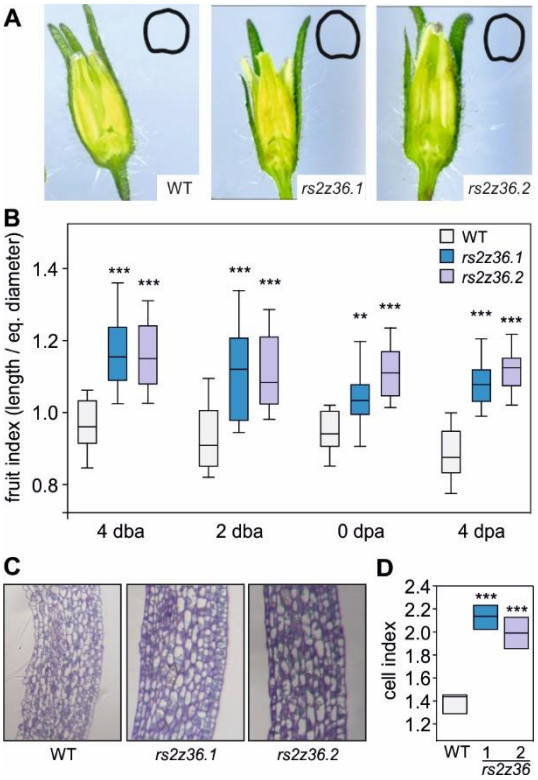



**Figure 3. The shape of ovaries from *rs2z36* mutant lines is more elliptical compared to wild type.** (A) Representative flowers of WT and *rs2z36* lines. An illustration of the ovary shape is shown on the top right. (B) Fruit index at 4 and 2 days before anthesis (dba), at anthesis (0) and at 4 days post anthesis (dpa). Box plots are based on 15-20 fruits. (C) Representative images of the equatorial region from longitudinal sections of fruits of WT and mutants at anthesis after staining with Toluidine blue. (D) Cell index from the same region as shown in C, from 4-6 independent fruits (>40 cells from each fruit from 3 sections). Statistically significant differences between the WT and the mutants are indicated by two ($p < 0.01$) or three asterisks ($p < 0.001$) based on t-tests.

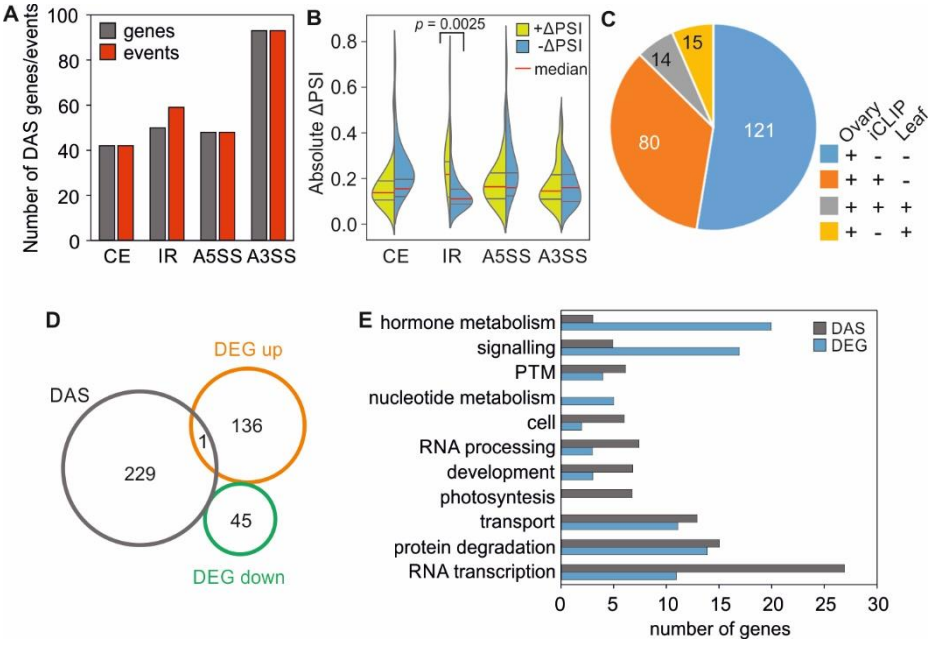



**Figure 4. Regulation of alternative splicing and gene expression by RS2Z36 in tomato ovaries**. (A) Number of DAS events and genes in four main alternative splicing categories. CE: cassette exon, IR: intron retention, A5SS and A3SS: alternative 5´ and 3´ splice site selection, respectively. (B) Violin plot of the absolute ΔPSI values of DAS events. Red line is the median and the grey lines above and below are the quartiles. (C) Pie chart showing the overlap of DAS genes identified in ovaries with DAS genes and iCLIP targets reported in leaves by Rosenkranz *et al*. (2024). (D) Number of up- and downregulated DEGs and comparison to DAS genes. (E) List of top biological processes that DAS genes and DEGs are involved in. A full list is provided in Supplemental Table 5.

