The Molecular Dialog between Flowering Plant Reproductive Partners Defined by SNP-Informed RNA-Sequencing

Alexander R. Leydon,a Caleb Weinreb,a Elena Venable,a Anke Reinders,b John M. Ward,b and Mark A. Johnsona

ªDepartment of Molecular Biology, Cell Biology, and Biochemistry, Brown University, Providence, Rhode Island 02912
bDepartment of Plant and Microbial Biology, University of Minnesota, St. Paul, Minnesota 55108-6106

ORCID IDs: 0000-0003-3034-1482 (A.R.L.); 0000-0001-6100-6084 (C.W.); 0000-0001-8463-0533 (E.V.); 0000-0001-5351-3940 (A.R.); 0000-0001-8185-0434 (J.M.W.); 0000-0002-8508-2724 (M.A.J.)

The molecular interactions between reproductive cells are critical for determining whether sexual reproduction between individuals results in fertilization and can result in barriers to interspecific hybridization. However, it is a challenge to define the complete molecular exchange between reproductive partners because parents contribute to a complex mixture of cells during reproduction. We unambiguously defined male- and female-specific patterns of gene expression during Arabidopsis thaliana reproduction using single nucleotide polymorphism-informed RNA-sequencing analysis. Importantly, we defined the repertoire of pollen tube-secreted proteins controlled by a group of MYB transcription factors that are required for sperm release from the pollen tube to the female gametes, a critical barrier to interspecific hybridization. Our work defines the pollen tube gene products that respond to the pistil and are required for reproductive success; moreover, we find that these genes are highly evolutionarily plastic both at the level of coding sequence and expression across A. thaliana accessions.

INTRODUCTION

Flowering plants have evolved a unique reproductive system that depends on a series of interactions between pollen and pistil that ensure two sperm cells are delivered to each ovule. A pair of genetically identical sperm develops within a pollen grain and is transported within the cytoplasm of the pollen tube, a highly polar extension of the pollen cell (Berger and Twell, 2011). The pollen tube penetrates and extends through floral tissue and responds to guidance cues (Márton et al., 2005; Okuda et al., 2009; Takeuchi and Higashiyama, 2012, 2016; Mizukami et al., 2016; Wang et al., 2016) that direct it into an ovule, where the pollen tube encounters a synergid cell, one of two supportive cells that flank the egg. The pollen tube then bursts just as one of the two synergids degenerates (Denninger et al., 2014; Ngo et al., 2014; Leydon et al., 2015); pollen tube rupture propels the pair of sperm to a position where they contact the egg and central cell (Hamamura et al., 2011; Ngo et al., 2014). Two simultaneous, or nearly simultaneous, plasma membrane fusion events then occur; each releases a sperm nucleus into either the egg or central cell cytoplasm (Denninger et al., 2014; Hamamura et al., 2014) for nuclear fusion. Double fertilization is the defining feature of flowering plants and initiates development of an embryo and a supportive endosperm, the major components of seeds.

To understand the molecular mechanisms that mediate pollen tube delivery of sperm and double fertilization, it is important to define the gene expression patterns of the critical cells as they interact with each other. Genomics of individual cells and tissues has provided key insights into the transcriptomes of pollen (Becker et al., 2003; Honys and Twell, 2003; Pina et al., 2005), embryo sac (Yu et al., 2005; Johnston et al., 2007; Jones-Rhoades et al., 2007; Steffen et al., 2007), pistil (Scutt et al., 2003; Tung et al., 2005; Swanson et al., 2005), sperm (Borges et al., 2008), and isolated egg and central cells (Dresselhaus et al., 1994; Wuest et al., 2010). Furthermore, comparing the transcriptomes of pollen tubes grown in vitro (Wang et al., 2008) with pollen tubes collected after they had grown through a pistil explant (semi-in vivo condition [SIV]; Qin et al., 2009) identified dynamic regulation of pollen tube transcription during growth through the pistil. However, none of these studies provide a comprehensive and integrated analysis of the timing and relationships between pollen tube and pistil cell types as they interact over the 8 h required for the pollen tube to reach an ovule. Importantly, microarray profiling of pollinated pistils unveiled dynamic changes over 8 h of pollen tube growth in vivo (Boavida et al., 2011); however, in this study, it was impossible to distinguish pollen tube from pistil gene expression.

In Arabidopsis thaliana, pollen tubes grow through a complex extracellular matrix and form hundreds of cell-cell contacts en route to the female gametes within an ovule. It is likely that numerous cell-cell recognition and signaling exchanges are required to promote successful fertilization by appropriate pollen genotypes and to block exotic pollen genotypes. For example, there is strong evidence that species-specific cell-cell interactions are required for pollen tube guidance (Márton et al., 2012; Takeuchi and Higashiyama, 2012, 2016; Wang et al., 2016) into the ovule and for pollen tube reception resulting in sperm release (Williams et al., 1986; Escobar-Restrepo et al., 2007; Leydon et al., 2014; Müller et al., 2016). How the pollen tube and pistil recognize each...
other and transduce identity information into alterations in the transcriptomes of critical cells is unknown. However, a group of three pollen tube–expressed MYB transcription factors have been shown to be induced by growth through the pistil and are required for the ability of pollen tubes to burst and release their sperm upon arrival at the female gametophyte (Leydon et al., 2013; Liang et al., 2013). These findings suggest a complex network of transcriptional regulation critically important for cellular interactions responsible for flowering plant reproduction.

We set out to identify sex-specific (pollen tube versus pistil) gene expression patterns during pollination when pollen tubes are enmeshed within the pistil tissue and actively engaged in signaling processes required for fertility. We sought a method that allowed us to define changes in pollen and pistil transcript abundance that was not invasive, did not require an in vitro system, and did not use transgenic markers. A major challenge for genomics has been to develop techniques that define the transcriptomes of individual cells or cell types within complex tissues. This type of analysis was pioneered in the A. thaliana root using cell-type–specific marker lines that facilitated cell sorting followed by microarray analysis (Birnbaum et al., 2005). More recently, expression of pollen tube–specific tagged ribosomal protein facilitated the biochemical isolation of pollen tube transcripts associated with ribosomes (Lin et al., 2014).

We took advantage of the modular nature of pollen by applying pollen of one A. thaliana accession to the pistil of another. We reasoned that polymorphisms in transcripts would allow us to distinguish RNA-sequencing (RNA-seq) reads derived from the pollen tube from those expressed in the pistil. This approach was used to successfully differentiate monoallelic from biallelic gene expression during embryo and endosperm development (Aptula et al., 2011; Gehring et al., 2011; Nordborg et al., 2005). We created a custom database of Cvi SNPs that had been validated by genome sequence (Gehring et al., 2011; Nordine and Bartel, 2012; Pignatta et al., 2014). More recently, this approach was applied to analyze patterns of gene expression associated with pollen–pistil compatibility in interspecific crosses of tomato (Solanum lycopersicum) species (Pease et al., 2016). Using this approach, we have analyzed pollen tube and pistil gene expression in crosses between Cape Verde Island–0 (Cvi) pistils and Columbia–0 (Col) pollen. Interestingly, analysis of gene expression in hybrid embryos revealed that the onset of paternal allele expression varied depending upon the combination of parental accessions (Aptula et al., 2011; Nodine and Bartel, 2012; Del Toro–De León et al., 2014). Our application of this approach analyzes gene expression in reproductive tissues before formation of any hybrid cells (embryo and endosperm). Nonetheless, we were curious to determine whether pollen tube gene expression was altered depending on the genotype of the pistil through which it grew.

Our study provides a comprehensive analysis of pollen tube and pistil gene expression in their native context and has revealed transcriptional changes that occur specifically in either the pistil or pollen tube component of a gene expression profile: changes not identifiable by other methods. Furthermore, we applied this approach to define gene expression controlled by pollen tube MYB transcription factors that control pollen tube reception signaling (Leydon et al., 2013; Liang et al., 2013) and identified a wealth of small, secreted proteins that may be critical for signaling from the pollen tube to the pistil. Our analysis revealed differences in sex-specific gene expression patterns across accessions and prompted us to analyze the patterns of selection that have shaped the evolution of gene expression during pollination. We found a higher rate of polymorphisms in the coding sequence of pollen tube– and pistil–expressed genes compared with the background rate and that a family of MYB-dependent, secreted, cysteine-rich proteins is diversifying rapidly in A. thaliana via expansion and positive selection. These findings are consistent with the hypothesis that MYB-dependent pollen tube–expressed genes diversify rapidly and reinforce boundaries to interspecific hybridization by generating a species–distinct molecular signature.

