Running head: Domain-specific transcriptomic analysis

Robert G. Franks

Department of Plant and Microbial Biology

Campus Box 7614

North Carolina State University,

Raleigh, NC, 27695-7614, USA

919 513-7705

rgfranks@ncsu.edu

Research Area:

Development - Flower Development

Secondary Research Area:

Development - Meristem Maintenance

Keywords:

Flower development, transcriptomics, FACS, Arabidopsis, meristem, *SHATTERPROOF2*, cell-type specific, *REPRODUCTIVE MERISTEM*, *ABI3/VP1*

Temporal and spatial domain-specific transcriptomic analysis of a vital reproductive meristem in *Arabidopsis thaliana*

Gonzalo. H. Villarino[1], Qiwen Hu[2], Silvia Manrique[3], Miguel Flores-Vergara[1], Bhupinder Sehra[1], Linda Robles[1], Javier Brumos[1], Anna N. Stepanova[1], Lucia Colombo[3], Eva Sundberg[4], Steffen Heber[2] and Robert G. Franks[1*]

[1] North Carolina State University, Department of Plant and Microbial Biology, Raleigh, NC, 27606, USA

[2] North Carolina State University, Department of Computer Science and the Bioinformatics Research Center, NCSU, Raleigh, NC, 27606, USA

[3] Università degli Studi di Milano Dip. di BioScienze - Sezione di Botanica Generale Via Celoria, Milano, Italy

[4] Department of Plant Biology, Swedish University of Agricultural Sciences, Uppsala, Sweden

[*]Corresponding author

Gonzalo. H. Villarino: ghvillar@ncsu.edu; Qiwen Hu: qhu@ncsu.edu; Silvia Manrique: silvia.manrique@unimi.it; Miguel Flores-Vergara: maflores@ncsu.edu; Bhupinder Sehra: bsehra@ncsu.edu; Linda Robles: lrobles1@uncc.edu; Javier Brumos: jbrumos@ncsu.edu; Anna Stepanova: atstepan@ncsu.edu; Lucia Colombo: lucia.colombo@unimi.it; Eva Sundberg: eva.sundberg@slu.se; Steffen Heber: sheber@ncsu.edu; Robert G. Franks: rgfranks@ncsu.edu

One Sentence Summary: A new method to assay transcriptional profiles of spatially- and temporally-restricted cell populations from the *Arabidopsis* gynoecium reveals the meristematic nature of the gynoecial medial domain.

Footnotes:

Financial source: This work was funded by a grant from the National Science Foundation to RGF and SH (NSF IOS-1355019) and the FP7–PEOPLE–2013–IRSE FRUIT LOOK programme (to ES, LC and RGF).

Corresponding Author:

Robert G. Franks - rgfranks@ncsu.edu

ABSTRACT

Plant meristems, like animal stem cell niches, maintain a pool of multi-potent, undifferentiated cells that divide and differentiate to give rise to organs. In *Arabidopsis thaliana*, the carpel margin meristem is a vital meristematic structure that generates ovules from the medial domain of the gynoecium, the female floral reproductive structure. The molecular mechanisms that specify this meristematic region and regulate its organogenic potential are poorly understood. Here, we present a novel approach to analyse the transcriptional signature of the medial domain of the Arabidopsis gynoecium, highlighting the developmental stages that immediately proceed ovule initiation, the earliest stages of seed development. Using a floral synchronization system and a *SHATTERPROOF2* domain-specific reporter, paired with fluorescence-activated cell sorting and RNA sequencing, we assayed the transcriptome of the gynoecial medial domain with temporal and spatial precision. This analysis reveals a set of genes that are differentially expressed within the *SHATTERPROOF2* expression domain including genes that have been shown previously to function during the development of medial domain-derived structures, including the ovules, thus validating our approach. Global analyses of the transcriptomic dataset indicate a similarity of the *pSHP2*-expressing cell population to previously characterized meristematic domains, further supporting the meristematic nature of this gynoecial tissue. Our method identifies additional genes including novel isoforms, cis-natural antisense transcripts and a previously unrecognized member of the *REPRODUCTIVE MERISTEM* family of transcriptional regulators that are potential novel regulators of medial domain development. This data set provides genome-wide transcriptional insight into the development of the carpel margin meristem in Arabidopsis.

INTRODUCTION

The seedpod of flowering plants develops from the gynoecium, the female reproductive structure of the flower (Seymour et al., 2013). The gynoecium generates the ovules (the precursors of the seeds) and develops into the edible fruit in many fruiting species. As an estimated two-thirds of the calories of humankind's' diet are derived from gynoecia and seeds, the gynoecium is a globally vital structure (Oram and Brock, 1972), (Singh and Bhalla, 2007).

In the flowering plant *Arabidopsis thaliana,* the gynoecium is a morphologically complex, multi-organ structure with a diversity of tissues and cell types (Sessions and Zambryski, 1995; Bowman et al., 1999b; Seymour et al., 2013). The mature gynoecium displays morphological and functional differentiation along apical-basal, medio-lateral and adaxial-abaxial (inner-outer) axes. Stigmatic and stylar tissue form at the apex of the gynoecium, where the pollen grains are received and germinate. The stigma and style also comprise the apical-most portion of the transmitting tract, a structure that allows the pollen tube cell and sperm cells to reach the internally-located female gametophytes (Sessions and Zambryski, 1995; Sessions, 1999; Crawford and Yanofsky, 2008) (Fig. 1a,b,c). Located basal to the stigmatic and stylar tissue is the ovary portion of the gynoecium.

Ovules form within the ovary from a meristematic structure termed the medial ridge or carpel margin meristem (CMM), located in medial portions of the gynoecium (Bowman et al., 1999b; Alvarez and Smyth, 2002; Reyes-Olalde et al., 2013) (Fig. 1a,b,c). Plant meristems are analogous to animal stem cell niches as they maintain a set of undifferentiated cells that can divide and differentiate into numerous tissues and cell types (Aichinger et al., 2012). Early during floral development, patterning events divide the gynoecial primordium into medial domain that contains the carpel margin meristem and lateral domains that will form the walls of the gynoecium (Bowman et al., 1999b). These domains express different sets of transcriptional regulators from early developmental time points.

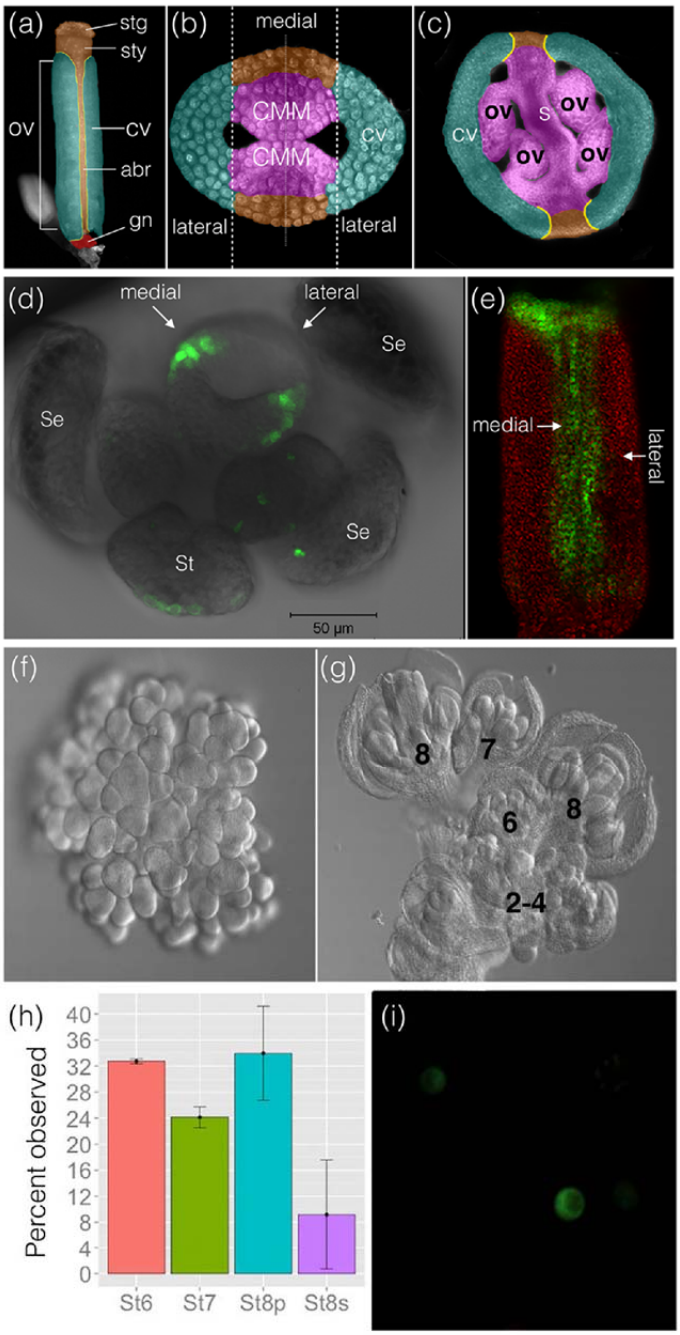

Figure 1. A system for the collection of temporally- and spatially-restricted cell populations from the *Arabidopsis thaliana* gynoecium. **(a)** Microscopic image of a mature wild type Arabidopsis gynoecium. The stigma (stg), style (sty), carpel valve (cv), abaxial replum (abr), gynophore (gn), and ovary (ovy) are false colored. **(b)** False-colored confocal cross section of a stage-8 gynoecium. Medial and lateral domains of the Arabidopsis gynoecium are indicated. The carpel margin meristem/medial ridge (CMM) is false colored pink. **(c)** False-colored stage-11 cross-section. Ovules (ov), septum (s) and carpel valves (cv) are indicated. **(d)** Confocal microscope image of the *pSHP2-YFP* two-component reporter in the *ap1; cal; pAP1::AP1:GR* background. YFP expression from the *pSHP2-YFP* reporter is chiefly confined to the medial domain of the gynoecium at late stage 7/early stage 8, although weak, non-medial domain expression can be detected in portions of the stamens. Sepals (se) and stamens (st) are labeled. **(e)** Z-stack composite 3D projection image of a gynoecium isolated from the flower at mid-stage 8. YFP expression from the *pSHP2-YFP* reporter is detected in the medial domain and at the apex of the gynoecium. **(f)** Chloral hydrate image of an inflorescence of an *ap1; cal; pAP1::AP1:GR* plant after mock treatment. Inflorescence-like meristems do not transition to floral meristems. **(g)** Chloral hydrate image of an inflorescence of an *ap1; cal; pAP1::AP1:GR* plant 125 hours after spray application of Dexamethasone synthetic hormone (Dex). Samples were enriched for stages 6-8. **(h)** Percentage of flowers at a given stage from inflorescences used for FACS-sorting. Stages 6, 7, 8p (pre-ovules) and 8s (post-ovules) are indicated in the Xaxis as St6, St7, St8p, St8s, respectively. Stage 8p is before any visible morphological manifestation of ovule primordia upon observation under DIC microscopy. Stage 8s ovule primordia were observed and were at ovule stage 1-I or 1-II according the Schneitz *et al.* (Schneitz et al., 1995). **(i)** Confocal microscopy of YFP fluorescence of protoplasted cells after FACS. Panels **(a)**, **(b)** and **(c)** are adapted from Azhakanandam *et al.* (Azhakanandam et al., 2008) (with permission).

Many genes that play a role in the development of the CMM and in the generation of ovules from this structure have been previously analyzed (Reyes-Olalde et al., 2013). However, due to the complexity of the developing gynoecium and the heterogeneity of the gynoecial tissues,

the ability to analyze the transcriptomic signature of the developing CMM or even other specific developing gynoecial structural domains has been limited. Wynn *et al*. previously evaluated the transcriptional properties of the gynoecial medial domain using hand-dissected gynoecial samples from the *seuss aintegumenta* (*seu ant*) double mutants that display a loss of many medial-domain-derived structures including ovules (Wynn et al., 2011). They identified 210 genes displaying reduced expression in *seu ant* gynoecia from floral stages 8-10 (Stages according to Smyth *et al.* (Smyth et al., 1990)). Many of these genes were shown via *in situ* hybridization to be preferentially expressed in the developing medial domain of the wild-type gynoecium and several of these genes have been shown to function during the development of ovules from the medial domain (Reyes-Olalde et al., 2013). It is, however, difficult with this approach to obtain samples from gynoecia younger that stage 8 and thus to assay the earliest gynoecial patterning events.

An alternative approach to investigate the transcriptional properties of specific cellular populations utilizes Fluorescence-Activated Cell Sorting (FACS) of protoplasted cells to isolate specific-cell populations based on patterns of gene expression. This approach has been successfully applied to the Arabidopsis Shoot Apical Meristem (SAM) (Yadav et al., 2009; Yadav et al., 2014) and roots (Birnbaum et al., 2003; Birnbaum et al., 2005), (Carter et al., 2013), (Lan et al., 2013) as well as to developing cell lineages within the Arabidopsis leaf epidermis (Adrian et al., 2015).

Here, we developed a novel FACS-based system for the transcriptomic analysis of a specific cellular population from the developing gynoecium, specifically the population of cells expressing the transcriptional regulator *SHATTERPROOF2* (*SHP2*). *SHP2* encodes a MADS-domain transcription factor that is expressed early within the developing CMM and thus functions as a marker for the meristematic population of cells that generate the transmitting tract and ovules (Ma et al., 1991; Savidge et al., 1995; Colombo et al., 2010; Larsson et al., 2014). In order to focus our analysis on early stages of gynoecium development during which key patterning events occur, we generated a *SHP2*-domain-specific reporter in a genetic background that allowed the synchronization of floral development. This, coupled with FACS-based protoplast sorting procedures and RNA sequencing, provided a unique temporal and spatial precision to assay the transcriptional signature of the gynoecial *SHP2*-expression domain.

Our system provides the ability to isolate a large numbers of cells from a temporally- and spatially-restricted gynoecial domain. We apply this method to investigate the transcriptomic signature of the medial domain of the gynoecium at the developmental stages when key patterning events and ovule initiation occur. Our analysis reveals many genes that are expressed preferentially within the developing medial portions of the gynoecium including members of the *REPRODUCTIVE MERISTEM* (*REM)* family of transcriptional regulators (Swaminathan et al., 2008; Romanel et al., 2009). We also take advantage of strand-specific RNA sequencing technology to find coding protein genes and non-coding RNAs (ncRNAs) as well as to examine isoforms and naturally occurring antisense transcripts that are preferentially expressed in the medial domain. This work complements and extends previous analyses of medial domain development and generates a list of potential novel regulators of medial domain development that are strong candidates for future functional analyses. Furthermore, global analyses of the transcriptomic dataset indicate a similarity of the *pSHP2*-expressing cell population to previously characterized meristematic domains, further supporting the meristematic nature of this gynoecial tissue.

RESULTS AND DISCUSSION

FACS-based protoplast sorting allows the collection of the *SHP2*-expressing cell population from a temporally restricted inflorescence sample

The transcriptional regulator *SHP2* is preferentially expressed in the medial domain of the gynoecium and in a subset of the medial-domain derived tissues (Savidge et al., 1995; Colombo et al., 2010; Larsson et al., 2014) (Fig. 1d,e). *SHP2* plays an important role in the development of the medial domain and in the specification of ovule identity (Liljegren et al., 2000; Favaro et al., 2003; Pinyopich et al., 2003; Colombo et al., 2010; Galbiati et al., 2013). To better characterize the molecular mechanisms of the medial domain and ovule development, we sought to identify transcripts that are differentially expressed within the medial domain of the *Arabidopsis thaliana* gynoecium relative to the rest of the inflorescence. To enable this, we generated a transgenic line containing a two-component reporter system, in which a *pUAS-3xYPET* reporter was driven by a *pSHP2-GAL4* driver construct (Methods). Throughout this manuscript we refer to this two-component reporter as *pSHP2-YFP*.