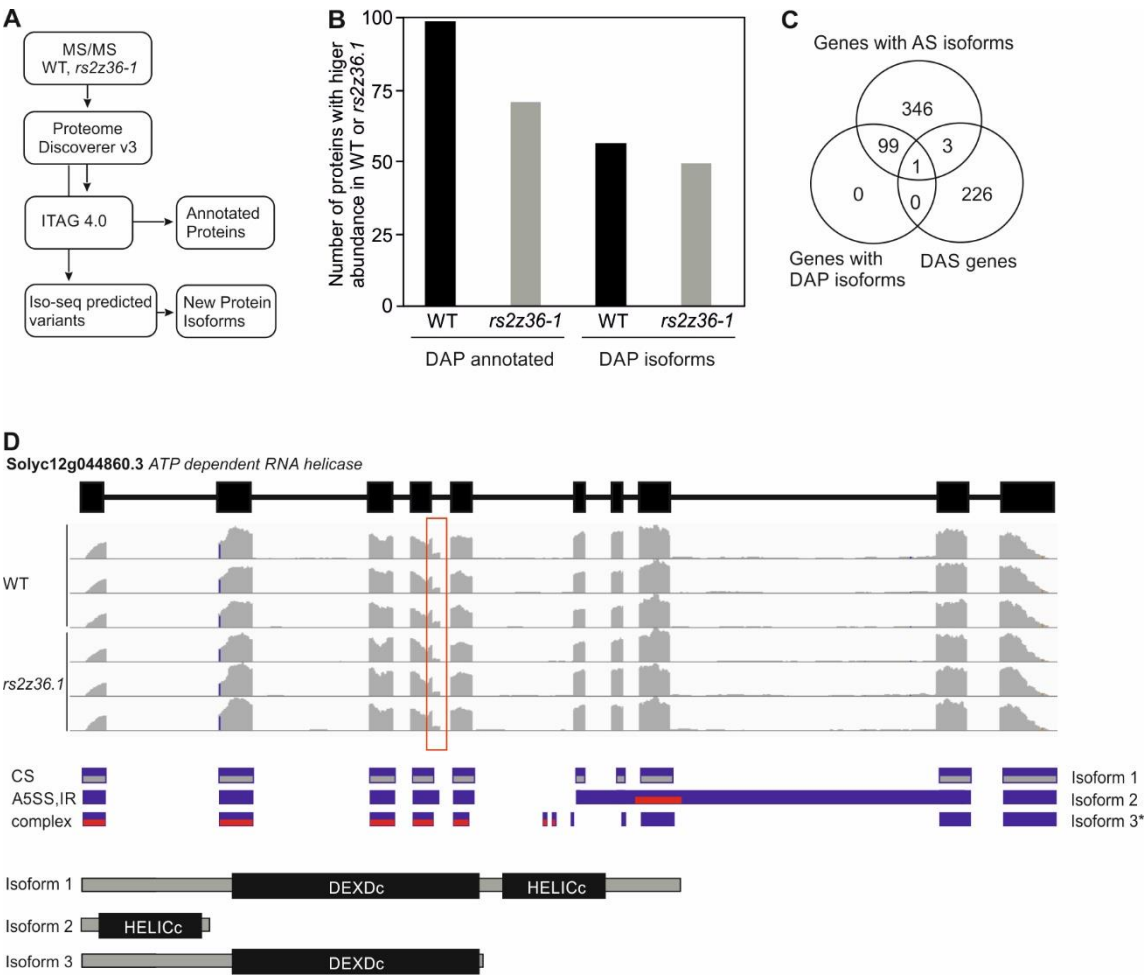



**Figure 5. Changes in the proteome of ovaries due to *rs2z36* mutation.** (A) Proteome analysis workflow. (B) Number of differentially abundant annotated proteins (DAP annotated) and peptides to be the results of alternative splicing (DAP isoforms), expressed at higher levels either in the WT or *rs2z36.1* ovaries. (C) Venn diagram of DAS genes and genes corresponding to differentially abundant annotated proteins and isoforms. (D) Gene structure, read coverage, RNA splice variants, and protein isoforms with the indicated domains of *Solyc12g044860* gene. Grey boxes in the splice variants indicate the coding regions. Red boxes indicate the coding regions for the identified unannotated peptides. Violet colour indicates RNA variants. The asterisk indicates a splice variant identified in the Iso-seq library but is not an RS2Z36-dependent DAS event.