RESULTS

A SNP Database to Determine Whether RNA-Seq Reads from Pollinated Pistils Are Derived from the Pollen or Pistil Parent

There are ~100 to 200 pollen tube cells and thousands of pistil cells in a pollinated pistil, so we estimated ~1% of the pollinated pistil transcriptome would be pollen tube derived. To maximize the likelihood that we would detect RNA-seq reads unambiguously defining pollen tube transcripts, we chose two A. thaliana accessions, Col and Cvi, with a high single nucleotide polymorphism (SNP) frequency (~1 SNP every 200 bp; Schmid et al., 2003; Nordborg et al., 2005). We created a custom database of Cvi SNPs derived from RNA-seq of unpollinated and self-pollinated Cvi pistils (8 h after hand self-pollination; Figure 1A). Variant analysis of unpollinated and self-pollinated Cvi pistils added 15,995 new SNPs to the previously identified 576,722 Cvi SNPs (Pignatta et al., 2014), resulting in a comprehensive variant list used in all subsequent analyses (1 SNP every 201 bp; Supplemental Data Set 1). As 36.2% of these SNPs occur in exons (coding sequences and untranslated regions), they are useful for determining the parent of origin of individual RNA-seq reads from pollinated pistils (Figure 1B; Supplemental Figure 1). Importantly, 75% of annotated A. thaliana genes had at least one SNP in a transcribed gene region, providing the potential to quantify pollen tube and pistil transcript abundance unambiguously (24,441/32,678 = 74.7%).

Eight Percent of Total mRNA within the Pollinated Pistil Is Derived from the Pollen Tube

We used bioinformatic approaches previously developed to define allele–specific gene expression in the developing embryo and endosperm (Gehring et al., 2011; Pignatta et al., 2014) to determine whether RNA-seq reads contained SNPs and whether reads with SNPs were derived from the pollen tube or the pistil (Supplemental Figure 2, analysis workflow; see Methods). We used methods that count reads mapping to a single genome location (HTseq) or that count reads mapping to multiple locations (EasyRNASeq; Supplemental Figure 2, Supplemental Data Sets 2 and 3, comparison of methods). In pollinated pistils, 12% of 131 million singly mapped reads analyzed contained SNPs (Figure
1C), and these reads were distributed across the transcribed regions of genes similarly to reads without SNPs (Supplemental Figure 3 and Supplemental Data Set 2). Inclusion of reads that were mapped to multiple locations did not significantly alter the proportion of reads with useful polymorphisms (Supplemental Data Set 2); however, multiply mapped reads facilitated enhanced detection of members of gene families expressed in pollen. So, multiply mapped reads were used for subsequent analyses.

These data suggest that in a typical experiment, ~17 million reads with SNPs are available to quantify patterns of transcript abundance in the pollen tube and pistil. Importantly, we determined that ~60% of genes expressed in the pollinated pistil have at least one SNP, allowing unambiguous determination of parent of origin (Figure 1) in crosses between Col and Cvi. This analysis shows that we will not be able to unambiguously determine the parent of origin for all genes using Cvi pistils pollinated with Col pollen. Further bioinformatic analysis of sequenced A. thaliana accessions showed that crosses between Col and Pu2-8 would provide the greatest number of additional useful SNPs (~20% more genomic features; Supplemental Figure 4).

Figure 1. Pollen- and Pistil-Specific Gene Expression Can Be Distinguished in Pollinated Pistils Using SNPs in RNA-Seq Reads.

(A) Experimental samples prepared for RNA-seq.
(B) The distribution of SNPs across indicated gene regions in our Cvi versus Col SNP database (percentage is given above bars; black represents transcribed regions; gray represents untranscribed regions of transposable elements [TE]).
(C) Reads from three biological replicates of a pollinated pistil RNA-seq experiment were analyzed to determine the number with Col or Cvi SNPs; the numbers of total and sex-specific reads are given. Venn diagram is not proportional to number of included elements.
(D) and (E) Scatterplots of mean reads/gene (three biological replicates).
(D) Pistil reads (y axis) are plotted against pollen reads (x axis) for all genes with SNPs detected in pollinated pistils. The y axis is 10-fold greater than the x axis.
(E) Total reads/gene are plotted for unpollinated pistils (y axis) and pollinated pistils (x axis). All genes are plotted as colored dots indicating sex-specific expression: blue, pollen specific; red, pistil specific, gray, genes without SNPs.
(F) Quantification of genes expressed in pollen, pistil, or both tissues. Supporting bioinformatic analysis is shown in Supplemental Figures 1, 3, and 4.
To determine the relative contributions of pollen tube and pistil to the pollinated pistil transcriptome (Figure 1A), we analyzed the total number of reads from Cvi pistils pollinated with Col pollen that contained SNPs (Figure 1C) and found that 92% of these reads had the Cvi allele (pistil reads), while 8% had the Col allele (pollen tube reads; Figure 1C). These data show that pollen tube RNA constitutes ~9% of the total RNA extracted from a pollinated pistil. Moreover, a scatterplot comparing the number of pistil reads (Cvi, y axis) with pollen tube reads (Col, x axis) for all genes with SNPs shows that pistil transcripts are much more abundant for the vast majority of genes (y axis is 10-fold greater than x axis). This is not surprising given the abundance of pistil cells compared with pollen tubes. However, genes with expression values along the x axis point to a significant fraction only expressed from the pollen tube and readily detected using this method (Figure 1D). Analysis of these data defines groups of pollen-expressed and pistil-expressed genes.

Sex-Specific Transcript Analysis Reveals the Parent of Origin of Genes Expressed during Pollination

A scatterplot comparing the number of total reads/gene in unpollinated pistils with pollinated pistil identifies genes with expression patterns that changed upon pollination and highlights a group of genes that are only detected upon pollination (Figure 1E, points along x axis). The simplest hypothesis is that these genes are expressed in the pollen tube; however, it is possible that some of these genes are pistil-expressed genes induced in response to pollen tube growth. To distinguish between these possibilities, we color-coded the scatterplot as follows: pollen tube-expressed genes are shaded blue (>10 average pollen reads), pistil-expressed reads are shaded red (>10 average pistil reads), and genes expressed in both tissues (>10 average pollen and pistil reads) are assigned a color between red and blue based upon the ratio of pistil to pollen reads (Figure 1E; Supplemental Data Set 3). Pollen gene expression is primarily responsible for the group of genes that are highly induced during pollination (Figure 1E, blue dots).

To define pistil genes that responded to pollination, we focused on reads containing Cvi SNPs in unpollinated (Cvi) versus pollinated pistils (Cvi x Col; Supplemental Figure 2). We used the DESeq2 software package (Love et al., 2014), which uses read counts and estimates variance-mean dependence and tests differential gene expression using the negative binomial distribution and does not normalize expression to the length of genes within libraries. This feature of DESeq2 was advantageous because we wanted to compare the expression of individual genes across conditions using the subset of reads that included SNPs. While a majority of pistil-expressed genes remain unchanged by pollination (Figure 1E, red dots along diagonal), it is clear that a number of pistil-expressed genes (pistil reads only) are differentially expressed in response to pollination; 476 genes are significantly below the diagonal (induced by pollination) or above the diagonal (higher in unpollinated pistils, DESeq2; see Methods; Supplemental Data Set 4) and could play roles in modulating the pistil environment to promote directed pollen tube growth.

To identify pollen-specific and pistil-specific genes, we normalized the number of sex-specific reads over each gene to the number of SNPs/gene. Then, we applied a 2-fold enrichment criterion for pollen-specific genes: >20 Col reads/SNP/gene, and <10 Cvi reads/SNP/gene (average of three replicates). The reciprocal criteria were used to identify pistil-specific genes. Normalization by SNPs/gene accounts for the large variation in the SNP frequency across genes (1 to >300; Supplemental Data Set 3). These stringent thresholds identified 277 genes expressed specifically from the pollen tube and 6838 genes expressed specifically from the pistil (Figure 1F). The ~25-fold higher number of pistil-specific genes is due, at least in part, to the vast excess of pistil tissue.

To determine whether SNP-informed determination of sex-specific gene expression identifies genes known to be expressed in the pollen tube or pistil, we analyzed genes defined in published microarray data. We graphed read coverage over genes and color-coded reads to indicate whether they contained a pollen- or pistil-derived SNP (red, pistil SNP; blue, pollen SNP) or whether they had no SNP (gray, Figures 2A to 2F). AT1G747530, AT2G13690, and AT5G46800 had been shown to be abundant in pistils (Boavida et al., 2011), but absent or minimally detectable in pollen grain and pollen tube microarray experiments (Becker et al., 2003; Honys and Twell, 2003; Pina et al., 2005; Qin et al., 2009). All of the reads with SNPs carried the Cvi allele for these genes (Figures 2A to 2C, red), indicating that in pollinated pistils all of the RNA transcribed from these genes is transcribed in pistil cells.