To better understand the early specification of medial and lateral gynoecial domains and in the earliest stages of ovule primordium initiation, we focused our transcriptomic analysis on floral stages 6-8, when these key developmental events occur (Bowman et al., 1999b). In order to increase our ability to collect a large number of *pSHP2-YFP*-expressing cells from this specific bracket of developmental stages, we crossed the *pSHP2-YFP* reporter into an *ap1 cal*-based floral synchronization system that allows the collection of large numbers of semi-synchronized flowers at roughly the same developmental stage (Wellmer et al., 2006; Ó'Maoiléidigh and Wellmer, 2014). The expression of the *pSHP2-YFP* reporter in the floral synchronization system was largely similar to that observed in wild-type inflorescences (Ma et al., 1991; Colombo et al., 2010; Larsson et al., 2014), and was confined chiefly to the medial domain and medial domain-derived tissues (Fig. 1d,e). Some expression was observed in non-medial domain tissues. The most apparent of this was expression in the apex of the developing gynoecium where both medial and lateral domains express the *pSHP2-YFP* reporter. Additionally, expression could be observed in a small number of cells within the stamens (Fig. 1d) and occasionally in the edges of sepals that appeared to have undergone a homeotic transformation toward a carpelloid fate (data not shown). Thus, the vast majority of the *pSHP-YFP* reporter expression reflected the endogenous *pSHP2* expression domain (in the

206 medial and apical portions of the gynoecium). A minority of the expression outside of the
207 gynoecium may reflect ectopic expression of the reporter due to genetic background or
208 transgene insertion site or limitations of the regulatory sequences used in the *pSHP-YFP*
209 reporter construct.
210
211 Microscopic examination of our semi-synchronized inflorescence samples indicated that
212 flowers ranged between floral stages 1 and early stage 8, with a strong enrichment for floral
213 stages 6 through early 8 (Fig. 1g,h). Flowers that had developed beyond late stage 8 were not
214 detected in our samples. Thus, our biological sample is strongly enriched for transcripts that
215 are expressed during early patterning of the gynoecium and the earliest stages of ovule
216 development (initiation) and does not include later floral developmental stages where *SHP2* is
217 expressed in stigma, style and valve margin tissues. Additionally, as the initial expression of
218 the *pSHP-YFP* reporter is detected at late stage 5 or early stage 6 (Larsson et al., 2014), we
219 expect that the population of YFP-expressing protoplasts derived from this material will be
220 highly enriched with cells from the stage 6-8 medial domain.
221
222 FACS-sorting of protoplasts derived from these inflorescences yielded three populations of
223 sorted cells (collected in biological quadruplicate): "YFP-positive", "YFP-negative" and "all-
224 sorted" (Figure S1). The "all-sorted" sample included all protoplasts recovered (regardless of
225 YFP expression) after sorting gates were applied to remove debris and broken cells
226 (Methods). We additionally collected (also in biological quadruplicate) "non-sorted" samples
227 from entire non-protoplasted inflorescences to measure the abundance of transcripts in the
228 biological starting material before protoplast generation and FACS-sorting. In order to
229 evaluate the purity of the YFP-positive protoplasts during a preliminary FACS run, YFP-
230 positive cells were resorted. Ninety six percent of the YFP-positive cells were found to resort
231 into the YFP-positive gate, indicating a high degree of enrichment and purity in the YFP-
232 positive sample (Figure S1). Confocal microscopy also revealed an enriched population of
233 intact YFP-positive protoplasts after FACS (Fig. 1i).
234
235 We used real time PCR (qRT-PCR) to estimate the degree of enrichment of the endogenous
236 *SHP2* and *NGATHA1* (*NGA1*) transcripts in RNA samples derived from the YFP-positive and

YFP-negative samples. *NGA1* is expressed in the adaxial portions of the gynoecium starting at stage 7 in a domain that partially overlaps with the *SHP2* expression domain (Alvarez et al., 2009; Trigueros et al., 2009) and thus provides an additional benchmark to estimate the enrichment of medial domain-expressed transcripts. The normalized level of the *SHP2* transcript was ~30 fold higher in the YFP-positive samples relative to the YFP-negative samples (p <0.001) while the *NGA1* transcript was ~4 fold higher in the YFP-positive sample (p <0.05). The difference in the levels of the *TUBLIN6* was not found to be statistically significant (p = 0.4) between the YFP-positive and YFP-negative samples (Figure S2).

Transcriptomic analysis of the gynoecial *SHP2* expression domain and identification of candidate regulators of gynoecial medial domain development

To investigate the transcriptomic profile of the gynoecial *SHP2* expression domain, we performed high-throughput RNA-sequencing (RNA-seq) from the collected protoplasts and non-protoplasted inflorescences samples. We expect that the identification of differentially expressed genes (DEGs) between the YFP-positive and YFP-negative samples (referred to as "YFP-positive/YFP-negative" or "YFP+/-") will provide insight into the set of transcripts differentially expressed in the gynoecial medial domain relative to the rest of the inflorescence. Additionally, DEGs identified in the all-sorted and non-sorted comparison (referred to as "all-sorted/non-sorted") are expected to reveal transcripts that are differentially represented as a result of the protoplasting/FACS-sorting protocol.

Two lanes of the HiSeq2500 Illumina sequencing platform yielded 320 million raw reads with an average of 20 million reads (MR) per library. Nearly 11 MR were filtered out after removing barcode-adapters and low quality sequences. The remaining 306 MR were aligned against the *Arabidopsis thaliana* TAIR10 reference genome (Lamesch et al., 2012) with more than 90% of them successfully mapping to the genome sequence. Among the mapped reads, 244 MR mapped uniquely to only one location and were used for subsequent analyses. A detailed breakdown is shown in Table S1.

We used three different programs to determine expressed and differentially expressed protein coding genes in our dataset: Cufflinks (Trapnell et al., 2012), edgeR (Robinson et al., 2010)

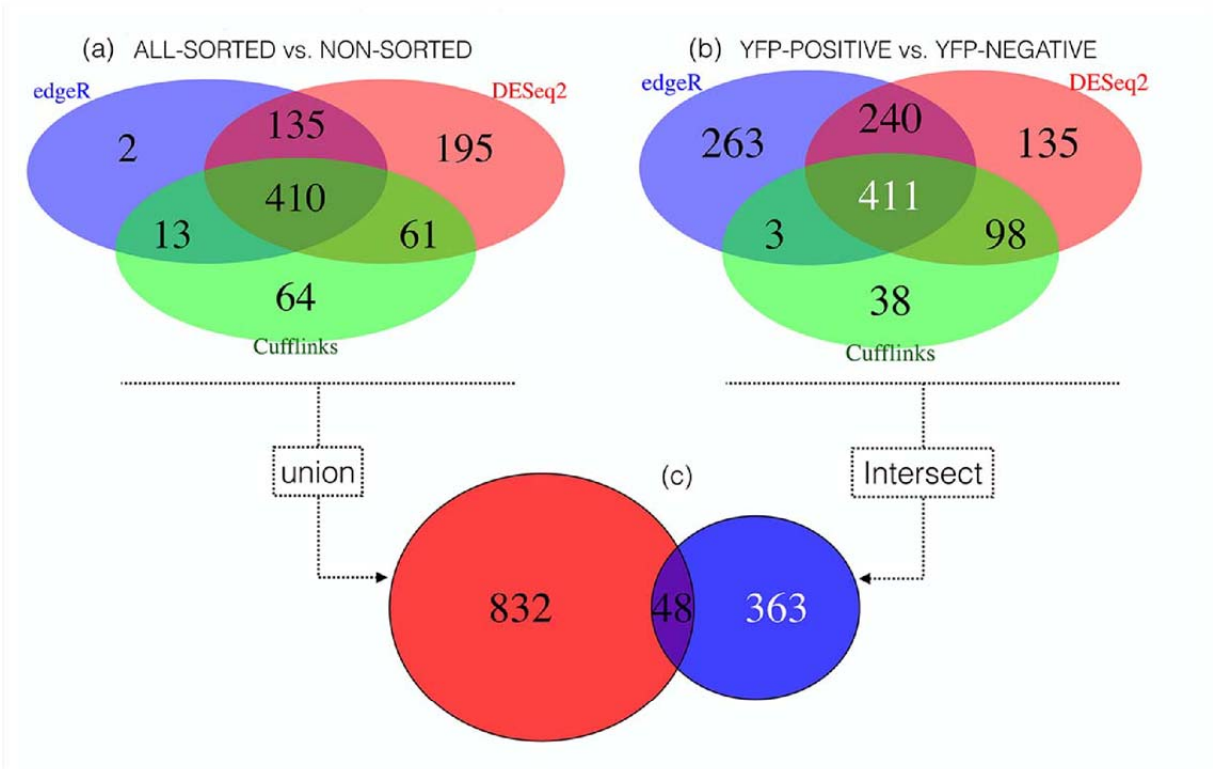

Figure 2. Venn diagram of differentially expressed genes (DEGs) using Cufflinks, edgeR and DESeq2 (FDR<0.001, Fold Change >4). **(a)** Venn showing DEGs identified between the all-sorted/non-sorted samples with the 3 programs used for differential expression analysis of RNA-seq expression profiles. **(b)** Venn showing DEGs between YFP+/- samples identified in the 3 programs. **(c)** Intersection of the DEGs (48) from both datasets **(a)** and **(b)**. DEGs (363), after removing DEGs induced by the protoplsting/FACS-sorting stress, were used for downstream analysis.

1

and DESeq2 (Love et al., 2014) (See Methods) (Non-protein coding gene models were considered separately and are presented below). Here, the term "differentially expressed gene (DEG)" is used to indicate a gene whose steady-state transcript level differs significantly at a false discovery rate (FDR) of <0.001 and shows a fold change of four or more between the two compared RNA samples. To identify potential regulators of gynoecial medial domain development, a 'stringent' criteria was used to select a subset of the YFP+/- DEGs for

downstream analysis. For a gene to be selected from the YFP+/- comparison, we required that the transcript is identified as differentially expressed by all three independent software packages (Fig. 2b). Alternatively, to identify DEGs in response to the protoplasting/FACS-sorting procedure, a 'less stringent' criterion was used. Transcripts in the union set of all the non-sorted/all-sorted DEGs were considered to be potential protoplast-induced genes even if they were identified by only one software program (Fig. 2a). Only 48 transcripts were found in common between the YFP+/- DEGs and the all-sorted/non-sorted DEGs (Fig. 2c), indicating a high degree of specificity in the DEGs identified in each comparison. We then removed these 48 transcripts from our analysis to eliminate any that might be differentially expressed as a result of the protoplast generation or FACS-sorting procedures, leaving 363 "cleaned" protein coding DEGs (Fig. 2c). The expression profiles of these 363 YFP+/- DEGs, including data from the all-sorted and non-sorted samples, are represented in a heatmap (Figure S3). This gene set includes 95 DEGs whose transcript levels were higher in the YFP-positive samples ("enriched") and 268 DEGs whose transcript levels were lower ("depleted") in the YFP-positive samples, relative to the YFP-negative samples (Table S2).

For the 95 DEGs that were enriched in the YFP-positive sample (at a fold change > 4), we expected many to be preferentially expressed in the medial portions of the gynoecium at floral stages 6-8. To test this, we examined the literature to determine the expression patterns of members of this gene set. From the top 15 of the 95 YFP-positive enriched DEGs (ranked by fold change), five have previously been reported to be preferentially expressed in the gynoecial medial domain via *in situ* or reporter gene analysis [i.e. *HECATE1* (*HEC1)*, *HEC2*, *SHP1*, *SHP2* and *STYLISH1* (*STY1)*] (Ma et al., 1991; Savidge et al., 1995; Kuusk et al., 2002; Gremski et al., 2007; Colombo et al., 2010) and three others are previously described as enriched in medial domain-derived tissues in published transcriptomic datasets (i.e., AT1G66950, AT5G14180, and AT1G03720) (Skinner and Gasser, 2009; Wuest et al., 2010) (Table I). An additional gene from this list, *CRABS CLAW* (*CRC*), has been shown via *in situ* hybridization to be expressed in portions of the medial gynoecial domain as well as non-medial portions of the gynoecium (Bowman and Smyth, 1999; Azhakanandam et al., 2008). The expression pattern of the remaining six genes from this gene list have not yet been assayed in the gynoecium. Thus, as predicted, the set of 95 genes enriched in the YFP-positive

sample is enriched for genes that are preferentially expressed in the gynoecial medial domain.

Published functional analyses of *HEC1*, *HEC2*, *SHP1*, *SHP2* and *STY1* indicate that these genes function during the development of the medial domain or medial domain-derived tissues (Kuusk et al., 2002; Favaro et al., 2003; Pinyopich et al., 2003; Gremski et al., 2007; Colombo et al., 2010). Many other genes in the set of 95 DEGs enriched in the YFP-positive sample have been previously shown to play a role in medial domain development (e.g. *NGA* family members (Alvarez et al., 2009; Trigueros et al., 2009), *SPT* (Heisler et al., 2001), and *CUC2* (Kamiuchi et al., 2014). Other genes within this list are interesting candidates for future functional studies. This includes members of the *REM* family of transcriptional regulators (Swaminathan et al., 2008; Romanel et al., 2009), several auxin synthesis or signaling-related genes such as *LIKE AUXIN RESISTANT 1* (*LAX1)* (AT5G01240) (Bennett et al., 1996) and *YUCCA4* (*YUC4)* (AT5G11320) (Cheng et al., 2006), as well as transcription factors regulating other developmental processes such as *MATERNAL EFFECT EMBRYO ARREST 3* (*MEE3)* (AT2G21650) (Pagnussat et al., 2005) and *GLABROUS 3* (AT5G41315) (Payne et al., 2000).

It is important to note that the 48 DEGs that were identified in both the YFP+/- and all-sorted/non-sorted comparisons (Fig. 2c) should not be discounted as potential medial domain regulators. These genes may be both preferentially expressed in the YFP-positive cell population as well as induced in response to the protoplasting procedure (Table S2). Indeed, some of these genes, including the transcription factors *HECATE3* and *BR-ENHANCED EXPRESSION1* (*BEE1*), have been reported to be preferentially expressed in medial domain-derived tissues and to function in gynoecium development (Gremski et al., 2007), (Crawford and Yanofsky, 2011). However, we chose to use the "cleaned" set of 363 YFP+/- DEGs for downstream analyses in order to reduce the likelihood of the inclusion of genes whose expression was significantly altered by the protoplasting process.

REPRODUCTIVE MERISTEM family members are differentially expressed in the _SHP2_-expression domain

In order to look for enriched categories of transcription factors within the set of "cleaned" 363 YFP+/- DEGs (Fig. 2c), we used the online Transcription Factor Enrichment Calculator (Desai, Jigar, Dmitry Grinevich and Colleen Doherty). Members of the *ABI3/VP1* transcription factor family that includes the *REM* and *NGA* family TFs were found to be statistically over-represented (Table S7) (corrected p <9.97E-06). The *REMs* belong to the plant-specific B3 superfamily of transcription factors and expression of many REM family members is observed in meristematic tissues such as the inflorescence meristem, floral meristem and the CMM (Franco-Zorrilla et al., 2002; Swaminathan et al., 2008; Romanel et al., 2009; Wynn et al., 2011; Mantegazza et al., 2014a; Mantegazza et al., 2014b). The numerical designations used to describe the *REM* family members in this manuscript are taken from Romanel *et al.* (Romanel et al., 2009). In our study, six *REM* members were amongst the 363 statistically significant YFP+/- DEGs; five were found to have enriched expression in the YFP-positive sample, while one, *REM25* (*AT5G09780*), was ~4 fold less abundant in the YFP-positive sample. *REM13* (At3g46770) transcript level is enriched ~12 fold in the *pSHP2-YFP* expressing cells. *REM13* was previously predicted to be preferentially expressed in the inner integument, ovule primordia and medial domain based on transcriptomic data (Skinner and Gasser, 2009). We employed *in situ* hybridization to assay the expression pattern of the *REM13* transcript during gynoecial development (Figure 3). Using a *REM13* antisense probe, we detected signal in the medial portions of the gynoecium corresponding to the carpel margin meristem as early as stage 7. Expression was also observed in the initiating ovule primordia in stage 8 gynoecia and then continued to be detected in portions of the ovules at later developmental stages.