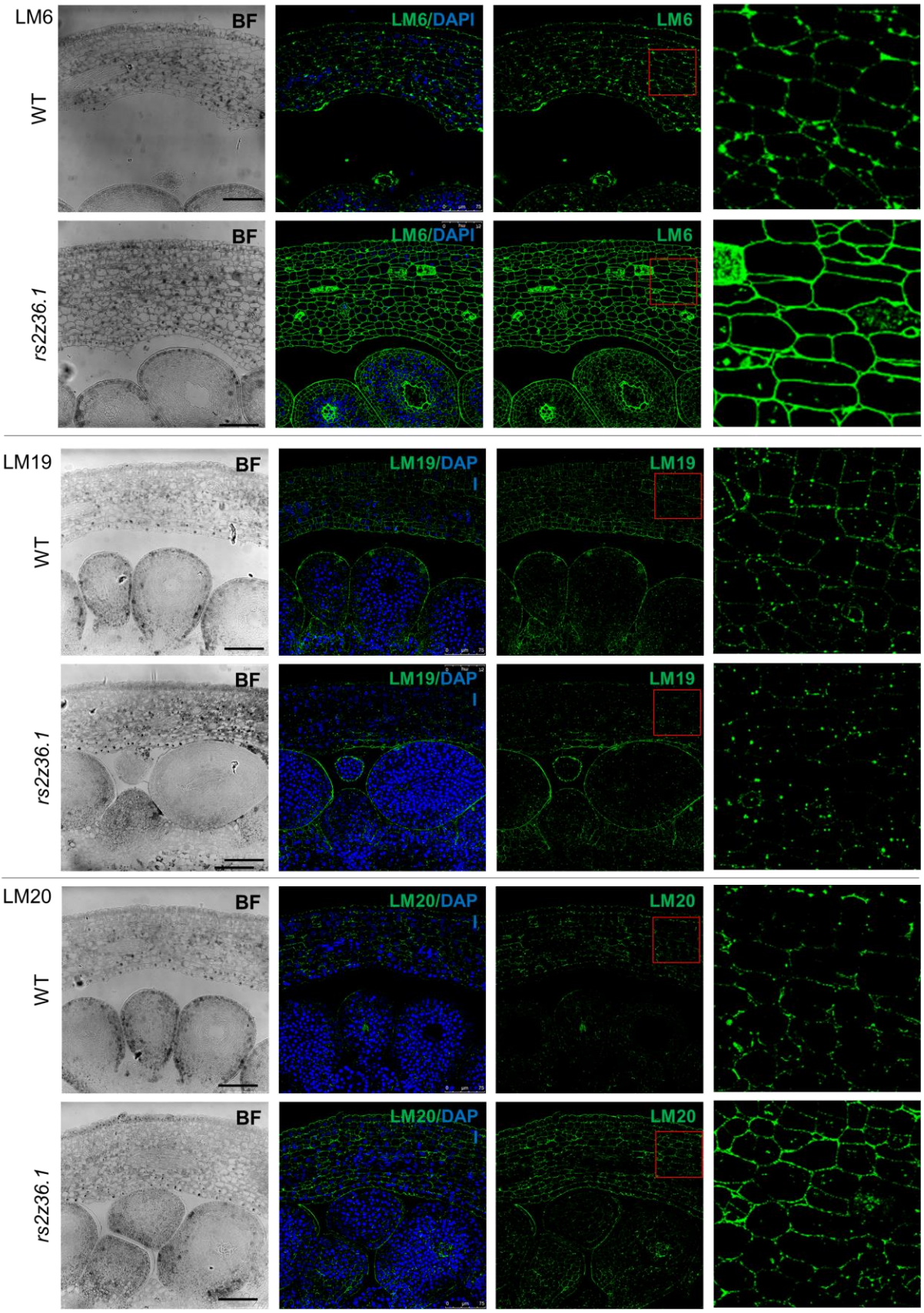
