RESEARCH ARTICLE

Genome-wide identification of pistil-specific genes expressed during fruit set initiation in tomato (*Solanum lycopersicum*)

Kentaro Ezura[1], **Kim Ji-Seong**[2], **Kazuki Mori**[3], **Yutaka Suzuki**[4], **Satoru Kuhara**[3], **Tohru Ariizumi**[1,2]*, **Hiroshi Ezura**[1,2]*

1 Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan, **2** Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan, **3** Faculty of Agriculture, Kyushu University, Higashi-ku, Fukuoka, Japan, **4** Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan

* ariizumi.toru.ge@u.tsukuba.ac.jp (TA); ezura.hiroshi.fa@u.tsukuba.ac.jp (HE)

Abstract

Fruit set involves the developmental transition of an unfertilized quiescent ovary in the pistil into a fruit. While fruit set is known to involve the activation of signals (including various plant hormones) in the ovary, many biological aspects of this process remain elusive. To further expand our understanding of this process, we identified genes that are specifically expressed in tomato (*Solanum lycopersicum* L.) pistils during fruit set through comprehensive RNA-seq-based transcriptome analysis using 17 different tissues including pistils at six different developmental stages. First, we identified 532 candidate genes that are preferentially expressed in the pistil based on their tissue-specific expression profiles. Next, we compared our RNA-seq data with publically available transcriptome data, further refining the candidate genes that are specifically expressed within the pistil. As a result, 108 pistil-specific genes were identified, including several transcription factor genes that function in reproductive development. We also identified genes encoding hormone-like peptides with a secretion signal and cysteine-rich residues that are conserved among some *Solanaceae* species, suggesting that peptide hormones may function as signaling molecules during fruit set initiation. This study provides important information about pistil-specific genes, which may play specific roles in regulating pistil development in relation to fruit set.

Introduction

The pistil is a single reproductive organ that develops into a fruit after fruit set. The efficiency of fruit set is one of the most important traits that determine yield in many fruit-bearing crops such as tomato (*Solanum lycopersicum* L.). Because of its worldwide production and availability, tomato has been widely accepted as a model system for investigating fruit set. In general, fruit set is induced after successful development of the pistil upon pollination and following fertilization [1]. Through conventional molecular, genetic, and biochemical analyses of tomato, plant hormones such as auxin and gibberellic acid (GA) have been shown to play

important roles in various plant developmental processes, including inducing fruit set in the pistil [1–5]. Mimicking fruit set signals by exogenous application of these hormones and mutation of the genes related to hormone signaling or metabolism induce fruit set without pollination/fertilization, a process known as parthenocarpy [6]. Furthermore, endogenous induction of auxin biosynthesis in ovules through genetic engineering is one of the most effective approaches for inducing parthenocarpy [7]. However, the key mechanisms and signals that induce fruit set in conjunction with plant hormones in the pistil remain largely unknown. To investigate this issue, it would be useful to obtain transcriptome profiles in the pistil to uncover genes regulated by signals related to fruit set.

Microarray and next generation sequencing of transcripts (RNA-Seq) are two major transcriptome profiling systems that have been widely used in molecular biology [8]. One of the benefits of transcriptome analysis is that it allows the global gene expression profiles of thousands to nearly 40,000 genes to be investigated in a single experiment. Recently, RNA-seq has become more popular than microarray analysis for obtaining transcriptome profiles and the associated quantitative data. Comparative transcriptomics by RNA-seq produces massive amounts of accurate information about differentially expressed genes between various biological events and among related individuals, providing many clues about the mechanisms underlying plant development, growth, responses to various environmental signals, and the evolution of plant species [9–15]. In studies investigating fruit development, RNA-seq-based transcriptome analyses have revealed important biological pathways and gene sets associated with fruit development and ripening [16,17–22]. However, only a limited number of transcriptome studies have targeted pistils during fruit set in tomato [20,23–25]. These studies have identified various gene sets that appear to be expressed during fruit set, such as genes related to plant hormone metabolism and sensitivity, transcription factors regulating meristem differentiation and floral organ development, and those involved in carbohydrate metabolism [20,26]. Because of their multiple effects on various aspects of plant development, it is still difficult to narrow down candidate genes or biological pathways that directly influence the induction and completion of fruit set downstream of plant hormone signaling.

Pistil comprises a mixture of heterogeneous tissues consisting of ovules, style, placenta, and pericarp (ovary wall), which often hinders the elucidation of the detailed mechanism of early fruit development due to this inherent complexity. The development of each tissue may directly influence the success of fruit set and subsequent fruit growth. After pollination, pollen enters the ovule through the style. The fertilized ovules become seeds, which provide growth signals to the entire fruit, while the rate of cell division in the ovary wall and placenta determines the final size of the fruit [1]. Recently, cell-type-specific transcriptomes of the pistil during fruit set were uncovered by two independent groups using wild tomato *S. pimpinellifolium* and tomato cultivar 'Moneymaker', providing important information about cell type-specific transcriptomes during fruit set [23,27]. In addition, several individual pistil-specific genes (PSGs) were identified, which play important roles in processes such as pollen tube extension, pollen-pistil interactions, and ovule development, highlighting the importance of PSGs in the regulation of tissue-specific development in the pistil, including two polygalacturonase genes (*PG7* and *TAPG4*) in tomato [28], one extensin-like glycoprotein gene (*PELP3*) in *Nicotiana tabacum* [29–31], one endo-1,4-β-D-glucanase gene, and one MADS box transcription factor gene (*SEEDSTOCK/AGL11*) in *Arabidopsis thaliana* [32,33]. Nonetheless, few studies have focused on the isolation of PSGs due to technical difficulties such as the small size of the tissue. Recently, anther-specific genes were identified in various species using a transcriptomic approach, which play important roles in tissue differentiation and specification [34,35]. The isolation of genes expressed in specific tissues not only provides new insights into the development of each tissue, but it also provides genetic engineering tools for molecular breeding [36].

Therefore, to extend our understanding of the molecular mechanism underlying fruit set and to generate new tools for pistil-specific regulation of fruit set-associated genes, it is important to identify PSGs that are specifically expressed during fruit set initiation.

In this study, we conducted genome-wide analysis of PSGs in tomato by RNA-seq and compared the results with publicly available data. As a result, we identified about one hundred of PSGs including genes encoding signaling-related proteins, several transcription factors, and peptide hormone-like proteins, in addition to many genes of unknown function. Further analysis of these mined genes would increase our understanding of the mechanisms underlying of pistil development and fruit set and would be useful for generating genetic engineering tools, such as tissue-specific promoters.

Material and methods

Plant materials, hormone treatment, and cDNA synthesis

Tomato cv 'Micro-Tom' was used in this study. The seeds were incubated on wet filter paper in a Petri dish at 25˚C to stimulate germination, followed by growth in a cultivation room under a 16 h/8 h light/dark cycle at 25˚C/22˚C (day/night). Total RNA was extracted using an RNeasy Plant Mini Kit (Qiagen, USA) from 17 samples of different organs at different developmental stages: pistil and fruit samples (#1–8): pistils of 2–2.5 mm buds (#1), 3–4 mm buds (#2), 1 day before flowering (1 DBF) (#3), at anthesis (#4), 5 days after flowering (5 DAF) (#5), 5 mm ovaries of 7 days after flowering (7 DAF) (#6), mature green fruits at 33 days after flowering (MG) (#7), red fruits at 44 days after flowering (RED) (#8); stamens and other floral organ samples (#9–11): stamens of 3–4 mm buds (#9), 1 DBF (#10) and at anthesis (#11), sepals at anthesis (#12), petals at anthesis (#13), vegetative organs (#14–17 samples): 3-week-old leaves (#14), mature leaves (#15), stems (#16), and roots (#17). The total RNA was treated with DNase to remove contaminating DNA using a DNA-free RNA Kit (Zymo Research, USA). The cDNA was synthesized with 2 μg of total RNA using SuperScript VILO MasterMix (Thermo Fisher, USA) according to the manufacturer's instructions. The cDNA libraries for RNA-seq were prepared using a TruSeq RNA Sample Prep Kit v2 (Illumina) according to manufacturer's protocol.

RNA-seq, processing, mapping of Illumina reads, and detection of PSGs

The 35-nt and 100-nt single-end sequencing analysis was conducted on the Illumina Genome Analyzer IIx system and Illumina HiSeq 2000, respectively. To identify the transcriptome of each tissue, "direct-mapping method" was conducted.

For the direct-mapping method, the quality of Illumina raw FASTQ data was checked by FastQC before and after trimming with Trimmomatic according to the instruction manual (S1 Fig) [37]. After trimming, only sequences with a minimum length of 20 bp were retained. The trimmed sequence data were imported into CLC Genomics Workbench ver 7.0.4 (QIAGEN, Germany) and mapped to the tomato genome SL2.50 and gene model SL2.40. Gene expression data were obtained as gene length by reads per kilobase of exon per million mapped reads (RPKM) values [38]. The data were normalized by the quantile method to reduce obscuring of variation among samples, and a logarithmic transformation part 2 subjects the normalized data after adding 1 to each values to logarithmic transformation for heat map analysis [39,40]. To narrow down the candidate genes expressed specifically in pistils, an RPKM value of 0.5 was used as the cutoff value to determine specific expression in each sample. Based on this criterion, candidate tomato PSGs with values higher than 0.5 in pistils and lower than 0.5 other tissues were identified by comparing the transcriptome data for each tissue.

Data mining of publically available RNA-seq data

To examine the expression patterns of the identified genes in tissues other than pistils, publically available data were downloaded from transcriptome analyses of tomato from the Tomato Functional Genomics Database (http://ted.bti.cornell.edu/cgi-bin/TFGD/digital/home.cgi). Data from nine different vegetative samples from tomato cv. Heinz and wild tomato species *S. pimpinellifolium* were extracted and investigated to determine whether the candidate genes were expressed in these tissues.

To estimate the regions in the pistil in which the candidate genes are expressed, tissue-specific transcriptome data from the pistils of tomato wild relative *S. pimpinellifolium* [27] were used to identify genes with expression levels higher than RPM (reads per million mapped reads) = 2 in at least one sample. The expression levels of the top-ten genes in each tissue were then examined. To confirm the expression patterns of the candidate genes in the pistil, their expression levels were also investigated using transcriptome data from tomato cv. 'Moneymaker' [23]. If the expression level was higher than FPKM (Fragments Per Kilobase of exon per Million mapped fragments) 0.5 in at least one sample, it was judged to be an expressed gene. To investigate the responses of the genes to plant hormone treatment, a publically available dataset from the transcriptomes of pollinated or parthenocarpic fruit induced by hormone treatment was utilized [20]. To compare the list of differentially expressed genes with the candidate genes, unigene numbers were converted to ITAG IDs using the Unigene converter in the SGN database.

Gene ontology analysis

ITAG IDs of the candidate PSGs were used as input with the AgriGO agricultural gene ontology (GO) analysis tool (http://bioinfo.cau.edu.cn/agriGO/analysis.php) to elucidate enriched GO terms. A false discovery rate (FDR; e-value corrected for list size) of ≤ 0.05 was used as the criterion to obtain enriched GO terms.

Gene expression analysis by RT-PCR

To confirm the expression patterns of the candidate genes by RNA-seq analysis, RT-PCR was performed using cDNA samples derived from vegetative and reproductive organs, including young leaves, mature leaves, mature stems, mature roots, flower buds, and flower from 3-week-old plants. To analyze the expression patterns of the genes in ovaries or fruits before/after pollination, RT-PCR was performed using cDNA samples from tomato pistils and fruits at the corresponding developmental stages: A, pistils from 2–2.5 mm flower buds at 10 days before flowering (10 DBF); B, pistils from 3–4 mm flower buds at 7 days before flowering (7 DBF); C, pistils at 1 day before flowering (1 DBF); D, pistils at anthesis/pollination (0 DAF); E, pistils at 5 days after flowering (5 DAF); F, 5 mm ovaries at 7 days after flowering (7 DAF); G, mature green fruits at 33 days after flowering (MG); H, red fruits at 44 days after flowering (RED). Semi-quantitative reverse transcription polymerase chain reaction (RT-PCR) analysis was performed with Mastercycler ProS (Eppendorf, Germany) using an ExTaq Kit (TaKaRa Bio, Japan) and the primer sets listed in S5 Table. As an internal control for expression analysis in different organs, *SAND* expression was monitored using the primers *SAND*-F (5′ − TTGCTTGGAGGAACAGACG −3′) and *SAND*-R (5′ − GCAAACAGAACCCCTGAATC −3′) [41].

Sequence analysis of genes with unknown functions

Protein sequences were downloaded from the Sol genomic network. Sequence alignments were conducted using ClustalW in DDBJ (http://clustalw.ddbj.nig.ac.jp/). Phylogenetic trees were generated using CLC Genomic Workbench. The presence of secretion signals in the

small proteins was investigated using SignalP 4.1 Server (http://www.cbs.dtu.dk/services/SignalP/). The conserved domains and motifs within the identified proteins were searched using NCBI's Conserved Domain Database (CDD) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?) [42].

Availability of RNA-seq dataset

Transcriptome data are available at the GEO database under accession number DRA005810.

Results and discussion

Transcriptome analysis of various tomato tissues

To obtain transcriptome profiles of various tomato organs in order to identify PSGs, we performed RNA-seq analysis of 17 different floral and vegetative samples at different developmental stages (Fig 1A). We initially selected six different stages for the pistil samples (P1 to P6) and three different stages for the anther samples (A1 to A3). P1 to P3 and A1 to A2 represent samples at pre-anthesis; P1 corresponds to pistils in 2–2.5 mm flower bud, P2 and A1 correspond to pistils and anthers, respectively, in 3–4 mm flower bud, and P3 and A2 correspond to those in flower buds 1 day before flowering (1 DBF), while P4 and A3 correspond to those in flower buds at anthesis (0 DAF). P5 and P6 represent samples from post-anthesis stages: P5 and P6 correspond to pistils/fruits in flowers at 5 days after flowering (5 DAF) and in 5 mm ovaries at 7 days after flowering (7 DAF) samples, respectively. In addition, we used eight samples from different tissues. We conducted 35 nt and 50 nt single reads sequencing by Illumina GAIIx and Hiseq 2000, respectively (S1 Table). We used the "direct-mapping method" to identify sets of PSGs (Fig 1B). In the direct-mapping method, whole sequenced short reads were directly mapped onto the tomato reference genome.

The direct-mapping method is a common approach for transcriptome analysis in which sequence reads are mapped onto the reference genome of a target organism [8,38]. Our RNA-seq generated different amounts of raw data ranging from 8.32 to 39.42 million reads. After quality checking and trimming of low quality reads and adapter sequences, we obtained 7.28 to 35.23 million clean reads for mapping (S1 Table). We analyzed the reads using CLC Genomic Workbench ver. 7.0.4, a user-friendly mapping tool; 82.3% to 90.5% of the clean reads from each sample were mapped to the tomato genome SL2.40 [43] (S1 Table). An RPKM cut-off value of 0.5 was utilized to declare a locus expressed, resulting in an average of approximately 25,000 genes above the expression threshold in 17 samples (S2A Fig).