REM34/ATREM1 (At4g31610) (Franco-Zorrilla et al., 2002; Romanel et al., 2009), *REM 36* (At4g31620) (Mantegazza et al., 2014b), and *VERDANDI* (*VDD/REM20*) (Matias-Hernandez et al., 2010; Mantegazza et al., 2014b) also displayed enriched expression levels in the YFP-positive sample of ~8 fold, ~9 fold and ~6 fold, respectively. Published *in situ* hybridization patterns indicate enriched medial domain expression patterns for *REM34/ATREM1* and *VDD/REM20* (Franco-Zorrilla et al., 2002; Matias-Hernandez et al., 2010; Wynn et al., 2011). Additionally, expression of At5g60142, a previously unnamed member of the *REM* family, is enriched ~11 fold in the YFP-positive sample (Table S3). At5g60142 is an interesting

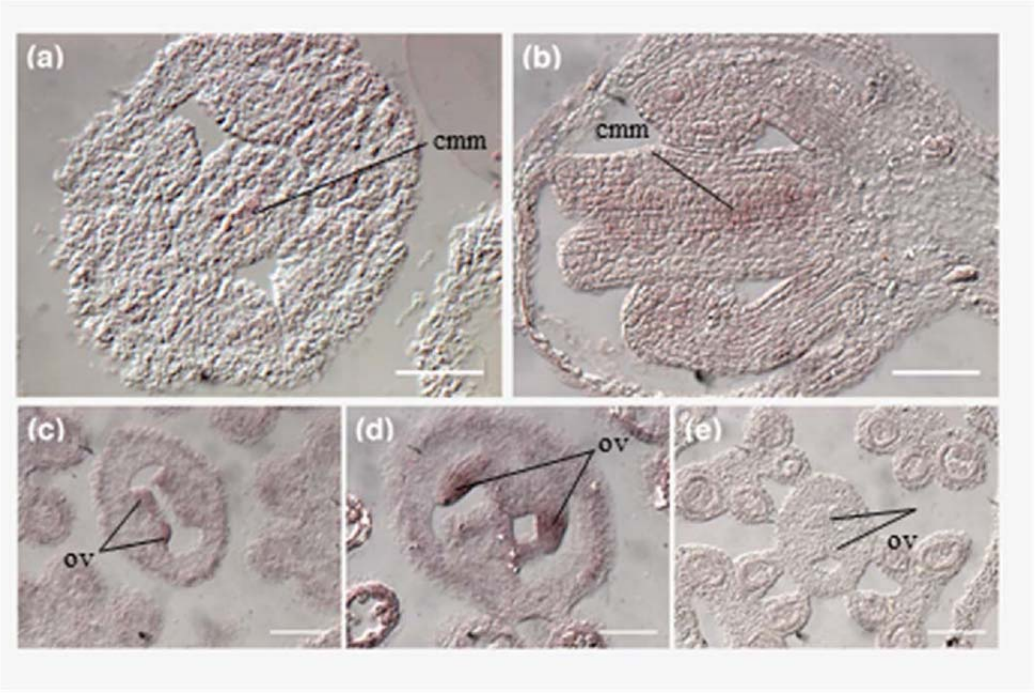

Figure 3. Candidate medial domain regulator *REM13* (At3g46770) is expressed within the medial gynoecial domain and developing ovules. Results from an RNA *in situ* hybridization with *REM13* probe. **(a-d)** antisense probe. **(e)** sense strand probe. **(a)** Hybridization signal is detected in the carpel margin meristem (adaxial portions of the medial gynoecial domain) in the stage 7 longitudinal section. **(b), (c)** and **(d)** In transverse gynoecial sections *REM13* expression is detected in the ovule primordia; stage 7 (panel b) stage 8 (panel c) and stage 9 (panel d) gynoecia. **(e)** A stage 8 section hybridized with a *REM13* sense strand probe. (ov) - ovules, (cmm) - carpel margin meristem. Scale bars for each panel represent 50 microns.

1

366 candidate for functional studies that is located on chromosome V in tandem to *REM11*
367 (At5g60140) and *REM12* (At5g60130) and shares a high degree of sequence similarity with

these two genes, as well as *REM13* (Romanel et al., 2009; Mantegazza et al., 2014b). We propose to designate At5g60142 as *REM46*.

Gene Set Enrichment Analysis

To gain global insights into underlying biological mechanisms of medial domain development and function, Gene Set Enrichment Analysis (GSEA) was performed for the 95 YFP-positive enriched DEGs in the medial domain. This analysis identified 147 GO terms that were statistically overrepresented ($p < 0.01$), including "gynoecium development" (GO:0048467) and "flower development" (GO:0009908), "response to gibberellin" (GO:0009739) and "auxin homeostasis" (GO:0010252) (Fig. 4 and Table S6). This GSEA analysis further suggests that the set of 95 genes enriched in the YFP-positive sample function as regulators of medial domain development.

In contrast, when performing GSEA with DEGs identified between the all-sorted/non-sorted samples, a different set of 304 overrepresented GO terms were identified, including "response to stress" (GO:0006950) and "response to wounding" (GO:0009611), suggesting that many of the genes identified as differentially expressed between the all-sorted/non-sorted samples reflect stress-induced changes in gene expression during protoplast/FACS-sorting.

The transcriptomic signature of the *SHP2*-expressing cell population shares commonalities with transcriptional signatures of other meristematic samples

In order to gain insight into the characteristics of the 363 YFP+/- DEGs identified from the *SHP2* expression domain, we compared the expression profile of this set of genes across several different tissues. Using Spearman rank correlation analysis, we compared our dataset to existing Arabidopsis RNA-seq transcriptomic datasets from whole flowers (Mizzotti et al., 2014), aerial seedlings tissues (GEO accession: GSE54125), as well as from Laser Capture Microdissected (LCM) inflorescence meristems, floral meristems and stage-3 flowers (Mantegazza et al., 2014a). In the sample-wise hierarchical clustering (Fig. 5a), the transcriptomic profiles from the *SHP2*-expressing (YFP-positive) sample clustered more closely with the meristematic samples, while the YFP-negative and all-sorted samples clustered more closely with the whole-flower and whole-seedling samples. This suggests that

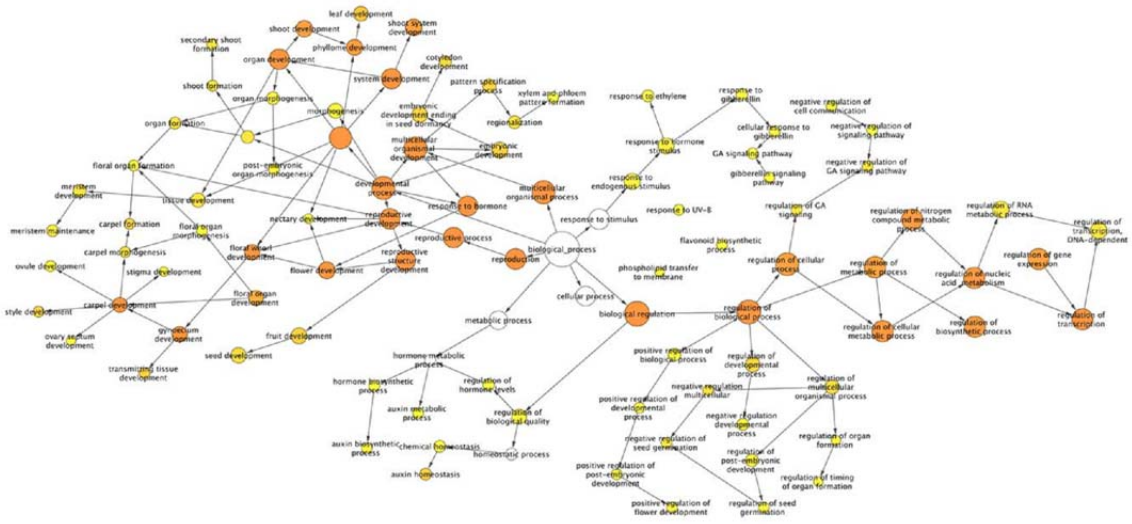

Figure 4. GO term overrepresentation of *SHP2*-domain enriched genes suggests a role for this set of genes in floral, gynoecial and ovule development. BiNGO/Cytoscape representation of overrepresented GO terms from the 95 YFP+/- DEGs displaying enriched expression in the YFP-positive samples. Edges represent the parent/child relationships of the GO terms (Ashburner et al., 2000), while color of the nodes indicates the degree of statistical significance (p <0.01) as reported by BiNGO (Maere et al., 2005). To unclutter the figure, given the large number of significant GO terms, selected nodes and edges have been removed from this graphical representation.

1

the expression signature of the YFP-positive sample is more similar to that of the floral and inflorescence meristems and young flowers, than it is to whole flowers or young vegetative seedlings (Fig. 5a).

Further supporting the similarity of the *SHP2*-expressing domain to other meristematic samples, the expression levels of *GA20OX1* (AT4G25420) and *GA20OX2* (AT5G51810) were both significantly depleted in the YFP-positive sample, relative to the YFP-negative sample

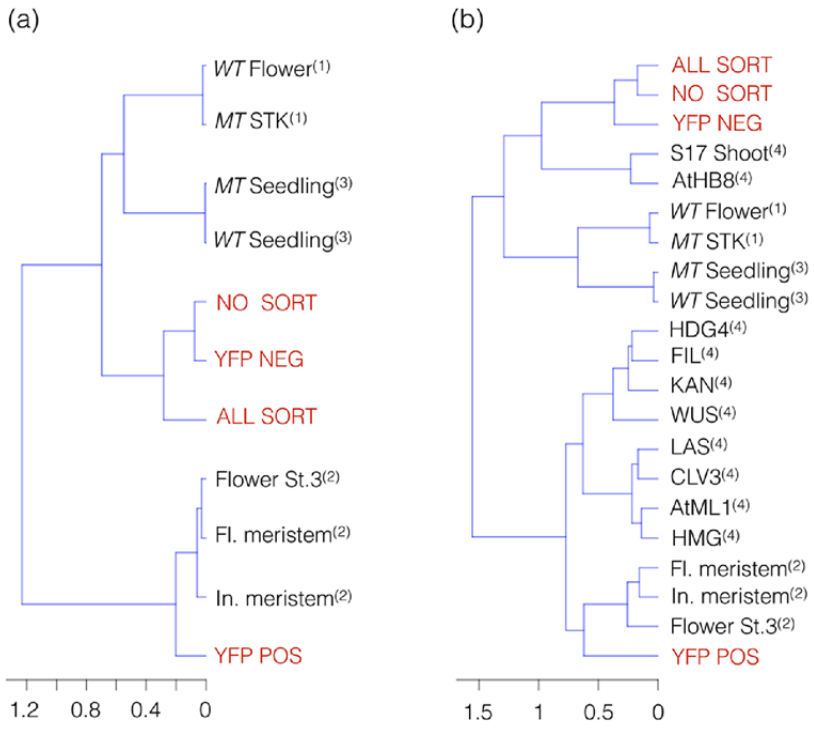

Figure 5. The transcriptomic signature of the *SHP2*-expressing domain is more similar to the transcriptomes of other meristematic samples than it is to whole flower. **(a)** Dendrogram based on hierarchical clustering using the Spearman rank correlation using RNA-seq (RPKM) expression values from flowers and other tissues. **(b)** Comparison of RNA-seq and affymetrix ATH1 arrays samples including transcriptomic data from whole flower, shoot apical meristem and seedling.
WT = wild type, MT= mutants. Data from Mizzotti *et al.* (Mizzotti et al., 2014)[(1)], Mantegazza *et al.* (Mantegazza et al., 2014b)[(2)], GEO accession: GSE54125[(3)] and Yadav *et al.* (Yadav et al., 2009; Yadav et al., 2014)[(4)] were used for comparison. Samples corresponding to this study are color coded red in both dendrograms.

406 (Table S10). *GA20OX1* and *GA20OX2* encode key biosynthetic enzymes of the plant hormone
407 gibberillic acid (GA) (Phillips et al., 1995). Levels of expression of *GA20OX1* and *GA20OX2*
408 are low in the shoot apical meristem (SAM) relative to expression in the juxtaposed young

organ primordia and high levels of GA synthesis interfere with the maintenance of meristematic fate in the SAM (Hay et al., 2002; Jasinski et al., 2005). These data suggest that low levels of GA may also be associated with the meristematic nature of the carpel margin meristem. Although not discussed here, expression values of genes annotated with a role in ethylene signaling are found in Table S10.

We additionally compared the medial domain transcriptional signature to datasets generated with the Affymetrix ATH1 array allowing comparisons to transcriptomic signatures of a variety of cell types including vascular and meristematic cell types from the Arabidopsis SAM isolated via FACS (Yadav et al., 2009; Yadav et al., 2014). When these additional samples are included, the hierarchical clustering dendrogram (Fig. 5b) shows the YFP-positive sample is more similar to the SAM cell-types, rather than to the vascular procambium (AtHB8) and phloem cell types (S17). This again suggests the meristematic character of the YFP-positive sample (Fig. 5b). One should be cautious, however, to interpret the results of this (or any) cross-platform (array/RNA-seq) comparison until validated cross-platform comparisons methods are available. To the best of our knowledge, there is no clear consensus in the literature of a standard cross-platform comparison practice (Guida et al., 2011), (Bradford et al., 2010), (Mudge et al., 2008), (Nookaew et al., 2012). Indeed, many researchers have used both platforms (array/RNA-seq) in the same experiment comparing final results rather than finding a way to directly compare the two technologies (Wang et al., 2014), (Xu et al., 2013), (Zhao et al., 2014), (Nookaew et al., 2012), (Marioni et al., 2008). Here, we employ a Spearman rank correlation as it is less sensitive than the Pearson correlation to strong outliers, makes no assumptions about data distribution, and does not inflate type I error rates. This approach fits well with the data in this work as samples do not cluster based on technology platforms but rather cluster based on the apparent cell-type similarities of gene RPKM (Reads Per Kilobase of transcript per Million mapped reads) expression levels.

Transcriptomic analysis of the *SHP2* expression domain complements existing medial domain and CMM data sets

Wynn *et al*. previously carried out a related transcriptomic study and identified many genes that were shown via *in situ* hybridization to be preferentially expressed in the developing

medial domain of the wild-type gynoecium (Wynn et al., 2011). When comparing the 95 enriched DEGs from our RNA-seq experiment (Table S2 and Figure S3) with a set of 210 medial domain enriched genes from Wynn *et al*., 23 genes were found in common (Table III). The 24% overlap of these two gene sets is significantly higher than expected by chance (hypergeometric test; $p = 3.15 \times 10^{-30}$) (Halbritter et al., 2012). Members of the *REM, HECATE* and *NGA* gene families, as well as several auxin-homeostasis-related genes were among the set of 23 genes identified in both experiments (Table III).

Reyes-Olalde *et al.* recently performed a comprehensive literature survey of genes that function during CMM development (Reyes-Olalde et al., 2013). They reported 86 protein-coding genes corresponding to transcription factors, hormonal pathways, transcriptional co-regulators, and others of widely diverse functions. While all 86 are expressed in our dataset, fifteen of these CMM developmental regulators are found within the set of 363 YFP-positive DEGs (hypergeometric test; $p = 3.3 \times 10^{-13}$) (Halbritter et al., 2012) (Fig. 6). The expression profiles of the 86 genes reported by Reyes-Olalde *et al*. within the medial domain-enriched dataset from this work, as well as within data from floral meristem enriched samples (Mantegazza et al., 2014a), is displayed in a heatmap in Figure 6 (RPKM values can be found in Table S9).

Transcript isoforms in the *Arabidopsis* medial domain

One utility of transcriptome analysis through RNA-seq is the identification of novel alternative spliced transcripts, alternative transcription start sites (TSS), and instances of isoform switching (Sims et al., 2014). To further characterize the transcriptome of the *SHP2*-expression domain at the isoform level, we first selected isoforms that showed a significant ($\alpha < 0.01$) change in their expression between YFP+/- samples using Cufflinks/Cuffdiff. For this analysis we did not apply a fold magnitude cutoff, thus capturing all isoforms with $\alpha < 0.01$. To avoid transcripts that were affected by the cell-sorting procedure, we removed all isoforms that showed a significant ($\alpha < 0.01$) expression level change between all-sorted/non-sorted samples. This resulted in 4555 YFP+/- differentially expressed isoforms (Table S8). Within this set of isoforms differentially expressed between the YFP+/- samples, we sought to

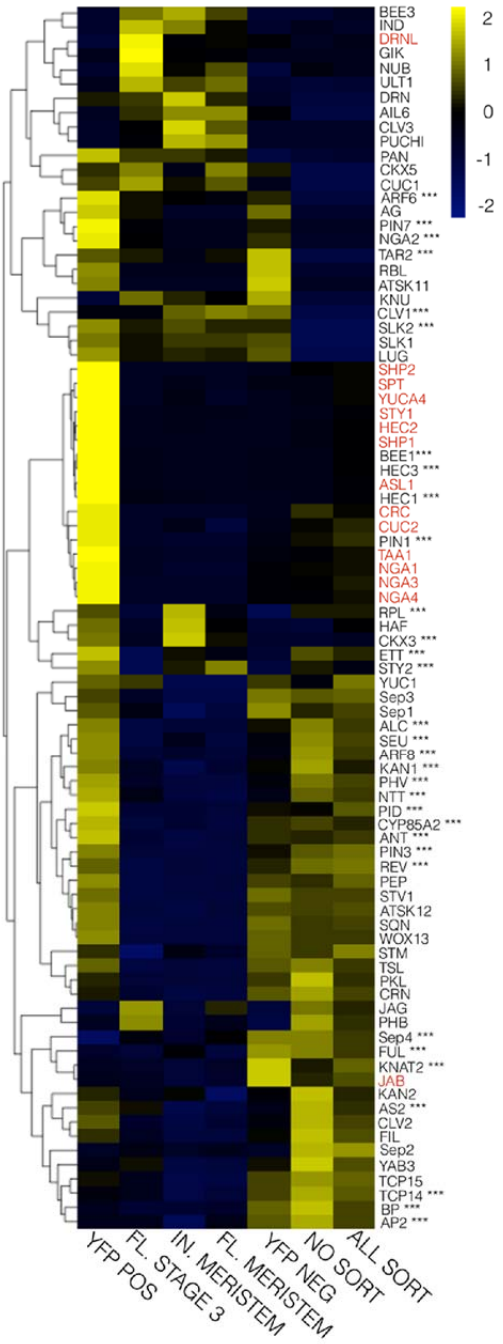

Figure 6. Heatmap representation of the expression profiles of previously identified regulators of Carpel Margin Meristem development. Expression profiles in Reads Per Kilobase of transcript per Million mapped reads (RPKM) of the 86 genes reported by Reyes-Olalde *et al. (Reyes-Olalde et al., 2013)* with functional role during CMM development. Transcriptional profiles from this study (YFP POS = YFP-positive, YFP NEG = YFP-negative, ALL SORT = all-sorted, and NO SORT = no-sorted) as well as Mantegazza *et al.* (Mantegazza et al., 2014a) corresponding to flower stage 3 (FL.STAGE 3), floral meristem (FL.MERISTEM) and inflorescence meristem (IN.MERISTEM) are included. Genes color-coded in red are those identified as DEGs between YFP-positive and YFP-negative samples (FC >4 and FDR <0.001) while genes that displayed a statistically significant expression level (FDR <0.01) between YFP-positive and YFP-negative (regardless of their fold change) are indicated with ***.

highlight multi-isoform genes that showed major changes in the relative frequency of individual isoforms between the YFP-positive and YFP-negative samples. To this end, we estimated the relative frequency of each isoform as a percentage of the total expression for the

gene. Among the 4555 significantly differentially expressed isoforms, only 52 isoforms from multi-isoform genes displayed changes of 20% or more in their relative frequency. The major isoform (most highly expressed isoform) differed between YFP+/- samples for only 15 genes (Table II). Remarkably, the transcriptional co-regulator SEU (At1g43850), previously implicated in medial domain development (Franks et al., 2002), (Azhakanandam et al., 2008), showed a significant increase of isoform At1g43850.1 in the YFP-positive samples, while its second isoform At1g43850.2 did not significantly change between samples. As a result, isoform 1 was the major (predominant) isoform in YFP-positive cells, and isoform 2 was the major (predominant) isoform in the other samples. The functional significance, if any, of this isoform switching is currently unknown.