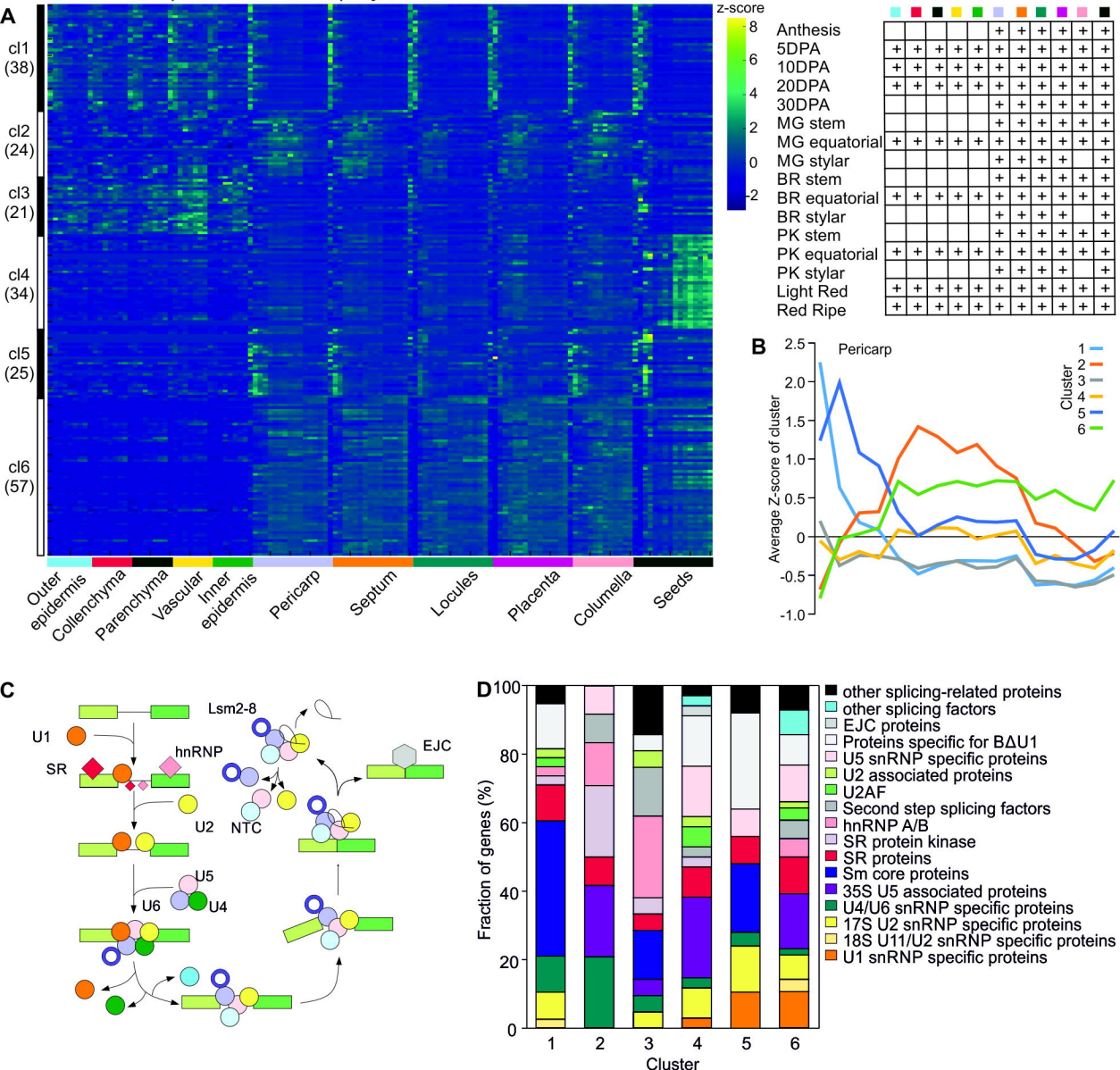
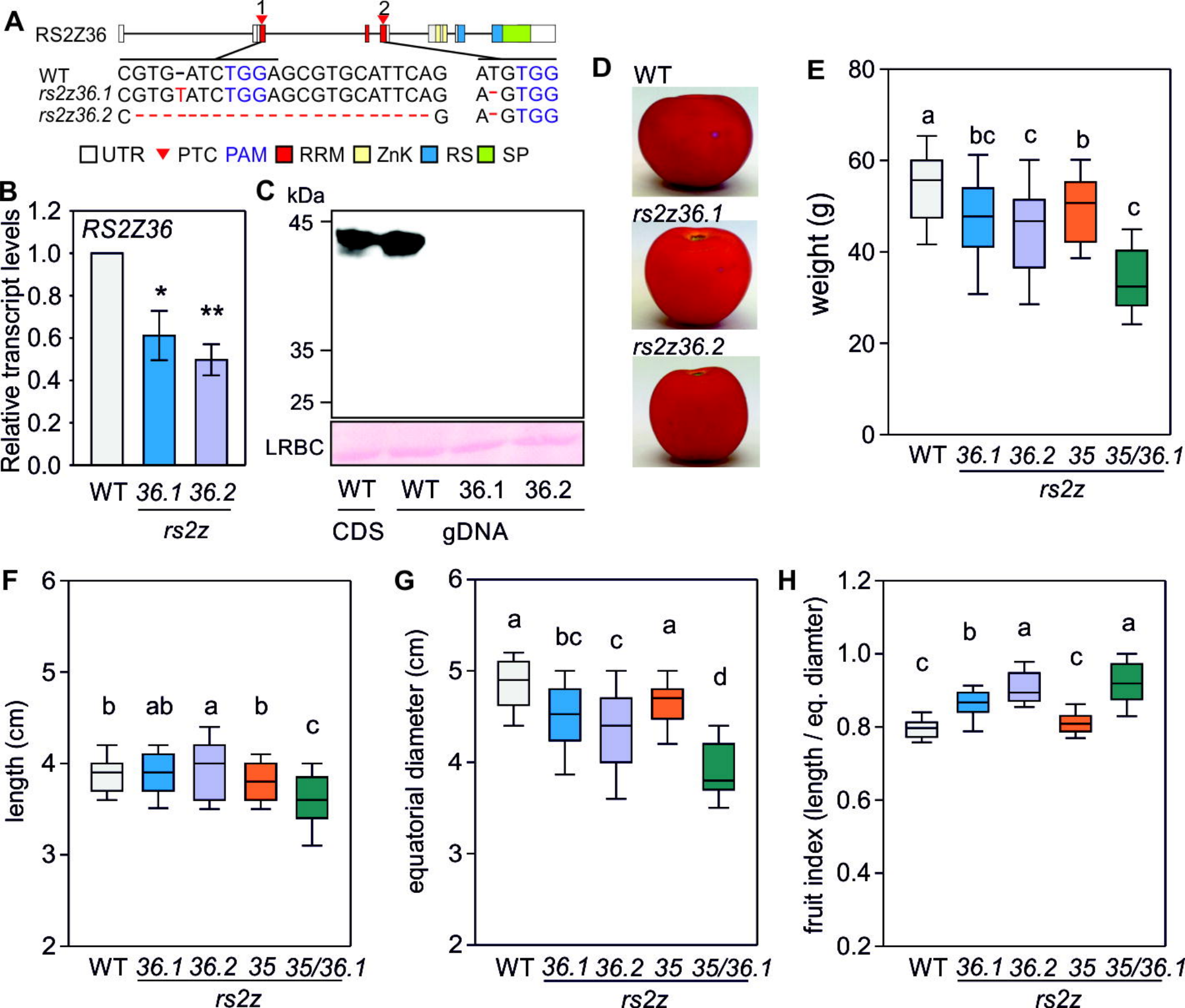