Before isolating PSGs, we examined the quality of our transcriptome data, as we used only one replicate per sample. Initially, to characterize the transcriptome data, we conducted principal component analysis (PCA) with CLC Genomic Workbench ver. 7.0.4. Component 1 explained 91% of the variation, while component 2 explained 2% of the variation, indicating that the two components together explained 93% of the variation of the 17 original variables. Samples from vegetative and reproductive organs were separated into two groups, with samples such as petals and sepals (which are components of reproductive organ but are composed of vegetative cells) located in the middle of the two groups (Fig 2A), indicating specialized transcriptomes. We investigated the expression patterns of the homologs that had already been identified as tissue-specific genes, such as YABBY transcription factor genes. In Arabidopsis, two YABBY transcription factors, *CRABS CLAW* (*CRC*) and *INNER NO OUTER* (*INO*), show pistil-specific expression and are involved in the pistil and early fruit development [44–46]. Thus, we expected the expression patterns of their homologs in tomato to show pistil-specific expression, and we therefore examined this possibility. Three of nine YABBY transcription factor genes found in the tomato genome, *SlCRCa* (*Solyc01g0101240*), *SlCRCb*
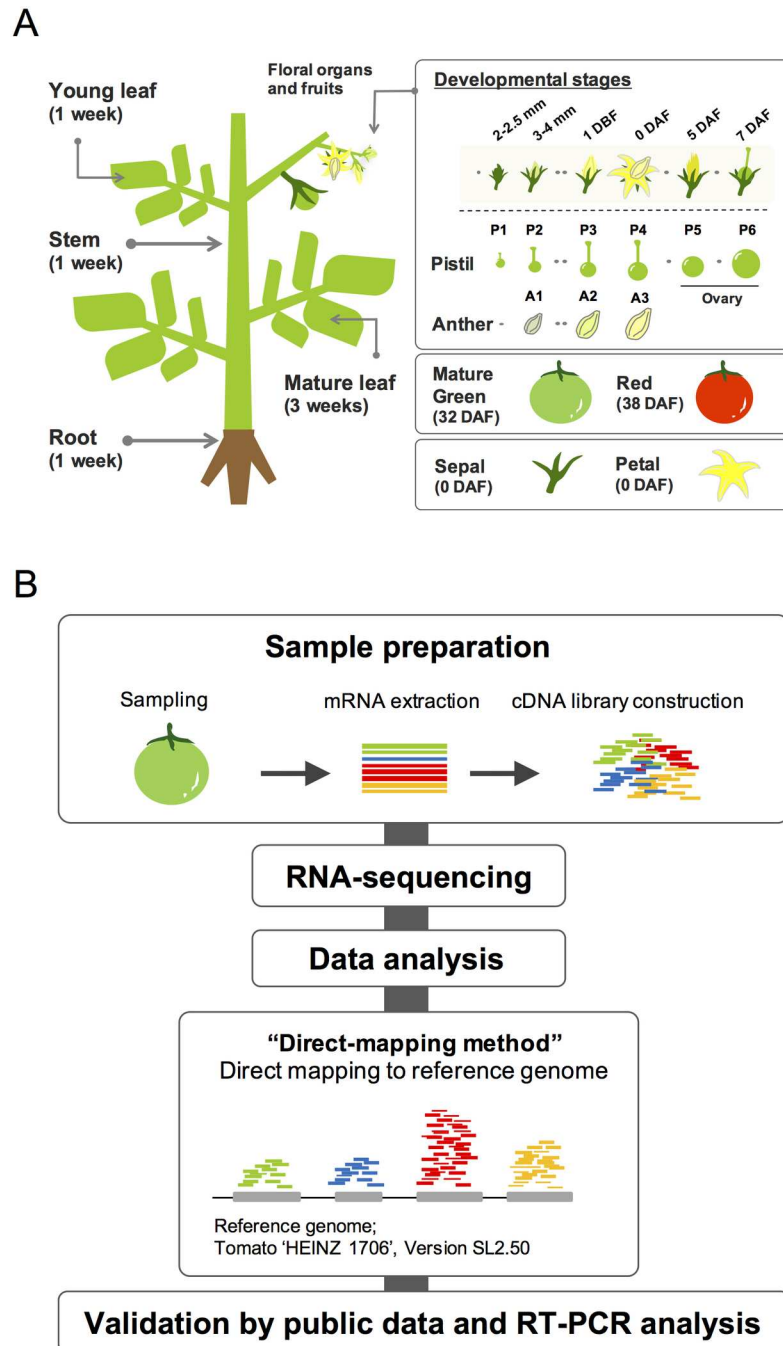
Fig 1. Experimental design for RNA-seq analysis. (A) The 17 samples used for transcriptome analysis. For vegetative organs, four samples were collected, including mature leaves from 3-week-old plants and young leaves, stems, and roots from 1-week-old plants. For reproductive organs and fruits, 13 samples were collected, including pistils and anthers of 2–2.5 mm buds, pistils and anthers of 3–4 mm buds, pistils at 1 day-before-flowering (1 DBF), pistils and anthers at anthesis (0 DAF), ovaries of 5-days after flowering (5 DAF), 5 mm ovaries (7 DAF), sepals and petals at anthesis, mature green fruits (MG), and red fruits (RED). (B) Work flow of transcriptome analyses. For the direct-mapping method, whole transcriptome data from short reads were obtained, which were directly mapped onto the tomato reference genome, and expressed genes were identified.
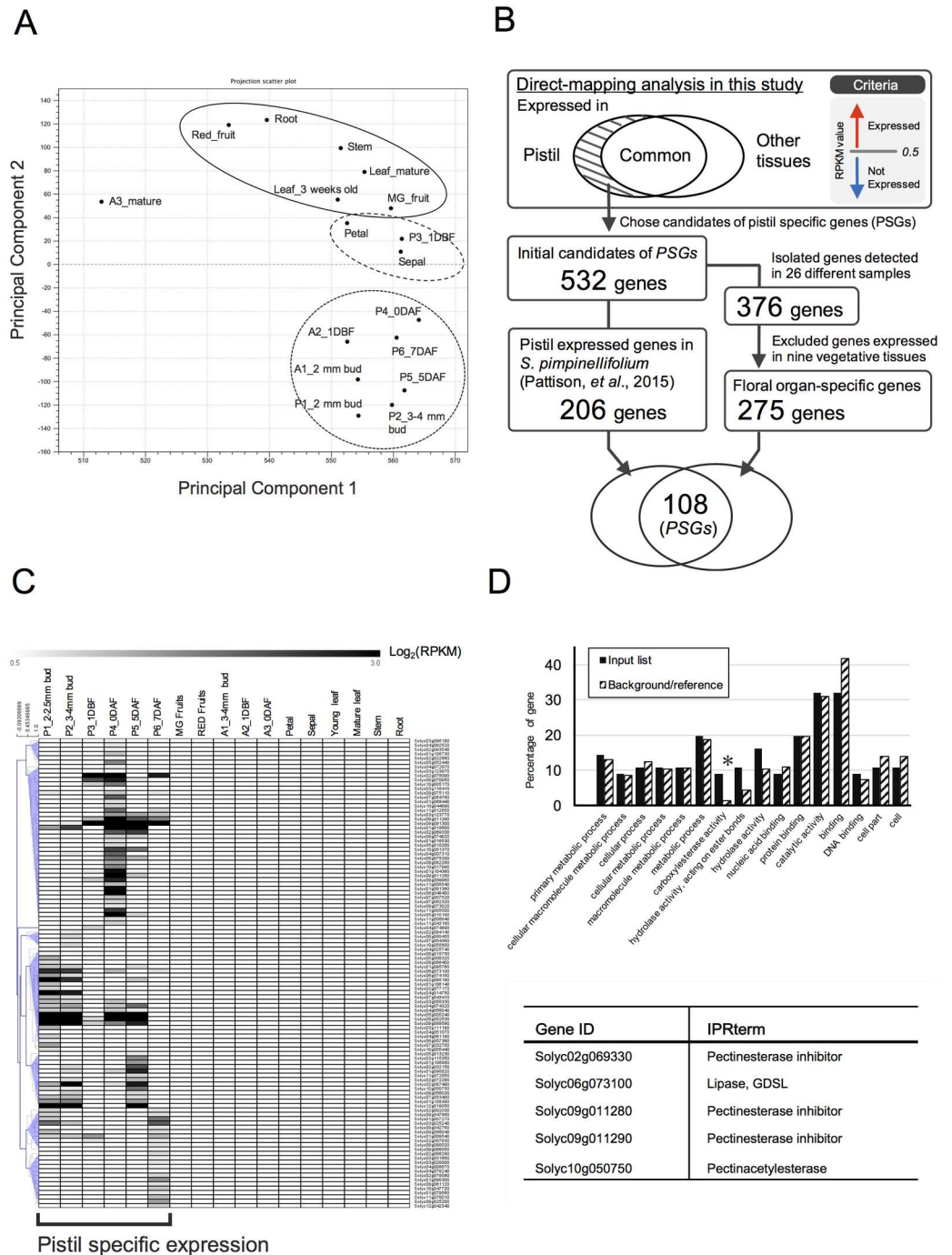
Fig 2. Identification of genes preferentially expressed in pistils based on the direct-mapping method.
(A) PCA analysis of the RNA-seq data. (B) Identification and validation of the expression of 532 candidate genes with RPKM values greater than 0.5 in at least one pistil sample and less than 0.5 in the other tissue samples using publically available data. Out of 532 genes, 206 were found to be expressed in the transcriptome produced by Pattison et al. (2015) [27]. On the other hand, the expression of 376 genes was detected in at least one sample from many different tissues and conditions, and 275 of these showed RPKM values less than 1 in nine vegetative samples (Floral organ specific genes). Finally, by comparing the two gene sets, 108 genes were found in both sets, identified as pistil-specific genes (PSGs). (C) Heatmap of the expression of 108 genes in 17 different samples. Normalized Log2-transformed expression data were visualized by constructing a heatmap using MeV software. Hierarchical clustering by Pearson correlation was conducted. (D) Gene ontology analysis was performed using AgriGO (http://bioinfo.cau.edu.cn/agriGO/). Only one category, carboxylesterase activity

(GO:0004091), was significantly (FDR<0.05) represented in the gene set, while 56 genes were not annotated and were not assigned to GO terms. bottom table; Gene ID and functional annotation described in the SGN database.

(*Solyc05g012050*), and *SlINO* (*Solyc05g005240*), were specifically expressed in flower buds and flower at the anthesis stage, which is consistent with the results obtained in a previous study [47] (S2B Fig). *SlCRCa* was expressed in the early stage of pistil development, and *SlCRCb* and *SlINO* were expressed through all stages of pistil development, while they were barely expressed in the other tissues (S2B Fig). These data support the quality of the transcriptome dataset. Next, according to RPKM values, we narrowed down the list of genes to those with RPKM values greater than 0.5 in at least one pistil sample and less than 0.5 in the other tissues, resulting in the identification of 532 of the initial candidate PSGs obtained by the direct-mapping method (Fig 2B).

Validation of the expression specificity of the candidate genes using publically available datasets

To reconfirm the tissue-specific expression of the 532 candidate genes, we performed comparative analyses between our transcriptome dataset and two publicly available transcriptome datasets (Experiment 1 and Experiment 2) from 26 samples, including vegetative tissues and floral tissues derived from tomato cv. Heinz and wild relative *S. pimpinellifolium* (strain. LA1589) available in the Tomato Functional genomics database (http://ted.bti.cornell.edu/); Experiment 1 (Exp1; Tomato Genome Consortium, 2012), Experiment 2 (Exp2; accession no. PRJNA179156). As a result, 376 of the 532 candidate genes were detected in at least one of the 26 samples from the public data, suggesting that these genes are most likely expressed in tomato plants. We investigated the expression levels of the 376 genes in nine different vegetative samples. We then excluded genes whose RPKM values were >1 in any of nine vegetative samples and identified 275 genes as "Floral organ-specific genes" (Fig 2B).

Alternatively, to obtain information about the cell types in which the candidate genes are expressed, we investigated their expression patterns in cell-type-specific transcriptome data from pistils of wild tomato (*S. pimpinellifolium*) [27]. We then selected genes expressed in pistils based on the criterion used by Pattison et al. [27]; genes with RPM values > 2 in at least one sample were chosen. In total, 206 genes were defined as "Pistil expressed genes"; their expression was evident in the pistil, especially after anthesis, while the other 326 genes excluded by this step may not be expressed in the pistil or may be expressed only at the earlier stages than 1 DBF (Fig 2B).

We compared "Floral organ-specific genes" and "Pistil expressed genes" and selected redundant genes, ultimately identifying 108 genes as PSGs by the direct-mapping method (Fig 2B and 2C, Table 1 and S2 Table). Among these, 56 genes had not been characterized. Public transcriptome data analysis provided information about both the organs and cell types in which the 108 *PSG*s were expressed. Using cell-type-specific transcriptome dataset from pistils of wild relative *S. pimpinellifolium* [27], we obtained spatial information about the expression of *PSGs* within the pistil (S3 Table). Hierarchical heat mapping clearly showed their cell-type-specific expression profiles (Fig 3A). Remarkably, roughly two-thirds of the genes appeared to show highly tissue-specific expression in the ovule and/or seed tissues (embryo, endosperm, seed coat). While many genes were preferentially expressed in the ovule and the seed tissues except seed coat, several genes were preferentially expressed in the pericarp at anthesis, in the placenta, and in the seed coat after pollination (Fig 3A). For example, five genes were preferentially expressed in the pericarp before pollination: genes encoding cinnamoyl CoA reductase-

Table 1. List of 108 pistil-specific genes (*PSGs*) identified by the direct-mapping-based method.

#	ITAG ID	Description in ITAG2.40	Homologue in Arabidopsis	length (aa)	Identities (%)	
1	Solyc01g007270	Cytokinin riboside 5'-monophosphate phosphoribohydrolase LOG (AHRD V1 **—LOG_ORYSJ)	AT5G06300	217	56/ 68	82	
2	Solyc01g008540	Cinnamoyl CoA reductase-like protein (AHRD V1 ***-B9HNY0_POPTR); Interpro domain(s) IPR016040 NAD(P)-binding domain	AT5G19440	326	223/ 315	71	NAD(P)-binding Rossmann-fold superfamily protei
3	Solyc01g010600	Homeobox-leucine zipper-like protein (AHRD V1 *-*-Q3HRT1_PICGL); contains In contains terpro domain(s) IPR001356 Homeobox	AT1G69780	294	150/ 300	50	ATHB13
4	Solyc01g016530	Unknown Protein (AHRD V1); contains Interpro domain(s) IPR008507 Protein of unknown function DUF789	AT1G73210	314	32/ 69	46	Protein of unknown function (DUF789)
5	Solyc01g068440	Os06g0207500 protein (Fragment) (AHRD V1 ***- Q0DDQ9_ORYSJ); contains Interpro domain(s) IPR004253 Protein of unknown function DUF231, plant	AT2G42570	367	166/ 341	49	TBL39 (TRICHOME BIREFRINGENCE-LIKE 39)
6	Solyc01g079560	B3 domain-containing protein Os11g0197600 (AHRD V1 ***-Y1176_ORYSJ); contains Interpro domain(s) IPR003340 Transcriptional factor B3	AT3G18990	341	30/ 92	33	VRN1, REM39
7	Solyc01g081360	Unknown Protein (AHRD V1)	-	-	-	-	
8	Solyc01g090300	Ethylene responsive transcription factor 1b (AHRD V1 *-*-C0J9I8_9ROSA); contains Interpro domain(s) IPR001471 Pathogenesis-related transcriptional factor and ERF, DNA-binding	AT2G44840	226	69/ 107	64	ATERF13, EREBP, ERF13
9	Solyc01g090820	Expansin B1 (AHRD V1 ***-C8CC40_RAPSA); contains Interpro domain(s) IPR007112 Expansin 45, endoglucanase-like	AT1G65680	273	119/ 249	48	ATEXPB2, EXPB2, ATHEXP BETA 1.4
10	Solyc01g095760	UDP-glucosyltransferase (AHRD V1 ***- Q8LKG3_STERE); contains Interpro domain(s) IPR002213 UDP-glucuronosyl/UDP-glucosyltransferase	AT5G49690	460	164/ 471	35	UDP-Glycosyltransferase superfamily protein
11	Solyc01g104390	Blue copper protein (AHRD V1 **—B6TT37_MAIZE); contains Interpro domain(s) IPR003245 Plastocyanin-like	AT1G17800	129	49/ 116	42	ARPN
12	Solyc01g106140	F-box protein family-like (AHRD V1 *-*- Q6ZCS3_ORYSJ); contains Interpro domain(s) IPR005174 Protein of unknown function DUF295	AT3G25750	348	41/ 162	25	F-box family protein with a domain of unknown function (DUF295)
13	Solyc01g106730	MADS box transcription factor 1 (AHRD V1 *-*- D9IFM1_ONCHC); contains Interpro domain(s) IPR002100 Transcription factor, MADS-box	AT5G60440	299	95/ 160	59	AGL62

(*Continued*)

Table 1. (*Continued*)

#	ITAG ID	Description in ITAG2.40	Homologue in Arabidopsis	length (aa)	Identities (%)		
14	Solyc01g106980	Endo-1 4-beta-xylanase (AHRD V1 *—B6SW51_MAIZE); contains Interpro domain(s) IPR013781 Glycoside hydrolase, subgroup, catalytic core	AT4G33840	576	276/ 545	51	Glycosyl hydrolase family 10 protein
15	Solyc01g108380	Protease inhibitor protein (AHRD V1 -**- B3FNP9_HEVBR); contains Interpro domain(s) IPR000864 Proteinase inhibitor I13, potato inhibitor I	AT2G38900	88	27/ 61	44	Serine protease inhibitor, potato inhibitor I-type family protein
16	Solyc02g022860	FAD-binding domain-containing protein (AHRD V1 **—D7MFI0_ARALY); contains Interpro domain(s) IPR006094 FAD linked oxidase, N-terminal	AT4G20820	532	243/ 532	46	FAD-binding Berberine family protein
17	Solyc02g032150	Unknown Protein (AHRD V1)	-	-	-	-	-
18	Solyc02g067630	Polygalacturonase 1 (AHRD V1 ***-O22311_SOLLC); contains Interpro domain(s) IPR000408 Regulator of chromosome condensation, RCC1 IPR000743 Glycoside hydrolase, family 28	AT2G43860	384	232/ 389	60	Pectin lyase-like superfamily protein
19	Solyc02g069330	Unknown Protein (AHRD V1); contains Interpro domain(s) IPR006501 Pectinesterase inhibitor	AT5G64620	180	26/ 80	33	C/VIF2, ATC/VIF2
20	Solyc02g072280	Subtilisin-like protease (AHRD V1 **—Q9LWA3_SOLLC); contains Interpro domain(s) IPR015500 Peptidase S8, subtilisin-related	AT5G67360	757	335/ 761	44	ARA12
21	Solyc02g077170	X1 (Fragment) (AHRD V1 *—Q7FSP8_MAIZE); contains Interpro domain(s) IPR005379 Region of unknown function XH	AT1G15910	634	96/ 259	37	XH/XS domain-containing protein
22	Solyc02g078090	Unknown Protein (AHRD V1)	-	-	-	-	-
23	Solyc02g079080	F-box family protein (AHRD V1 ***-B9GFH4_POPTR); contains Interpro domain(s) IPR001810 Cyclin-like F-box	AT5G02930	469	108/ 440	25	F-box/RNI-like superfamily protein
24	Solyc02g084140	Unknown Protein (AHRD V1)	-	-	-	-	-
25	Solyc02g085190	GATA transcription factor 19 (AHRD V1 *-** B6TS85_MAIZE); contains Interpro domain(s) IPR000679 Zinc finger, GATA-type	AT3G50870	295	127/ 292	43	MNP, HAN, GATA18
26	Solyc02g086290	Receptor serine/threonine kinase (AHRD V1 ***- Q9FF31_ARATH)	AT1G66940	332	76/ 267	28	protein kinase-related
27	Solyc02g087490	Prolyl 4-hydroxylase alpha subunit-like protein (AHRD V1 ***-Q9LSI6_ARATH); contains Interpro domain(s) IPR006620 Prolyl 4-hydroxylase, alpha subunit	AT3G28490	288	176/ 265	66	Oxoglutarate/iron-dependent oxygenase
28	Solyc02g092030	Cbs domain containing protein expressed (Fragment) (AHRD V1 *—A6N095_ORYSI); contains Interpro domain(s) IPR002550 Protein of unknown function DUF21	AT2G14520	423	283/ 423	67	CBS domain-containing protein with a domain of unknown function (DUF21)