AT1G29140, AT1G55560, and AT2G26450 were shown to be expressed in SIV-grown pollen tubes (Qin et al., 2009) but were absent in unpollinated pistils (Boavida et al., 2011). We found that all of the reads with SNPs carried the Col allele for these genes (Figures 2D to 2F, blue), indicating that in pollinated pistils all of the RNA transcribed from these genes is transcribed in pollen tubes.

Finally, we took the 10 most highly expressed genes detected in microarray analysis of pollinated pistils (Boavida et al., 2011) and graphed the distribution of reads for SNP-containing genes (≥10 genes contained useful SNPs). Four representative genes are shown (AT4G10340, AT4G35100, AT5G02380, and AT1G72280; Figures 2G to 2J). These genes highlight one of the chief advantages of SNP-informed RNA-seq analysis of reproductive gene expression; they clearly show that the male and female components of gene expression can be determined unambiguously for genes expressed in both the pollen tube and the pistil.

SNP-Informed RNA-Seq Correlates with Previously Defined Reproductive Tissue Transcriptomes

Comparison of sex-specific gene expression defined by SNP-informed RNA-seq with previous microarray analysis of component cell and tissue types indicates overall similarity, but also identifies genes missed in prior analyses. One of the principal advantages of SNP-informed RNA-seq analysis of pollination is that the pollen tube and pistil components of gene expression can be determined without invasive procedures that potentially distort native expression patterns. Prior to this analysis, it was necessary to piece together male and female gene expression patterns from experiments that interrogated pollen (Honys and Twell, 2004; Wang et al., 2008; Qin et al., 2009) and pistil separately (Swanson et al., 2005; Boavida et al., 2011). We performed a Spearman’s correlation analysis to test whether sex-specific transcriptomes
determined using SNP-informed RNA-seq correlate with those of component cell and tissue types previously determined by microarray analysis (Figure 2K; Supplemental Figure 5). Correlations with microarray experiments were limited because the lists of gene identifiers are not completely overlapping. Nonetheless, we found that in Cvi x Col pollinations, pollen-expressed genes (reads containing Col SNP) demonstrated a high correlation with pollen transcriptomes, especially SIV-grown pollen tubes (Figure 2K, left column; dark red is the highest Pearson coefficient observed). Conversely, genes from Cvi x Col pollinations identified as

Figure 2. SNP-Informed RNA-Seq Defines the Male and Female Contributions to Gene Expression during Reproduction.

(A) to (F) Plots of reads over gene models at representative loci. Coverage (y axis) = log₁₀ number of reads counted over each nucleotide. Gray color represents reads without SNPs. In regions with SNPs (green line), reads are divided into sex-specific components: red = reads with the Cvi SNP (pistil parent) and blue = reads with the Col SNP (pollen parent).

(A) to (C) Pistil-expressed genes.

(D) to (F) Pollen-expressed genes.

(G) to (J) Genes that are abundant in pollinated pistils and are expressed in both pollen and pistil. Exons are represented by black bars, introns by black lines, and polarity of transcription by arrowheads.

(K) Pearson’s R correlation values are graphed with the indicated scale comparing pollen tube- and pistil-expressed genes defined by SNP-informed RNA sequencing of pollinated pistils with previously published microarray experiments (references given in text). Only genes with one unambiguous microarray ID were analyzed. Supporting bioinformatic analysis is shown in Supplemental Figure 5.
pistil-expressed (reads containing Cvi SNP) demonstrated a high correlation with pistil transcriptomes, especially pistils pollinated for 3.5 or 8 h (Figure 2K, right column). As expected, pistil-expressed genes defined by SNP-informed RNA-seq did not correlate with previously identified pollen development, tube growth (in vitro and SIV pollen tubes), or vegetative tissue profiles (Figure 2K, right column). Therefore, SNP-informed RNA-seq analysis provides a noninvasive, in vivo analysis of gene expression that occurs while the pollen tube is growing through pistil tissue and thus provides a more biologically relevant view of gene expression during pollen tube-pistil interactions.

We compared pollen tube-expressed genes defined by SNP-informed RNA-seq analysis with lists of genes previously defined by microarray analysis of SIV-grown pollen tubes (Qin et al., 2009), unpollinated and pollinated pistils (Boavida et al., 2011), and by translating ribosome affinity purification of pollen tubes growing through pistil tissue (Lin et al., 2014). All comparisons are reported in Supplemental Data Set 5. The SIV transcriptome included 7275 expressed genes, whereas SNP-informed RNA seq identified 4375 genes with RNA-seq reads that included the Col SNP (pollen reads); the overlap was 28% (2540 genes). This discrepancy was not due to a lack of SNP coverage: Only 31 SIV genes were not detectable because they lack a Col-Cvi SNP. Conversely, we identified a much larger group of genes (4704) with useful SNPs that were not detected as pollen tube-expressed genes in vivo but that were detected as pollen tube-expressed genes using the SIV method. Some of these genes may have been induced by the manipulations of the SIV method in which pollen tubes exit pistil tissue and are exposed to air while growing on agar media (Qin et al., 2009). We also identified 1735 genes with clear evidence for pollen tube expression in the pollinated pistil (>10 Col reads) that were not detected by the SIV method. These newly identified pollen tube expressed genes include genes not present on the ATH1 microarray and genes that require native interaction with pistil tissue for their expression.

We found substantial overlap between the 267 pollen tube-specific genes identified by SNP-informed RNA-seq (Supplemental Data Set 5) and 328 previously identified pollination-induced genes (Boavida et al., 2011; overlap = 58%; 123 in common). This is not surprising given that both analyses began with RNA isolated from intact pollinated pistils. By contrast, there was no overlap between the genes defined as pollen tube specific by SNP-informed RNA-seq and by translating ribosome affinity purification (n = 469; Lin et al., 2014; Supplemental Data Set 5). Comparison of all genes identified as pollen expressed in our study (n = 3939) with the pollen tube-specific genes identified by translating ribosome affinity purification identified a modest overlap of 69 genes. The higher degree of overlap between our method and microarray analysis of SIV-grown pollen tubes (Qin et al., 2009) and pollinated pistils (Boavida et al., 2011) is likely because these techniques begin with total RNA rather than RNA associated with a tagged ribosomal subunit that must compete with endogenous ribosomal subunits (Lin et al., 2014).

The Sex-Specific Components of the Pollination Transcriptome Comprise Distinct Functional Categories

We found that pollen tube-specific genes were enriched for Gene Ontology (GO) terms suggesting functions in pectin modification (14-fold; P value = 2.04 × 10^{-8}, all GO enrichment P values were determined using the Benjamini method), cell-cell signaling (25-fold; P value = 1.4 × 10^{-7}), calcium-mediated signaling (14-fold; P value = 1.23 × 10^{-9}), and pollen tube growth (8-fold; P value = 5.5 × 10^{-5}). Analysis of GO cellular component annotations revealed 27.9% of transcripts encoded genes predicted to enter the endomembrane system (n = 74; P value = 2.72 × 10^{-10}), which includes nine members of the rapid alkalization factor gene family (27-fold; P value = 1.18 × 10^{-7}; Supplemental Data Set 6), which have been found to participate in extracellular signaling (Hanuta et al., 2014; Du et al., 2016; Stegmann et al., 2017). On the other hand, there were 476 pistil-expressed genes that were differentially regulated upon pollination: 216 were upregulated upon pollination, 259 decreased abundance upon pollination (Supplemental Data Set 7). Secreted glycoproteins (n = 38; P value = 7.25 × 10^{-12}) and secreted cell wall localized proteins (n = 26, P value = 6.31 × 10^{-9}) both increase upon pollination as well as a small but significant number of genes involved in response to chitin (n = 4, P value = 0.001) and response to auxin (n = 8, P value = 0.005). These enriched GO categories may indicate parallels between the pistil’s response to pollen tube growth and previously characterized biotic stress responses. By contrast, there is widespread decrease in expression of cytochrome p450 (n = 12; P value = 4 × 10^{-9}) and oxidative stress genes (n = 10, P value = 4 × 10^{-4}) upon pollination. Of the 476 pistil-expressed genes, 168 were found to be pistil specific by our criteria, of which 46 are predicted to be secreted (Supplemental Data Set 7). These observed changes in expression likely indicate a major shift in the transcriptional program of female tissue during pollination.