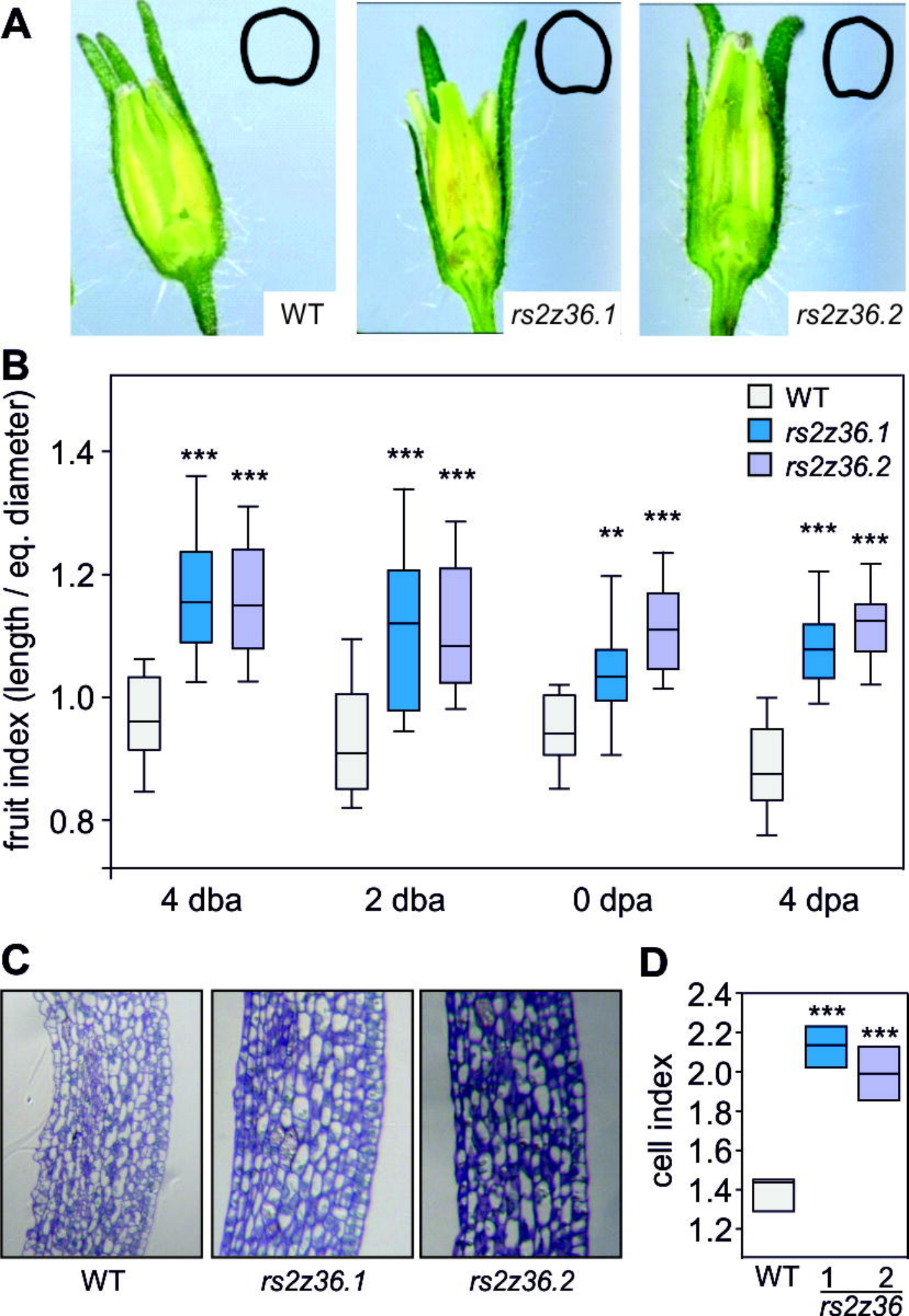