(*Continued*)

Table 1. (*Continued*)

#	ITAG ID	Description in ITAG2.40	Homologue in Arabidopsis	length (aa)	Identities (%)		
29	Solyc02g093540	Cytochrome P450	AT3G50660	513	193/ 470	41	DWF4, CYP90B1, CLM, SNP2, SAV1, PSC1
30	Solyc03g020000	Pentatricopeptide repeat-containing protein (AHRD V1 *-*- D7L041_ARALY); contains Interpro domain(s) IPR002885 Pentatricopeptide repeat	AT2G22410	681	181/ 487	37	SLO1
31	Solyc03g025240	Multidrug resistance protein mdtK (AHRD V1 *—MDTK_YERP3); contains Interpro domain(s) IPR002528 Multi antimicrobial extrusion protein MatE	AT4G25640	514	273/ 398	69	DTX35
32	Solyc03g031660	Unknown Protein (AHRD V1)	-	-	-	-	-
33	Solyc03g058330	Unknown Protein (AHRD V1)	AT5G06760	158	57/ 144	40	LEA4-5
34	Solyc03g096190	Receptor like kinase, RLK	AT3G47570	1010	441/ 1003	43	Leucine-rich repeat protein kinase family protein
35	Solyc03g111190	Auxin-independent growth promoter-like protein (AHRD V1 ***-Q9FMW3_ARATH); contains Interpro domain(s) IPR004348 Protein of unknown function DUF246, plant	AT5G63390	559	343/ 557	62	O-fucosyltransferase family protein
36	Solyc03g115350	Expansin 2 (AHRD V1 ***-C0KLG9_PYRPY); contains Interpro domain(s) IPR002963 Expansin	AT5G39280	259	146/ 223	65	ATEXPA23, ATEXP23, ATHEXP ALPHA 1.17
37	Solyc03g116410	Zinc finger CCCH domain-containing protein 39 (AHRD V1 ***-C3H39_ARATH); contains Interpro domain(s) IPR000571 Zinc finger, CCCH-type	AT3G19360	386	54/ 199	27	Zinc finger (CCCH-type) family protein
38	Solyc03g123770	Unknown Protein (AHRD V1)	-	-	-	-	-
39	Solyc03g123970	Lipid-binding serum glycoprotein family protein (AHRD V1 *-*-D7LAX8_ARALY)	AT3G20270	722	26/ 51	51	lipid-binding serum glycoprotein family
40	Solyc04g007310	Thaumatin-like protein (AHRD V1 ***- C1K3P2_PYRPY); contains Interpro domain(s) IPR001938 Thaumatin, pathogenesis-related	AT4G38670	253	108/ 252	43	Pathogenesis-related thaumatin superfamily protein
41	Solyc04g008670	Gibberellin 2-beta-dioxygenase 7 (AHRD V1 **** B6SZM8_MAIZE); contains Interpro domain(s) IPR005123 Oxoglutarate and iron-dependent oxygenase	AT4G21200	336	166/ 302	55	ATGA2OX8, GA2OX8
42	Solyc04g014750	TNFR/CD27/30/40/95 cysteine-rich region (AHRD V1 ***-Q2HT38_MEDTR)	AT1G12064	109	34/ 73	47	Unkown protein
43	Solyc04g025740	Homeobox-leucine zipper protein ROC3 (AHRD V1 ***-ROC3_ORYSJ); contains Interpro domain(s) IPR001356 Homeobox	AT1G73360	722	52/ 125	42	HDG11, EDT1, ATHDG11
44	Solyc04g051070	Unknown Protein (AHRD V1)	-	-	-	-	-
45	Solyc04g058040	Laccase (AHRD V1 ***-Q9AUI3_PINTA); contains Interpro domain(s) IPR011707 Multicopper oxidase, type 3	AT5G09360	569	82/ 212	39	LAC14

(*Continued*)

Table 1. (*Continued*)

#	ITAG ID	Description in ITAG2.40	Homologue in Arabidopsis	length (aa)	Identities (%)		
46	Solyc04g072870	Beta-D-xylosidase (AHRD V1 **** Q8W011_HORVU); contains Interpro domain(s) IPR001764 Glycoside hydrolase, family 3, N-terminal	AT1G78060	767	445/756	59	Glycosyl hydrolase family protein
47	Solyc04g074320	Zinc finger protein (AHRD V1 *—D7KHP2_ARALY); contains Interpro domain(s) IPR007087 Zinc finger, C2H2-type	AT1G34790	303	143/200	72	TT1, WIP1
48	Solyc04g074890	Unknown Protein (AHRD V1)	-	-	-	-	-
49	Solyc04g078240	Natural resistance associated macrophage protein (AHRD V1 *—B3W4E1_BRAJU); contains Interpro domain(s) IPR001046 Natural resistance-associated macrophage protein	AT1G47240	530	73/95	77	NRAMP2, ATNRAMP2
50	Solyc04g081180	Unknown Protein (AHRD V1)	-	-	-	-	-
51	Solyc04g082520	Ring zinc finger protein (Fragment) (AHRD V1 *—A6MH00_LILLO); contains Interpro domain(s) IPR008166 Protein of unknown function DUF23	AT4G37420	588	233/500	47	Domain of unknown function (DUF23)
52	Solyc05g005240	YABBY-like transcription factor CRABS CLAW-like protein (AHRD V1 **-* Q6SRZ7_ANTMA); contains Interpro domain(s) IPR006780 YABBY protein	AT1G23420	231	100/184	54	INO
53	Solyc05g008320	Fasciclin-like arabinogalactan protein (AHRD V1 ***-B9N201_POPTR); contains Interpro domain(s) IPR000782 FAS1 domain	AT5G40940	424	114/328	35	FLA20
54	Solyc05g010190	Unknown Protein (AHRD V1)	AT3G42565	119	48/121	40	ECA1 gametogenesis related family protein
55	Solyc05g010200	Unknown Protein (AHRD V1)	-	-	-	-	-
56	Solyc05g013230	Unknown Protein (AHRD V1)	AT3G23880	364	21/57	37	F-box and associated interaction domains-containing protein
57	Solyc05g052440	Os03g0291800 protein (Fragment) (AHRD V1 **—Q0DSS4_ORYSJ); contains Interpro domain(s) IPR004253 Protein of unknown function DUF231, plant	AT2G40320	425	279/411	68	TBL33
58	Solyc05g052530	Endoglucanase 1 (AHRD V1 ***-B6U0J0_MAIZE); contains Interpro domain(s) IPR001701 Glycoside hydrolase, family 9	AT2G44550	490	292/476	56	ATGH9B10
59	Solyc06g007380	Os08g0119500 protein (Fragment) (AHRD V1 *-*- Q0J8C9_ORYSJ)	AT5G01710	513	258/510	51	methyltransferases
60	Solyc06g048400	Unknown Protein (AHRD V1); contains Interpro domain(s) IPR008502 Protein of unknown function DUF784, Arabidopsis thaliana	AT3G30387	115	34/97	35	Protein of unknown function (DUF784)

(*Continued*)

Table 1. (*Continued*)

| #	ITAG ID	Description in ITAG2.40	Homologue in Arabidopsis	length (aa)	Identities (%)		
61	Solyc06g060450	Transmembrane emp24 domain-containing protein 10 (AHRD V1 ***-B6SSF8_MAIZE); contains Interpro domain(s) IPR000348 emp24/gp25L/p24	AT1G2190	216	108/210	51	emp24/gp25L/p24 family/GOLD family protein
62	Solyc06g070950	ATP-binding cassette (ABC) transporter 17 (AHRD V1 ***-Q4H493_RAT); contains Interpro domain(s) IPR003439 ABC transporter-like	AT3G47780	935	503/937	54	ATATH6, ATH6
63	Solyc06g073100	GDSL esterase/lipase At3g27950 (AHRD V1 ***- GDL54_ARATH); contains Interpro domain(s) IPR001087 Lipase, GDSL	AT3G27950	361	197/375	53	GDSL-like Lipase/Acylhydrolase superfamily protein
64	Solyc06g074160	B3 domain-containing protein Os03g0212300 (AHRD V1 ***-Y3123_ORYSJ); contains Interpro domain(s) IPR003340 Transcriptional factor B3	AT3G06160	374	38/131	29	AP2/B3-like transcriptional factor family protein
65	Solyc06g075200	Unknown Protein (AHRD V1)	AT5G37474	80	28/83	34	Putative membrane lipoprotein
66	Solyc07g007520	Unknown Protein (AHRD V1)	-	-	-	-	-
67	Solyc07g032700	Unknown Protein (AHRD V1)	-	-	-	-	-
68	Solyc07g043410	UDP-glucosyltransferase family 1 protein (AHRD V1 **** C6KI43_CITSI); contains Interpro domain(s) IPR002213 UDP-glucuronosyl/UDP-glucosyltransferase	AT2G15480	484	166/487	34	UGT73B5
69	Solyc07g053400	Unknown Protein (AHRD V1)	-	-	-	-	-
70	Solyc07g054360	Unknown Protein (AHRD V1)	-	-	-	-	-
71	Solyc07g062320	Unknown Protein (AHRD V1)	-	-	-	-	-
72	Solyc07g064780	Unknown Protein (AHRD V1)	-	-	-	-	-
73	Solyc08g015750	F-box family protein (AHRD V1 ***-B9I6K2_POPTR); contains Interpro domain(s) IPR001810 Cyclin-like F-box	AT5G02920	469	58/200	31	F-box/RNI-like superfamily protein
74	Solyc08g061120	Unknown Protein (AHRD V1)	-	-	-	-	-
75	Solyc08g066400	Protein kinase (Fragment) (AHRD V1 *-*- A2Q5N5_MEDTR)	AT2G25760	676	217/333	65	Protein kinase family protein
76	Solyc08g074920	Aspartic proteinase nepenthesin I (AHRD V1 **—A9ZMF9_NEPAL); contains Interpro domain(s) IPR001461 Peptidase A1	AT5G33340	437	206/437	47	CDR1
77	Solyc08g080020	Serine protease inhibitor potato inhibitor I-type family protein (AHRD V1 ***- D7LT19_ARALY); contains Interpro domain(s) IPR000864 Proteinase inhibitor I13, potato inhibitor I	AT3G46860	85	32/86	37	Serine protease inhibitor, potato inhibitor I-type family protein
78	Solyc08g082260	Integrin-linked kinase-associated serine/threonine phosphatase 2C (AHRD V1 **** ILKAP_RAT); contains Interpro domain(s) IPR015655 Protein phosphatase 2C	AT2G29380	362	134/298	45	HAI3

(*Continued*)

Table 1. (*Continued*)

#	ITAG ID	Description in ITAG2.40	Homologue in Arabidopsis	length (aa)	Identities (%)		
79	Solyc09g011280	Unknown Protein (AHRD V1); contains Interpro domain(s) IPR006501 Pectinesterase inhibitor	AT3G17220	173	31/131	24	ATPMEI2
80	Solyc09g011290	Invertase inhibitor homolog (AHRD V1 ***- O49603_ARATH); contains Interpro domain(s) IPR006501 Pectinesterase inhibitor	AT5G64620	180	52/173	30	C/VIF2, ATC/VIF2
81	Solyc09g025200	Ribosomal protein L18 (AHRD V1 *-*- B7FMF5_MEDTR); contains Interpro domain(s) IPR000039 Ribosomal protein L18e	AT3G05590	187	31/50	62	RPL18
82	Solyc09g042760	ZIP4/SPO22 (AHRD V1 **—A5Y6I6_ARATH); contains Interpro domain(s) IPR013940 Meiosis specific protein SPO22	AT5G48390	936	527/936	56	ATZIP4
83	Solyc09g047860	HAT family dimerisation domain containing protein (AHRD V1 *-*- Q2R1C3_ORYSJ); contains Interpro domain(s) IPR008906 HAT dimerisation	AT5G33406	509	52/173	30	hAT dimerisation domain-containing protein / transposase-related
84	Solyc09g056030	Unknown Protein (AHRD V1)	AT4G12570	873	17/44	39	UPL5
85	Solyc09g056040	Ubiquitin-protein ligase 1 (AHRD V1 ***- Q5CHN2_CRYHO); contains Interpro domain(s) IPR000569 HECT	AT4G12570	873	153/413	37	UPL5
86	Solyc09g066050	Homeodomain-containing transcription factor FWA (AHRD V1 **-* B5BQ02_ARASU); contains Interpro domain(s) IPR002913 Lipid-binding START	AT1G73360	722	211/587	36	HDG11, EDT1, ATHDG11
87	Solyc09g073020	Unknown Protein (AHRD V1)	-	-	-	-	-
88	Solyc09g075110	Unknown Protein (AHRD V1)	-	-	-	-	-
89	Solyc09g089590	Ramosa1 C2H2 zinc-finger transcription factor (AHRD V1 *-*- D0UTY8_ZEAMM); contains Interpro domain(s) IPR007087 Zinc finger, C2H2-type	AT3G23130	204	78/192	78	SUP, FON1, FLO10
90	Solyc09g089960	Unknown Protein (AHRD V1)	-	-	-	-	
91	Solyc09g091300	Self-incompatibility protein (Fragment) (AHRD V1 -**- C8C1B5_9MAGN); contains Interpro domain(s) IPR010264 Plant self-incompatibility S1	AT3G26880	161	35/135	33	Plant self-incompatibility protein S1 family
92	Solyc10g005170	Purine permease (AHRD V1 *—* B6TET5_MAIZE); contains Interpro domain(s) IPR004853 Protein of unknown function DUF250	AT1G30840	382	208/330	63	ATPUP4, PUP4
93	Solyc10g005440	Serine/threonine-protein kinase receptor (AHRD V1 **** B6U2B7_MAIZE); contains Interpro domain(s) IPR002290 Serine/threonine protein kinase	AT4G21390	849	440/858	51	B120, S-locus lectin protein kinase family protein

(*Continued*)

Table 1. (*Continued*)