**SNP-Informed RNA-Seq Analysis Defines the Role of MYB Transcription Factors in Regulating Pollen Tube Gene Expression and the Pistil’s Response to Pollination**

Pollen tubes lacking three closely related MYB transcription factors (Stracke et al., 2001), MYB97, MYB101, and MYB120, fail to release their cargo of two sperm upon entering an ovule. Consequently, myb triple-mutant pollen tubes continue to grow within the female gametophyte producing a coil of pollen tube that fails to fertilize female gametes (Leydon et al., 2013; Liang et al., 2013). It was proposed that pollen tube MYBs respond to the pistil environment and in turn activate the expression of genes critical for pollen tube reception signaling (Leydon et al., 2013, 2014), the interaction between the male and female gametophyte that results in sperm release (Kessler and Grossniklaus, 2011). This model was explored using microarray analysis of triple-mutant pollen grains (Liang et al., 2013) or pollinated pistils (Leydon et al., 2014). It is impossible to know the parent of origin of transcripts using microarray analysis, so neither of these approaches was able to determine the pollen- and pistil-expressed genes with altered transcript abundance in mutant compared with wild-type pollinations. Importantly, microarray analysis failed to identify differential expression of the pollen component of a gene expressed in both the pollen and the pistil because changes in the pollen tube component were masked by the much larger pool of pistil transcripts. SNP-informed RNA-seq provides an opportunity to determine the changes in gene expression due to loss of critical components of the pollen tube.
Figure 3. SNP-Informed RNA-Seq Defines MYB-Responsive Genes in the Pollen Tube and Pistil.

(A) Experimental samples prepared for RNA-seq.
(B) and (C) Scatterplots of all average reads (ignoring SNPs) in wild type versus myb triple-mutant pollinations. (B) Differentially expressed genes identified before division of reads based upon SNPs are indicated (purple, 318 genes, adjusted P value < 0.05). (C) The color code of the scatterplot in (B) was altered to indicate genes with significant changes in the pollen (reads with Col SNP, light blue, n = 278, P value < 0.05, DESeq2) or pistil (reads with Cvi SNP, pink, n = 34, adjusted P value < 0.05) components of their expression profile.
(D) A comparison of the differentially expressed genes identified with or without SNP analysis (P value < 0.05, DESeq2).
transcriptional regulators of the pollen tube cell as it interacts with the pistil tissue it has evolved to navigate.

First, we compared Cvi x Col pollinations to Cvi x myb triple mutant (accession = Col) pollinations and analyzed differential expression using all mapped reads (Figures 3A and 3B). This analysis is similar to the previous microarray analysis because it ignores the parent-of-origin of transcripts. As expected, RNA-seq identified the majority of the differentially expressed genes previously identified by microarray analysis of pollinated pistils (~63%, 30/48, DESeq2; Leydon et al., 2013); however, the overlap with microarray analysis of MYB-dependent gene expression in pollen grains was much lower (16%, 4/25, DESeq2; Liang et al., 2013). This difference between the MYB regulatory networks of the pollen grain and the pollinated pistil is consistent with our previous observation that pollen tube MYBs are induced by growth through the pistil and likely regulate unique targets as the pollen tube grows through pistil tissue (Leydon et al., 2013). Importantly, RNA-seq analysis uncovered an additional 288 MYB-dependent genes (318 total, DESeq2; Supplemental Figure 6 and Supplemental Data Set 8; Figure 3B). Genes with significantly lower expression in myb triple mutants were enriched in several functional categories including transmembrane transporters, carbohydrate active proteins, and secreted proteins, as was previously defined (Supplemental Data Set 8; Leydon et al., 2013).

Next, we implemented sex-specific read identification on RNA-seq data from wild-type and mutant pollinations to define the patterns of differential expression resulting from loss of MYB97, MYB101, and MYB120 in pollen tubes and the pistil. This analysis identified 278 pollen-expressed genes and 34 pistil-expressed genes that are differentially expressed in myb triple-mutant pollinations (Supplemental Data Set 9). When the mutant versus the wild type total read scatterplot (Figure 3B) is color-coded to highlight genes with significant changes in the pollen tube (blue) or pistil (pink; Figure 3C) component of their expression profile, it becomes clear that the total abundance of many potentially MYB-regulated transcripts does not change even though the pollen component does (blue dots along diagonal). Interestingly, this analysis also defines a small number of pistil-expressed genes that change significantly when the pistil is pollinated by mutant compared with wild-type pollen tubes (pink dots; Figure 3C).

By comparing the differentially expressed genes identified using total reads (ignoring parent-of-origin) with those determined following analysis of sex-specific reads, we find that 218 differentially expressed genes were unmasked when differential gene expression was determined for pollen and pistil reads independently (Figure 3C). This finding shows that SNP-informed RNA-seq is successful at defining changes in the pollen- and pistil-specific components of gene expression for those genes expressed in both tissues and indicates that many genes are under regulatory control in the pollen tube that is distinct from that of the pistil. However, since we used a different bioinformatic procedure to initially align total reads than that used to process reads with SNPs (see Methods), we repeated this analysis and compared sex-specific reads with all reads lacking a SNP; the result was similar to our previous analysis (Figure 3D). Many pollen tube-expressed genes were determined to be differentially expressed in myb triple mutants whether or not reads with SNPs were analyzed (Figure 3E, dark-blue intersection, n = 68); however, an additional 210 genes are identified by determining whether pollen reads (Col SNP) are differentially abundant (P < 0.05, Wald test, DESeq2) in the mutant compared with the wild type. Pollen tube-specific MYB-dependent genes comprise several functional annotation categories, the most enriched being transmembrane transporters (n = 26, P value = 4.46 × 10^-5, Benjamini test; Supplemental Data Set 9). By contrast, only a few pistil-expressed genes (34 genes; Figure 3E) were found to be differentially expressed in mutant compared with wild-type pollinations. Thirteen of these were also identified upon analysis of pistil (Cvi) reads. This small but interesting gene set is enriched in chitin-responsive genes (n = 6, P value = 9.08 × 10^-5, Benjamini test) and may represent pistil genes that respond to myb mutant pollen tubes, which are perceived as different from the wild type (Supplemental Data Set 9).

To illustrate the consequence of MYB loss of function on the pollen tube component of gene expression, we plotted read coverage for two genes identified as differentially regulated in myb triple-mutant pollen tubes (Figure 3F; Supplemental Data Set 9). In wild-type pollinations (Figure 3F, top graph) the RNA-seq reads covering SNPs comprise both pistil (Cvi, red) and pollen (Col, blue) reads, with a 10-fold greater contribution from the pistil. In myb triple-mutant pollinations, the pollen component (Col, blue) disappears, yet the much larger pistil component is unchanged. This analysis shows the utility of SNP-informed RNA-seq analysis of reproduction, which facilitates independent analysis of the pollen tube and pistil components of gene expression as the pollen tube interacts with the pistil.

**Genetic Reporter Constructs Corroborate SNP-Informed RNA-Seq Analysis**

We produced transgenic plants expressing GFP fusion constructs to determine whether bioinformatic separation of pollen tube

---

Figure 3. (continued).

(E) Differential expression analysis was performed again comparing genes identified after analysis of reads without SNPs with those identified based on differential abundance (P value < 0.05, DESeq2) of pollen (Col) or pistil (Cvi) reads. Differentially expressed genes identified by both analyses are highlighted in dark colors, and genes only identified by SNP-informed RNA-seq are highlighted in light colors (blue, pollen reads; red, pistil reads). The total number of differentially expressed genes is shown next to each category, along with the number that were downregulated (down) and upregulated (up) in mutant compared with wild-type pollinations.

(F) Representative read coverage graphs (as in Figure 2) of MYB-dependent genes (AT1G04310 and AT1G32470). Top row, Cvi x Col-0 (wild-type) pollinations; bottom row, Cvi x myb triple-mutant pollinations. Supporting bioinformatic analysis is shown in Supplemental Figure 2. Supporting bioinformatic analysis is shown in Supplemental Figure 6.
reads from pistil reads accurately identifies pollen tube-specific genes. We chose three MYB-dependent genes for which only the pollen (Col) allele was detected in reads over SNPs and were thus deemed pollen tube specific (Figures 4A, 4H, and 4O). We proposed that MYB expression, and thus target gene activation, is induced by growth through pistil tissue (Leydon et al., 2013, 2014). To test this hypothesis and to determine whether RNA-seq analysis was validated by reporter construct analysis, we analyzed GFP fusion proteins in the SIV pollen tube growth assay and in wild-type pistils.