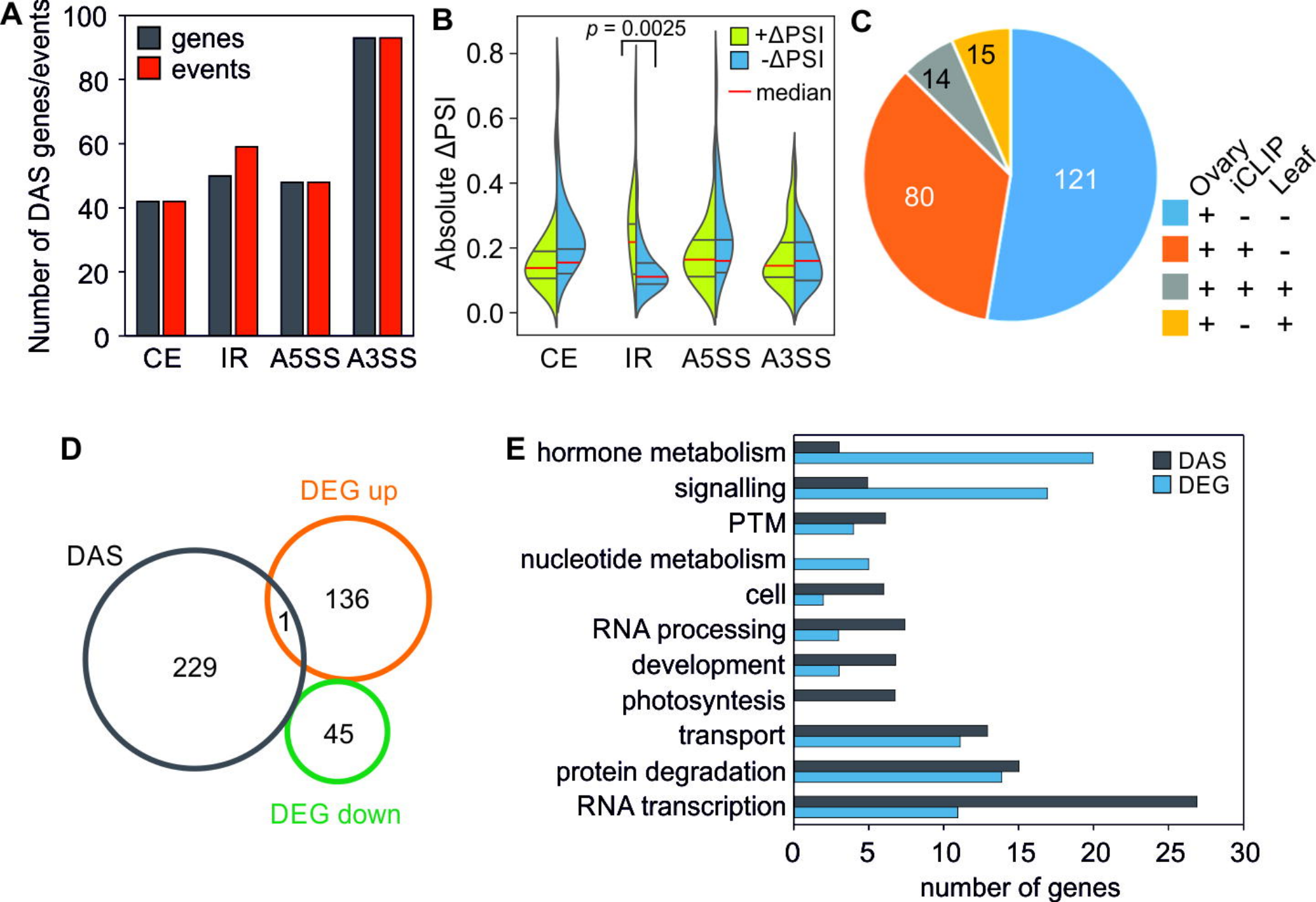

**Figure 6. Immunolocalization of AGP epitopes in ovaries at anthesis.** Longitudinal sections of ovaries from WT and *rs2z36.1* mutant line were immunolabeled with LM6 (top) recognizing (1→5)-α-L-arabinan epitopes in rhamnogalacturonan-I (RG-I) and arabinogalactan protein (AGPs), and LM19