#	ITAG ID	Description in ITAG2.40	Homologue in Arabidopsis	length (aa)	Identities (%)		
94	Solyc10g017990	Cytokinin oxidase/dehydrogenase 2 (AHRD V1 *-** C0LPA7_SOLTU); contains Interpro domain(s) IPR015345 Cytokinin dehydrogenase 1, FAD and cytokinin binding	AT2G41510	575	214/525	41	ATCKX1, CKX1
95	Solyc10g044690	Annexin (AHRD V1 ***-D2D2Z9_GOSHI); contains Interpro domain(s) IPR009118 Annexin, type plant	AT5G12380	316	173/316	55	ANNAT8
96	Solyc10g047720	Unknown Protein (AHRD V1)	AT5G26805	156	44/163	27	unknown protein
97	Solyc10g050750	Pectinacetylesterase like protein (Fragment) (AHRD V1 *—Q56WP8_ARATH); contains Interpro domain(s) IPR004963 Pectinacetylesterase	AT4G19420	397	234/381	61	Pectinacetylesterase family protein
98	Solyc10g051370	LRR receptor-like serine/threonine-protein kinase, RLP	AT2G16250	915	105/198	53	Leucine-rich repeat protein kinase family protein
99	Solyc10g055600	S-phase kinase-associated protein 1A (AHRD V1 **—B2VUU5_PYRTR); contains Interpro domain(s) IPR001232 SKP1 component	AT4G34210	152	38/47	81	ASK11, SK11
100	Solyc11g005500	ECA1 protein (AHRD V1 *-*-Q53JF8_ORYSJ); contains Interpro domain(s) IPR010701 Protein of unknown function DUF1278	AT1G76750	158	63/124	51	EC1.1
101	Solyc11g005540	ECA1 protein (AHRD V1 *-*-Q53JF8_ORYSJ); contains Interpro domain(s) IPR010701 Protein of unknown function DUF1278	AT2G21750	125	61/130	47	EC1.3
102	Solyc11g006840	Unknown Protein (AHRD V1)	-	-	-	-	-
103	Solyc11g012650	TPD1 (AHRD V1 *-*-Q6TLJ2_ARATH)	AT1G32583	179	66/112	59	TPD1-like
104	Solyc11g043160	Endo-1 4-beta-xylanase (AHRD V1 ***-B6SW51_MAIZE); contains Interpro domain(s) IPR013781 Glycoside hydrolase, subgroup, catalytic core	AT4G33840	576	217/545	40	Glycosyl hydrolase family 10 protein
105	Solyc11g070010	F8A5.6 protein (AHRD V1 **—Q9ZP57_ARATH)	AT1G60500	669	117/391	30	DRP4C
106	Solyc11g072650	Trans-2-enoyl CoA reductase (AHRD V1 **—C5MRG3_9ROSI); contains Interpro domain(s) IPR002085 Alcohol dehydrogenase superfamily, zinc-containing	AT3G45770	375	215/335	64	Polyketide synthase, enoylreductase
107	Solyc12g019050	Exostosin-like (AHRD V1 ***-A4Q7M8_MEDTR); contains Interpro domain(s) IPR004263 Exostosin-like	AT3G42180	470	203/349	57	Exostosin family protein
108	Solyc12g042340	Genomic DNA chromosome 5 P1 clone MAC9 (AHRD V1 ***-Q9FLS4_ARATH)	AT5G61865	417	136/368	35	unknown protein
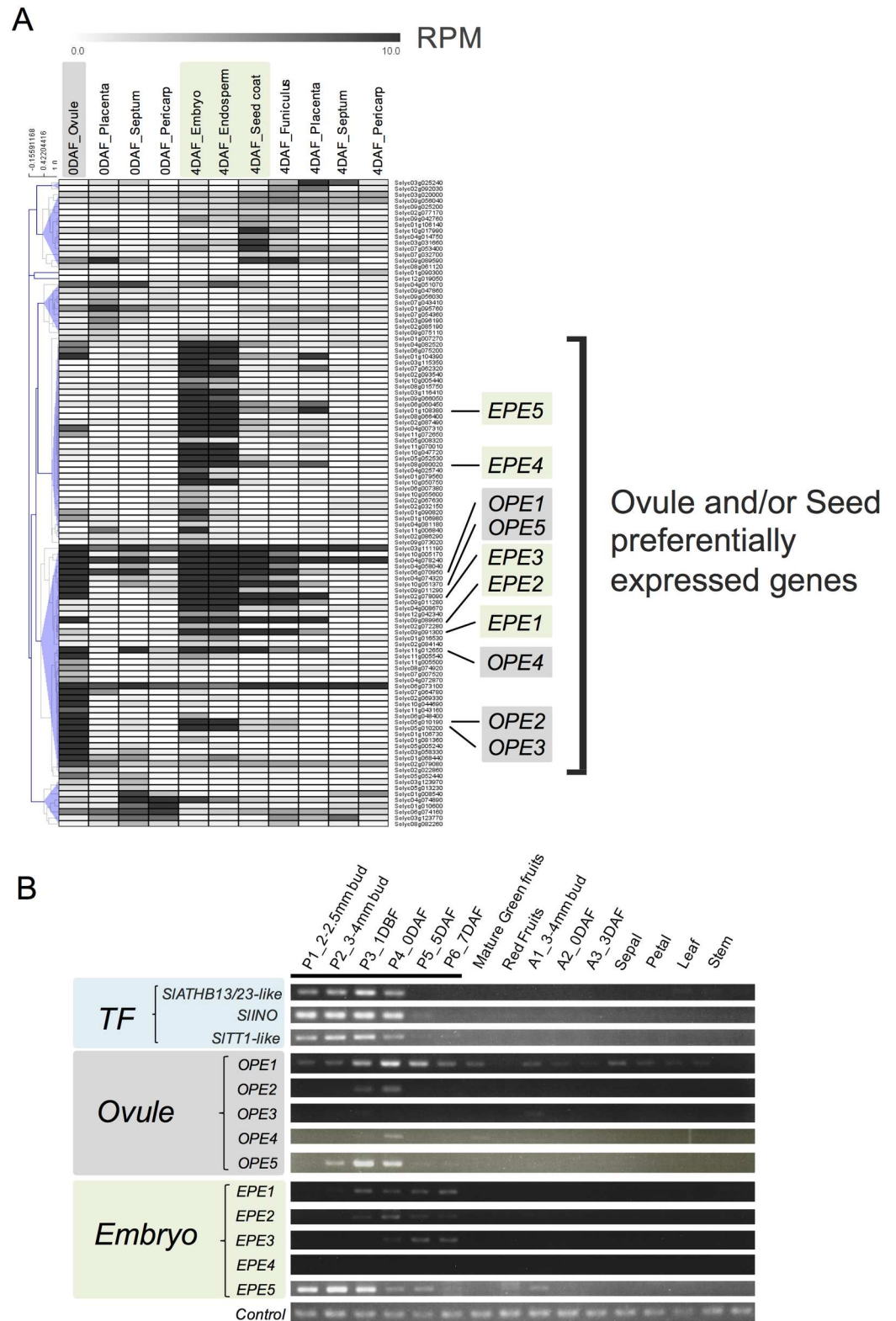
Fig 3. Characterization and validation of the 108 genes. (A) Many of the 108 PSGs were predominantly expressed in ovules and/or seeds in the pistil. Reads per million (RPM) values of PSGs in *S. pimpinellifolium* were visualized by constructing a heatmap using MeV software. Hierarchical clustering by Pearson correlation was conducted. The transcriptome data were obtained from [27]. *OPE*; ovule preferentially expressed genes, *EPE*;

embryo preferentially expressed genes. (B) Validation of the expression of ovule preferentially expressed (*OPE*) genes, embryo preferentially expressed (*EPE*) genes, and several transcription factor genes by RT-PCR. Most of the genes were specifically expressed in the pistil. Three pistil-specific transcription factor genes, *SlATHB13/23-like* (*Solyc01g010600*), *SlINO* (*Solyc05g005240*), and *SlTT1* (*Solyc10g051370*) showed pistil-specific expression before anthesis. Bottom one represents the expression of the internal control gene SAND [41].

like protein (*Solyc01g008540*), Unknown Protein (*Solyc04g074890*), homeobox-leucine zipper-like protein (*Solyc01g010600*), B3 domain-containing protein Os03g0212300–like protein (*Solyc06g074160*), and Unknown Protein (*Solyc03g123770*). Furthermore, the gene encoding cytokinin oxidase/dehydrogenase 8 (*SlCKX8*, *Solyc10g017990*), TNFR/CD27/30/40/95 cysteine-rich region (*Solyc04g014750*), Unknown Protein (*Solyc03g031660*), Unknown Protein (*Solyc07g053400*), and Ramosa1 C2H2 zinc-finger transcription factor (*Solyc09g089590*) were preferentially expressed in the seed coat. *Solyc09g089590* encodes one of two homologous proteins of Arabidopsis SUPERMAN (SUP), which regulates auxin biosynthesis [48]. In addition, the expression of 63 out of 108 genes was detected also in the recently published ovary transcriptome dataset derived from cultivated tomato 'Moneymaker' [23], in which RNA-seq analyses against ovule and ovary wall tissue were conducted; their average expression levels were over FPKM of 0.5 [23] (S4 Table). The 55 other genes were not detected in that dataset, indicating that these 55 genes were barely expressed in cultivated tomato or were only expressed in other type of tissues such as the placenta and septum, which were excluded from their experiment.

Validation of gene expression patterns by RT-PCR

We then verified the expression patterns of the PSGs by RT-PCR analysis. Since many of these genes were highly expressed in the ovule and/or seed, especially the embryo (Fig 3A), we initially focused on genes specifically expressed in these tissues. Among the 108 PSG candidates, the top-five PSGs highly expressed in 0 DAF ovules were designated O̲vule P̲referentially E̲xpressed genes 1–10 (*OPE1–5*) (S5 Table). We verified the tissue-specific expression of five of these genes by RT-PCR analysis (Fig 3B). *OPE1* was preferentially but not exclusively expressed in the pistil at anthesis, *OPE2* and *OPE5* were specifically expressed in the pistil at both 1 DBF and 0 DAF, and the expression of *OPE3* in the pistil was not detected in this experiment. *OPE4* was expressed in the pistil at 0 DAF and mature green fruits. We also designated the top-five PSGs that were highly expressed in 4 DAF embryos as E̲mbryo P̲referentially E̲xpressed genes 1–5 (*EPE1–5*) (S5 Table). *EPE1*, encoding a self-incompatibility protein-like protein according to SGN, might function in pollen-pistil interactions, while most of the *EPEs* had not been functionally characterized or annotated in previous studies. Like the *OPEs*, we investigated the expression of the five *EPEs* (*EPE1–5*) by RT-PCR to validate their tissue-specific expression patterns. Three genes, *EPE1-EPE3*, were specifically expressed in the pistil and EPE5 was preferentially expressed in the pistil especially before anthesis, although we failed to detect the expression of *EPE4* in our RT-PCR analysis (Fig 3B). *EPE1* was specifically expressed in the pistil throughout pistil/fruit development but was not expressed in mature red fruits. *EPE3* was also specifically expressed in the pistil, but only after anthesis. *EPE2* was expressed exclusively during fruit set initiation between 1 DBF and 0 DAF (Fig 3B). In summary, three *OPEs* and four *EPEs* were specifically expressed in pistils, confirming their tissue-specific expression in the pistil (Fig 3B). Therefore, we confirmed the tissue-specific expression of *PSGs* in the pistil. These results indicate that the direct-mapping method also successfully identified true PSGs.

GO analysis using AgriGO

To elucidate the enriched functional categories of the 108 identified PSGs, we performed GO analysis using AgriGO. A false discovery rate (FDR; e-value corrected for list size) of < 0.05 was used as the criterion to obtain enriched GO terms. Consequently, only one category, Carboxylesterase activity (GO:0004091), showed significant abundance (p-value = 0.0017, FDR = 0.037) (Fig 2D). This category includes five genes (listed in Fig 2D), three of which (*Solyc02g069330*, *Solyc09g011280*, and *Solyc09g011290*) were assigned to the sub-term "Pectinesterase inhibitor". Even though *Solyc09g011290* was classified as a "Pectinesterase inhibitor", it was labeled as an "invertase inhibitor homolog" in the SGN database and has higher sequence homology with the invertase inhibitor group that includes *invertase inhibitor 1* (*INVINH1*, *Solyc12g099200*), which specifically regulates cell wall invertase activity in early developing fruits [49].

Pectin, a major component of the primary cell walls of higher plants, is methyl-esterified by pectin methyltransferase (PMT) before its transport to the cell wall following its biosynthesis in Golgi bodies [50,51], whereas pectin methylesterase (PME) catalyzes the removal of methyl esters from pectin [52–54]. The removal of methyl group from pectin allows carboxyl groups to form Ca^{2+}- and Mg^{2+}-mediated linkages, leading to the hardening of pectin [55,56]. In addition, pectin methylesterase inhibitors (PMEIs) directly interact with PME and inhibit its activity, affecting pectin composition in the cell wall. Lionetti et al. [57] reported that overexpressing Arabidopsis *PMEI* increased the degree of pectin methylesterification by approximately 16%, resulting in longer roots due to the promotion of cell elongation. Therefore, the degree of methylation and demethylation of pectin determines the balance between extensibility and rigidity, affecting growth and cell shape. In tomato, *PMEU1*, a ubiquitously expressed pectin methylesterase gene, is expressed during early fruit development [58]. Terao et al. [59] recently reported the occurrence of rapid pectin metabolism during the early stage of fruit development in tomato: immunolocalization analysis demonstrated that methyl-esterified pectin levels in the ovary increased from 1 DBF to 3 DAF [59]. During fruit set, the transition of cell state from cell division to cell expansion occurs during a short period of time, and the regulation of this process is important for determining the size of the fruit. Therefore, it would be interesting to investigate whether PMEI plays a role in the post-translational regulation of PME and cell wall state during fruit set.

In addition, the pectinesterase inhibitor protein family includes several enzyme inhibitors such as invertase (Beta-fructofuranosidase) inhibitors, each of which has a specific target [60,61]. *Solyc09g011290* was annotated as an invertase inhibitor homolog in the SGN database. Invertase inhibitors regulate specific invertases in a post-translational manner, negatively affecting the enzyme activity of their targets [49,61]. We found that *Solyc09g011290* was specifically and highly expressed in the ovule/seed (S3 Table). The expression of *Solyc09g011290* was induced during anthesis and remained at high levels in the absence of pollination but was down-regulated by pollination and hormone treatments (S6 Fig). Several studies on the cell wall invertase (CWIN) and INVINH1 in tomato suggest that these proteins play important roles in seed set and fruit set by regulating the unloading of sugar from the phloem during the ovary-to-fruit transition [4,49,62,63]. Thus, the expression pattern of *Solyc09g011290*, the up-regulation during flowering and the down-regulation by the fruit-set stimulus (S6 Fig), suggests that *Solyc09g011290* may also participate in the modulation of the sugar unloading to unpollinated pistil via post-translational inhibition of invertase activity.

Identification of pistil-specific transcriptional regulators

Pistil-specific transcription factors. Next, we performed similar RT-PCR analyses of several transcription factor genes listed among the PSGs and confirmed the tissue-specific expression of three transcription factor genes (Fig 3B).

Solyc04g074320 (SlTT1-like) shares high homology (71.5%) with Arabidopsis zinc-finger protein TRANPARENT TESTA1 (TT1) [64]. Arabidopsis *TT1* expression is restricted to developing ovules and young seeds and functions in the accumulation of proanthocyanidin pigments in the seed coat [65,66], while in the current study, tomato *SlTT1-like* transcripts were exclusively detected in the ovule, embryo, and endosperm but not in the seed coat ([Fig 3B](#)). Mazzucato *et al.* [67] provided evidence that higher anthocyanin content is associated with increased early fruit growth in non-pollinated flowers. Furthermore, there is an evidence that the alteration of the flavonoid pathway via the regulation of biosynthesis genes induces seedless fruit development in both a pollination-dependent and pollination-independent manner [68,69]. Further elucidation of the function of *SlTT1-like* in the control of flavonoid-related genes may provide insight into the role of flavonoids during fruit set initiation.

SlINO (*Solyc05g005240*) was identified as a pistil-specific *YABBY* transcription factor gene (S2B Fig and [Fig 3B](#)). YABBY family proteins contain two conserved domains, i.e., a C2C2 zinc-finger-like domain in their N-termini and a helix-loop-helix domain known as the YABBY domain [70]. In Arabidopsis, two YABBY genes, *INO* and *CRC*, show tissue-specific expression in the pistil and are involved in pistil and early fruit development [44–46]. Nine YABBY genes were previously identified in tomato, three of which (*SlCRCa*, *Solyc01g0101240*; *SlCRCb*, *Solyc05g012050*; *SlINO*, *Solyc05g005240*) are specifically expressed in the flower bud and in open flowers at anthesis [50]. In the current study, we found that *SlCRCa*, *SlCRCb*, and *SlINO* were preferentially expressed in the pistil (S2B Fig). *SlCRCa* was expressed in the early stage of pistil development, while *SlCRCb* and *SlINO* were expressed during all stage of pistil development. Furthermore, we confirmed the tissue-specific expression of *SlINO* in pistils by RT-PCR analysis ([Fig 3B](#)), suggests its role in the regulation of pistil development [46].

Solyc01g010600 (*SlATHB13/23-like*), which encodes a homeodomain leucine zipper 1 transcription factor (HD-Zip TF), shares similarity with *Arabidopsis* class-1 HD-Zip genes *AtHB13* and *AtHB23* and was specifically expressed in the pistil before anthesis ([Fig 3B](#)). The HD-Zip TF family forms a large gene family that is divided into four classes; 58 HD-Zip proteins found in both Arabidopsis and tomato are listed in PlantTFDB version 3.0 (http://planttfdb.cbi.pku.edu.cn) [71,72]. Although the molecular functions of class-1 HD-Zip proteins in the regulation of pistil development remain elusive, AtHB13 and AtHB23 were shown to play negative roles in inflorescence stem elongation by affecting cell division, and AtHB13 also regulates pollen hydration and development [73]. In tomato, class-1 HD-Zip SlHZ24 functions as a transcriptional activator of *SlGMP3* (encoding GDP-$_D$-mannosepyrophosphorylase), which plays an important role in the production of the antioxidant ascorbate [74]. In addition, virus-induced gene silencing of class-1 HD-Zip *LeHB1* reduced the mRNA accumulation of *LeACO1* and inhibited ripening [75]. Further, Lin et al. [75] also reported that ectopic overexpression of *LeHB1* led to the conversion of sepals into carpel-like structures. We also found that *SlATHB13/23-like* was highly expressed in the ovary wall in the pistil at anthesis (S3 and S4 Tables), suggesting its regulatory role in carpel and ovary wall development.