The regulation of gene expression through alternative promoter usage or use of alternative TSS is frequently observed in multicellular organisms (Ayoubi and Van De Ven, 1996). Using the same pipeline and criteria we employed to select differentially expressed isoforms in the YFP+/- samples, we identified 93 isoforms that were differentially expressed as a result of the use of alternative promoter/transcriptional start sites (Table S8). Interestingly, one such promoter/transcriptional start site switch was found for the *REVERSIBLY GLYCOSYLATED POLYPEPTIDE 5* (*RGP5*) gene (isoform). Members of the *RGP* family (*RGP1* and *RGP2*) involved in sugar metabolism are expressed in other Arabidopsis meristematic tissues, such as the root tip and the apical meristem of young seedlings (Drakakaki et al., 2006). In our work, the transcript level of *RGP5* isoform 2 (At5g16510.2) in the YFP positive sample is 61% higher relative to the level of this isoform in the YFP-negative sample, while the level of isoform 1 (At5g16510.1) is 75% lower (Fig. 7a and Table S8).

Auxin homeostasis and the development of the gynoecial medial domain

Auxins are a class of plant hormones that regulate growth and development (Woodward and Bartel, 2005; Sauer et al., 2013). The most common plant auxin is Indole-3-Acetic Acid (IAA). The regulation of auxin homeostasis (including synthesis, response, transport, inactivation and degradation) plays an essential role in patterning the gynoecium and other lateral organs (Woodward and Bartel, 2005; Sehra and Franks, 2015). The role of auxin during the development of the medial and lateral domains of the gynoecium is less clearly

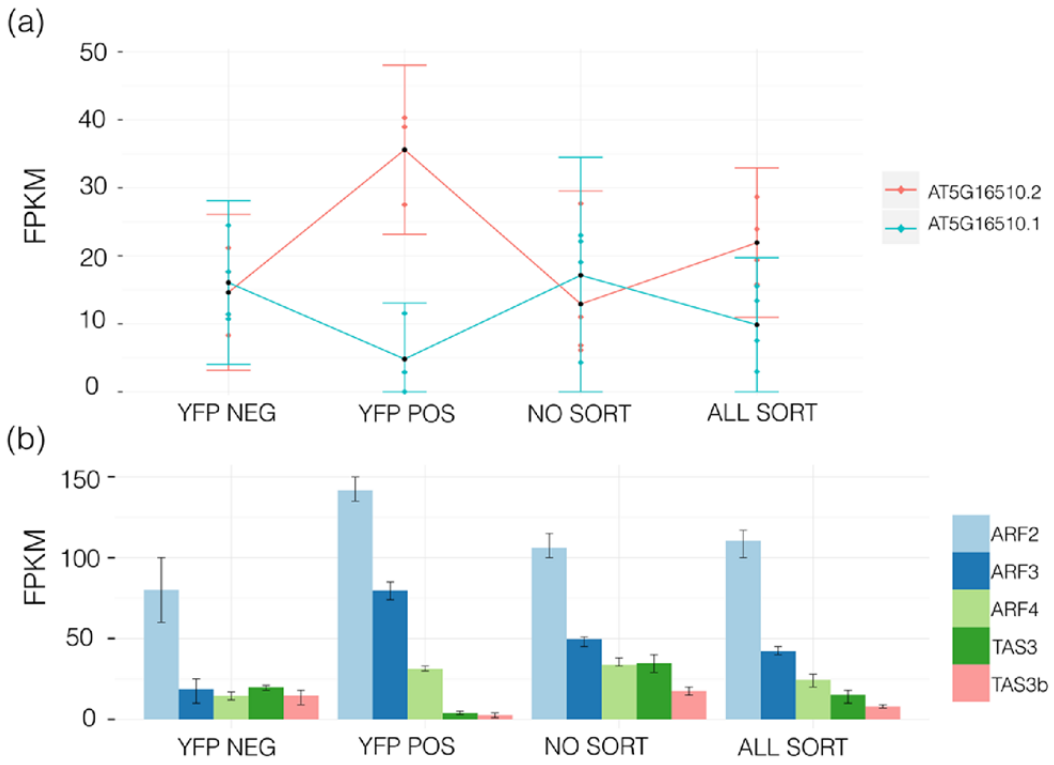

Figure 7. Differential expression of *REVERSIBLY GLYCOSYLATED POLYPEPTIDE 5* (*RGP5)* isoforms as well as *TRANS-ACTING siRNA3* (*TAS3*) and *AUXIN RESPONSE FACTOR* genes. **(a)** Promoter/transcriptional start site switch found for the *RGP5* gene (At5g16510). The isoform 2 (At5g16510.2) increases its expression in the YFP-positive domain while isoform 1 (At5g16510.1) of the same gene decreases its expression in the same domain. **(b)** Expression of the *AUXIN RESPONSE FACTORS* (*ARFs*) (*ARF2*, *ARF3*, *ARF4*) and *TAS3* transcripts. Expression levels of *ARF2*, *ARF3*, *ARF4* are significantly enriched in the YFP-positive sample at FDR <0.01. Expression levels of the *TRANS ACTING siRNA3* (*TAS3*) genes At5g49615 and At3g17185, that negatively regulate the expression of *ARF2*, *ARF3*, and *ARF4* expression (Williams et al., 2005), are significantly reduced (FDR <0.01) in the YFP-positive sample.

1

505 defined, however recent studies suggest that auxin homeostasis mechanisms are likely to be
506 distinct in medial and lateral domains (Larsson et al., 2014; Moubayidin and Ostergaard,
507 2014; Sehra and Franks, 2015).

To better analyze auxin homeostatic mechanisms during medial domain development, we examined the expression of 127 genes with an annotated function in auxin homeostasis. Of these 127 genes, 80 were expressed in our dataset and 60 were differentially expressed at a FDR of < 0.01 in the YFP +/- comparison, without applying a fold enrichment filter (Table S10). The expression levels of *TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS 1* (*TAA1)* and *YUC4,* two genes encoding proteins in the auxin synthetic pathway, were strongly enriched (> 4 fold) in the YFP positive samples as was predicted from previously published expression patterns indicating enriched expression within the medial portions of the gynoecium (Zhao et al., 2001; Cheng et al., 2006; Stepanova et al., 2008; Tao et al., 2008; Trigueros et al., 2009; Martinez-Fernandez et al., 2014). Within the *PINFORMED* (*PIN*) family of polar auxin transporters, the expression levels of *PIN1*, *PIN3* and *PIN7* were significantly enriched in the YFP-positive sample (Table S10). This is consistent with the reported expression patterns at the protein level of these *PIN* transporters within the medial domain of the gynoecium (Benkova et al., 2003; Blilou et al., 2005; Larsson et al., 2014; Moubayidin and Ostergaard, 2014).

Auxin induces gene expression through a family of transcription factors called *AUXIN RESPONSE FACTORS* (*ARFs*) (Woodward and Bartel, 2005). At a fold change level of 1.5 fold and FDR of < 0.01, ten *ARFs* were enriched in the YFP-positive sample (*ARF1, ARF2, ARF3/ETTIN*, *ARF4*, *ARF5*, *ARF6*, *ARF7*, *ARF8*, *ARF16* and *ARF18*), while no ARFs were identified as depleted in the YFP-positive sample (Table S10). Our data suggests these *ARF* family members may be preferentially expressed in the medial domain and play a role during development of this meristematic tissue. Previous studies have documented gynoecial developmental defects in *arf3/ettin* mutants (Sessions and Zambryski, 1995) as well as *arf6 arf8* (Nagpal et al., 2005; Wu et al., 2006) double mutants. Interestingly, the levels of the precursor transcripts for two *TRANS-ACTING SIRNA3* (*TAS3*) genes (At5g49615 and At3g17185) were significantly reduced (FDR <0.01) in the YFP-positive sample (Fig. 7b and Table S8). The trans-acting siRNAs that are encoded by the *TAS3* genes negatively regulate the levels of *ARF2*, *ARF3*, and *ARF4* transcripts (Williams et al., 2005). Thus the enrichment of *ARF2 ARF3* and *ARF4* transcript levels in the *SHP2*-expression domain may in part be due

to a reduction in the level of expression of the *TAS3*-encoded tasi-RNAs in the medial domain.

SQUAMOSA PROMOTER BINDING PROTEIN-LIKE3 (SPL3) and the cis-NAT antisense gene At2g33815

The *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) genes function in the regulation of the transition from juvenile to adult growth phases, and regulation of shoot regenerative capacity (Wu and Poethig, 2006; Wang et al., 2008; Wang et al., 2009; Zhang et al., 2015). In our study, the expression of *SPL3* (At2g33810) was more than four-fold lower in the YFP-positive sample relative to the YFP-negative sample. *SPL3* encodes a DNA-binding protein directly regulating *APETALA1* (At1g69120), a key regulator of floral-meristem-identity specification (Leal Valentim et al., 2015). Interestingly, the expression of the cis-NAT antisense gene At2g33815, complementary to portions of the *SPL3* gene, was also significantly reduced ~4.5 fold in the YFP-positive sample (Cufflinks data in Table S3 and Table S8). This is perhaps in contrast to the expected pattern of expression, where the expression levels of the targeted *SPL3* transcript might be expected to go up as the levels of cis-NAT antisense At2g33815 go down. The expression of another regulator of *SPL3* activity, the *miRNA157D* (At1g48742), was also significantly reduced in the YFP-positive samples. The *miRNA157D* reduces translation of the *SPL3* transcript by acting through a *miRNA156/157*-responsive element in the *SPL3* 3'UTR (Gandikota et al., 2007; Wang et al., 2009). These data suggests that the *miRNA156/157/SPL* module may act during medial domain development and may be regulated by the cis-NAT antisense gene At2g33815. A complete list of differentially expressed natural-antisense, transposable-element and other non-protein coding transcripts identified as differentially expressed by Cufflinks, DESeq2 and edgeR is found in Table S8.

Protoplasting-induced stress genes

While the predominant focus of this work was to perform transcriptomic analysis in medial-domain-enriched cells (YFP+/-), transcripts induced by the protoplasting and sorting process (all-sorted/non-sorted) were also identified (Methods). To facilitate the visualization of all samples, we generated an interactive 6-way Venn diagram using the web-based tool

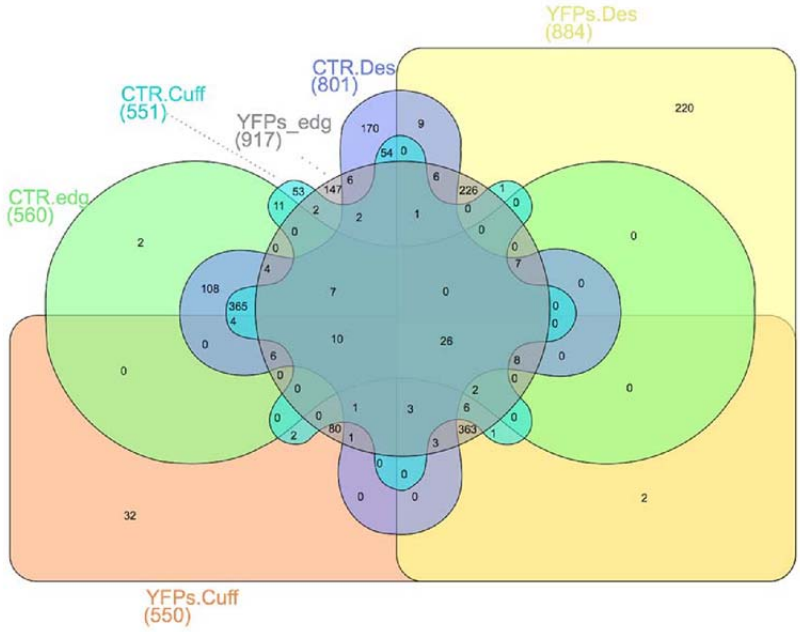

Figure 8. Six-way venn diagram image showing detailed overlap from all the differentially expressed gene (DEGs) datasets. The total number of DEGs under each condition and for each program are indicated in parentheses. CTR= DEGs between all-sorted/non-sorted and YFPs= DEGs between YFP+/-. Cuff= Cufflinks, edg=edgeR, Des=DESeq2. The interactive tool can be accessed online using the 'InteractiVenn' web-tool (InteractiVenn) and uploading Additional file Table S11.

1

570 'InteractiVenn' (Heberle et al., 2015). By uploading the Additional file S11 to InteractiVenn
571 (InteractiVenn), mousing over, and clicking on the numbers in the Venn diagram, researchers
572 will find gene ID from DEGs between YFP+/- and all-sorted/non-sorted samples (3 programs

and two comparisons). As expected, when comparing such different types of samples (all-sorted/non-sorted and YFP+/-), few DEGs (26) overlapped across the 6 samples (Fig. 8). The lack of overlap of DEGs across the entire experiment indicates that the YFP+/- DEGs reported here are not a result of protoplasting-stress-induced processes.

When comparing the protoplasting-induced gene set from this work (all-sorted/non-sorted DEGs) with those induced due to FACS-sorting methodology in shoot apical meristem by Yadav *et al.* (Yadav et al., 2009) and in roots as reported by Birnbaum *et al*. (Birnbaum et al., 2005) few DEGs were found in common (seven across all datasets) (Figure S4) indicating that different tissues and/or different protoplasting techniques generate different sets of protoplast-induced gene-expression changes. Thus, appropriate controls should be included to control for condition-specific protoplasting-induced gene-expression changes.

CONCLUSIONS

Despite the importance of the gynoecial medial domain in ovule development, no domain-specific transcriptome has been previously reported, mainly, due to the difficulty of isolating the meristematic cells from which ovules are derived. In this work, we developed a novel FACS-based system using the *SHP2*-expression-domain-specific using a GAL4/pUAS-based two-component system that, when combined with flower synchronization and flow cytometry, allowed for the efficient isolation of medial-domain cells expressing *SHP2*. The quality and quantity of biological samples that can be recovered with our system enables cell-type and strand-specific RNA-seq transcriptomic analysis and opens up possibilities for small RNAs, metabolomic and proteomic analyses (Petersson et al., 2009; Breakfield et al., 2012; Petricka et al., 2012; Li et al., 2013; Moussaieff et al., 2013). This approach, coupled with high-throughput RNA-sequencing, has yielded a unique and novel snapshot of the gynoecial medial domain transcriptome and a set of candidate regulators of medial domain development for future functional analysis.