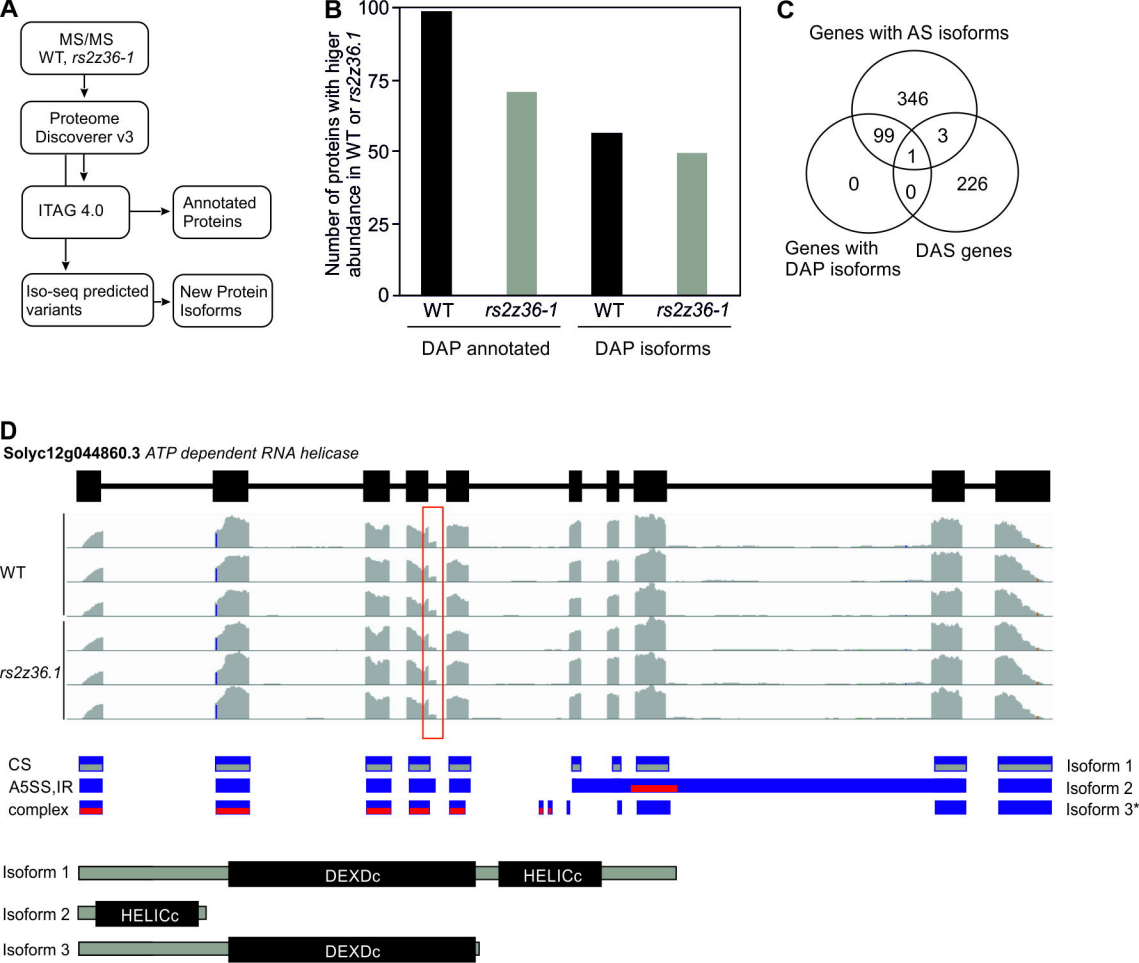
(middle) and LM20 (bottom) detecting partially and highly methylesterified homogalacturonan (HG) epitopes. In blue, DAPI staining indicates nuclei. On the left, micrographs of bright field controls are shown. Bars are 75 μm. In the left panel, a magnification of the area indicated by the red box is shown. A negative control is provided in Supplemental Figure 1.