Besides, we identified the type-I MADS box gene, *Solyc01g106730* among PSGs ([Table 1](#)). MADS box genes encode transcription factors which are generally involved in homeotic regulation of reproductive organs and can be divided into two subfamilies (type-I and type-II) according to the presence of conserved domains [75, 76]. Some type-II MIKC MADS box genes play key roles as regulators of meristem identity, flowering time, and fruit and seed development [77,78], whereas little information is available about the function of type-I MADS box proteins. Most (38 out of 61) Arabidopsis type-I MADS box genes are expressed in the female gametophyte [79], whereas others exhibit highly specific expression such as in the central cell and embryo sac [80]. For example, Arabidopsis type-I *AGAMOUS-LIKE61*(*AGL61*)/*DIANA* (*DIA*) is expressed exclusively in the central cell and early endosperm and plays crucial role in

endosperm development after fertilization through transcriptional control in the central cells [79–82]. *AtAGL62* also plays important role in the endosperm and seed coat development [83–85]. Gene expression analysis using tissue-specific transcriptome data from wild tomato *S. pimpinellifolium* [27] revealed that the type-I MADS box gene *Solyc01g106730* is preferentially expressed in the ovule (S3 and S4 Tables). In addition to the relationship between type-I MADS box genes and seed development, there is an evidence that down-regulation or mutation in type-II MADS box genes, such as *TM29*, *TAP3*, *TM8*, *SlAGL11* or *SlAGL6*, results in parthenocarpy [86–90]. Thus, it would be important to investigate the roles of *Solyc01g106730* in pistil, seed, and fruit development.

Pistil-specific peptide hormone-like small peptide genes and receptor-like proteins. The role of peptide hormones in plant signaling pathways is a popular focus of study [91–95]. The peptide hormone signaling system involves two main components: (1) small ligand proteins such as small cysteine-rich peptides (CRPs) and (2) receptor proteins such a leucine-rich receptor-like kinases (LRR-RLKs) [96]. CRPs function as signaling molecules (peptide hormones) in various plant species, which are required for many aspects of development including antimicrobial defense, pollen tube guidance, stomatal patterning, and early embryo patterning [97–104]. CRPs contain four, six, or eight conserved cysteine residues at their C-termini in addition to a secretion signal at their N-termini. Interestingly, a substantial number of PSGs identified in this study encode small proteins (44 out of 108 genes identified by the mapping-based method [40%]) less than 200 amino acids in length (Table 2). Small proteins are defined as proteins smaller than 200 amino acids according to previous reports [94,97,105]. Since peptide hormone-like small proteins share a conserved structure, we performed a sequence similarity search of the 44 identified small proteins and one TAPETUM DETERMINANT 1 (TPD1)-like protein (204 aa) by BLAST analysis and SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/) manually to investigate whether they have conserved residues or functional domains. Roughly half of these proteins also have a secretion signal in their N-termini (Table 2). Notably, through subsequent sequence analysis of these small proteins, four tissue-specific CRPs including an unknown gene (*Solyc06g075200*) and two LRR-RLK-like proteins were identified (Listed in Table 2, S5 Fig).

OPE4 (*Solyc11g012650*) was homologous to Arabidopsis *TAPETUM DETERMINANT* (*AtTPD1*, AT4G24972) (6e-37), encoding a peptide hormone that functions as a ligand molecule to regulate the specification of tapetum cells in coordination with receptor protein EMS/EXS [106, 107], while BLAST searches of the tomato genome identified three other homologs, designated *SlTPD1L1* (*Solyc03g097530*), *SlTPD1L2/OPE4* (*Solyc11g012650*), and *SlTPD1L3* (*Solyc11g006850*), based on sequence similarity to *AtTPD1*, with 59.4% (1e-49), 55% (6e-37), and 50% (4e-33) sequence similarity, respectively. Like AtTPD1, we confirmed the presence of a secretion signal in the N-terminal region and conserved cysteine residues at the C-terminus among the three deduced proteins (S3A Fig). Although the sequence of N-terminal secretion signal region varied among the three proteins, an alignment of each SlTPD1L compared to amino acids 26–179 of AtTPD1 revealed a high degree of similarity (48–56%) (S3B Fig). Although it is known that *AtTPD1* is also expressed in inflorescence meristems, floral meristems, carpel primordia, and ovule primordia, its function in these tissues remains unknown [106, 107]. The notion that *SlTPD1L2/OPE4* (*Solyc11g012650*) showed pistil-specific expression (Fig 3B), and that *OPE4* is shown to be preferentially expressed in the ovule both in *S. pimpinellifolium* and tomato cultivar 'Moneymaker' (S3 and S4 Tables), it was suggested that *SlTPD1L2/OPE4* might play a tissue-specific role in ovule development.

Solyc11g005540, *Solyc11g005500*, and *Solyc05g010190* share sequence similarity with Arabidopsis ECA1-like protein (EC1) genes, which involved in gamete fusion [108] (S4B and S4C Fig). mRNA from the five Arabidopsis *EC1* genes (*EC1.1* to *EC1.5*; belonging to the ECA1

Table 2. List of pistil-specific or preferentially expressed small proteins.

#	ITAG ID	Description in ITAG2.40	length (aa)	*Presence of predicted secreted signal (aa)	Homologue in Arabidopsis	length (aa)	Identities (%)		Description	"Expression in pistil of Moneymaker (Ovule and/or ovary wall) from Zhang et al (2016)[23]"
PSSP1	Solyc01g016530	Unknown Protein (AHRD V1); contains Interpro domain(s) IPR008507 Protein of unknown function DUF789	87	-	AT1G73210	314	32/69	46	Protein of unknown function (DUF789)	Ovule
PSSP2	Solyc01g081360	Unknown Protein (AHRD V1)	151	1–29	-	-	-	-	-	Ovule
PSSP3	Solyc01g108380	Protease inhibitor protein (AHRD V1 -**-B3FNP9_HEVBR); contains Interpro domain(s) IPR000864 Proteinase inhibitor I13, potato inhibitor I	77	-	AT2G38900	88	27/61	44	Serine protease inhibitor, potato inhibitor I-type family protein	Ovule
PSSP4	Solyc02g069330	Unknown Protein (AHRD V1); contains Interpro domain(s) IPR006501 Pectinesterase inhibitor	180	1–19	AT5G64620	180	26/80	33	C/VIF2, ATC/VIF2	Ovule
PSSP5	Solyc03g058330	Unknown Protein (AHRD V1)	108	-	AT5G06760	158	57/144	40	LEA4-5	Ovule
PSSP6	Solyc04g081180	Unknown Protein (AHRD V1)	79	-	-	-	-	-	-	Ovule
PSSP7	Solyc05g010200	Unknown Protein (AHRD V1)	115	1–25	-	-	-	-	-	Ovule
PSSP8	Solyc06g048400	Unknown Protein (AHRD V1); contains Interpro domain(s) IPR008502 Protein of unknown function DUF784, Arabidopsis thaliana	155	-	AT3G30387	115	34/97	35	Protein of unknown function (DUF784)	Ovule
PSSP9	Solyc06g075200	Unknown Protein (AHRD V1)	81	1–22	AT5G37474	80	28/83	34	Putative membrane lipoprotein	Ovule
PSSP10	Solyc07g062320	Unknown Protein (AHRD V1)	79	-	-	-	-	-	-	Ovule
PSSP11	Solyc08g080020	Serine protease inhibitor potato inhibitor I-type family protein (AHRD V1 ***-D7LT19_ARALY); contains Interpro domain(s) IPR000864 Proteinase inhibitor I13, potato inhibitor I	104	1–19	AT3G46860	85	32/86	37	Serine protease inhibitor, potato inhibitor I-type family protein	Ovule
PSSP12	Solyc09g011280	Unknown Protein (AHRD V1); contains Interpro domain(s) IPR006501 Pectinesterase inhibitor	178	1–23	AT3G17220	173	31/131	24	ATPMEI2	Ovule
PSSP13	Solyc09g089590	Ramosa1 C2H2 zinc-finger transcription factor (AHRD V1 *-*-DOUTY8_ZEAMM); contains Interpro domain(s) IPR007087 Zinc finger, C2H2-type	197	-	AT3G23130	204	78/192	78	SUP, FON1, FLO10	Ovule
PSSP14	Solyc11g005500	ECA1 protein (AHRD V1 *-*-Q53JF8_ORYSJ); contains Interpro domain(s) IPR010701 Protein of unknown function DUF1278	130	1–26	AT1G76750	158	63/124	51	EC1.1	Ovule
PSSP15	Solyc11g005540	ECA1 protein (AHRD V1 *-*-Q53JF8_ORYSJ); contains Interpro domain(s) IPR010701 Protein of unknown function DUF1278	136	1–16	AT2G21750	125	61/130	47	EC1.3	Ovule
PSSP16	Solyc11g006840	Unknown Protein (AHRD V1)	126	-	-	-	-	-	-	Ovule
PSSP17	Solyc09g025200	Ribosomal protein L18 (AHRD V1 *-*-B7FMF5_MEDTR); contains Interpro domain(s) IPR000039 Ribosomal protein L18e	72	1–22	AT3G05590	187	31/50	62	RPL18	Ovary wall
PSSP18	Solyc09g056030	Unknown Protein (AHRD V1)	82	-	AT4G12570	873	17/44	39	UPL5	Ovary wall

(Continued)

Table 2. (Continued)

| #	ITAG ID	Description in ITAG2.40	length (aa)	* Presence of predicted secreted signal (aa)	Homologue in Arabidopsis	length (aa)	Identities (%)		Description	"Expression in pistil of Moneymaker (Ovule and/or ovary wall) from Zhang et al (2016)[23]"
PSSP19	Solyc01g007270	Cytokinin riboside 5'-monophosphate phosphoribohydrolase LOG (AHRD V1 **—LOG_ORYSJ)	70	-	AT5G06300	217	56/68	82	-	Not detected
PSSP20	Solyc01g079560	B3 domain-containing protein Os11g0197600 (AHRD V1 ***-Y1176_ORYSJ); contains Interpro domain(s) IPR003340 Transcriptional factor B3	109	-	AT3G18890	341	30/92	33	VRN1, REM39	Not detected
PSSP21	Solyc02g032150	Unknown Protein (AHRD V1)	147	-	-	-	-	-	-	Not detected
PSSP22	Solyc02g084140	Unknown Protein (AHRD V1)	132	-	-	-	-	-	-	Not detected
PSSP23	Solyc03g116410	Zinc finger CCCH domain-containing protein 39 (AHRD V1 ***-C3H39_ARATH); contains Interpro domain(s) IPR000571 Zinc finger, CCCH-type	117	-	AT3G19360	386	54/199	27	Zinc finger (CCCH-type) family protein	Not detected
PSSP24	Solyc04g025740	Homeobox-leucine zipper protein ROC3 (AHRD V1 ***- ROC3_ORYSJ); contains Interpro domain(s) IPR001356 Homeobox	148	-	AT1G73360	722	52/125	42	HDG11, EDT1, ATHDG11	Not detected
PSSP25	Solyc04g051070	Unknown Protein (AHRD V1)	80	-	-	-	-	-	-	Not detected
PSSP26	Solyc04g078240	Natural resistance associated macrophage protein (AHRD V1 *—B3W4E1_BRAJU); contains Interpro domain(s) IPR001046 Natural resistance-associated macrophage protein	161	-	AT1G47240	530	73/95	77	NRAMP2, ATNRAMP2	Not detected
PSSP27	Solyc05g013230	Unknown Protein (AHRD V1)	118	-	AT3G23880	364	21/57	37	F-box and associated interaction domains-containing protein	Not detected
PSSP28	Solyc07g054360	Unknown Protein (AHRD V1)	142	-	-	-	-	-	-	Not detected
PSSP29	Solyc08g061120	Unknown Protein (AHRD V1)	190	-	-	-	-	-	-	Not detected
PSSP30	Solyc09g073020	Unknown Protein (AHRD V1)	50	-	-	-	-	-	-	Not detected
PSSP31	Solyc09g075110	Unknown Protein (AHRD V1)	63	-	-	-	-	-	-	Not detected
PSSP32	Solyc10g047720	Unknown Protein (AHRD V1)	172	-	AT5G26805	156	44/163	27	unknown protein	Not detected
PSSP33	Solyc10g055600	S-phase kinase-associated protein 1A (AHRD V1 **—B2VUU5_PYRTR); contains Interpro domain(s) IPR001232 SKP1 component	51	-	AT4G34210	152	38/47	81	ASK11, SK11	Not detected
PSSP34	Solyc01g104390	Blue copper protein (AHRD V1 **—B6TT37_MAIZE); contains Interpro domain(s) IPR003245 Plastocyanin-like	122	1–27	AT1G17800	129	49/116	42	ARPN	Both
PSSP35	Solyc02g078090	Unknown Protein (AHRD V1)	105	1–26	-	-	-	-	-	Both
PSSP36	Solyc03g123770	Unknown Protein (AHRD V1)	112	-	-	-	-	-	-	Both
PSSP37	Solyc03g123970	Lipid-binding serum glycoprotein family protein (AHRD V1 *-*-D7LAX8_ARALY)	116	1–17	AT3G20270	722	26/51	51	lipid-binding serum glycoprotein family	Both

(Continued)

Table 2. (Continued)

#	ITAG ID	Description in ITAG2.40	length (aa)	* Presence of predicted secreted signal (aa)	Homologue in Arabidopsis	length (aa)	Identities (%)	Description	"Expression in pistil of Moneymaker (Ovule and/or ovary wall) from Zhang et al (2016)[23]"
PSSP38	Solyc04g014750	TNFR/CD27/30/40/95 cysteine-rich region (AHRD V1 ***-Q2HT38_MEDTR)	105	1–32	AT1G12064	109	34/73	47	Unkown protein	Both
PSSP39	Solyc05g005240	YABBY-like transcription factor CRABS CLAW-like protein (AHRD V1 **-* Q6SRZ7_ANTMA); contains Interpro domain(s) IPR006780 YABBY protein	192	-	AT1G23420	231	100/184	54	INO	Both
PSSP40	Solyc05g010190	Unknown Protein (AHRD V1)	138	1–23	AT3G42565	119	48/121	40	ECA1 gametogenesis related family protein	Both
PSSP41	Solyc07g032700	Unknown Protein (AHRD V1)	120	-	-	-	-	-	-	Both
PSSP42	Solyc07g053400	Unknown Protein (AHRD V1)	97	-	-	-	-	-	-	Both
PSSP43	Solyc09g011290	Invertase inhibitor homolog (AHRD V1 ***-O49603_ARATH); contains Interpro domain(s) IPR006501 Pectinesterase inhibitor	188	1–24	AT5G64620	180	52/173	30	C/VIF2, ATC/VIF2	Both
PSSP44	Solyc09g091300	Self-incompatibility protein (Fragment) (AHRD V1 -**-C8C1B5_9MAGN); contains Interpro domain(s) IPR010264 Plant self-incompatibility S1	148	1–23	AT3G26880	161	35/135	33	Plant self-incompatibility protein S1 family	Both
PSSP45	Solyc11g012650	TPD1 (AHRD V1 *-*-Q6TLJ2_ARATH)	204	1–28	AT1G32583	179	66/112	59	TPD1-like	Ovule

* Presence of secreted signal sequence in each protein was predicted by SignalP 4.1 Server with default setting.

https://doi.org/10.1371/journal.pone.0180003.t002
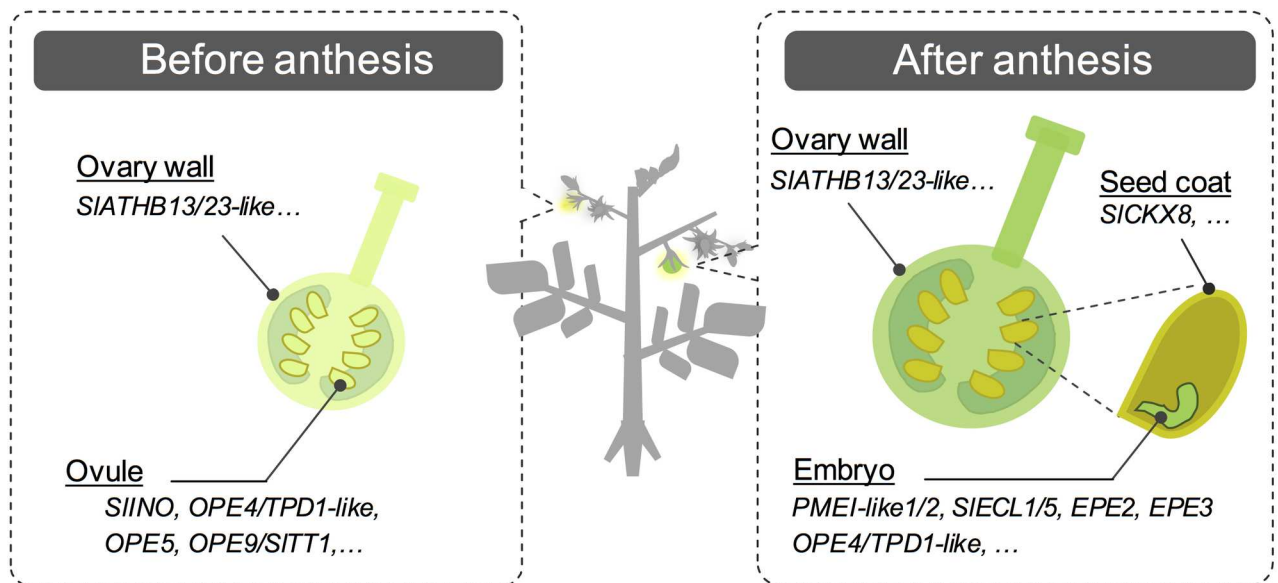
Fig 4. PSGs identified from this study. Many candidate transcriptional regulators in tomato pistils were identified, including cysteine-rich peptide (CRP)-like proteins, indicating their roles in the development of specific types of cells in the pistil. Ovule- and embryo-specific genes: *SlINO*, *SlTT1-like*, *SlECL1*, *SlECL5*. Pericarp-specific gene; *SlATHB13/23-like*. Seed coat-specific gene; *SlCKX8*.