Full genomic loci containing the predicted promoter regions and coding sequences of Sucrose-Proton Symporter9 (SUC9), Pectin Lyase (PL), and Alpha Carbonic Anhydrase4 (ACA4) were cloned in frame with GFP to produce translational fusions (see Methods). We introduced these reporter constructs into Arabidopsis thaliana (Col) carrying the pollen-specific LAT52:DsRed marker (Twell et al., 1989; Francis et al., 2007), which drives expression of soluble DsRed in the pollen tube cytoplasm. Mature pollen from each reporter was analyzed by confocal microscopy, but GFP was not detectable, consistent with the hypothesis that these MYB-dependent genes are induced by growth through pistil tissue. By contrast, after analyzing at least 21 independent primary transformants (T1) for each construct in SIV growth experiments, we found that pollen-tube expression was observed in the vast majority of transformants (SUC9:SUC9:GFP, 36/36; PL:PL:GFP, 18/21; ACA4:ACA4:GFP, 32/36; Figures 4B, 4L, and 4P).

We determined the subcellular localization of each GFP fusion protein in growing pollen tubes and each displayed a localization pattern consistent with the predicted localization. SUC9:GFP appeared to localize predominantly to the plasma membrane and to vesicles in the pollen tube tip (Figure 4C; Supplemental Movie 1). PL:GFP was observed in the endomembrane system (puncta) and what is likely the pollen tube cell wall (Figure 4J; Supplemental Movie 2). ACA4:GFP was observed in the endomembrane system (puncta, Figure 4Q; Supplemental Movie 3). In all SIV experiments where GFP fluorescence was observed (86/93), expression was apparent in extending pollen tubes and not pollen grains.

To determine whether these genes are pollen tube specific, as they had been designated by SNP-informed RNA-seq, we examined three independent transformants for the expression of GFP unpollinated pistils using confocal microscopy. For each transgenic line, we observed only autofluorescence in the stigma, style, transmitting tract, and all other somatic tissues of the pistil (Figures 4D, 4K, and 4R). Surprisingly, we observed expression of SUC9:GFP in the synergid cells (Figures 4D and 4E; Supplemental Figures 7A, 7B, and 7I) and ACA4:GFP in all cells of the female gametophyte (Figures 4R and 4S; Supplemental Figures 7C, 7D, and 7I); as expected, PL:GFP was undetectable in any pistil tissue. We then analyzed self-pollinated pistils for each reporter construct (Figures 4F, 4M, and 4T). Each reporter was most strongly detected in pollen tubes, especially in the style region where pollen tubes are concentrated (arrows). Additionally, each reporter fusion was detected in ovules clearly targeted by a pollen tube (Figures 4G, 4N, and 4U), supporting a possible role for these MYB-dependent genes in pollen tube reception (Leydon et al., 2013).

Expression of SUC9:SUC9:GFP and ACA4:ACA4:GFP reporter constructs in the female gametophyte was surprising in light of our SNP-based RNA-seq results, where we detected no SUC9 or ACA4 Cvi SNP-containing transcripts. Transgenic lines were generated in the Col accession, whereas our RNA-seq used Cvi as the pistil donor. So, we tested whether expression of SUC9, PL, and ACA4 in the unpollinated pistil was accession-specific using RT-qPCR. Of the three genes tested, only SUC9 transcripts were detected at low levels in Col unpollinated pistils; none were detected in either Cvi or Ler (ms1) unpollinated pistils (Supplemental Figure 8). This observation is consistent with SUC9:SUC9:GFP expression being detected in synergid cells of transgenic Col plants, but not in Cvi, the female used in our RNA-seq experiments. Interestingly, ACA4 transcripts were undetectable in unpollinated pistils of all three accessions, suggesting that gametophytic expression of ACA4:ACA4:GFP may be due to inclusion or exclusion of cis-regulatory regions in the promoter region of this transgene. Pectate lyase transcripts were also undetectable in unpollinated pistils, consistent with our RNA-seq results.

The observation that SUC9 was expressed in the synergid cells of Col but not other accessions (Ler and Cvi), prompted us to test whether accession affected SUC9, PL, and ACA4 expression in self- and cross-pollinations. Col self-pollinations result in higher levels of each transcript compared with Cvi (Supplemental Figure 8). All three genes behaved similarly in Col and Ws self-pollinations, except for SUC9, which was induced at lower levels in Ws. Furthermore, Cvi x Col inter-accession pollinations induce SUC9, PL, and ACA4 transcripts at lower levels than Col self-pollinations, suggesting that pollen-pistil interactions, not pollen genotype alone, affect the strength of MYB-dependent pollen tube transcriptional activation in vivo (Supplemental Figure 8). This accession-specific activation of MYB-dependent genes led us to examine the role of accession in reproductive gene expression in our RNA-seq data sets.

**Gene Expression Changes Are Induced by Inter-Accession Pollination**

Col and Cvi are interfertile; however, inter-accession crosses show subtle defects in pollen tube reception that resemble pollen tube overgrowth defects in fer, ire, or myb triple-mutant pollinations (Escobar-Restrepo et al., 2007; Capron et al., 2008; Leydon et al., 2014; Müller et al., 2016). These observations along with the finding that pollen tube-expressed genes are differentially expressed between accessions prompted us to determine whether there are genome-wide, accession-specific changes in pollen and pistil gene expression. First, we used hierarchical clustering to assess whether the transcriptomes of inter-accession pollinations were distinct from intra-accession pollinations or unpollinated pistils. As expected, hierarchical clustering demonstrated that pollinated pistils are distinct from unpollinated Cvi pistils (Figure 5A). We compared the transcriptomes of self-pollinations (Cvi x Cvi and Col x Col) to inter-accession pollinations (Cvi x Col) and found a core set of 16,838 genes expressed at detectable levels in each pollination type (>100 reads; Figure 5B). However, differential expression analysis identified hundreds of genes with significantly different levels of expression between Col and Cvi self-pollinations (Figure 5C; Supplemental Data Set 10).
Figure 4. Pollen-Specific Gene Expression of Three MYB-Dependent Genes Detected by SNP-Informed RNA-Seq.

(A), (H), and (O) Sequencing reads over SNPs identifying potential MYB target genes as pollen specific (as in Figure 2).

(B) to (F) SUC9:SUC9:GFP (AT5G06170) and LAT52:DsRed.

(I) to (M) PL or PL:PL:GFP (AT5G09280) and LAT52:DsRed.

(P) to (T) ACA4:ACA4:GFP (AT4G20990) and LAT52:DsRed.

(B), (I), and (P) Representative SIV pollen tube growth experiments on ms1 pistil explants (epifluorescent micrographs). In each, the left image is the red channel (LAT52:DsRed) signal, which is present in the pollen grain and the tube. The right image is the green channel (GFP) signal. The autofluorescence of the ms1 stigma explant is visible in each panel. Bars = 500 μm.

(C), (J), and (Q) Confocal micrographs of SIV-grown pollen tubes: left panel, GFP; middle panel, DsRED; right panel, merge. Bars = 20 μm.

(D), (K), and (R) Representative confocal optical sections of unpollinated pistils; the merge of differential interference contrast and GFP channels is shown. Arrows indicate expression of the GFP fusion protein in the female gametophyte. Bars = 100 μm.

(E), (L), and (S) Close up of unpollinated ovules from (D), (K), and (R). Bars = 100 μm, projection of red and green channels.

(F), (M), and (T) Projections of confocal micrographs of pollinated pistils. Red and green channels are merged, and arrows indicate regions of strong coexpression of DsRed and the indicated GFP fusion protein (yellow) in the style. Twelve hours after pollination. Bars = 100 μm.

(G), (N), and (U) Close-up of ovules targeted by GFP reporter constructs from (F), (M), and (T). Bars = 100 μm.

Supporting confocal micrographs are shown in Supplemental Figure 7. Supporting RT-qPCR analysis is presented in Supplemental Figure 8.
Figure 5. Reproductive Gene Regulation Is Highly Divergent between Col and Cvi.