MATERIALS AND METHODS

Construction of *pSHP2-GAL4//pUAS-3xYpet* dual construct lines

The *SHP2* promoter fragment was amplified from Columbia wild type genomic DNA using the primers proSHP2gwF1 (5'CACCATCTCCAACGCATTGTTACG3') and proSHP2gwR1 (5'CATTTCTATAAGCCCTAGCTGAAG3'). This fragment contains the sequences from -2170 to +1 relative to the *SHP2* ATG and includes the 5'UTR, the first intron and the first Met codon of *SHP2*. This promoter previously was shown to mimic the endogenous *SHP2* expression pattern (Colombo et al., 2010). This genomic fragment was cloned into the pENTR/D-TOPO vector (Invitrogen) to create plasmid LJ001, and then shuttled via gateway LR reaction (Invitrogen) into the destination vector JMA859 (i.e. pEarleygate303-GAL4) to create plasmid AAS003. Transgenic Arabidopsis lines were created by *Agrobacterium*-mediated transformation of the AAS003 plasmid into the S. No. 1880 seed stock that contained the pGWB2-pUAS-3xYpet responder construct (see below) generating the pSHP2-GAL4; pUAS-3xYpet dual construct line (S. No. 1896), referred to as pSHP2-YFP. The pSHP2-YFP plants were crossed to the *ap1 cal1* Wellmer floral induction system (Wellmer et al., 2006) as described below.

JMA859 (pEarleygate303-GAL4) is a modified pEarleygate303 (Earley et al., 2006) plasmid in which the reporter was replaced by the coding sequences from the GAL4 yeast transcriptional activator. To achieve this, pEarleygate303 was cut with NcoI (New England Biolabs) and SpeI (New England Biolabs). Then fusion PCR was used to create the insert that fused the *GAL4* sequences to the deleted portions of pEarlygate303. This required three PCR reactions: 1st PCR with primers pEarl303NcoIFor (5'TGGCCAATATGGACAACTTCT3') and pEarl303Rev_GAL4 (tale) (5'ATGGAGGACAGGAGCTTCATACACAGATCTTCTTCAGAGA 3'); 2nd PCR with primers GAL4F_pEarl303(tale) (5'TCTCTGAAGAAGATCTGTGTATGAAGCTCCTGTCCTCCAT3') and GAL4Rev_SpeI (5' CCGGACTAGTCTACCCACCGTACTCGTCAA3'), and then a fusion PCR joining these two fragments using the external primers to amplify. The product of the fusion PCR was double-digested with NcoI/SpeI and ligated into NcoI/SpeI-cut pEarleygate303.

JMA382 (pUAS-pGWB2) was created from pGWB2 (Nakagawa et al., 2007) by replacing the p35S sequences in pGWB2 with *pUAS* sequences (HindII/XbaI sites used). A Gateway LR

reaction was then used to move the *3xYpet* cassette from JMA710 (pENTR/D-TOPO-3xYpet) into JMA382, creating vector JMA721 (i.e. pGWB2-pUAS-3xYpet). Homozygous single insertion-site transgenic lines harboring JMA721 were then generated (S.No 1880).

Plant material

In a wild-type inflorescence, cells expressing *SHP2* represent a small percentage of the total cells. Additionally, wild-type inflorescence contains a full range of developmental series of floral stages. The Wellmer floral synchronization system (Wellmer et al., 2006) was used to maximize the amount of gynoecial tissue from floral stages 6-8 (Smyth et al., 1990). The Wellmer group kindly provided *pAP1-AP1::GR*; *ap1; cal* seeds (KanR in Ler background - S.No. 1927). The *pSHP2-GAL4; pUAS-3xYpet* dual construct plants (S.No. 1896) were crossed to *pAP1-AP1::GR*; *ap1; cal*. Lines homozygous for *er*, *ap1*, *cal* and the transgenes were selected in F2 and F3 generations (generating S. No. 2060). Because of the mixed ecotype cross (Col and Ler), lines that were *erecta* homozygous mutant and gave consistent YFP expression pattern and consistent inducibility of the *AP1-GR* activity were selected before the generation of protoplasts. Plants were grown under constant light and temperature at 22 °C to minimize circadian transcriptional fluctuations. To induce flowering in the transgenic plants, 20 μm of the synthetic steroid hormone dexamethasone (DEX) (Sigma, USA) in 0.015 % silwet was applied directly (spray application) ~30 days after planting (Wellmer et al., 2006). Inflorescences were collected for protoplast generation ~120 h after DEX-induced floral synchronization. When collecting samples for protoplast preparation, 5-6 inflorescence heads were fixed for chloral hydrate clearing and DIC microscopy to determine the developmental stages of the flowers of the inflorescence samples. Additionally, before protoplasting, whole inflorescences were also collected and frozen immediately in liquid nitrogen for analysis of the transcriptional starting state of the non-protoplasted tissue (non-sorted samples, see Experimental design).

Experimental design

Material for RNA samples was gathered from batches of plants grown at one-week intervals to generate biological replicates (material from each week was considered as a biological replicate). To reduce variability between bioreplicates due to environmental heterogeneity

within the growth chamber, each bioreplicate was drawn from a pool that contained plants grown within three different chamber positions. Four biological replicates of each of four tissue samples (YFP-positive, YFP-negative, all-sorted, and non-sorted) were collected (16 samples total). Whole inflorescences were collected for non-sorted samples and immediately frozen in liquid nitrogen before RNA isolation (i.e. these samples were not subjected to protoplasting nor FACS-sorting). The all-sorted samples represented the total population of protoplasts that come off the FACS machine after debris and broken cells are removed based on sorting gates (Figure S1). The YFP-positive and YFP-negative protoplast populations are processed equivalently to the all-sorted samples except that a final FACS-sorting gate is used to divide the all-sorted protoplasts into YFP-positive and YFP-negative samples (Figure S1). RNA was isolated from these three protoplast populations, as well as from entire non-protoplasted inflorescences ("non-sorted"). The YFP-positive, YFP-negative and all-sorted samples were prepared and collected as described below (Protoplast recovery and cell sorting).

Protoplast recovery and cell sorting

Protoplasts from the S. No. 2060 plants were generated according to the protocol of Birnbaum *et al*. (Birnbaum et al., 2005), with adaptations for inflorescence plant material. Inflorescences (~200) were hand-collected with forceps and/or scissors and chopped with a "Personna double edge prep blade" (American Safety razor company; 74-002) within a 15 min period. Cell-wall polysaccharides were digested by immersing the chopped plant material in 10 ml of filter-sterilized solution B in a 50 ml falcon tube. Solution B (prepared according to Birnbaum *et al*.) is prepared from Solution A (10 mM KCl, 2 mM $MgCl_2$, 0.2M MES, 600 mM Mannitol) to which cell wall digesting enzymes were added [final concentrations of 1.5% Cellulase (Yakult, Japan), 1% Pectolyase (Yakult, Japan) and 1% Hemicellulase (Sigma, USA)]. This mixture is then dissolved by gently swirling, covered in foil, and warmed in a water bath at 55 °C for ten minutes to inactivate DNAses and proteases. After cooling to room temperature, $CaCl_2$ (2 mM final) and BSA (0.1% final) were added and the solution was filter-sterilized through a 25-micron filter.

After 1 h of incubation at room temperature with occasional gentle agitation, 10 ml of the protoplast-rich solution B was filtered through a 70-micron filter basket to a 50 ml falcon tube. A 10 ml rinse of solution A was applied directly to the material left in the 70-micron filter basket to rinse through any protoplasts left behind. Protoplasts were spun at 500 g, 10 °C for 10 min; the majority of the supernatant was removed by aspiration being careful not to disturb the protoplast pellet which is typically not tightly compacted. Protoplasts were resuspended in 25 ml of Solution A as a rinse step to remove cell-wall-digesting enzymes. Protoplasts were filtered again through a 50-micron filter mesh to a new tube adding 8 ml of solution A to again rinse through any protoplasts stuck in the filter. Protoplasts were then spun again at 500 g for 10 min. The majority of the supernatant was removed leaving 2 ml of the protoplasts in solution after the second centrifugation step. Propidium Iodide (5 micrograms/ml final) was added to the protoplasts (to allow separation of broken protoplasts) and a final filtering step though a 30-micron mesh filter (CellTrics, Partec) was carried out before loading onto the FACS machine.

Flow cytometry through FACS-sorting (Moflo XDP; Beckman Coulter Inc.) was used to isolate the YFP expressing cells from the total pool of cells. The FACS machine was equipped with a cooling device (set to 10 °C) and fitted with a 100-µm nozzle. Protoplasts were sorted at a rate of up to 10,000 events per second at a fluid pressure of 25 psi. Four sorting gates were set in an effort to collect the cleanest set of protoplasts and to eliminate debris and broken cells. A first gate based on size and granularity using side-scatter (SS) and forward-scatter (FS) parameters was used to select for intact protoplasts. Then a second gate was used to select for single cells and remove “doublets”. A third gate was used to select for cells that were negative for propidium iodide (PI) signal, as broken protoplasts and debris are preferentially stained by PI, which is excited by the 488 nm laser and emits at 617 nm. The total population of protoplasts that came off the FACS sorter machine after these gates constituted the all-sorted sample. In parallel, the YFP-positive protoplasts and YFP-negative protoplasts were separated into two collection tubes using the gates described above and one additional sorting gate based on the level of emission intensity in the green channel (529nm/28nm filter). Preliminary experiments with protoplasts that did not express the YFP transgene were used to set this gate and determine the levels of auto-fluorescence of the

protoplasts. Protoplasts were collected directly into 14 ml tubes containing 4 ml of Trizol (Invitrogen/Life Technologies) and occasionally agitated during the approximately 40 min of sorting required to collect the protoplasts. Trizol was the method of choice as it maintains a high level of RNA integrity during tissue homogenization while also disrupting and breaking down cells and cell components. In order to minimize artifactual changes to transcript levels, the entire process of cell wall digestion, protoplast generation and FACS-sorting was kept under three hours. This procedure typically yielded between 300,000 and 500,000 YFP-positive protoplasts. These YFP-positive protoplasts typically represented approximately 0.5% of the total FACS sorting events. On average from four sorting trials representing four biological replicates, the number of cells collected and processed for each sample was: 575K for the YFP-positive, 1000K for the YFP negative and 493K for all-sorted samples.

RNA extraction and quantitative RT-PCR

Total RNA was extracted from sorted protoplasts collected in Trizol (keeping a 3:1 ratio of Trizol to sorted cells) and by modifying the Plant RNeasy Mini Kit, Qiagen protocol, as follow: collected cells in Trizol (4 ml total) were vortexed for 5 min at room temperature (RT) and 1 ml of chloroform (Sigma) was then added. The solution was vortexed again for 1 min at RT and centrifuged at 4,000 rpm for 10 min at 4 °C to separate phases; RNA from the aqueous phase (top layer) was carefully sucked up and mixed with 700 µl of Qiagen RLT buffer (Plant RNeasy Mini Kit, Qiagen) and 7µl of B-Mercaptoethanol (Sigma). 500 µl of 100% ethanol was added, solution was then transferred to a Qiagen MinElute column (Plant RNeasy Mini Kit, Qiagen) and spun in a 2 ml microfuge tube for 15 sec at ~10,000 rpm. 500 µl of RPE (Plant RNeasy Mini Kit, Qiagen) was added to the spin column, spun for 15 sec at ~10,000. 750 µl of 80% ethanol was added to the MinElute column and spun at ~10,000 rpm for 15 sec (twice) to ensure removal of all guanidine salts that may inhibit downstream applications. A final 5 min spin at top speed with the cap off was performed to remove trace amounts of ethanol. Total RNA was then eluted with 10 µl of RNAse-free water. A second elution was performed with another 10 ul of RNAse-free water. It is worth noting that one biological replicate (4^{th} biological replicate) from the YFP-positive protoplasts was lost at this point, leaving only 3 biological replicates for this tissue sample and yielding a total number of 15 samples sequenced in two lanes and used for the experiment.

Prior to high-throughput sequencing, quantitative RT-PCR (qRT-PCR) was conducted on YFP-positive and YFP-negative samples using the $2^{-\Delta\Delta CT}$ method as suggested by Schmittgen and Livak (Schmittgen and Livak, 2008) to assess relative gene expression of specific medial domain markers, *SHATTERPROOF2* and *NGATHA1*. Total isolated RNA was quantified using fluorometric quantitation (Qubit RNA Assay Kit, Life Technologies, Inc.) for both YFP-positive and YFP-negative samples [~100 ng]. SuperScript III First-Strand Synthesis System (Invitrogen/Life Technologies) was used to generate cDNA (cDNA diluted 1:4 prior qRT-PCR analysis) from total RNA. qRT-PCR experiment assay was performed (Thermal Cyclers from Applied Biosystems) using a SYBR green mix (QuantiTect SYBR Green PCR Kits, Qiagen). Three biological replicates of the YFP-positive and YFP-negative samples were included and each biological replicate was assayed in triplicate. The expression levels of the *ADENINE PHOSPHORIBOSYL TRANSFERASE1* (*APT1*) (At1g27450) gene was used for normalization.

Barplots

Barplots graphs were constructed using the 'R' package bear (R package bear) and plyr (CRAN - plyr package) to calculate mean, standard error and confidence intervals and ggplot2 (CRAN - ggplot2 package) to generate the plots.

Library preparation and mRNA sequencing

Total RNA isolated was quantified using fluorometric quantitation (Qubit RNA Assay Kit, Life Technologies, Inc.) and RNA quality was assessed using Agilent 2100 Bioanalyzer (Agilent). The RNA integrity number (RIN) for the 15 samples was higher than 7.3, which is above the Illumina threshold for library construction (> RIN 7). Strand-specific cDNA libraries were constructed from approximately 100 ng of total RNA using a NEB Ultra Directional Library Prep Kit for Illumina (New England Biolabs). The average size of the cDNA fragments was ~ 250 bp. The 15 bar-coded libraries were pooled and single-end sequencing was performed in a HiSeq 2500 Illumina (Illumina, Inc.) with 'HiSeq SR Cluster Kit v4' for the flow-cell and 'HiSeq SBS v4' for sequencing reagents. cDNA libraries were sequenced in 125-cycle plus 7-cycle for multiplexed samples. Sequencing was performed in

two lanes of a flow-cell; all 15 libraries were sequenced twice and the results from the two independent lanes were analyzed as technical replicates. As no lane-specific effects were observed during data analysis, the reads from each lane were pooled for analysis of DEGs (see Table counts and technical replicates).

Bioinformatics analysis

All bioinformatics analyses were performed on a server cluster with 128 GB (gigabytes) of RAM, 16 cores (CPUs) and Ubuntu Linux-Distribution 12.04 operating system using 'Simple Linux Utility Resource Management' (SLURM) queue management system at the Bioinformatics Research Center (BRC) at the North Carolina State University, Raleigh, NC, USA.

Read Processing

Quality control and preprocessing of metagenomic data was performed using FastQC software (Schmieder and Edwards, 2011). Adapters and low quality sequences were filtered out with Ea-Utils software (Lindgreen, 2012). Reads with phred-like quality score (Q-score) > 30 and read length > 50-bp were kept and aligned against the TAIR10 Arabidopsis reference genome.

Sequence alignment to the Arabidopsis genome

Splice junction mapper TopHat2 (version 2.0.10) (Trapnell et al., 2009) was used to align filtered RNA-seq reads to the *Arabidopsis thaliana* TAIR10 genome (Ensembl annotation) downloaded from the iGenome database (iGenomes Database.). Default parameters for TopHat2 were used except for strand specificity (--library-type=fr-firststrand) to match to the first strand of cDNA synthesized (anti-sense to the mRNA) and maximal intron length (--I 2000), as it has been shown that the large majority of the known introns are smaller than the selected threshold (Li et al., 2013). To align reads solely and exclusively against TAIR10 annotated gene models, the arguments '--T' (transcriptome only) and '--no-novel-juncs' (no novel junction) were also included. Uniquely mapped reads were extracted from the TopHat2 output binary (BAM) file using samtools (Li et al., 2009) and selecting for the "NH:i:1" two-character string-tag. Only uniquely mapped reads were used for downstream analysis.

Table counts and technical replicates

The 'HTSeq: Analyzing high-throughput sequencing data with Python' software (Anders et al., 2015) was used with default parameters except for the 'stranded=reverse' mode to generate tables-counts for downstream differential expression analysis for the 'R' packages edgeR (Robinson et al., 2010) and DESeq2 (Love et al., 2014).

Using edgeR, we assessed the gene level variance versus log gene expression level among technical replicates (corresponding to two lanes in the flow-cell of the Illumina HiSeq 2500). A linear-dependent Poisson distribution was observed for technical replicates (Figure S5), in accordance with several studies (Robinson et al., 2010), (Marioni et al., 2008), (Anders and Huber, 2010). Thus, differential gene expression analysis was performed using pooled technical replicates.

Gene expression and differential gene expression

Gene expression and differential gene expression analysis was carried out using 'R' packages edgeR (Robinson et al., 2010) and DESeq2 (Love et al., 2014) and the Linux-based Cufflinks program (v2.2.1) (- G option) (Trapnell et al., 2012), for differentially expressed genes and transcripts (Trapnell et al., 2012). To facilitate future use of these datasets, all the expressed genes identified and their expression values (F/RPKM) in YFP+/- (Table S3) and all-sorted/non-sorted (Table S4) are included as supplementary material.