[Early Culture Abundant] gametogenesis-related cysteine-rich protein subfamily) is present only in egg cells before fertilization. And, small proteins encoded by *EC1* genes are secreted into the outer region of the egg cell to redundantly regulate the fusion of the germ cell to the sperm cell during double fertilization [82,108]. Three tomato EC1-like protein homologs are expressed in the ovule/seed in *S. pimpinellifolium* (S3 and S4 Tables), while in the current study, we confirmed the pistil-specific expression of *Solyc11g005500* and *Solyc11g005540* (S4A Fig). Sequence similarity analysis of these protein sequences with AtEC1s identified five tomato homologs, which share six conserved cysteine residues in the C-terminal region and an N-terminal secretion signal peptide (S4B Fig). Therefore, these genes were designated *S. lycopersicum* SlEC1-like genes (SlECLs; e.g., *SlECL1* [*Solyc11g005500*], *SlECL3* [*Solyc05g010190*], and *SlECL5* [*Solyc11g005540*]), suggesting that they play a conserved role in the pistil at fertilization.

In addition to several ligand-like, pistil-specific CRPs, we also identified two proteins with RLK-related domain; *OPE5* (*Solyc10g051370*) and *Solyc03g096190*. *OPE5* was expressed in ovule and seed, and *Solyc03g096190* was expressed in placenta and septum (S3 Table). In tomato, the LRR-RLK family represents the largest family of RLKs [109]. Molecular genetic studies have shown that LRR-RLKs are involved in a wide range of plant development, such as stem cell maintenance [110,111], cell fate determination and patterning [103,112], and brassinosteroid signaling [113–115]. Generally, LRR-RLK proteins, such as EMS1/EXS and CLAVATA1 (CLV1), function as receptors by interacting with small ligand proteins at LRR domain and subsequently transmit external signals into cells by activating kinase domain, inducing various cellular responses [116–118]. While *Solyc03g096190* possesses the LRR domain, one transmembrane domain and the cytoplasmic kinase domain, *OPE5* possesses only small LRR domain composed of five repeats (S5 Fig). Some receptor-like proteins with extracellular LRR domains lacking the internal kinase domain also have been identified in other plants, such as *Arabidopsis* CLAVATA2 (CLV2) [119], which regulates various developmental and immunity signaling pathways by interacting with other RLKs including CLV1 [119–122]. Unraveling the role of pistil-specific RLK like proteins may help understand the detailed mechanism of signal transduction during fruit set.

Relationship between PSGs and fruit set signaling

We further characterized and identified PSGs associated with fruit set in the pistil. Specifically, we investigated their expression in a dataset of differentially expressed genes (DEGs) in the pistils of plants treated with 2,4-D and GA$_3$ from Tang et al. [20]. Among total 4764 and 6875 DEGs in ovaries undergoing fruit set after induction by plant hormones (auxin and GA) or pollination at 4 DAF compared to unpollinated ovaries at 2 DBF and unpollinated ovaries at 4 DAF, respectively, three genes, including *invertase inhibitor homolog* (*Solyc09g011290*), *SlATHB13/23-like* (*Solyc01g010600*), *SlCKX8* (*Solyc10g017990*) were included out of 108 PSGs, suggesting their roles in pistil development and/or fruit set initiation (S6 Fig). The mRNA levels of *SlATHB13/23-like* (*Solyc01g010600*) were significantly reduced at 4 DAF by either plant hormone (auxin or GA) treatment or pollination compared to that at 2 DBF, while no significant difference was found compared to ovaries at 6-days post-emasculation (6 DPE). By contrast, the mRNA levels of *SlCKX8* (*Solyc10g017990*), *invertase inhibitor homolog* (*Solyc09g011290*) were significantly lower in ovaries undergoing fruit set compared to those at 6 DPE, indicating that these genes are up-regulated in the absence of fruit set signaling after flowering.

Regulatory regions of PSGs as genetic engineering tools

Tissue-specific promoters are useful tools for regulating the expression of genes of interest in a spatial and temporal manner, which could provide new insights into various biological mechanisms and facilitate their application to molecular breeding, such as generating genetically modified organisms [123]. Constitutive promoters such as the 35S promoter from *Cauliflower mosaic virus* (CaMV) is widely used to regulate the expression of target genes in plants [124,125]. However, the constitutive regulation of target genes is not always useful, since it sometimes induces additional undesirable effects that hamper agronomic applications. Thus, in the past several decades, many tissue-specific promoters have been isolated, such as fruit-specific promoters [126–129], root-specific promoters [130–132], seed-specific promoters [133–135], many pollen and/or anther-specific promoters [36,136–139], and so on. Several individual promoters targeting the stigma and/or style in the pistil have also been evaluated in several plant species, such as the stigma- and style-specific thaumatin/PR5-like protein(*PsTL1*) promoter from Japanese pear (*Pyrus serotina*) [140–142]. Importantly, the use of the ovule-specific *DefH9* and *INO* promoters allows parthenocarpy to be efficiently induced via the ovule-specific activation of auxin signaling without producing substantial undesirable fruit traits in tomato [143]. Moreover, several reproductive tissue-specific promoters, such as sperm cell- and egg cell-specific promoters, have been successfully used to induce targeted mutagenesis by CRISPR/Cas9 [36,144]. For example, Wang et al. [144] used the promoter of egg cell-specific *EC1.2* (homolog of *SlEC1.2* and *SlEC1.3* identified in the current study) for CRISPR/Cas9 to generate homozygous mutants for multiple target genes in a single generation in Arabidopsis. These findings indicate that pistil-specific promoters, especially ovule-specific promoters, can be highly effective tools for plant breeding. Here, we identified various types of PSGs in tomato, e.g., *SlINO* and two homologs of Arabidopsis *EC1.2*, which were specifically expressed in the ovule (Table 2). Thus, these promoters could contribute to tissue-specific regulation or genome editing of target genes to improve fruit set by inducing parthenocarpy.

Conclusion

In conclusion, in this study, we conducted global analysis of tomato PSGs in tomato, which might be involved in the ovary development or fruit set process, by performing RNA-seq analysis and comparisons with publically available data. This study successfully identified several genes encoding signaling-related transcription factor and peptide hormone-like

proteins, in addition to many genes with unknown functions (Fig 4). Although their biological functions remain to be determined, our findings lay the foundation for further analysis of the precise gene regulatory network and developmental mechanisms underlying fruit set, in addition to the usage of promoter region of PSGs for genetic engineering and molecular breeding.

Supporting information

S1 Fig. Results of trimming and filtering of raw read data from each sample obtained by RNA-seq. Trimming was performed using FastQC. #1 and #2 represent FastQC analysis of original and trimmed data from petals, respectively. Pistil and fruit samples (#3–10): pistils of 2–2.5 mm buds (#3), 3–4 mm buds (#4), 1 DBF (#5), anthesis (#6), 5 DAF (#7), 5 mm ovaries at 7 DAF (#8), mature green fruits (#9), and red fruits (#10); Stamen and other floral organ samples (#12–14): stamens of 3–4 mm buds (#12), 1 DBF (#13) and anthesis (#14). sepals; Vegetative organs (#15–18): 3-week-old leaves (#15), mature leaves (#16), stems (#17), roots (#18), from left to right, respectively.
(TIF)

S2 Fig. Identification of pistil-specific genes based on the direct-mapping method. (A) Number of expressed genes in different tissue/stages. Genes with RPKM values greater than 0 and 0.5 are shown in the top and bottom panels, respectively. (B) Expression of tomato YABBY transcription factor family genes. The expression of nine YABBY transcription factor genes was examined. *SlCRCa*, *SlCRCb*, and *SlINO* appeared to be preferentially expressed in the pistil. Vertical axis represents the expression value (RPKM). Horizontal axis represents the 17 samples used for RNA-seq analysis.
(TIF)

S3 Fig. TPD1-like cysteine-rich peptides specifically expressed in pistils. (A) Alignment of five TPD1-like proteins in tomato, including SlTPD1 (Solyc11g005500), SlTPD1-like1 (Solyc12g009850), TPD1-like2 (Solyc05g010190), TPD1-like3 (Solyc04g071640), and TPD1 (AT4G24972). rice MIL2/TPD1A (Os12g0472500), maize MAC1 (JN247438). (B) Phylogenetic tree of three tomato TPD1-like proteins and several orthologs of Arabidopsis, rice and maize. Numbers above the branches indicate bootstrap values (10,000 replicates).
(TIF)

S4 Fig. EC1-like cysteine-rich peptides specifically expressed in pistils. (A) Expression of tomato EC1-like (ECL) genes. Both *SlECL1* and *SlECL5* were specifically expressed in the pistil at anthesis. Bottom one represents the expression of the internal control gene SAND [41]. (B) Alignment of five ECA1-like proteins in tomato, including SlECL1 (Solyc11g005500), SlECL2 (Solyc05g010190), SlECL3 (Solyc12g009850), SlECL4 (Solyc04g071640), and SlECL5 (Solyc11g005540). Arabidopsis EC1s and Tomato ECLs share six conserved cysteine residues at their C-termini (asterisk). (C) Neighbor-joining tree of amino acid sequences of five tomato ECLs proteins and five Arabidopsis EC1s. Numbers above the branches indicate bootstrap values (10,000 replicates).
(TIF)

S5 Fig. Second structure of CRPs and LRR-RLKs identified in this study. Conserved domains and motifs were searched using CDD in NCBI. The presence of secretion signal and transmembrane region was investigated using SignalP 4.1 Server and TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/).
(TIF)

S6 Fig. Five *PSGs* that are differentially regulated during fruit set. The data were obtained from Tang et al. 2015. Vertical axis represents the expression values normalized to transcripts per million (TPM) 6. Horizontal axis represents the pistil sample types.
(TIF)

S1 Table. Tomato tissues subject to RNA-seq and summary of sequencing results.
(XLSX)

S2 Table. Gene expression patterns of the 108 genes in various tissues of Micro-Tom.
(XLSX)

S3 Table. Gene expression patterns of the 108 PSGs in the pistils of wild tomato relative *S. pimpinellifolium* based on published data from Pattison et al. (2015).
(XLSX)

S4 Table. Gene expression patterns of the 108 PSGs based on published data from Zhang et al. (2016).
(XLSX)

S5 Table. List of oligonucleotide primers used for RT-PCR in this study. List of ovule preferentially expressed genes (*OPEs*) and embryo preferentially expressed genes (*EPEs*).
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: Kentaro Ezura, Tohru Ariizumi.

Data curation: Yutaka Suzuki, Satoru Kuhara.

Formal analysis: Kentaro Ezura, Kazuki Mori, Tohru Ariizumi.

Funding acquisition: Yutaka Suzuki, Tohru Ariizumi, Hiroshi Ezura.

Investigation: Kentaro Ezura, Kim Ji-Seong, Kazuki Mori, Tohru Ariizumi.

Methodology: Kentaro Ezura, Kim Ji-Seong, Kazuki Mori, Tohru Ariizumi.

Project administration: Yutaka Suzuki, Satoru Kuhara, Tohru Ariizumi, Hiroshi Ezura.

Resources: Yutaka Suzuki, Satoru Kuhara, Hiroshi Ezura.

Software: Satoru Kuhara.

Supervision: Tohru Ariizumi, Hiroshi Ezura.

Validation: Kentaro Ezura, Tohru Ariizumi.

Visualization: Kentaro Ezura, Tohru Ariizumi.

Writing – original draft: Kentaro Ezura, Tohru Ariizumi, Hiroshi Ezura.

Writing – review & editing: Kentaro Ezura, Tohru Ariizumi.

References

1. Gillaspy G, Ben-David H, and Gruissem W. Fruits: A Developmental Perspective. Plant Cell 1993; 5: 1439–1451. https://doi.org/10.1105/tpc.5.10.1439 PMID: 12271039

2. Azzi L, Deluche C, Gévaudant F, Frangne N., Delmas F., Hernould M, et al. Fruit growth-related genes in tomato. J Exp Bot 2015; 66: 1075–1086. https://doi.org/10.1093/jxb/eru527 PMID: 25573859

3. Pesaresi P, Mizzotti C, Colombo M, Masiero S. Genetic regulation and structural changes during tomato fruit development and ripening. Front Plant Sci 2014; 5: 124. https://doi.org/10.3389/fpls.2014.00124 PMID: 24795731

4. Ruan YL, Patrick JW, Bouzayen M, Osorio S, Fernie AR. Molecular regulation of seed and fruit set. Trends Plant Sci 2012; 17: 656–665. https://doi.org/10.1016/j.tplants.2012.06.005 PMID: 22776090

5. Sotelo-Silveira M, Marsch-Martínez N, de Folter S. Unraveling the signal scenario of fruit set. Planta 2014; 239: 1147. https://doi.org/10.1007/s00425-014-2057-7 PMID: 24659051

6. Ariizumi T, Amagai M, Shibata D, Hatakeyama K, Watanabe M. Comparative study of promoter activity of three anther-specific genes encoding lipid transfer protein, xyloglucan endotransglucosylase/hydrolase and polygalacturonase in transgenic Arabidopsis thaliana. Plant Cell Rep 2002; 21: 90–96. https://doi.org/10.1007/s00299-002-0487-3

7. Mezzetti B, Landi L, Pandolfini T, Spena A. The defH9-iaaM auxin-synthesizing gene increases plant fecundity and fruit production in strawberry and raspberry. BMC Biotechnol 2004; 4. https://doi.org/10.1186/1472-6750-4-4 PMID: 15113427

8. Mutz KO, Heilkenbrinker A, Lönne M, Walter JG, Stahl F. Transcriptome analysis using next-generation sequencing. Curr. Opin. Biotechnol 2013; 24: 22–30. https://doi.org/10.1016/j.copbio.2012.09.004 PMID: 23020966

9. Di C, Yuan J, Wu Y, Li J, Lin H, Hu L, et al. Characterization of Stress-responsive lncRNAs in Arabidopsis thaliana by Integrating Expression, Epigenetic and Structural Features. Plant J 2014; 80: 848–861. https://doi.org/10.1111/tpj.12679 PMID: 25256571

10. Ju C, Van de Poel B, Cooper ED, Thierer JH, Gibbons TR, Delwiche CF, et al. Conservation of ethylene as a plant hormone over 450 million years of evolution. Nature Plants 2015; 1: 14004. https://doi.org/10.1038/nplants.2014.4 PMID: 27246051

11. Kim G, LeBlanc ML, Wafula EK, de Pamphilis CW, Westwood JH. Plant science. Genomic-scale exchange of mRNA between a parasitic plant and its hosts. Science 2014; 345: 808–811. https://doi.org/10.1126/science.1253122 PMID: 25124438

12. Koenig D, Jiménez-Gómez JM, Kimura S, Fulop D, Chitwood DH, Headland LR, et al. Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. Proc Natl Acad Sci U S A 2013; 110: E2655–E2662. https://doi.org/10.1073/pnas.1309606110 PMID: 23803858

13. O'Rourke JA, Yang SS, Miller SS, Bucciarelli B, Liu J, Rydeen A, et al. An RNA-Seq transcriptome analysis of orthophosphate-deficient white lupin reveals novel insights into phosphorus acclimation in plants. Plant Physiol 2013; 161: 705–724. https://doi.org/10.1104/pp.112.209254 PMID: 23197803

14. Shen X, Wang Z, Song X, Xu J, Jiang C, Zhao Y, et al. Transcriptomic profiling revealed an important role of cell wall remodeling and ethylene signaling pathway during salt acclimation in Arabidopsis. Plant Mol Biol 2014; 86: 303–317. https://doi.org/10.1007/s11103-014-0230-9 PMID: 25092201

15. Voß U, Wilson MH, Kenobi K, Gould PD, Robertson FC, Peer WA, et al. The circadian clock rephases during lateral root organ initiation in Arabidopsis thaliana. Nat Commun 2015; 6: 7641. https://doi.org/10.1038/ncomms8641 PMID: 26144255

16. Sweetman C, Wong DC, Ford CM, Drew DP. Transcriptome analysis at four developmental stages of grape berry (Vitis vinifera cv. Shiraz) provides insights into regulated and coordinated gene expression. BMC Genomics 2012; 13: 691. https://doi.org/10.1186/1471-2164-13-691 PMID: 23227855

17. Gómez MD, Vera-Sirera F, Pérez-Amador MA. Molecular programme of senescence in dry and fleshy fruits. J Exp Bot 2014; 65: 4515–4526. https://doi.org/10.1093/jxb/eru093 PMID: 24874021