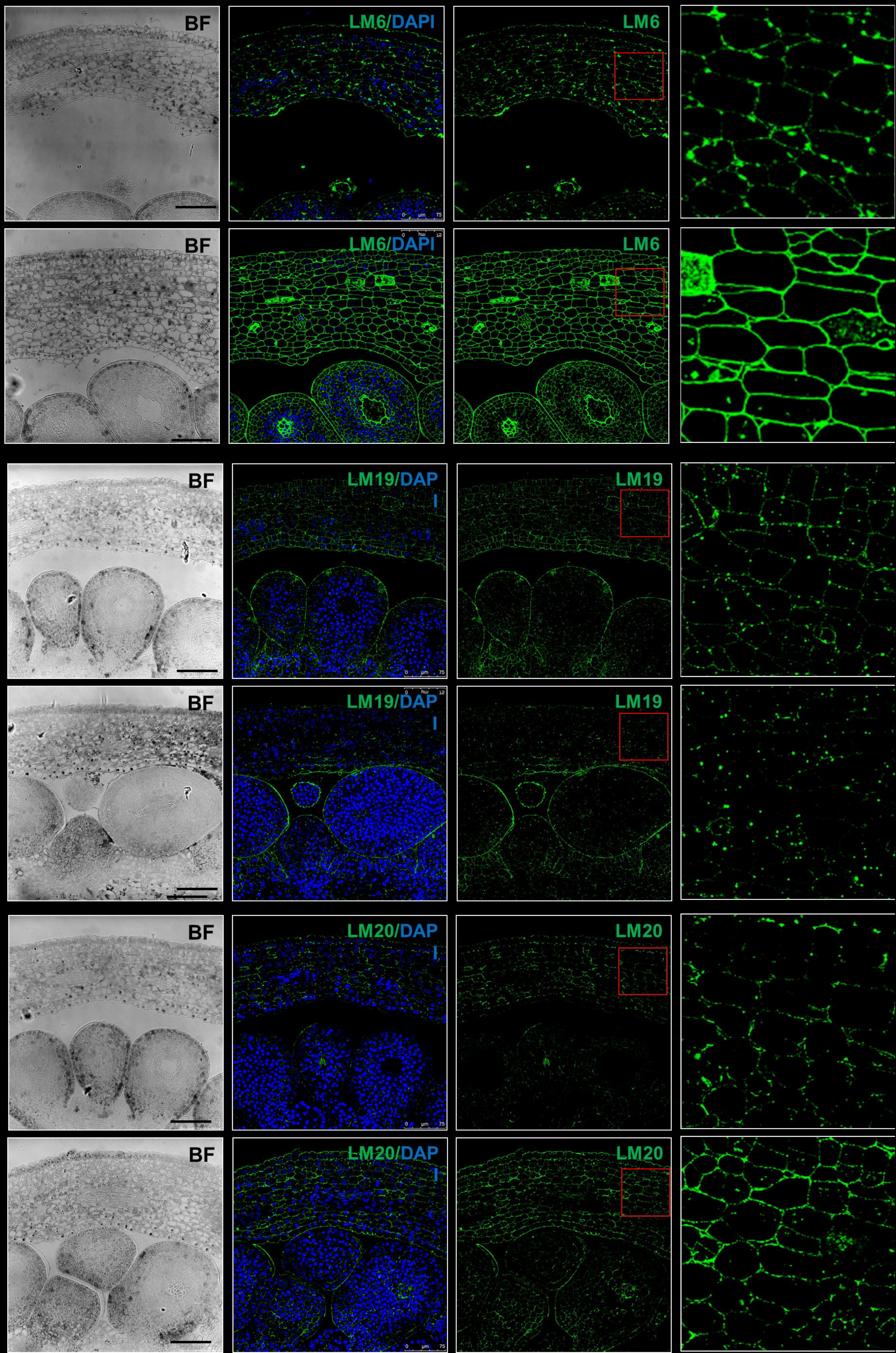

A
z-score
8
6
4
2
0
-2
cl1 (38)
cl2 (24)
cl3 (21)
cl4 (34)
cl5 (25)
cl6 (57)
Outer epidermis
Collenchyma
Parenchyma
Vascular
Inner epidermis
Pericarp
Septum
Locules
Placenta
Columella
Seeds
Anthesis
5DPA
10DPA
20DPA
30DPA
MG stem
MG equatorial
MG stylar
BR stem
BR equatorial
BR stylar
PK stem
PK equatorial
PK stylar
Light Red
Red Ripe
B
Pericarp
Average Z-score of cluster
Cluster
1
2
3
4
5
6
C
Lsm2-8
U1
SR
hnRNP
EJC
U2
NTC
U5
U6
U4
D
Fraction of genes (%)
Cluster
other splicing-related proteins
other splicing factors
EJC proteins
Proteins specific for BΔU1
U5 snRNP specific proteins
U2 associated proteins
U2AF
Second step splicing factors
hnRNP A/B
SR protein kinase
SR proteins
Sm core proteins
35S U5 associated proteins
U4/U6 snRNP specific proteins
17S U2 snRNP specific proteins
18S U11/U2 snRNP specific proteins
U1 snRNP specific proteins


A
RS2Z36
1
2
WT CGTG-ATCTGGAGCGTGCATTCAG ATGTGG
rs2z36.1 CGTGTATCTGGAGCGTGCATTCAG A-GTGG
rs2z36.2 C G A-GTGG
UTR PTC PAM RRM ZnK RS SP
B
RS2Z36
Relative transcript levels
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
**
WT 36.1 36.2
rs2z
C
kDa
45
35
25
LRBC
WT WT 36.1 36.2
CDS gDNA
D
WT
rs2z36.1
rs2z36.2
E
weight (g)
80
60
40
20
0
a bc c b c
WT 36.1 36.2 35 35/36.1
rs2z
F
length (cm)
6
5
4
3
2
b ab a b c
WT 36.1 36.2 35 35/36.1
rs2z
G
equatorial diameter (cm)
6
5
4
3
2
a bc c a d
WT 36.1 36.2 35 35/36.1
rs2z
H
fruit index (length / eq. diamter)
1.2
1.0
0.8
0.6
0.4
0.2
c b a c a
WT 36.1 36.2 35 35/36.1
rs2z


A
WT
rs2z36.1
rs2z36.2
B
fruit index (length / eq. diameter)
1.4
1.2
1.0
0.8
WT
rs2z36.1
rs2z36.2
***
***
***
***
**
***
***
***
4 dba
2 dba
0 dpa
4 dpa
C
WT
rs2z36.1
rs2z36.2
D
cell index
2.4
2.2
2.0
1.8
1.6
1.4
1.2
***
***
WT
1
2
rs2z36


A
Number of DAS genes/events
genes
events
CE
IR
A5SS
A3SS
B
Absolute ΔPSI
p = 0.0025
+ΔPSI
-ΔPSI
median
C
121
80
14
15
Ovary
iCLIP
Leaf
D
DAS
DEG up
DEG down
229
136
1
45
E
hormone metabolism
signalling
PTM
nucleotide metabolism
cell
RNA processing
development
photosyntesis
transport
protein degradation
RNA transcription
DAS
DEG
number of genes


A
MS/MS
WT, rs2z36-1
Proteome
Discoverer v3
ITAG 4.0
Annotated
Proteins
Iso-seq predicted
variants
New Protein
Isoforms
B
Number of proteins with higer
abundance in WT or rs2z36.1
100
75
50
25
0
WT
rs2z36-1
WT
rs2z36-1
DAP annotated
DAP isoforms
C
Genes with AS isoforms
346
99
3
1
0
0
226
Genes with
DAP isoforms
DAS genes
D
Solyc12g044860.3 ATP dependent RNA helicase
WT
rs2z36.1
CS
A5SS,IR
complex
Isoform 1
Isoform 2
Isoform 3*
Isoform 1
DEXDc
HELICc
Isoform 2
HELICc
Isoform 3
DEXDc


BF
LM6/DAPI
LM6
BF
LM6/DAPI
LM6
BF
LM19/DAP
I
LM19
BF
LM19/DAP
I
LM19
BF
LM20/DAP
I
LM20
BF
LM20/DAP
I
LM20