(A) Hierarchical clustering of all RNA-seq samples identifies accession-specific clusters (DESeq2).
(B) Intersection of all expressed genes (mean expression >100 reads) from self and inter-accession crosses. Venn diagram is not proportional to number of included elements.
(C) Differential gene expression observed between Col and Cvi 8 h after self-pollination. Differentially expressed genes are colored purple (adjusted P value < 0.005, DESeq2).
(D) to (F) Heat maps of Col-specific gene expression; colors represent read volume, with scales below indicating the number of reads for color value.
(D) Defensin-like genes.
(E) Low-molecular-weight cysteine-rich genes.
(F) Toll-interleukin receptor-nucleotide binding site-leucine rich repeat (TIR-NBS-LRR) type disease resistance proteins.
(G) Differential gene expression between self-pollinated Col and Cvi pistils. A total of 2877 differentially expressed were identified following pollination; 1640 transcripts were higher in inter-accession (Cvi x Col) pollinations, and 1238 transcripts were higher in self-pollinations (Cvi x Cvi). Adjusted P value < 0.005 (DESeq2).
(H) Sex-specific transcript identity of (A). Red dots, Cvi-specific reads; blue dots, Col-specific reads; gray dots, no SNPs for assignment.
Accession-specific transcriptional responses have been previously observed (Autran et al., 2011; Nodine and Bartel, 2012; Del Toro-De León et al., 2014), so it is not surprising that Col and Cvi accessions display variation in transcriptional profiles of reproductive cells and tissues. Interestingly, we identified members of gene families that were enriched in the Col pollination transcriptome that were absent from Cvi pollinations such as small secreted defensin-like proteins (Figure 5D), low molecular weight cysteine-rich proteins (Figure 5E), and NBS-LRR genes (TIR-NBS-LRR; Figure 5F). By contrast, analysis of genes enriched in Cvi pollinations, but absent in Col, did not show enrichment for any particular GO enrichment or gene families (see Methods). However, despite the lack of enriched GO terms suggesting roles in extracellular signaling, further analysis revealed that 44% of Cvi-specific pollination genes were predicted to contain signal peptides, similar to ~31% found in Col-specific pollinations (SignalP). These data suggest that Cvi may express its own repertoire of secreted proteins during pollination, but that unlike Col, the gene families being expressed are not as well annotated.

**The Pistil Transcriptome Responds to the Pollen Genotype during Pollination**

We set out to determine whether the pistil transcriptome changes in response to pollen accession in the pollinated pistil. Our analysis of *myb* triple-mutant pollinations suggested the pistil responds differently to mutant versus wild-type pollen tubes (Figures 3C and 3E), and we reasoned the pistil may also mount unique responses to different pollen accessions. We defined pollen tube and pistil changes in gene expression (differentially expressed genes) that occur during intra-accession (Cvi x Cvi) versus inter-accession (Cvi x Col) pollinations (Figure 5F, purple dots). First, we ascertained genes that were differentially expressed between self-crosses (Cvi x Cvi) and inter-accession crosses (Cvi x Col; Figure 5G; Supplemental Data Set 10). This analysis identified 2877 differentially expressed genes; 1238 are significantly more abundant in intra-ecotypic pollinations, 1640 are significantly higher in inter-ecotypic pollinations (632 up in inter-accession pollinations and 418 up in intra-accession pollinations). We then determined the parent of origin for these genes and identified 1050 pistil-expressed genes that are differentially expressed depending on the pollen genotype (Figure 5H, red dots; Supplemental Data Set 10). Most of the accession-sensitive transcriptional changes were subtle (<4-fold). The 632 pistil genes that were differentially up-regulated in inter-accession pollinations were enriched for responses to various molecules (22.38 enrichment score for this cluster of responses), including chitin (*n* = 37, *P* = 2.6 × 10⁻²⁵), carbohydrates (*n* = 39, *P* = 2.26 × 10⁻²⁰), and organic substances (*n* = 83, *P* = 1.2 × 10⁻¹⁵). Transcription factors were also enriched (*n* = 61, *P* = 1.6 × 10⁻⁶). Interestingly, genes induced in the inter-accession (Cvi x Col) pistil, included 14 ethylene-responsive transcription factors that are implicated in pathogenesis-related transcriptional response (*n* = 14, *P* = 4.7 × 10⁻³; Supplemental Data Set 11). The 418 pistil-expressed genes that were enriched in intra-accession crosses (Cvi x Cvi) were not strongly enriched for these or other functional categories. This subtle difference in transcriptional profiles indicates that pollen accession is perceived by the pistil, activating genes associated with defense responses during inter-accession pollinations.

**Expression of a Large Family of Thionin-Like CRPs Is MYB Dependent and Accession Specific**

Previous analysis of *myb* triple-mutant pollinations using microarray analysis had identified a single thionin-like gene (AT5G36720) as being among the most differentially regulated between mutant and wild-type pollinations (Leydon et al., 2013). This prompted us to propose that these proteins may be secreted by pollen tubes to female gametophytes to mediate pollen tube reception signaling. However, analysis of this family (cysteine-rich protein 2460 [CRP2460]; Silverstein et al., 2007) indicated that many members of this large group of genes were not annotated and therefore were not represented on the ATH1 array. We analyzed this family using our RNA-seq data to determine whether multiple members were potentially MYB regulated and to determine whether their expression varied among accessions. We plotted the RNA expression of each family member alongside the CRP2460 phylogenetic tree (Figure 6; Supplemental Data Set 12). Differential expression analysis identifies 22/28 CRP2460 transcript levels as statistically lower in pollinations with *myb* triple-mutant pollen tubes (Figure 6B, asterisk) compared with the wild type. CRP2460 thionin proteins were expressed accession preferentially: High expression was observed in Col pollen tubes, while Cvi pollen tubes expressed fewer members and at lower levels (Figure 6B). Of the 28 thionin-like proteins analyzed, 15 contained polymorphisms that allowed us to determine whether expression of these genes was from the pollen tube or pistil (Figure 6B, +). In each case, all reads over SNPs carried the Col allele, indicating that thionins are pollen tube expressed (Figure 6C). Thionin proteins therefore represent a pollen tube expression signature that differentiates *A. thaliana* accessions from each other and may be involved in signaling between the pollen tube and the female gametophyte important for sperm release.

**MYB-Dependent, Pollen Tube-Expressed Genes Show Elevated Polymorphism Rates, Gene Family Expansion, and Positive Selection**

*myb* triple-mutant pollen tubes fail to release sperm and continue to grow within the female gametophyte (Leydon et al., 2013; Liang et al., 2013). This defect, first identified in the *A. thaliana* feronia mutant (Huck et al., 2003; Rotman et al., 2003), had been characterized as a prezygotic barrier to hybridization in interspecific cross-pollinations (Williams et al., 1986; Escobar-Restrepo et al., 2007; Müller et al., 2016) and suggests that MYB-dependent genes may be involved in shaping the species identity of the pollen tube that is recognized by the female gametophyte (Leydon et al., 2014). Molecules involved in gametegamete recognition have been shown to evolve rapidly and to be responsible for species-level discrimination in mating interactions (Metz et al., 1998; Clark et al., 2006). Interestingly, it was recently shown that the pollen tube expresses a larger fraction of species-limited genes than other tissues (Arunkumar et al., 2013; Gossmann et al., 2014), which is consistent with the idea that pollen tubes may express...
Figure 6. MYB Transcription Factors Control Expression of Thionin CRPs (CRP2460).
species-limited proteins that determine their species-level identity. To address whether groups of genes expressed in reproductive tissues maintain higher rates of sequence diversity, we calculated Col versus Cvi exonic SNP rates and compared them to control rates simulated using 100 sets of randomly selected genes of the same number. We used the ratio of the measured SNP rate to the simulated rate to identify groups of genes with elevated SNP rates (Table 1; Supplemental Data Set 13).

Analysis of groups of genes defined by their expression pattern during pollination without determining parent of origin (analysis of all reads) revealed that the ~10,000 genes expressed in the pollinated pistil have SNP rates that are significantly below the simulated rate. However, genes induced by pollination contain significantly higher control rates (Table 1). The 292 genes that are differentially expressed in myb triple-mutant pollinations (identified from all mapped reads) do not show elevated SNP rates. When analyzed in bulk, we only observe increased polymorphism rates in genes induced by pollination.

However, when sex-specific reads are analyzed to focus on changes occurring in pollen and/or pistil gene expression, three groups were found to have polymorphism rates elevated compared with the simulated rate: pollen-expressed genes, pistil-specific pollination-induced genes, and MYB-dependent genes (Table 1). To gain insight into MYB-dependent genes with elevated polymorphism rates, we focused on a MYB-dependent CRP gene family (Figure 6). We started by analyzing the NBS-LRR gene family, which was previously found to have high rates of major effect SNPs in A. thaliana accessions indicative of diversifying selection (Clark et al., 2007; Schmitz et al., 2013) and found a ratio of measured SNP rate to control rate (1.24; Table 1) that was significantly above one. A significantly high SNP rate (P < 0.01, Poisson distribution) was also observed for a group of 825 CRP encoding genes, but not for the thionin subgroup (n = 71). We also determined the frequency that SNPs in these groups of genes resulted in missense mutations. Interestingly, the thionin subclade of CRPs had the highest rate of missense mutations (Table 1; Supplemental Data Set 14), suggesting this group of genes may be under diversifying selective pressure.