Filters were applied to determine if a gene was detected, abiding by the suggestions of statisticians and bioinformaticians (Rau et al., 2013), (Soneson and Delorenzi, 2013), (Bourgon et al., 2010), (Pimentel et al., 2014), (Love et al., 2014), (Seyednasrollah et al., 2015) as a means to enrich for true DEGs, to reduce type I error and to improve *P*-value adjustment. The edgeR function (Robinson et al., 2010) 'cpm' (counts per million) was used to discard those genes whose cpm was lower than a threshold of 2 reads per gene in at least 3 biological replicates, as suggested in the edgeR vignette. For cufflinks, a minimum RPKM of 5 was set for a gene to be expressed, following Suzuki *et al*. criteria (Suzuki et al., 2015). According to Sims *et al.* 80% of genes can be accurately quantified with FPKM > 10 (Sims et al., 2014). DESeq2 performs independent filtering using the 'results' function, as described in

the DESeq2 vignette (Love et al., 2014). An FDR cutoff of < 0.01 was used to determine differentially expressed genes in all three programs. The Gene Regulatory Information Server (AGRIS) was used to identify transcription families in the dataset (Yilmaz et al., 2011). Enriched categories of transcription factors within the set of "cleaned" 363 YFP+/- DEGs was assessed with the online Transcription Factor Enrichment Calculator tool (Desai, Jigar, Dmitry Grinevich and Colleen Doherty).

Venn Diagrams and heatmaps

Venn diagrams were constructed using the 'R' package VennDiagram (Chen and Boutros, 2011) and the web-based tool package InteractiVenn (Heberle et al., 2015). Heatmaps were produced using the 'R' package pheatmap (CRAN - pheatmap package). RPKM normalization by gene length and library size values were produced using the 'rpkm' function from edgeR (Robinson et al., 2010). To calculate gene length, a TAIR10 gene length list (CDS plus UTRs) was constructed by extracting length information from the TAIR10 GFF file with homemade Perl script. Genes with multiple isoforms were collapsed and length was calculated using the longest one. RPKM values were then calculated for clustering purposes and to have an intermediate point of comparison between Cufflinks, edgeR and DESeq2. Samples were clustered (default clustering) with parameters provided in the software. The 'R' package colorRamp (CRAN - colorRamp package) was used to produce a gradient of color values corresponding to gene-fold change values.

Gene set enrichment analysis

Gene Ontology (GO) enrichment tests were performed using the 'R' package topGO (Alexa and Rahnenführer, 2009), with the 'classic' algorithm (where each GO category is tested independently) and the 'fisher' statistic test for 'biological processes', 'molecular function' and 'cellular component'. Enrichment analysis was performed separately for all the genes that were differentially expressed between the YFP+/- samples and between the all-sorted/non-sorted samples. Network analysis of GO terms was performed using BiNGO (Maere et al., 2005a) plugin for Cytoscape (Shannon et al., 2003). GO terms for the 268 genes identified as depleted in the YFP-positive sample, as well cellular component (CC) and molecular function (MF) for the YFP+/- sample can be found in Table S6.

882

883 **Dendrograms**

884 The ‘R’ Dist function was used to compute a distance matrix using the spearman method
885 (Spearman test rank correlation) and the ‘R’ Cor function to compute the variance of the
886 matrix. To perform hierarchical clustering, the hclust function in ‘R’ was used. All statistical
887 analyses were performed in ‘R’ v.3.0.2. Dendrogram plots were built using the ‘R’ ape
888 package with edge.color = "blue".

889

890 **Confocal microscopy**

891 Confocal microscopy was performed using a Zeiss LSM 710 (Carl Zeiss, Inc. Thornwood,
892 NY), microscope model (Zeiss Axio Observer Z.1), objective type Plan-Apochromat 20x/0.8
893 M27. Z-stack intervals were set to 2 μm and the total thickness of the stack was 62 μm.

894

895 **Chloral hydrate clearing and Differential Contrast (DIC) Microscopy**

896 Inflorescence samples were fixed in a solution of 9 parts ethanol: 1 part acetic acid for two
897 hours at room temperature, and then washed twice in 90% ethanol for 30 min each wash.
898 Inflorescences were transferred to Hoyer’s solution (70% Chloral hydrate w/v, 4% glycerol, 5
899 % gum Arabic) and allowed to clear for several hours to overnight. Samples were then
900 dissected in Hoyer’s solution. The dissected inflorescence heads were mounted in Hoyer’s
901 under coverslips and examined with DIC optics on a Zeiss Axioskop 2 to determine the floral
902 stages.

903

904 ***In situ* hybridization**

905 For *in situ* hybridization analysis, *Arabidopsis thaliana* Col-0 flowers were fixed and
906 embedded in paraffin as described previously (Franks et al., 2002; Wynn et al., 2011).
907 Sections of plant tissue were probed with digoxigenin-labeled antisense and sense RNA
908 probes (Roche). Probes corresponded to nucleotides +686 to +920 of *REM13* relative to the
909 transcriptional start site of the CDS using the following oligos to amplify the template:
910 REM13_ISH_Fwd 5’ AAAATAGAACGCGCATACCG 3’ and REM13_ISH_Rev 5’
911 TCGTGAACCAAACCGTGATA 3’. Hybridization and immunological detection were
912 performed as described previously (Franks et al., 2002; Wynn et al., 2011).

ACCESSION NUMBERS

Illumina sequencing raw data (fastq) have been submitted to the Gene Expression Omnibus (GEO) database (accession GSE74458).

ACKNOWLEDGEMENTS

We thank Sarah Schuett (CVM, flow cytometry facility, NCSU) and the Genomic Sciences Laboratory (GSL) research facility (NCSU) for library preparation and Illumina sequencing. We also would like to thank William Thomson (NCSU) and Emily Wear (NSCU) for FACS-sorting assistance; Frank Wellmer and Diarmuid O'Maoileidigh (Smurfit Institute, Trinity College of Dublin, Ireland) for the *pAP1:AP1:GR ap1 cal1* floral synchronization system; Colleen Doherty (NCSU) for help with Cytoscape/BiNGO analysis; Maria Angels De Luis Balaguer (NCSU) for assistance with cross-platform expression correlation approaches; José Alonso and Ross Sozzani (NCSU) for thoughtful comments on the manuscript; Jigar Desai, Dmitry Grinevich and Colleen Doherty (NCSU) for help with the transcription factor enrichment analysis. Aureliano Bombarely (Virginia Tech) for help with GO terms analysis; and Eva Johannes (CMIF, Molecular Imaging Facility, NCSU) for laser scanning confocal microscope assistance.

AUTHOR CONTRIBUTIONS

RGF, GHV, ES and SH coordinated and conceived of the study. AS, JV, LR built the GAL4/pUAS-based two-component reporter system. GHV, MFV and BS performed FACS-sorting, wet-lab and microscopy analyses. SM and LC performed *in situ* hybridization. GHV, QH and SH performed RNA-seq and related bioinformatic and computational analysis. GHV and RGF interpreted the data and drafted the manuscript. All of the authors revised and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

TABLES

944

945 **Table 1. Top 15 differentially expressed genes enriched in the YFP-positive sample as**

946 **ranked by fold change.** Arabidopsis gene ID is shown in the 1st column, gene name (TAIR10

947 annotation) is shown in the 2nd column and the 3rd column reference for available reporter lines

948 and/or *in situ* hybridization. Average Reads Per Kilobase of transcript per Million mapped

949 reads (RPKM) values are indicated for each sample (YFP_NEG = YFP-negative, YFP_POS=

950 YFP-positive).

951

Gene ID	Gene name	Reference for expression	YFP_NEG average expression	YFP_POS average expression	Fold Change	P Value	FDR
AT3G50330	*HECATE 2*	(Gremski et al., 2007)	1.267	80.148	67.0	$1.26E^{-81}$	$5.36E^{-78}$
AT5G67060	*HECATE 1*	(Gremski et al., 2007)	0.708	40.298	59.4	$1.44E^{-105}$	$2.44E^{-101}$
AT1G66950	*ATPDR11*	(Wuest et al., 2010)	0.136	5.931	45.2	$9.47E^{-76}$	$2.02E^{-72}$
AT3G58780	*SHP1*	(Ma et al., 1991)	1.678	54.233	33.9	$1.06E^{-52}$	$7.54E^{-50}$
AT5G17040	unknown	NA	0.272	7.453	28.0	$1.06E^{-29}$	$1.24E^{-27}$
AT1G06920	*OFP4*	NA	0.453	11.267	25.9	$6.18E^{-34}$	$1.04E^{-31}$
AT3G51060	*STY1*	(Kuusk et al., 2002)	3.141	73.227	24.5	$2.22E^{-98}$	$1.89E^{-94}$
AT5G14180	*MPL1*	(Skinner and Gasser, 2009)	1.348	25.806	20.1	$9.86E^{-52}$	$6.22E^{-49}$
AT2G42830	*SHP2*	(Ma et al., 1991)	1.642	31.392	20.1	$1.58E^{-74}$	$2.99E^{-71}$
AT1G03720	unknown	(Skinner and Gasser, 2009)	1.898	31.968	17.7	$6.66E^{-52}$	$4.54E^{-49}$
AT2G33850	unknown	NA	0.676	10.904	16.7	$1.85E^{-19}$	$7.15E^{-18}$
AT1G69180	*CRC*	(Bowman et al., 1999a)	19.435	296.675	16.0	$2.60E^{-78}$	$8.84E^{-75}$

AT2G22460	unknown	NA	0.727	10.899	15.6	$2.85E^{-17}$	$8.52E^{-16}$
AT2G21650	*MEE3*	(Pagnussat et al., 2005; Baxter et al., 2007)	1.333	17.643	13.9	$1.65E^{-16}$	$4.37E^{-15}$
AT2G24540	*AFR*	NA	0.84	11.029	13.8	$2.27E^{-24}$	$1.61E^{-22}$

952

953

954

955

956

957 **Table 2.** List of genes displaying isoform switching behavior between the *SHP2*-positive and

958 *SHP2*-negative cell populations. Genes for which the most highly expressed isoform differed

959 between YFP+ and YFP- samples. Match between the Cufflinks transcripts and TAIR10

960 genome are indicated with class code (cc): '=' for complete transcript match and 'j' for

961 potentially novel isoform (fragment) (Trapnell et al., 2012). TSS (transcriptional start site).

Isoform ID	TSS group ID	cc	Nearest reference ID	isoform fpkm. YFP NEG	isoform fpkm YFP POS	isoform fpkm. NO SORT	isoform fpkm. ALL SORT	isoform. rel.fpkm .YFP NEG	isoform. rel.fpkm YFP POS	isoform rel.fpkm NO SORT	isoform rel.fpkm ALL SORT
TCONS_00000149	TSS88	=	AT1G02110.1	7.85	32.40	6.29	18.10	0.12	0.35	0.09	0.22
TCONS_00010172	TSS6382	=	AT1G23340.2	1.45	1.87	2.38	1.60	0.21	0.58	0.24	0.24
TCONS_00019427	TSS12315	=	AT2G42890.1	3.37	2.74	0.93	4.94	0.24	0.41	0.10	0.36
TCONS_00057346	TSS36662	j	AT5G53050.2	3.69	2.53	0.68	5.27	0.57	0.58	0.26	0.72
TCONS_00004108	TSS2634	=	AT1G43850.2 (*SEUSS*)	13.76	14.84	20.31	13.50	0.69	0.48	0.73	0.52
TCONS_00044799	TSS28620	j	AT4G32250.3	2.98	5.95	6.48	6.08	0.20	0.24	0.32	0.29
TCONS_00001095	TSS637	=	AT1G10570.2	6.28	3.49	4.13	4.63	0.27	0.26	0.18	0.30

TCONS_00005695	TSS3643	=	AT1G61820.1	9.73	0.70	5.34	3.81	1.00	0.39	0.76	0.81
TCONS_00009214	TSS5796	=	AT1G14170.1	10.39	2.78	6.04	5.73	0.49	0.30	0.45	0.48
TCONS_00019421	TSS12312	j	AT2G42830.2	4.64E-06	3.02	0.41	0.64	3.42E-06	0.12	0.10	0.11
TCONS_00001793	TSS1060	j	AT1G16710.2	1.63	0.52	1.23	1.47	0.21	0.06	0.13	0.19
TCONS_00009282	TSS5837	j	AT1G14690.2	16.72	7.94	27.05	7.78	0.24	0.19	0.58	0.20
TCONS_00054221	TSS34554	j	AT5G19130.1	0.34	0	3.69	8.65E-04	0.04	0	0.20	5.30E-05
TCONS_00046341	TSS29519	j	AT5G06610.1	0.74	0.34	0.23	0.92	0.05	0.05	0.02	0.1
TCONS_00043885	TSS28059	j	AT4G23660.3	2.23E-04	1.41E-03	2.20E-04	0	2.98E-05	1.10E-04	1.80E-05	0

962

963

964 **Table 3.** Overlapping differentially expressed genes (DEGs) between the 95 DEGs from this
965 study that display enriched in *SHP2*-expressing cells and 210 DEGs from Wynn *et al*.
966 displaying reduced expression in the *seuss aintegumenta* (*seu ant*) double mutant relative to
967 other genotypes.

968

		RNA-seq (this study)					Array (log_2 scaled expression values) (Wynn et al., 2011)			
Annotation/Gene symbol	GeneID	log_2FC	FC	P Value	FDR		WT	*ant*	*seu*	*seu ant*
HEC1	AT5G67060	5.9	59.4	1.44E-105	2.44E-101		8.887	8.355	8.030	7.292
Cysteine proteinase	AT1G03720	4.1	17.7	6.66E-52	4.54E-49		8.998	8.165	8.409	7.053
Plant invertase	AT3G62820	3.7	12.9	4.76E-40	1.42E-37		8.614	8.900	8.857	7.833
REM13	AT3G46770	3.6	12.2	8.42E-20	3.36E-18		8.653	8.620	7.954	7.271
ATCEL2	AT1G02800	3.4	10.4	7.33E-90	4.16E-86		11.927	11.466	11.182	9.948
LAX1	AT5G01240	3.3	9.7	2.74E-59	2.45E-56		9.311	9.254	9.226	8.709

LTP2	AT2G38530	3.3	9.5	6.42E-54	4.97E-51		11.297	10.301	10.987	10.056
AMP-dependent synthetase	AT1G21540	3.2	9.3	3.45E-22	1.91E-20		8.547	8.859	7.986	7.917
AGO5	AT2G27880	3.0	8.0	1.20E-71	1.45E-68		10.194	9.653	9.977	8.937
UGT84A2	AT3G21560	2.9	7.6	2.77E-73	4.29E-70		9.735	9.707	8.868	8.370
TAA1	AT1G70560	2.8	6.9	1.34E-69	1.43E-66		9.608	9.072	9.068	7.623
ATDOF5.8	AT5G66940	2.7	6.6	2.26E-39	6.53E-37		10.380	10.135	9.761	8.007
VDD	AT5G18000	2.6	6.0	3.71E-31	5.10E-29		8.554	7.719	8.536	7.018
MCT1	AT1G37140	2.6	5.9	8.22E-10	7.74E-09		8.003	7.406	7.572	6.713
unknown	AT2G41990	2.5	5.6	5.70E-39	1.56E-36		9.690	9.448	9.068	8.450
Protein kinase superfamily protein	AT1G74490	2.5	5.5	6.27E-37	1.40E-34		7.798	7.727	7.312	6.976
NGA3	AT1G01030	2.3	5.1	4.07E-37	9.23E-35		7.543	7.666	7.214	7.007
NGA1	AT2G46870	2.3	4.8	9.56E-39	2.58E-36		8.423	8.516	7.849	7.168
Cystatin/monellin	AT1G03710	2.3	4.8	1.64E-34	2.98E-32		10.186	10.156	9.780	8.809
YUC4	AT5G11320	2.3	4.8	8.01E-21	3.70E-19		7.787	7.889	7.858	7.070
Heavy metal transport	AT1G56210	2.2	4.7	7.11E-49	3.46E-46		9.229	8.487	9.092	8.507
ANAC098	AT5G53950	2.1	4.2	1.88E-22	1.08E-20		8.719	8.453	8.270	7.780
unknown	AT1G51670	2.1	4.1	6.00E-30	7.25E-28		9.507	9.049	9.027	8.388

FIGURE LEGENDS

Figure 1. A system for the collection of temporally- and spatially-restricted cell populations from the *Arabidopsis thaliana* gynoecium. **a** Microscopic image of a mature wild type Arabidopsis gynoecium. The stigma (stg), style (sty), carpel valve (cv), abaxial replum (abr),