18. Gupta V, Estrada AD, Blakley I, Reid R, Patel K, Meyer MD, et al. RNA-Seq analysis and annotation of a draft blueberry genome assembly identifies candidate genes involved in fruit ripening, biosynthesis of bioactive compounds, and stage-specific alternative splicing. Gigascience 2015; 4: 5. https://doi.org/10.1186/s13742-015-0046-9 PMID: 25830017

19. Royo C, Carbonell-Bejerano P, Torres-Pérez R, Nebish A, Martínez Ó, Rey M, et al. Developmental, transcriptome, and genetic alterations associated with parthenocarpy in the grapevine seedless somatic variant Corinto bianco. J Exp Bot 2015; 67: 259–273. https://doi.org/10.1093/jxb/erv452 PMID: 26454283

20. Tang N, Deng W, Hu G, Hu N, Li Z. Transcriptome profiling reveals the regulatory mechanism underlying pollination dependent and parthenocarpic fruit set mainly mediated by auxin and gibberellin. PLoS ONE 2015; 10: e0125355. https://doi.org/10.1371/journal.pone.0125355 PMID: 25909657

21. Ye J, Hu T, Yang C, Li H, Yang M, Ijaz R, et al. Transcriptome Profiling of Tomato Fruit Development Reveals Transcription Factors Associated with Ascorbic Acid, Carotenoid and Flavonoid Biosynthesis. PLoS ONE 2015; 10: e0130885. https://doi.org/10.1371/journal.pone.0130885 PMID: 26133783

22. Zhu B, Yang Y, Li R, Fu D, Wen L, Luo Y, et al. RNA sequencing and functional analysis implicate the regulatory role of long non-coding RNAs in tomato fruit ripening. J Exp Bot 2015; 66: 4483–4495. https://doi.org/10.1093/jxb/erv203 PMID: 25948705

23. Zhang S, Xu M, Qiu Z, Wang K, Du Y, Gu L, et al. Spatiotemporal transcriptome provides insights into early fruit development of tomato (Solanum lycopersicum). Sci Rep 2016; 6: 23173. https://doi.org/10.1038/srep23173 PMID: 26988970

24. Vriezen WH, Feron R, Maretto F, Keijman J, Mariani C. Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. New Phytol 2008; 177: 60–76. https://doi.org/10.1111/j.1469-8137.2007.02254.x PMID: 18028300

25. Wang H, Schauer N, Usadel B, Frasse P, Zouine M, Hernould M, et al. Regulatory features underlying pollination-dependent and -independent tomato fruit set revealed by transcript and primary metabolite profiling. Plant Cell 2009; 21: 1428–1452. https://doi.org/10.1105/tpc.108.060830 PMID: 19435935

26. Ruiu F, Picarella ME, Imanishi S, Mazzucato A. A transcriptomic approach to identify regulatory genes involved in fruit set of wild-type and parthenocarpic tomato genotypes. Plant Mol Biol 2015; 89: 263–278. https://doi.org/10.1007/s11103-015-0367-1 PMID: 26319515

27. Pattison RJ, Csukasi F, Zheng Y, Fei Z, van der Knaap E, Catalá C. Comprehensive tissue-specific transcriptome analysis reveals distinct regulatory programs during early tomato fruit development. Plant Physiol 2015; 168 1684–1701. https://doi.org/10.1104/pp.15.00287 PMID: 26099271

28. Hong SB, and Tucker ML. Molecular characterization of a tomato polygalacturonase gene abundantly expressed in the upper third of pistils from opened and unopened flowers. Plant Cell Rep 2000; 19: 680–683. https://doi.org/10.1007/s002999900175

29. Eberle CA, Anderson NO, Clasen BM, Hegeman AD, Smith AG. PELPIII: the class III pistil-specific extensin-like Nicotiana tabacum proteins are essential for interspecific incompatibility. Plant J 2013; 74: 805–814. https://doi.org/10.1111/tpj.12163 PMID: 23461796

30. Goldman MH, Pezzotti M, Seurinck J, Mariani C. Developmental expression of tobacco pistil-specific genes encoding novel extensin-like proteins. Plant Cell 1992; 4: 1041–1051. https://doi.org/10.1105/tpc.4.9.1041 PMID: 1392607

31. de Graaf BH, Knuiman BA, Derksen J, Mariani C. Characterization and localization of the transmitting tissue-specific PELPIII proteins of Nicotiana tabacum. J Exp Bot 2003; 54: 55–63. https://doi.org/10.1093/jxb/erg002 PMID: 12456755

32. Brambilla V, Battaglia R, Colombo M, Masiero S, Bencivenga S, Kater MM, et al. Genetic and molecular interactions between BELL1 and MADS box factors support ovule development in Arabidopsis. Plant Cell (2007); 19: 2544–2556. https://doi.org/10.1105/tpc.107.051797 PMID: 17693535

33. Yung MH, Schaffer R, Putterill J. Identification of genes expressed during early Arabidopsis carpel development by mRNA differential display: characterisation of ATCEL2, a novel endo-1,4-beta-D-glucanase gene. Plant J 1999: 17; 203–208. https://doi.org/10.1046/j.1365-313X.1999.00359.x PMID: 10074717

34. Fasoli M, Dal Santo S, Zenoni S, Tornielli GB, Farina L, Zamboni A, et al. The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. Plant Cell 2012; 24: 3489–3505. https://doi.org/10.1105/tpc.112.100230 PMID: 22948079

35. Rhee SJ, Seo M, Jang YJ, Cho S, Lee GP. Transcriptome profiling of differentially expressed genes in floral buds and flowers of male sterile and fertile lines in watermelon. BMC Genomics 2015; 16: 914. https://doi.org/10.1186/s12864-015-2186-9 PMID: 26552448

36. Mao Y, Zhang Z, Feng Z, Wei P, Zhang H, Botella JR, et al. Development of germ-line-specific CRISPR-Cas9 systems to improve the production of heritable gene modifications in Arabidopsis. Plant Biotechnol J 2015; 14:519–532. https://doi.org/10.1111/pbi.12468 PMID: 26360626

37. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014; 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404

38. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature Methods 2008; 5: 621–628. https://doi.org/10.1038/nmeth.1226 PMID: 18516045

39. Hansen KD, Irizarry RA, Wu Z. Removing technical variability in RNA-seq data using conditional quantile normalization. Biostatistics 2012; 13: 204–216. https://doi.org/10.1093/biostatistics/kxr054 PMID: 22285995

40. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003; 19: 185–193. PMID: 12538238

41. Expósito-Rodríguez M, Borges AA, Borges-Pérez A, Pérez JA. Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. BMC Plant Biol 2008; 8: 131. https://doi.org/10.1186/1471-2229-8-131 PMID: 19102748

42. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. Nucleic Acids Res 2015; 43: D222–226. https://doi.org/10.1093/nar/gku1221 PMID: 25414356

43. Tomato Genome Consortium. The tomato genome sequence provides insights into fleshy fruit evolution. Nature 2012; 485: 635–641. https://doi.org/10.1038/nature11119 PMID: 22660326

44. Alvarez J, Smyth DR. CRABS CLAW and SPATULA, two Arabidopsis genes that control carpel development in parallel with AGAMOUS. Development 1999; 126: 2377–2386. PMID: 10225997

45. Balasubramanian S, Schneitz K. NOZZLE regulates proximal-distal pattern formation, cell proliferation and early sporogenesis during ovule development in Arabidopsis thaliana. Development 2000; 127: 4227–4238. PMID: 10976054

46. Villanueva JM, Broadhvest J, Hauser BA, Meister RJ, Schneitz K, Gasser CS. INNER NO OUTER regulates abaxial- adaxial patterning in Arabidopsis ovules. Genes Dev 1999; 13: 3160–3169. PMID: 10601041

47. Huang Z, Van Houten J, Gonzalez G, Xiao H, van der Knaap E. Genome-wide identification, phylogeny and expression analysis of SUN, OFP and YABBY gene family in tomato. Mol Genet Genomics 2013; 288: 111–129. https://doi.org/10.1007/s00438-013-0733-0 PMID: 23371549

48. Nibau C, Di Stilio VS, Wu HM, Cheung AY. Arabidopsis and Tobacco superman regulate hormone signalling and mediate cell proliferation and differentiation. J Exp Bot 2011; 62: 949–961. https://doi.org/10.1093/jxb/erq325 PMID: 20980362

49. Jin Y, Ni DA, Ruan TL. Posttranslational elevation of cell wall invertase activity by silencing its inhibitor in tomato delays leaf senescence and increases seed weight and fruit hexose level. Plant Cell 2009; 21: 2072–2089. https://doi.org/10.1105/tpc.108.063719 PMID: 19574437

50. Kauss H, Swanson AL, Hassid WZ. Biosynthesis of the methyl ester groups of pectin by transmethylation from S-adenosyl-L-methionine. Biochem Biophys Res Commun 1967; 26: 234–240. https://doi.org/10.1016/0006-291X(67)90240-9 PMID: 6030269

51. Vannier MP, Thoiron B, Morvan C, Demarty M. Localization of methyltransferase activities throughout the endomembrane system of flax (Linum usitatissimum L) hypocotyls. Biochem J 1992; 286: 863–868. PMID: 1329719

52. Jolie RP, Duvetter T, Van Loey AM, Hendrickx ME. Pectin methylesterase and its proteinaceous inhibitor: a review. Carbohydr Res 2010; 345: 2583–2595. https://doi.org/10.1016/j.carres.2010.10.002 PMID: 21047623

53. Micheli F. Pectin methylesterases: cell wall enzymes with important roles in plant physiology. Trends Plant Sci 2001; 6: 414–419. https://doi.org/10.1016/S1360-1385(01)02045-3 PMID: 11544130

54. Pelloux J, Rustérucci C, Mellerowicz EJ. New insights into pectin methylesterase structure and function. Trends Plant Sci 2007; 12: 267–277. https://doi.org/10.1016/j.tplants.2007.04.001 PMID: 17499007

55. Liners F, Thibault JF, Van Cutsem P. Influence of the degree of polymerization of oligogalacturonates and of esterification pattern of pectin on their recognition by monoclonal antibodies. Plant Physiol 1992; 99: 1099–1104. https://doi.org/10.1104/pp.99.3.1099 PMID: 16668976

56. Daher FB, Braybrook SA. How to let go: pectin and plant cell adhesion. Front Plant Sci 2015; 6: 523. https://doi.org/10.3389/fpls.2015.00523 PMID: 26236321

57. Lionetti V, Raiola A, Camardella L, Giovane A, Obel N, Pauly M, et al. Overexpression of pectin methylesterase inhibitors in Arabidopsis restricts fungal infection by Botrytis cinerea. Plant Physiol 2007; 143: 1871–1880. https://doi.org/10.1104/pp.106.090803 PMID: 17277091

58. Gaffe J, Tiznado ME, Handa AK. Characterization and functional expression of a ubiquitously expressed tomato pectin methylesterase. Plant Physiol 1997; 114: 1547–1556. https://doi.org/10.1104/pp.114.4.1547 PMID: 9276962

59. Terao A, Hyodo H, Satoh S, Iwai H. 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). J Plant Res 2013; 126: 719–728. https://doi.org/10.1007/s10265-013-0555-5 PMID: 23455617

60. Hothorn M, Wolf S, Aloy P, Greiner S, and Scheffzek K. Structural insights into the target specificity of plant invertase and pectin methylesterase inhibitory proteins. Plant Cell 2004; 16: 3437–3447. https://doi.org/10.1105/tpc.104.025684 PMID: 15528298

61. Hothorn M, Van den Ende W, Lammens W, Rybin V, Scheffzek K. Structural insights into the pH-controlled targeting of plant cell-wall invertase by a specific inhibitor protein. Proc Natl Acad Sci U S A 2010; 107: 17427–17432. https://doi.org/10.1073/pnas.1004481107 PMID: 20858733

62. Li Z, Palmer WM, Martin AP, Wang R, Rainsford F, Jin Y, et al. High invertase activity in tomato reproductive organs correlates with enhanced sucrose import into, and heat tolerance of, young fruit. J Exp Bot 2012; 63: 1155–1166. https://doi.org/10.1093/jxb/err329 PMID: 22105847

63. Palmer WM, Ru L, Jin Y, Patrick JW, Ruan YL. Tomato ovary-to-fruit transition is characterized by a spatial shift of mRNAs for cell wall invertase and its inhibitor with the encoded proteins localized to sieve elements. Mol Plant 2015; 8: 315–328. https://doi.org/10.1016/j.molp.2014.12.019 PMID: 25680776

64. Appelhagen I, Lu GH, Huep G, Schmelzer E, Weisshaar B, Sagasser M. TRANSPARENT TESTA1 interacts with R2R3-MYB factors and affects early and late steps of flavonoid biosynthesis in the endothelium of Arabidopsis thaliana seeds. Plant J 2011; 67: 406–419. https://doi.org/10.1111/j.1365-313X.2011.04603.x PMID: 21477081

65. Debeaujon I, Peeters AJ, Léon-Kloosterziel KM, Koornneef M. The TRANSPARENT TESTA12 gene of Arabidopsis encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. Plant Cell 2001; 13: 853–871. https://doi.org/10.1105/tpc.13.4.853 PMID: 11283341

66. Sagasser M Lu GH, Hahlbrock K, Weisshaar BA. thaliana TRANSPARENT TESTA 1 is involved in seed coat development and defines the WIP subfamily of plant zinc finger proteins. Genes Dev 2002; 16: 138–149. https://doi.org/10.1101/gad.212702 PMID: 11782451

67. Mazzucato A, Willems D, Bernini R, Picarella ME, Santangelo E, Ruiu F, et al. Novel phenotypes related to the breeding of purple-fruited tomatoes and effect of peel extracts on human cancer cell proliferation. Plant Physiol Biochem 2013; 72: 125–133. https://doi.org/10.1016/j.plaphy.2013.05.012 PMID: 23769702

68. Schijlen EGWM, de Vos CHR, Martens S, Jonker HH, Rosin FM, Molthoff JW, et al. RNA interference silencing of chalcone synthase, the first step in the flavonoid biosynthesis pathway, leads to parthenocarpic tomato fruits. Plant Physiol 2007; 144: 1520–1530. https://doi.org/10.1104/pp.107.100305 PMID: 17478633

69. Ingrosso I, Bonsegna S, De Domenico S, Laddomada B, Blando F, Santino A, et al. Over-expression of a grape stilbene synthase gene in tomato induces parthenocarpy and causes abnormal pollen development. Plant Physiol Biochem 2011 49 1092–1099. https://doi.org/10.1016/j.plaphy.2011.07.012 PMID: 21843947

70. Golz JF, Hudson A. Plant development: YABBYs claw to the fore. Curr Biol 1999; 9: R861–R863. https://doi.org/10.1016/S0960-9822(00)80047-0 PMID: 10574752

71. Ariel FD, Manavella PA, Dezar CA, Chan RL. The true story of the HD-Zip family. Trends Plant Sci 2007; 12: 419–426. https://doi.org/10.1016/j.tplants.2007.08.003 PMID: 17698401

72. Jin J, Zhang H, Kong L, Gao G, Luo J. PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. Nucleic Acids Res 2014; 42: D1182–D1187. https://doi.org/10.1093/nar/gkt1016 PMID: 24174544

73. Ribone PA, Capella M, Chan RL. Functional characterization of the homeodomain leucine zipper I transcription factor AtHB13 reveals a crucial role in Arabidopsis development. J Exp Bot 2015; 66: 5929–5943. https://doi.org/10.1093/jxb/erv302 PMID: 26136262

74. Hu T, Ye J, Tao P, Li H, Zhang J, Zhang Y, Ye Z. Tomato HD-Zip I transcription factor, SlHZ24, modulates ascorbate accumulation through positively regulating the D-mannose/L-galactose pathway. Plant J 2015; 85: 16–29. https://doi.org/10.1111/tpj.13085 PMID: 26610866

75. Lin Z, Hong Y, Yin M, Li C, Zhang K, Grierson D. A tomato HD-Zip homeobox protein, LeHB-1, plays an important role in floral organogenesis and ripening. Plant J 2008; 55: 301–310. https://doi.org/10.1111/j.1365-313X.2008.03505.x PMID: 18397374

76. Smaczniak C, Immink RG, Angenent GC, Kaufmann K. Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. Development 2012; 139: 3081–3098. https://doi.org/10.1242/dev.074674 PMID: 22872082

77. Smaczniak C, Immink RG, Muiño JM, Blanvillain R, Busscher M. Characterization of MADS-domain transcription factor complexes in Arabidopsis fl ower development. Proc Natl Acad Sci USA 2012; 109: 1560–1565. https://doi.org/10.1073/pnas.1112871109 PMID: 22238427

78. Dreni L, Kater MM. MADS reloaded: evolution of the AGAMOUS subfamily genes. New Phytol 2014; 201: 717–732. https://doi.org/10.1111/nph.12555 PMID: 24164649

79. Bemer M, Heijmans K, Airoldi C, Davies B, Angenent GC. An atlas of type I MADS box gene expression during female gametophyte and seed development in Arabidopsis. Plant Physiol 2010; 154: 287–300. https://doi.org/10.1104/pp.110.160770 PMID: 20631316

80. Colombo M, Masiero S, Vanzulli S, Lardelli P, Kater MM, Colombo L. AGL23, a type I MADS-box gene that controls female gametophyte and embryo development in Arabidopsis. Plant J 2008; 54: 1037–1048. https://doi.org/10.1111/j.1365-313X.2008.03485.x PMID: 18346189