To test for positive selection within thionins (CRP2200-2610), we first determined whether there were homologs for each gene in Arabidopsis lyrata, Populus trichocarpa, grape (Vitis vinifera), rice (Oryza sativa), and Physcomitrella patens using a recently published genome-wide analysis (Supplemental Data Set 15; Lloyd et al., 2015). Appreciable fractions of Actin and NBS-LRR gene families have homologs in distantly related species such as the monocot, rice, and the moss, P. patens (Figure 7A). By contrast, homologs were only identified in closely related A. lyrata for the vast majority of CRP2200-2610 genes (Figure 7A). For example, 30% of NBS-LRR genes have homologs in P. patens, whereas no homologs for CRP2200-2610 genes were identified (Lloyd et al., 2015). Moreover, only a small subset of the CRP2200-2610 genes have homologs in angiosperm species outside the Brassicaceae (P. trichocarpa, grape, and rice), suggesting a recent expansion of this group.

To profile diversification in these families, we plotted the mean Ka/Ks for genes with homologs; we found that Actin genes had lower rates than NBS-LRRs, consistent with the hypothesis that NBS-LRR genes are under diversifying selection, while the Actin genes are under purifying selection (Figure 7B). CRPs exhibit a Ka/Ks profile that is similar to NBS-LRR genes (Figure 7B). However, this analysis of mean Ka/Ks failed to highlight the elevated proportion of CRP gene family members (CRPs in general and CRP2200-2610 in particular) with a Ka/Ks above one in the comparison with closely related A. lyrata, which had the largest number of homologs (Figure 7C). This analysis revealed a dramatic expansion of the CRP2200-2610 thionin clade of CRPs in the Brassica lineage. Of particular interest is the CRP2460 (Figure 6) group, which has undergone an expansion in A. thaliana compared with the closely related A. lyrata. Only six of 28 genes within this group have homologs in A. lyrata, and none have homologs in any of the other species tested (Figure 6; Ka/Ks indicated). Furthermore, the Ka/Ks rate for At1g21835 compared with its homolog from A. lyrata was found to be above one (Ka/Ks = 1.06; Figure 6), suggesting selective pressure to maintain diversity of this gene. At1g21835 appears to be the founder of an expanded group of A. thaliana thionins because it is sister to 14 genes without homologs in A. lyrata (Figure 6; Supplemental Figure 9). Analysis of this expanded group revealed evidence for diversifying selection at specific sites within these ~100-amino acid proteins (Supplemental Figure 9 and Supplemental Data Set 15). These results are in keeping with our proposal that this group of secreted proteins convey species-unique information to the female gametophyte required for pollen tube reception.

**DISCUSSION**

Successful plant reproduction is the lynchpin of agricultural food production, so the molecular mechanisms responsible for pollen-pistil signaling pathways need to be established so they can be optimized for grain and seed production in a changing climate.
Moreover, it is clear that pollen tube-pistil interactions represent a key prezygotic barrier to interspecific hybridization; however, few molecules have been identified that facilitate sperm delivery in interspecific pollinations, while preventing sperm delivery in interspecific pollinations (Moyle et al., 2014). Determination of the pollen tube and pistil transcriptomes during pollination would facilitate identification of patterns of gene expression required for successful fertilization; however, this has been technically challenging because of the way that pollen tubes grow within complex pistil tissue. Here, we present a method to circumvent this challenge that takes advantage of SNPs to define pollen tube- and pistil-derived RNA-seq reads within the pollinated pistil. This method does not require transgenic plants or invasive procedures and is therefore translatable to crop species with sufficient diversity among cross-compatible accessions/cultivars. We validated this method using comparisons to previous data sets and by testing the expression patterns of pollen tube-specific genes using transgenic reporter constructs. We have successfully defined the relative contributions of the pollen tube and pistil to the pollination transcriptome and reveal pistil-expressed genes that respond specifically to the genotype of the pollen tube. Moreover, we defined patterns of pollen tube gene expression controlled by MYB97, MYB101, and MYB120, pistil-induced pollen tube transcription factors required for successful sperm release. This analysis revealed that pollen tube MYB transcription factors are required for the expression of secreted proteins, including small cysteine-rich proteins that may play critical roles in signaling to female cells to initiate sperm release. Finally, we present evidence that patterns of pollen tube and pistil gene expression are variable between A. thaliana accessions and that pollen tube-expressed genes are more highly polymorphic than other groups of genes. These data provide a framework for further analysis of the patterns of gene expression that promote successful sperm delivery in intraspecific pollinations while preventing sperm delivery in interspecific crosses.

The modular nature of pollen and pistil, which facilitates pollination between polymorphic parents, combined with the development of RNA-seq techniques provided a means to comprehensively define pollen and pistil transcriptomes during pollination. A. thaliana accessions are of great utility for this approach because pollen and pistil can be easily crossed in any combination. We chose Cvi and Col as the initial parental accessions because of their relatively high polymorphism rate (−60% of the genes expressed in the pollinated pistil) and found this combination was sufficient to track patterns of pollen- and pistil-specific gene expression for −60% of the genes expressed in the pollinated pistil (Figure 1F). This leaves a significant fraction of genes whose parent of origin cannot be defined in this particular cross (see Supplemental Figure 4 for analysis of accession combinations providing additional coverage). Despite this limitation, the ease of experimental setup and noninvasive nature of the technique argue for this approach, which does not require the extensive postharvesting purification steps required for alternative biochemical approaches such as TRAP-seq (Lin et al., 2014; Kim et al., 2015) or isolation of nuclei tagged in specific cell types (INTACT; Deal and Henikoff, 2011). Our analysis of Cvi x Col cross-pollinations revealed that −8% of pollinated A. thaliana pistil RNA is derived from the pollen tube and we found strong correlations between our analysis and previous microarray analyses of the component cell types (microarray experiments recently reviewed in Rutley and Twell [2015]; Figure 2). Importantly, we identified hundreds of novel pollen tube- and pistil-expressed genes and showed that SNP-informed RNA-seq analysis of pollination can be used to identify promoter elements that drive gene expression in the pollen tube in response to growth through the pistil (Figure 4).

It is becoming clear that the near simultaneous rupture of the pollen tube and receptive synergid cell (Leydon et al., 2015) are the

### Table 1. SNP Rates Are Elevated in MYB-Dependent Pollen-Specific Genes

| Family|--| Polinated pistil expressed| 10,029| 4.4 × 10⁻³| 3.1 × 10⁻³| 0.70*| 1.99 × 10⁻³| 2.19| 2.19| 1.90| 2.19| 1.90|
| Simulated Rate| Actual Rate| Ratio AR/SR| Actual Rate| Ratio AR/BR|
| NBS-LRR| 15,915| 4.3 × 10⁻³| 4.5 × 10⁻³| 1.03*| 1.83 × 10⁻³| 2.01|
| Pollen expressed (Col reads)| 4,375| 4.3 × 10⁻³| 6.6 × 10⁻³| 1.53*| 1.65 × 10⁻³| 1.81|
| Pistil-specific pollination induced (Cvi reads)| 254| 4.4 × 10⁻³| 7.0 × 10⁻³| 1.59*| 1.27 × 10⁻³| 1.39|
| MYB-dependent (Col reads)| 268| 4.4 × 10⁻³| 7.4 × 10⁻³| 1.71*| 1.16 × 10⁻³| 1.28|
| Randomly sampled SNP-containing genes| 268 (×10)| 4.4 × 10⁻³| 4.5 × 10⁻³| 1.06*| 1.27 × 10⁻³| 1.40|

For gene lists, see Supplemental Data Set 12. Rate = SNPs/nucleotide; ratio = see color scale. Simulated lists (100 randomly chosen gene lists of equal gene number to input set) were used to create a ratio of list rates to demonstrate SNP enrichments; see color scale. AR/SR, actual rate/simulated rate. Ratios that have an asterisk are statistically significant (P < 0.01, Poisson distribution). Missense mutation rate = missense mutation/nucleotide. Background rate for missense mutations = 9.1 × 10⁻⁴. AR/BR, actual rate/background rate.
consequence of highly coordinated cell-cell signaling requiring pollen tube and synergid transcriptional networks (Leydon et al., 2013; Liang et al., 2013; Mendes et al., 2016), expression of cell surface signaling molecules (Escobar-Restrepo et al., 2007; Kessler et al., 2010; Lindner et al., 2015; Liu et al., 2016), and Ca\(^{2+}\) fluxes that define which of two synergids receives the pollen tube (Denninger et al., 2014; Ngo et al., 2014). Using SNP-informed RNA-seq analysis of myb triple mutant compared with wild-type pollinations, we have provided a comprehensive set of genes expressed by pollen tubes that mediate interactions with female cells required for pollen tube burst, sperm release and synergid degeneration. Our analysis defined three main categories of genes regulated by MYB97, MYB101, and MYB120: transmembrane transporters, carbohydrate active proteins, and secreted proteins.