977 gynophore (gn), and ovary (ovy) are false colored. **b** False-colored confocal cross section of a
978 stage-8 gynoecium. Medial and lateral domains of the Arabidopsis gynoecium are indicated.
979 The carpel margin meristem/medial ridge (CMM) is false colored pink. **c** False-colored stage-
980 11 cross-section. Ovules (ov), septum (s) and carpel valves (cv) are indicated. **d** Confocal
981 microscope image of the *pSHP2-YFP* two-component reporter in the *ap1; cal;*
982 *pAP1::AP1:GR* background. YFP expression from the *pSHP2-YFP* reporter is chiefly
983 confined to the medial domain of the gynoecium at late stage 7/early stage 8, although weak,
984 non-medial domain expression can be detected in portions of the stamens. Sepals (se) and
985 stamens (st) are labeled. **e** Z-stack composite 3D projection image of a gynoecium isolated
986 from the flower at mid-stage 8. YFP expression from the *pSHP2-YFP* reporter is detected in
987 the medial domain and at the apex of the gynoecium. **f** Chloral hydrate image of an
988 inflorescence of an *ap1; cal; pAP1::AP1:GR* plant after mock treatment. Inflorescence-like
989 meristems do not transition to floral meristems. **g** Chloral hydrate image of an inflorescence of
990 an *ap1; cal; pAP1::AP1:GR* plant 125 hours after spray application of Dexamethasone
991 synthetic hormone (Dex). Samples were enriched for stages 6-8. **h** Percentage of flowers at a
992 given stage from inflorescences used for FACS-sorting. Stages 6, 7, 8p (pre-ovules) and 8s
993 (post-ovules) are indicated in the X-axis as St6, St7, St8p, St8s, respectively. Stage 8p is
994 before any visible morphological manifestation of ovule primordia upon observation under
995 DIC microscopy. Stage 8s ovule primordia were observed and were at ovule stage 1-I or 1-II
996 according the Schneitz *et al.* (Schneitz et al., 1995). **i** Confocal microscopy of YFP
997 fluorescence of protoplasted cells after FACS. Panels **a**, **b** and **c** are adapted from
998 Azhakanandam *et al*. (Azhakanandam et al., 2008) (with permission).
999

1000 **Figure 2.** Venn diagram of differentially expressed genes (DEGs) using Cufflinks, edgeR and
1001 DESeq2 (FDR<0.001, Fold Change >4). **a** Venn showing DEGs identified between the all-
1002 sorted/non-sorted samples with the 3 programs used for differential expression analysis of
1003 RNA-seq expression profiles. **b** Venn showing DEGs between YFP+/- samples identified in
1004 the 3 programs. **c** Intersection of the DEGs (48) from both datasets (a and b). DEGs (363),
1005 after removing DEGs induced by the protoplsting/FACS-sorting stress, were used for
1006 downstream analysis.
1007
1008 **Figure 3.** Candidate medial domain regulator *REM13* (At3g46770) is expressed within the
1009 medial gynoecial domain and developing ovules. Results from an RNA *in situ* hybridization
1010 with *REM13* probe. **a-d** antisense probe. **e** sense strand probe. **a** Hybridization signal is
1011 detected in the carpel margin meristem (adaxial portions of the medial gynoecial domain) in
1012 the stage 7 longitudinal section. **b, c** and **d** In transverse gynoecial sections *REM13* expression
1013 is detected in the ovule primordia; stage 7 (panel b) stage 8 (panel c) and stage 9 (panel d)
1014 gynoecia. **e** A stage 8 section hybridized with a *REM13* sense strand probe. (ov) - ovules,
1015 (cmm) - carpel margin meristem. Scale bars for each panel represent 50 microns.
1016
1017
1018 **Figure 4.** GO term overrepresentation of *SHP2*-domain enriched genes suggests a role for this
1019 set of genes in floral, gynoecial and ovule development. BiNGO/Cytoscape representation of
1020 overrepresented GO terms from the 95 YFP+/- DEGs displaying enriched expression in the
1021 YFP-positive samples. Edges represent the parent/child relationships of the GO terms
1022 (Ashburner et al., 2000), while color of the nodes indicates the degree of statistical

1023 significance (p <0.01) as reported by BiNGO (Maere et al., 2005b). To unclutter the figure,
1024 given the large number of significant GO terms, selected nodes and edges have been removed
1025 from this graphical representation.
1026
1027 **Figure 5.** The transcriptomic signature of the *SHP2*-expressing domain is more similar to the
1028 transcriptomes of other meristematic samples than it is to whole flower. **a** Dendrogram based
1029 on hierarchical clustering using the Spearman rank correlation using RNA-seq (RPKM)
1030 expression values from flowers and other tissues. **b** Comparison of RNA-seq and affymetrix
1031 ATH1 arrays samples including transcriptomic data from whole flower, shoot apical meristem
1032 and seedling.
1033 WT = wild type, MT= mutants. Data from Mizzotti *et al*. (Mizzotti et al., 2014)[(1)],
1034 Mantegazza *et al*. (Mantegazza et al., 2014b)[(2)], GEO accession: GSE54125[(3)] and Yadav *et*
1035 *al*. (Yadav et al., 2009; Yadav et al., 2014)[(4)] were used for comparison. Samples
1036 corresponding to this study are color coded red in both dendrograms.
1037
1038 **Figure 6.** Heatmap representation of the expression profiles of previously identified
1039 regulators of Carpel Margin Meristem development. Expression profiles in Reads Per
1040 Kilobase of transcript per Million mapped reads (RPKM) of the 86 genes reported by Reyes-
1041 Olalde *et al. (Reyes-Olalde et al., 2013)* with functional role during CMM development.
1042 Transcriptional profiles from this study (YFP POS = YFP-positive, YFP NEG = YFP-
1043 negative, ALL SORT = all-sorted, and NO SORT = no-sorted) as well as Mantegazza *et al*.
1044 (Mantegazza et al., 2014a) corresponding to flower stage 3 (FL.STAGE 3), floral meristem
1045 (FL.MERISTEM) and inflorescence meristem (IN.MERISTEM) are included. Genes color-

coded in red are those identified as DEGs between YFP-positive and YFP-negative samples (FC >4 and FDR <0.001) while genes that displayed a statistically significant expression level (FDR <0.01) between YFP-positive and YFP-negative (regardless of their fold change) are indicated with ***.

Figure 7. Differential expression of *REVERSIBLY GLYCOSYLATED POLYPEPTIDE 5* (*RGP5)* isoforms as well as *TRANS-ACTING siRNA3* (*TAS3*) and *AUXIN RESPONSE FACTOR* genes. **a** Promoter/transcriptional start site switch found for the *RGP5* gene (At5g16510). The isoform 2 (At5g16510.2) increases its expression in the YFP-positive domain while isoform 1 (At5g16510.1) of the same gene decreases its expression in the same domain. **b** Expression of the *AUXIN RESPONSE FACTORS* (*ARFs*) (*ARF2*, *ARF3*, *ARF4*) and *TAS3* transcripts. Expression levels of *ARF2*, *ARF3*, *ARF4* are significantly enriched in the YFP-positive sample at FDR <0.01. Expression levels of the *TRANS ACTING siRNA3* (*TAS3*) genes At5g49615 and At3g17185, that negatively regulate the expression of *ARF2*, *ARF3*, and *ARF4* expression (Williams et al., 2005), are significantly reduced (FDR <0.01) in the YFP-positive sample.

Figure 8. Six-way venn diagram image showing detailed overlap from all the differentially expressed gene (DEGs) datasets. The total number of DEGs under each condition and for each program are indicated in parentheses. CTR= DEGs between all-sorted/non-sorted and YFPs= DEGs between YFP+/-. Cuff= Cufflinks, edg=edgeR, Des=DESeq2. The interactive tool can be accessed online using the 'InteractiVenn' web-tool (InteractiVenn) and uploading Additional file 12.

Supplemental data files

The following additional data are available with the online version of this article.

Supplemental Figures S1: Sorting gates used to select YFP samples and the re-sorting of the YFP-positive cells to assess sample purity.

Supplemental Figures S2: qRT-PCR enrichment of medial domain genes *SHP2* and *NGA1* and the gene *TUB*.

Supplemental Figures S3: Expression profiles for the 363 differentially expressed genes (FC >4, FDR <0.001) across all 4 samples (YFP-positive, YFP-negative, all-sorted, non-sorted).

Supplemental Figures S4: Venn Diagram comparison of stressed induced genes due to protoplast/FACS-sorting procedure.

Supplemental Figures S5: Gene level variance versus log gene expression level among technical replicates.

Supplemental Table S1: Summary RNA-seq data (number of reads, mapped reads, uniquely mapped, etc.).

Supplemental Table S2: Differentially expressed genes (DEGs) from the YFP+/- and all-sorted/non-sorted comparison.

Supplemental Table S3 and Table S4: contain all the expressed genes identified with three different programs between all the YFP+/- samples and all-sorted/non-sorted samples, respectively.

Supplemental Table S5: Corresponds to raw high-throughput count data for YFP+/- and all-sorted/non-sorted comparison.

Supplemental Table S6: Gene Set Enrichment Analyses (GSEA) for YFP+/- and all-sorted/non-sorted comparison, including Biological Process (BP), Molecular Function (MF) and Cellular Component (CC).

Supplemental Table S7: Lists the transcription factors families identified in the DEGs from YFP+/- and their statistical enrichment.

Supplemental Table S8: contains isoforms expression, regulation of gene expression by alternative promoters and antisense transcripts identified by Cufflinks, edgeR and DESeq2.

Supplemental Table S9: Expression profile (RPKM) of the 86 genes described by Reyes-Olalde *et al.* (Reyes-Olalde et al., 2013) expressed in the medial domain.

Supplemental Table S10: Hormone (Auxin, GA, Ethylene) related-genes present in our dataset.

Supplemental Additional file S11: Data file to upload to the web-based tool package "InteractiVenn".

Parsed Citations

Adrian J, Chang J, Ballenger CE, Bargmann BOR, Alassimone J, Davies KA, Lau OS, Matos JL, Hachez C, Lanctot A, et al (2015) Transcriptome dynamics of the stomatal lineage: birth, amplification, and termination of a self-renewing population. Dev Cell 33: 107-118

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Aichinger E, Kornet N, Friedrich T, Laux T (2012) Plant stem cell niches. Annu Rev Plant Biol 63: 615-636

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Alexa A, Rahnenführer J (2009) topGO: Enrichment Analysis for Gene Ontology; 2010. Bioconductor package version.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Alvarez JP, Goldshmidt A, Efroni I, Bowman JL, Eshed Y (2009) The NGATHA distal organ development genes are essential for style specification in Arabidopsis. Plant Cell 21: 1373-1393

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Alvarez J, Smyth DR (2002) CRABS CLAW and SPATULA Genes Regulate Growth and Pattern Formation during Gynoecium Development in Arabidopsis thaliana. Int J Plant Sci 163: 17-41

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11: R106

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Anders S, Pyl PT, Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166-169

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25-29

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ayoubi TA, Van De Ven WJ (1996) Regulation of gene expression by alternative promoters. The FASEB Journal 10: 453-460

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Azhakanandam S, Nole-Wilson S, Bao F, Franks RG (2008) SEUSS and AINTEGUMENTA Mediate Patterning and Ovule Initiation during Gynoecium Medial Domain Development. Plant Physiol 146: 1165-1181

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Baxter CEL, Costa MMR, Coen ES (2007) Diversification and co-option of RAD-like genes in the evolution of floral asymmetry. Plant J 52: 105-113

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Benkova E, Michniewicz M, Sauer M, Teichmann T, Seifertova D, Jurgens G, Friml J (2003) Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell 115: 591-602

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bennett MJ, Marchant A, Green HG, May ST, Ward SP, Millner PA, Walker AR, Schulz B, Feldmann KA (1996) Arabidopsis AUX1 gene: a permease-like regulator of root gravitropism. Science 273: 948-950

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Birnbaum K, Jung JW, Wang JY, Lambert GM, Hirst JA, Galbraith DW, Benfey PN (2005) Cell type-specific expression profiling in plants via cell sorting of protoplasts from fluorescent reporter lines. Nat Methods 2: 615-619

Pubmed: Author and Title

CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN (2003) A gene expression map of the Arabidopsis root. Science 302: 1956-1960
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Blilou I, Xu J, Wildwater M, Willemsen V, Paponov I, Friml J, Heidstra R, Aida M, Palme K, Scheres B (2005) The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. Nature 433: 39-44
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bourgon R, Gentleman R, Huber W (2010) Independent filtering increases detection power for high-throughput experiments. Proceedings of the National Academy of Sciences 107: 9546-9551
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bowman JL, Baum SF, Eshed Y, Putterill J, Alvarez J (1999a) Molecular genetics of gynoecium development in Arabidopsis. Curr Top Dev Biol 45: 155-205
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bowman JL, Baum SF, Eshed Y, Putterill J, Alvarez J (1999b) 4 Molecular Genetics of Gynoecium Development in Arabidopsis. Curr Top Dev Biol 45: 155-205
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bowman JL, Smyth DR (1999) CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop- helix domains. Development 126: 2387-96.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bradford JR, Hey Y, Yates T, Li Y, Pepper SD, Miller CJ (2010) A comparison of massively parallel nucleotide sequencing with oligonucleotide microarrays for global transcription profiling. BMC Genomics 11: 282
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Breakfield NW, Corcoran DL, Petricka JJ, Shen J, Sae-Seaw J, Rubio-Somoza I, Weigel D, Ohler U, Benfey PN (2012) High-resolution experimental and computational profiling of tissue-specific known and novel miRNAs in Arabidopsis. Genome Res 22: 163-176
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Carter AD, Bonyadi R, Gifford ML (2013) The use of fluorescence-activated cell sorting in studying plant development and environmental responses. Int J Dev Biol 57: 545-552
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Cheng Y, Dai X, Zhao Y (2006) Auxin biosynthesis by the YUCCA flavin monooxygenases controls the formation of floral organs and vascular tissues in Arabidopsis. Genes Dev 20: 1790-1799
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Chen H, Boutros PC (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics 12: 35
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Colombo M, Brambilla V, Marcheselli R, Caporali E, Kater MM, Colombo L (2010) A new role for the SHATTERPROOF genes during Arabidopsis gynoecium development. Dev Biol 337: 294-302
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

CRAN - colorRamp package [http://stat.ethz.ch/R- manual/R -devel/library/grDevices/html/colorRamp.html] webcite.

CRAN - ggplot2 package [https://cran.r-project.org/web/packages/ggplot2/index.html] webcite.

CRAN - pheatmap package [http://cran.r- project.org/web/ packages/pheatmap/index.html] webcite.

CRAN - plyr package [https://cran.r-project.org/web/packages/plyr/index.html] webcite.

Crawford BCW, Yanofsky MF (2011) HALF FILLED promotes reproductive tract development and fertilization efficiency in Arabidopsis thaliana. Development 138: 2999-3009
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Crawford BCW, Yanofsky MF (2008) The Formation and Function of the Female Reproductive Tract in Flowering Plants. Curr Biol 18: R972-R978
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Desai, Jigar, Dmitry Grinevich and Colleen Doherty Transcription Factor Enrichment Calculator. https://dgrinevich.shinyapps.io/ShinyTF, https://dgrinevich.shinyapps.io/ShinyTF

Drakakaki G, Zabotina O, Delgado I, Robert S, Keegstra K, Raikhel N (2006) Arabidopsis Reversibly Glycosylated Polypeptides 1 and 2 Are Essential for Pollen Development. Plant Physiol 142: 1480-1492
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS (2006) Gateway-compatible vectors for plant functional genomics and proteomics. Plant J 45: 616-629
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Favaro R, Pinyopich A, Battaglia R, Kooiker M, Borghi L, Ditta G, Yanofsky MF, Kater MM, Colombo L (2003) MADS-Box Protein Complexes Control Carpel and Ovule Development in Arabidopsis. Plant Cell 15: 2603-2611
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Franco-Zorrilla JM, Cubas P, Jarillo JA, Fernandez-Calvin B, Salinas J, Martinez-Zapater JM (2002) AtREM1, a Member of a New Family of B3 Domain-Containing Genes, Is Preferentially Expressed in Reproductive Meristems. Plant Physiol 128: 418-427
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Franks RG, Wang C, Levin JZ, Liu Z (2002) SEUSS, a member of a novel family of plant regulatory proteins, represses floral homeotic gene expression with LEUNIG. Development 129: 253-263
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Galbiati F, Sinha Roy D, Simonini S, Cucinotta M, Ceccato L, Cuesta C, Simaskova M, Benkova E, Kamiuchi Y, Aida M, et al (2013) An integrative model of the control of ovule primordia formation. Plant J 76: 446-455
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Gandikota M, Birkenbihl RP, Höhmann S, Cardon GH, Saedler H, Huijser P (2007) The miRNA156/157 recognition element in the 3' UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. Plant J 49: 683-693
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Gremski K, Ditta G, Yanofsky MF (2007) The HECATE genes regulate female reproductive tract development in Arabidopsis thaliana. Development 134: 3593-3601
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Guida A, Lindstädt C, Maguire SL, Ding C, Higgins DG, Corton NJ, Berriman M, Butler G (2011) Using RNA-seq to determine the transcriptional landscape and the hypoxic response of the pathogenic yeast Candida parapsilosis. BMC Genomics 12: 628
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Halbritter F, Vaidya HJ, Tomlinson SR (2012) GeneProf: analysis of high-throughput sequencing experiments. Nat Methods 9: 7-8
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hay A, Angela H, Hardip K, Andrew P, Peter H, Sarah H, Miltos T (2002) The Gibberellin Pathway Mediates KNOTTED1-Type Homeobox Function in Plants with Different Body Plans. Curr Biol 12: 1557-1565
Pubmed: Author and Title
CrossRef: Author and Title

Google Scholar: Author Only Title Only Author and Title

Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R (2015) InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics. doi: 10.1186/s12859-015-0611-3
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Heisler M, Atkinson A, Bylstra YH, Walsh R, Smyth DR (2001) SPATULA, a gene that controls development of carpel margin tissues in Arabidopsis, encodes a bHLH protein. Development 128: 1089-1098
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

iGenomes Database. [http://support.illumina.com/sequencing/sequencing_software/igenome.html] webcite.