81. 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: 2088–2101. https://doi.org/10.1105/tpc.108.058958 PMID: 18713950

82. 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: 865–872. https://doi.org/10.1105/tpc.110.081737 PMID: 21378131

83. Figueiredo DD, Batista RA, Roszak PJ, Köhler C. Auxin production couples endosperm development to fertilization. Nat Plants 2015; 1: 15184. https://doi.org/10.1038/nplants.2015.184 PMID: 27251719

84. Figueiredo DD, Batista RA, Roszak PJ, Hennig L, Köhler C. Auxin production in the endosperm drives seed coat development in Arabidopsis. Elife 2016; 16: 5. pii: e20542. https://doi.org/10.7554/eLife.20542

85. Ocarez N., Mejía N. Suppression of the D-class MADS-box AGL11 gene triggers seedlessness in fleshy fruits. Plant Cell Rep 2016; 35: 239–254. https://doi.org/10.1007/s00299-015-1882-x PMID: 26563346

86. Ampomah-Dwamena C, Morris BA, Sutherland P, Veit B, Yao JL. Down-regulation of TM29, a tomato SEPALLATA homolog, causes parthenocarpic fruit development and floral reversion. Plant Physiol 2002; 130: 605–617. https://doi.org/10.1104/pp.005223 PMID: 12376628

87. de Martino G, Pan I, Emmanuel E, Levy A, Irish VF. Functional analyses of two tomato APETALA3 genes demonstrate diversification in their roles in regulating floral development. Plant Cell 2006; 18: 1833–1845. https://doi.org/10.1105/tpc.106.042978 PMID: 16844904

88. Daminato M, Masiero S, Resentini F, Lovisetto A, Casadoro G. Characterization of TM8, a MADS-box gene expressed in tomato flowers. BMC Plant Biol 2014; 14: 319. https://doi.org/10.1186/s12870-014-0319-y PMID: 25433802

89. Quinet M, Bataille G, Dobrev PI, Capel C. Gómez P, Capel J, et al. Transcriptional and hormonal regulation of petal and stamen development by STAMENLESS, the tomato (Solanum lycopersicum L.) orthologue to the B-class APETALA3 gene. J Exp Bot 2014; 65: 2243–2256. https://doi.org/10.1093/jxb/eru089 PMID: 24659487

90. Klap C, Yeshayahou E, Bolger AM, Arazi T, Gupta SK, Shabtai S, et al. Tomato facultative parthenocarpy results from SlAGAMOUS-LIKE 6 loss of function. Plant Biotechnol J 2016 https://doi.org/10.1111/pbi.12662 [Epub ahead of print] PMID: 27862876

91. Katsir L, Davies KA, Bergmann DC, Laux T. Peptide signaling in plant development. Curr Biol 2011; 21: R356–R364. https://doi.org/10.1016/j.cub.2011.03.012 PMID: 21549958

92. Matsubayashi Y. Post-translational modifications in secreted peptide hormones in plants. Plant Cell Physiol 2011; 52: 5–13. https://doi.org/10.1093/pcp/pcq169 PMID: 21071428

93. Tavormina P, De Coninck B, Nikonorova N, De Smet I, Cammue BP. The Plant Peptidome: An Expanding Repertoire of Structural Features and Biological Functions. Plant Cell 2015; 27: 2095–2118. https://doi.org/10.1105/tpc.15.00440 PMID: 26276833

94. Breiden M, Simon R. Q&A: How does peptide signaling direct plant development? BMC Biol 2016; 14: 58. https://doi.org/10.1186/s12915-016-0280-3 PMID: 27392694

95. Yamaguchi YL, Ishida T, Sawa S. CLE peptides and their signaling pathways in plant development. J Exp Bot 2016; 67: 4813–4826. https://doi.org/10.1093/jxb/erw208 PMID: 27229733

96. Lee JS, Torii KU. A tale of two systems: peptide ligand-receptor pairs in plant development. Cold Spring Harb Symp Quant Biol 2012; 77: 83–89. https://doi.org/10.1101/sqb.2012.77.014886 PMID: 23274454

97. Marshall E, Costa LM, Gutierrez-Marcos J. Cysteine-rich peptides (CRPs) mediate diverse aspects of cell-cell communication in plant reproduction and development. J Exp Bot 2011; 62: 1677–1686. https://doi.org/10.1093/jxb/err002 PMID: 21317212

98. Silverstein KA Jr. Moskal WA, Wu HC, Underwood BA, Graham MA, Town CD, et al. Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. Plant J 2007; 51: 262–280. https://doi.org/10.1111/j.1365-313X.2007.03136.x PMID: 17565583

99. Mergaert P, Nikovics K, Kelemen Z, Maunoury N, Vaubert D, Kondorosi A, et al. A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted

polypeptides with conserved cysteine motifs. Plant Physiol 2003; 132: 161–173. https://doi.org/10. 1104/pp.102.018192 PMID: 12746522

100. Silverstein KA, Graham MA, Paape TD, VandenBosch KA. Genome organization of more than 300 defensin-like genes in Arabidopsis. Plant Physiol 2005; 138: 600–610. https://doi.org/10.1104/pp. 105.060079 PMID: 15955924

101. Okuda S, Tsutsui H, Shiina K, Sprunck S, Takeuchi H, Yui R, et al. Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells. Nature 2009; 458: 357–361. https://doi.org/ 10.1038/nature07882 PMID: 19295610

102. Tesfaye M, Silverstein KA, Nallu S, Wang L, Botanga CJ, Gomez SK, et al. Spatio-temporal expression patterns of Arabidopsis thaliana and Medicago truncatula defensin-like genes. PLoS One 2013; 8: e58992. https://doi.org/10.1371/journal.pone.0058992 PMID: 23527067

103. Richardson LG, Torii KU. Take a deep breath: peptide signalling in stomatal patterning and differentiation. J Exp Bot 2013; 64: 5243–5251. https://doi.org/10.1093/jxb/ert246 PMID: 23997204

104. Costa LM, Marshall E, Tesfaye M, Silverstein KA, Mori M, Umetsu Y, et al. Central cell-derived peptides regulate early embryo patterning in flowering plants. Science 2014; 344: 168–172. https://doi. org/10.1126/science.1243005 PMID: 24723605

105. Yang X, Tschaplinski TJ, Hurst Jawdy GBS, Abraham PE, Lankford PK, Adams RM, et al. Discovery and annotation of small proteins using genomics, proteomics, and computational approaches. Genome Res. 2011; 21: 634–641. https://doi.org/10.1101/gr.109280.110 PMID: 21367939

106. Yang SL, Xie LF, Mao HZ, Puah CS, Yang WC, Jiang L, et al. Tapetum determinant1 is required for cell specialization in the Arabidopsis anther. Plant Cell 2003; 15: 2792–2804. https://doi.org/10.1105/ tpc.016618 PMID: 14615601

107. Jia G, Liu X, Owen HA, Zhao D. Signaling of cell fate determination by the TPD1 small protein and EMS1 receptor kinase. Proc. Natl. Acad. Sci. U S A 2008; 105: 2220–2225. https://doi.org/10.1073/ pnas.0708795105 PMID: 18250314

108. Sprunck S, Rademacher S, Vogler F, Gheyselinck J, Grossniklaus U, Dresselhaus T. Egg cell-secreted EC1 triggers sperm cell activation during double fertilization. Science 2012; 338: 1093–1097. https://doi.org/10.1126/science.1223944 PMID: 23180860

109. Wei Z, Wang J, Yang S, Song Y. Identification and expression analysis of the LRR-RLK gene family in tomato (Solanum lycopersicum) Heinz 1706. Genome 2015; 58: 121–134. https://doi.org/10.1139/ gen-2015-0035 PMID: 26207619

110. Schoof H, Lenhard M, Haecker A, Mayer KF, Jürgens G, Laux T. The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. Cell 2000; 100: 635–644. https://doi.org/10.1016/S0092-8674(00)80700-X PMID: 10761929

111. Hazak O, Hardtke CS. CLAVATA 1-type receptors in plant development. J Exp Bot 2016; 67: 4827–4833. https://doi.org/10.1093/jxb/erw247 PMID: 27340234

112. Huang J, Zhang T, Linstroth L, Tillman Z, Otegui MS, Owen HA, et al. Control of Anther Cell Differentiation by the Small Protein Ligand TPD1 and Its Receptor EMS1 in Arabidopsis.PLoS Genet 2016; 12: e1006147. https://doi.org/10.1371/journal.pgen.1006147 eCollection 2016. PMID: 27537183

113. Clouse SD. Brassinosteroid Signal Transduction: From Receptor Kinase Activation to Transcriptional Networks Regulating Plant Development. Plant Cell 2011; 23: 1219–1230. https://doi.org/10.1105/ tpc.111.084475 PMID: 21505068

114. Wang ZY, Bai MY, Oh E, Zhu JY. Brassinosteroid signaling network and regulation of photomorphogenesis. Annu Rev Genet 2012; 46: 701–724. https://doi.org/10.1146/annurev-genet-102209-163450 PMID: 23020777

115. Belkhadir Y, Jaillain Y. The molecular circuitry of brassinosteroid signaling. New Phytol 2015; 206: 522–540. https://doi.org/10.1111/nph.13269 PMID: 25615890

116. Canales C, Bhatt AM, Scott R, Dickinson H. EXS, a putative LRR receptor kinase, regulates male germline cell number and tapetal identity and promotes seed development in Arabidopsis. Curr Biol 2002; 12: 1718–1727. https://doi.org/10.1016/S0960-9822(02)01151-X PMID: 12401166

117. Clark SE, Williams RW, Meyerowitz EM. The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. Cell 1997; 89: 575–585. https://doi.org/10. 1016/S0092-8674(00)80239-1 PMID: 9160749

118. Osakabe Y, Yamaguchi-Shinozaki K, Shinozaki K, Tran LS. Sensing the environment: key roles of membrane-localized kinases in plant perception and response to abiotic stress. J Exp Bot 2013; 64: 445–458. https://doi.org/10.1093/jxb/ers354 PMID: 23307915

119. Jeong S, Trotochaud AE, Clark SE. The Arabidopsis CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. Plant Cell 1999; 11: 1925–1934. https://doi.org/10.1105/tpc.11.10.1925 PMID: 10521522

120. Fiers M, Golemiec E, Xu J, van der Geest L, Heidstra R, Stiekema W, et al. The 14-amino acid CLV3, CLE19, and CLE40 peptides trigger consumption of the root meristem in Arabidopsis through a CLA-VATA2-dependent pathway. Plant Cell 2005; 17: 2542–2553. https://doi.org/10.1105/tpc.105.034009 PMID: 16055633

121. Guo Y, Han L, Hymes M, Denver R, Clark SE. CLAVATA2 forms a distinct CLE-binding receptor complex regulating Arabidopsis stem cell specification. Plant J 2010; 63: 889–900. https://doi.org/10.1111/j.1365-313X.2010.04295.x PMID: 20626648

122. Pan L, Lv S, Yang N, Lv Y, Liu Z, Wu J, et al. The Multifunction of CLAVATA2 in Plant Development and Immunity. Front Plant Sci 2016; 7: 1573. https://doi.org/10.3389/fpls.2016.01573 PMID: 27822222

123. Dutt M, Dhekney SA, Soriano L, Kandel R, Grosser JW. Temporal and spatial control of gene expression in horticultural crops. Horti Res 2014; 1: 14047. https://doi.org/10.1038/hortres.2014.47 PMID: 26504550

124. Odell JT, Nagy F, Chua NH. Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 1985; 313: 810–812. https://doi.org/10.1038/313810a0 PMID: 3974711

125. Benfey PN, Chua NH. The Cauliflower Mosaic Virus 35S Promoter: Combinatorial Regulation of Transcription in Plants. Science1990; 250: 959–966. https://doi.org/10.1126/science.250.4983.959 PMID: 17746920

126. Hiwasa-Tanase K, Kuroda H, Hirai T, Aoki K, Takane K, Ezura H. Novel promoters that induce specific transgene expression during the green to ripening stages of tomato fruit development. Plant Cell Rep 2012; 31: 1415–1424. https://doi.org/10.1007/s00299-012-1257-5 PMID: 22481231

127. Deikman J, Kline R, Fischer RL. Organization of Ripening and Ethylene Regulatory Regions in a Fruit-Specific Promoter from Tomato (Lycopersicon esculentum). Plant Physiol 1992; 100: 2013–2017. https://doi.org/10.1104/pp.100.4.2013 PMID: 16653232

128. Kesanakurti D, Kolattukudy PE, Kirti PB. Fruit-specific overexpression of wound-induced tap1 under E8 promoter in tomato confers resistance to fungal pathogens at ripening stage. Physiol Plant 2012; 146: 136–148. https://doi.org/10.1111/j.1399-3054.2012.01626.x PMID: 22462603

129. Vaughan SP, James DJ, Lindsey K, Massiah AJ. Characterization of FaRB7, a near root-specific gene from strawberry (Fragariaxananassa Duch.) and promoter activity analysis in homologous and heterologous hosts. J Exp Bot 2006; 57: 3901–3910. https://doi.org/10.1093/jxb/erl185 PMID: 17032728

130. Hong YF, Liu CY, Cheng KJ, Hour AL, Chan MT, Tseng TH, et al. The sweet potato sporamin promoter confers high-level phytase expression and improves organic phosphorus acquisition and tuber yield of transgenic potato. Plant Mol Biol 2008; 67: 347–361. https://doi.org/10.1007/s11103-008-9324-6 PMID: 18389377

131. Jones MO, Manning K, Andrews J, Wright C, Taylor IB, Thompson AJ. The promoter fromSlREO, a highly-expressed, root-specificSolanum lycopersicum gene, directs expression to cortex of mature roots. Funct Plant Biol 2008; 35: 1224–1233. https://doi.org/10.1071/FP08139

132. Fiedler U, Filistein R, Wobus U, Bäumlein H. A complex ensemble of cis-regulatory elements controls the expression of a Vicia faba non-storage seed protein gene. Plant Mol Biol 1993; 22: 669–679. https://doi.org/10.1007/BF00047407 PMID: 8343602

133. Jeong HJ, Choi JY, Shin HY, Bae JM, Shin JS. Seed-specific expression of seven Arabidopsis promoters. Gene 2014; 553: 17–23. https://doi.org/10.1016/j.gene.2014.09.051 PMID: 25261846

134. Li ZT, Gray DJ. Isolation by improved thermal asymmetric interlaced PCR and characterization of a seed-specific 2S albumin gene and its promoter from grape (Vitis vinifera L.). Genome 2005; 48: 312–320. https://doi.org/10.1139/g04-110 PMID: 15838554

135. Miyoshi H, Usami T, Tanaka I. High level of GUS gene expression driven by pollen-specific promoters in electroporated lily pollen protoplasts. Sex plant reprod 1995; 8; 205–209. https://doi.org/10.1007/BF00228938

136. Ariizumi T, Shinozaki Y, Ezura H. Genes that influence yield in tomato. Breed Sci 2013; 63: 3–13. https://doi.org/10.1270/jsbbs.63.3 PMID: 23641176

137. Kato H, Xie G, Sato Y, Imai R. Isolation of Anther-specific Gene Promoters Suitable for Transgene Expression in Rice. Plant Mol Biol Rep 2010; 28: 381–387. https://doi.org/10.1007/s11105-009-0162-8

138. Gómez MD, Beltrán JP, Cañas LA. The pea END1 promoter drives anther-specific gene expression in different plant species. Planta 2004; 219: 967–981. https://doi.org/10.1007/s00425-004-1300-z PMID: 15221384

139. Sassa H, Hirano H. Style-specific and developmentally regulated accumulation of a glycosylated thaumatin/PR5-like protein in Japanese pear (Pyrus serotina Rehd.). Planta 1998; 205: 514–521. https://doi.org/10.1007/s004250050350 PMID: 9684355

140. Sassa H, Ushijima K, Hirano H. A pistil-specific thaumatin/PR5-like protein gene of Japanese pear (Pyrus serotina): sequence and promoter activity of the 5' region in transgenic tobacco. Plant Mol Biol 2002; 50: 371–377. https://doi.org/10.1023/A:1019866029629 PMID: 12369614

141. Kuboyama T. A novel thaumatin-like protein gene of tobacco is specifically expressed in the transmitting tissue of stigma and style. Sex plant reprod 1998; 11: 251–256. https://doi.org/10.1007/s004970050150

142. Ficker M, Wemmer T, Thompson RD. A promoter directing high level expression in pistils of transgenic plants. Plant Mol Biol 1997; 35: 425–431. https://doi.org/10.1023/A:1005898624425 PMID: 9349266

143. Martinelli F, Uratsu SL, Reagan RL, Chen Y, Tricoli D, Fiehn O, et al. Gene regulation in parthenocarpic tomato fruit. J Exp Bot 2009; 60: 3873–3890. https://doi.org/10.1093/jxb/erp227 PMID: 19700496

144. Wang ZP, Xing HL, Dong L, Zhang HY, Han CY, Wang XC, et al. Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biol 2015; 16: 144. https://doi.org/10.1186/s13059-015-0715-0 PMID: 26193878