InteractiVenn [http://www.interactivenn.net] webcite.

Jasinski S, Piazza P, Craft J, Hay A, Woolley L, Rieu I, Phillips A, Hedden P, Tsiantis M (2005) KNOX Action in Arabidopsis Is Mediated by Coordinate Regulation of Cytokinin and Gibberellin Activities. Curr Biol 15: 1560-1565
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kamiuchi Y, Yamamoto K, Furutani M, Tasaka M, Aida M (2014) The CUC1 and CUC2 genes promote carpel margin meristem formation during Arabidopsis gynoecium development. Front Plant Sci. doi: 10.3389/fpls.2014.00165
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kuusk S, Sohlberg JJ, Long JA, Fridborg I, Sundberg E (2002) STY1 and STY2 promote the formation of apical tissues during Arabidopsis gynoecium development. Development 129: 4707-4717
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, et al (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res 40: D1202-10
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lan P, Li W, Lin W-D, Santi S, Schmidt W (2013) Mapping gene activity of Arabidopsis root hairs. Genome Biol 14: R67
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Larsson E, Roberts CJ, Claes AR, Franks RG, Sundberg E (2014) Polar Auxin Transport Is Essential for Medial versus Lateral Tissue Specification and Vascular-Mediated Valve Outgrowth in Arabidopsis Gynoecia. Plant Physiol 166: 1998-2012
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Leal Valentim F, Sv M, Posé D, Kim MC, Schmid M, van Ham R (2015) A Quantitative and Dynamic Model of the Arabidopsis Flowering Time Gene Regulatory Network. PLoS One 10: e0116973
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF (2000) SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature 404: 766-770
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lindgreen S (2012) AdapterRemoval: easy cleaning of next-generation sequencing reads. BMC Res Notes 5: 337
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Li S, Liberman LM, Mukherjee N, Benfey PN, Ohler U (2013) Integrated detection of natural antisense transcripts using strand-specific RNA sequencing data. Genome Res 23: 1730-1739
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. doi: 10.1186/s13059-014-0550-8
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Maere S, Heymans K, Kuiper M (2005a) BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. Bioinformatics 21: 3448-3449
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Maere S, Heymans K, Kuiper M (2005b) BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. Bioinformatics 21: 3448-3449
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ma H, Yanofsky MF, Meyerowitz EM (1991) AGL1-AGL6, an Arabidopsis gene family with similarity to floral homeotic and transcription factor genes. Genes Dev 5: 484-495
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Mantegazza O, Gregis V, Chiara M, Selva C, Leo G, Horner DS, Kater MM (2014a) Gene coexpression patterns during early development of the native Arabidopsis reproductive meristem: novel candidate developmental regulators and patterns of functional redundancy. Plant J 79: 861-877
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Mantegazza O, Gregis V, Mendes MA, Morandini P, Alves-Ferreira M, Patreze CM, Nardeli SM, Kater MM, Colombo L (2014b) Analysis of the arabidopsis REM gene family predicts functions during flower development. Ann Bot 114: 1507-1515
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18: 1509-1517
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Martinez-Fernandez I, Sanchis S, Marini N, Balanza V, Ballester P, Navarrete-Gomez M, Oliveira AC, Colombo L, Ferrandiz C (2014) The effect of NGATHA altered activity on auxin signaling pathways within the Arabidopsis gynoecium. Front Plant Sci 5: 210
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Matias-Hernandez L, Battaglia R, Galbiati F, Rubes M, Eichenberger C, Grossniklaus U, Kater MM, Colombo L (2010) VERDANDI Is a Direct Target of the MADS Domain Ovule Identity Complex and Affects Embryo Sac Differentiation in Arabidopsis. Plant Cell 22: 1702-1715
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Mizzotti C, Ezquer I, Paolo D, Rueda-Romero P, Guerra RF, Battaglia R, Rogachev I, Aharoni A, Kater MM, Caporali E, et al (2014) SEEDSTICK is a Master Regulator of Development and Metabolism in the Arabidopsis Seed Coat. PLoS Genet 10: e1004856
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Moubayidin L, Ostergaard L (2014) Dynamic Control of Auxin Distribution Imposes a Bilateral-to-Radial Symmetry Switch during Gynoecium Development. Curr Biol 24: 2743-2748
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Moussaieff A, Rogachev I, Brodsky L, Malitsky S, Toal TW, Belcher H, Yativ M, Brady SM, Benfey PN, Aharoni A (2013) High-resolution metabolic mapping of cell types in plant roots. Proceedings of the National Academy of Sciences. doi: 10.1073/pnas.1302019110
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Mudge J, Miller NA, Khrebtukova I, Lindquist IE, May GD, Huntley JJ, Luo S, Zhang L, van Velkinburgh JC, Farmer AD, et al (2008) Genomic Convergence Analysis of Schizophrenia: mRNA Sequencing Reveals Altered Synaptic Vesicular Transport in Post-Mortem Cerebellum. PLoS One 3: e3625
Pubmed: Author and Title

CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Nagpal P, Ellis CM, Weber H, Ploense SE, Barkawi LS, Guilfoyle TJ, Hagen G, Alonso JM, Cohen JD, Farmer EE, et al (2005) Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. Development 132: 4107-4118
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Nakagawa T, Kurose T, Hino T, Tanaka K, Kawamukai M, Niwa Y, Toyooka K, Matsuoka K, Jinbo T, Kimura T (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. J Biosci Bioeng 104: 34-41
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Nookaew I, Papini M, Pornputtapong N, Scalcinati G, Fagerberg L, Uhlen M, Nielsen J (2012) A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in Saccharomyces cerevisiae. Nucleic Acids Res 40: 10084-10097
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ó'Maoiléidigh DS, Wellmer F (2014) A Floral Induction System for the Study of Early Arabidopsis Flower Development. In JL Riechmann, F Wellmer, eds, Flower Development. Springer New York, New York, NY, pp 307-314
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Oram RN, Brock RD (1972) Prospects for improving plant protein yield and quality by breeding. Australian Inst Agr Sci J
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Pagnussat GC, Yu H-J, Ngo QA, Rajani S, Mayalagu S, Johnson CS, Capron A, Xie L-F, Ye D, Sundaresan V (2005) Genetic and molecular identification of genes required for female gametophyte development and function in Arabidopsis. Development 132: 603-614
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Payne CT, Zhang F, Lloyd AM (2000) GL3 encodes a bHLH protein that regulates trichome development in arabidopsis through interaction with GL1 and TTG1. Genetics 156: 1349-1362
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Petersson SV, Johansson AI, Kowalczyk M, Makoveychuk A, Wang JY, Moritz T, Grebe M, Benfey PN, Sandberg G, Ljung K (2009) An auxin gradient and maximum in the Arabidopsis root apex shown by high-resolution cell-specific analysis of IAA distribution and synthesis. Plant Cell 21: 1659-1668
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Petricka JJ, Schauer MA, Megraw M, Breakfield NW, Thompson JW, Georgiev S, Soderblom EJ, Ohler U, Moseley MA, Grossniklaus U, et al (2012) The protein expression landscape of the Arabidopsis root. Proc Natl Acad Sci U S A 109: 6811-6818
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Phillips AL, Ward DA, Uknes S, Appleford NE, Lange T, Huttly AK, Gaskin P, Graebe JE, Hedden P (1995) Isolation and expression of three gibberellin 20-oxidase cDNA clones from Arabidopsis. Plant Physiol 108: 1049-1057
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Pimentel H, Parra M, Gee S, Ghanem D, An X, Li J, Mohandas N, Pachter L, Conboy JG (2014) A dynamic alternative splicing program regulates gene expression during terminal erythropoiesis. Nucleic Acids Res 42: 4031-4042
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, Yanofsky MF (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature 424: 85-88
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Rau A, Gallopin M, Celeux G, Jaffrezic F (2013) Data-based filtering for replicated high-throughput transcriptome sequencing experiments. Bioinformatics 29: 2146-2152

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Reyes-Olalde JI, Zuñiga-Mayo VM, Chávez Montes RA, Marsch-Martínez N, de Folter S (2013) Inside the gynoecium: at the carpel margin. Trends Plant Sci 18: 644-655
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139-140
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Romanel EAC, Schrago CG, Couñago RM, Russo CAM, Alves-Ferreira M (2009) Evolution of the B3 DNA Binding Superfamily: New Insights into REM Family Gene Diversification. PLoS One 4: e5791
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

R package bear [http://sourceforge.net/projects/yjlee-r-packages/files/bear/] webcite.

Sauer M, Robert S, Kleine-Vehn J (2013) Auxin: simply complicated. J Exp Bot 64: 2565-2577
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Savidge B, Rounsley SD, Yanofsky MF (1995) Temporal relationship between the transcription of two Arabidopsis MADS box genes and the floral organ identity genes. The Plant Cell Online 7: 721-733
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27: 863-864
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative CT method. Nat Protoc 3: 1101-1108
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schneitz K, Hülskamp M, Pruitt RE (1995) Wild-type ovule development in Arabidopsis thaliana: a light microscope study of cleared whole-mount tissue. Plant J 7: 731-749
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sehra B, Franks RG (2015) Auxin and cytokinin act during gynoecial patterning and the development of ovules from the meristematic medial domain. Wiley Interdiscip. Rev. Dev. Biol.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sessions A (1999) Piecing together the Arabidopsis gynoecium. Trends Plant Sci 4: 296-297
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sessions RA, Zambryski PC (1995) Arabidopsis gynoecium structure in the wild and in ettin mutants. Development 121: 1519-1532
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Seyednasrollah F, Laiho A, Elo LL (2015) Comparison of software packages for detecting differential expression in RNA-seq studies. Brief Bioinform 16: 59-70
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Seymour GB, Østergaard L, Chapman NH, Knapp S, Martin C (2013) Fruit development and ripening. Annu Rev Plant Biol 64: 219-241
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498-2504

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP (2014) Sequencing depth and coverage: key considerations in genomic analyses. Nat Rev Genet 15: 121-132
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Singh MB, Bhalla PL (2007) Control of male germ-cell development in flowering plants. Bioessays 29: 1124-1132
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Skinner DJ, Gasser CS (2009) Expression-based discovery of candidate ovule development regulators through transcriptional profiling of ovule mutants. BMC Plant Biol 9: 29
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Smyth DR, Bowman JL, Meyerowitz EM (1990) Early flower development in Arabidopsis. The Plant Cell Online 2: 755-767
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Soneson C, Delorenzi M (2013) A comparison of methods for differential expression analysis of RNA-seq data. BMC Bioinformatics 14: 91
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Stepanova AN, Robertson-Hoyt J, Yun J, Benavente LM, Xie D-Y, Doležal K, Schlereth A, Jürgens G, Alonso JM (2008) TAA1-Mediated Auxin Biosynthesis Is Essential for Hormone Crosstalk and Plant Development. Cell 133: 177-191
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Suzuki A, Matsushima K, Makinoshima H, Sugano S, Kohno T, Tsuchihara K, Suzuki Y (2015) Single-cell analysis of lung adenocarcinoma cell lines reveals diverse expression patterns of individual cells invoked by a molecular target drug treatment. Genome Biol. doi: 10.1186/s13059-015-0636-y
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Swaminathan K, Peterson K, Jack T (2008) The plant B3 superfamily. Trends Plant Sci 13: 647-655
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Tao Y, Ferrer JL, Ljung K, Pojer F, Hong F, Long JA, Li L, Moreno JE, Bowman ME, Ivans LJ, et al (2008) Rapid synthesis of auxin via a new tryptophan-dependent pathway is required for shade avoidance in plants. Cell 133: 164-176
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25: 1105-1111
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7: 562-578
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Trigueros M, Navarrete-Gomez M, Sato S, Christensen SK, Pelaz S, Weigel D, Yanofsky MF, Ferrandiz C (2009) The NGATHA Genes Direct Style Development in the Arabidopsis Gynoecium. THE PLANT CELL ONLINE 21: 1394-1409
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wang C, Gong B, Bushel PR, Thierry-Mieg J, Thierry-Mieg D, Xu J, Fang H, Hong H, Shen J, Su Z, et al (2014) The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. Nat Biotechnol 32: 926-932
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wang J-W, Czech B, Weigel D (2009) miR156-regulated SPL transcription factors define an endogenous flowering pathway in Arabidopsis thaliana. Cell 138: 738-749

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wang J-W, Schwab R, Czech B, Mica E, Weigel D (2008) Dual effects of miR156-targeted SPL genes and CYP78A5/KLUH on plastochron length and organ size in Arabidopsis thaliana. Plant Cell 20: 1231-1243
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wellmer F, Alves-Ferreira M, Dubois A, Riechmann JL, Meyerowitz EM (2006) Genome-Wide Analysis of Gene Expression during Early Arabidopsis Flower Development. PLoS Genet 2: e117
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Williams L, Carles CC, Osmont KS, Fletcher JC (2005) A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the Arabidopsis ARF2, ARF3, and ARF4 genes. Proc Natl Acad Sci U S A 102: 9703-9708
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Woodward AW, Bartel B (2005) Auxin: regulation, action, and interaction. Ann Bot 95: 707-735
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wuest SE, Vijverberg K, Schmidt A, Weiss M, Gheyselinck J, Lohr M, Wellmer F, Rahnenführer J, Mering C von, Grossniklaus U (2010) Arabidopsis female gametophyte gene expression map reveals similarities between plant and animal gametes. Curr Biol 20: 506-512
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wu G, Poethig RS (2006) Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. Development 133: 3539-3547
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wu M-F, Tian Q, Reed JW (2006) Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. Development 133: 4211-4218
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wynn AN, Rueschhoff EE, Franks RG (2011) Transcriptomic Characterization of a Synergistic Genetic Interaction during Carpel Margin Meristem Development in Arabidopsis thaliana. PLoS One 6: e26231
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Xu X, Zhang Y, Williams J, Antoniou E, McCombie WR, Wu S, Zhu W, Davidson NO, Denoya P, Li E (2013) Parallel comparison of Illumina RNA-Seq and Affymetrix microarray platforms on transcriptomic profiles generated from 5-aza-deoxy-cytidine treated HT-29 colon cancer cells and simulated datasets. BMC Bioinformatics 14: S1
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yadav RK, Girke T, Pasala S, Xie M, Reddy GV (2009) Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. Proceedings of the National Academy of Sciences 106: 4941-4946
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yadav RK, Tavakkoli M, Xie M, Girke T, Reddy GV (2014) A high-resolution gene expression map of the Arabidopsis shoot meristem stem cell niche. Development 141: 2735-2744
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E (2011) AGRIS: the Arabidopsis Gene Regulatory Information Server, an update. Nucleic Acids Res 39: D1118-D1122
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhang T-Q, Lian H, Tang H, Dolezal K, Zhou C-M, Yu S, Chen J-H, Chen Q, Liu H, Ljung K, et al (2015) An intrinsic microRNA timer regulates progressive decline in shoot regenerative capacity in plants. Plant Cell 27: 349-360
Pubmed: Author and Title
CrossRef: Author and Title

Google Scholar: Author Only Title Only Author and Title

Zhao S, Fung-Leung W-P, Bittner A, Ngo K, Liu X (2014) Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. PLoS One 9: e78644

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhao Y, Christensen SK, Fankhauser C, Cashman JR, Cohen JD, Weigel D, Chory J (2001) A role for flavin monooxygenase-like enzymes in auxin biosynthesis. Science 291: 306-309

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title