---

**Figure 7. Elevated Ka/Ks Rates in CRP Genes.**

(A) Percentage of genes in indicated gene families with homologs identified in given species identified by Lloyd et al. (2015).

(B) Average Ka/Ks of all homologs found between A. thaliana and the indicated species. Bars are SD (sample sizes vary and are defined in [A]; Lloyd et al., 2015).

(C) Percentage of homologous gene pairs out of the total number of family members with a Ka/Ks >1 (indicator of positive selection) for each respective species. The key pertains to (A) to (C).
of diverse function. This result was consistent with the enriched GO categories previously identified by microarray analysis (Leydon et al., 2013). However, RNA-seq analysis revealed many novel MYB-dependent genes in these categories that are members of large gene families and were not detectable by microarray analysis because hybridization-based probes could not differentiate closely related family members or because the genes were not annotated at the time of microarray production. The thionin-like family of small cysteine-rich proteins provides a striking example: A single MYB-dependent member was identified by microarray analysis (Leydon et al., 2013), while RNA-seq analysis identified 24 family members whose expression was significantly decreased in myb triple-mutant pollinations compared with the wild type (Figure 6).

MYB-dependent secreted proteins were of particular interest because recent analysis documented the rich diversity of the tobacco (*Nicotiana tabacum*) pollen tube secretome (Hafidh et al., 2016) and suggested that the pollen tube and the synergid cell (Punwani et al., 2007; Liu et al., 2015) both evolved highly active secretory systems. Interestingly, genes that require MYB transcription factors for expression in the pollen tube are enriched for secreted proteins including two alpha carbonic anhydrases (Figures 4O to 4T), 11 endopeptidases, 22/28 thionins in the CRP2460 gene family (Figure 6), 19/98 S-protein homologs (Ride et al., 1999), two cysteine-rich secretory, antigen 5, and pathogenesis-related 1 proteins (CAP; AT3G09590 and AT5G02730), and four rapid alkalinization factors (AT2G32835, AT2G32788, AT2G32775, and AT1G60815; Supplemental Data Sets 8 and 9). Interestingly, a comparison between the tobacco pollen tube and *A. thaliana* MYB-dependent pollen tube genes revealed significant overlap both in terms of specific gene homologs and for functional gene categories (Hafidh et al., 2016). These findings suggest a core group of proteins secreted by pollen tubes of dicot species that could be important for sperm release.

The pollen transcriptome is overrepresented by carbohydrate active enzymes involved in cell wall synthesis (Honys and Twell, 2004; Pina et al., 2005; Qin et al., 2009), an enzymatic function that was enriched in pollen tube-expressed genes identified in our RNA-seq experiments (Supplemental Data Set 6). Our analysis also found that the MYB-dependent transcriptome is enriched for secreted carbohydrate active enzymes such as pectate lyase (n = 6; Figure 4; Supplemental Data Sets 8 and 9) and pectin methylsterase inhibitors (n = 4; Supplemental Data Sets 8 and 9), both of which are predicted to modify pectin, a major component of the pollen tube wall. Regulation of pectin structure at the pollen tube tip has been shown to be essential for pollen tube growth; mutants lacking pectin methylsterases (enzymes facilitating pectin cross-linking) rupture in vitro and in the pistil (Bosch et al., 2005; Jiang et al., 2005). Unlike mutants defective in pectin methylsterases, *myb* triple-mutant pollen tubes do not have growth defects in vitro and do not rupture prematurely in the pistil; instead, they fail to burst upon contact with the synergid (Leydon et al., 2013; Liang et al., 2013). This phenotype suggests that carbohydrate active enzymes regulated by MYB transcription factors could be required for regulated cell wall instability upon contact with synergids, or they could determine a cell wall carbohydrate profile that is recognized by the synergid to facilitate sperm release by pollen tubes of the same species.

When we compared the Cvi and Col self-pollination transcriptomes, we found evidence for a core set of pollen tube-expressed genes, but also many examples of quantitative and qualitative differences in reproductive gene expression (Figure 5). These data are consistent with the idea that patterns of reproductive gene expression diversify rapidly when populations are geographically isolated. Defining the functions of accession-specific genes we identified here will inform efforts to understand the mechanisms of pollen tube growth, guidance, and reception, which appear to evolve rapidly and create barriers to interspecific hybridization (Escobar-Restrepo et al., 2007; Takeuchi and Higashiyama, 2012). Interestingly, we observed families of MYB-regulated, small secreted proteins that were overrepresented in Col self-pollinations, yet absent from Cvi pollinations. This suggests that mechanisms for pollen tube reception and sperm release may be tuned to a particular accession and are in keeping with our prior observation of pollen tube reception defects in inter-accession crosses (Leydon et al., 2014). Concordantly, we also observed many genes that were expressed in Cvi pollinations that were absent in Col pollinations and lacked any usable GO terminology. Additional analysis of these genes showed that many contain predicted signal peptides and may be members of as yet undefined gene families.

Comparison of self-pollinations with inter-accession pollinations revealed a small set of Cvi pistil-expressed genes that are induced by Col pollen tube growth, but not by Cvi pollen tube growth. These genes were enriched for functions associated with defense responses, especially an ethylene-responsive transcriptional response (Supplemental Data Set 11). Therefore, we propose that accession-specific expression of a complement of secreted proteins controls the interaction of the pollen tube with pistil cells and directs the pistil's response between reproduction and defense depending on the pollen tube genotype.

Reproductive genes have been shown to be rapidly evolving in many species (Clark et al., 2006). Some of the best examples include sperm proteins with high rates of nonsynonymous substitutions in aquatic species with motile sperm that spawn at the same time as closely related sympatric species (Metz et al., 1998). It has been proposed that rapid coevolution between sperm surface proteins and their egg-expressed binding partners provides a barrier to interspecific gamete fusion. In flowering plants, the pollen tube is motile and performs the required cell-cell interactions with female tissues (the pistil) that determine male reproductive recognition and fitness, not the sperm as in animals (Clark et al., 2006). Indeed, we find that SNPs are enriched 1.5× over controls in pollen-expressed genes (Col to Cvi), and these SNPs are 1.7 to 1.9× more likely to cause a missense mutation (Table 1). Strikingly, the MYB-dependent pollen tube-specific genes that contain SNPs carry 1.7× more SNPs than the background rate, providing sequence diversity on which evolution can act to select for prezygotic barriers to interspecific reproduction. Pollen-specific genes that are pollination responsive also display increased SNP rate and missense frequencies (Table 1).

We analyzed a group of small secreted, MYB-dependent CRPs (CRP2460) to determine whether their patterns of evolution are consistent with a role in determination of species-level identity of the pollen tube during interactions with the synergid cell. A MYB-dependent member of this group (At5g36720) was previously
shown to be expressed in the pollen tube and localized to the pollen tube secretory system (Leydon et al., 2013). This group is expanded in A. thaliana relative to A. lyrata and homologs were not found outside of the Brassicaceae (Figures 6 and 7; Supplemental Figure 9). Moreover, a comparison of the AtIG21835 gene and its A. lyrata homolog revealed a Ka/Ks value >1, suggesting selective pressure to maintain amino acid diversity. Intriguingly, this gene is sister to a clade of 14 A. thaliana genes with no homologs (Supplemental Figure 9A) in A. lyrata, consistent with a very recent A. thaliana expansion. Analysis of individual residues revealed evidence for diversifying selection interspersed with purifying selection to maintain the position of cysteines likely to be important for maintaining structural integrity of the proteins (Supplemental Figure 9B and Supplemental Data Set 15). Future experiments will be directed toward determining whether this group of secreted proteins is a critical component of the MYB-controlled pollen tube-pistil recognition system required for sperm release.

SNP-informed RNA-seq analysis of pollination provides a powerful method to quantify pollen-pistil interactions over any time course to define sex-specific transcriptional responses. We have used this method to define the set of genes controlled by pollen tube MYB transcription factors MYB97, MYB101, and MYB120 that mediate signaling between the pollen tube and synergid cell required for sperm release. However, this method could be used to analyze any pollen tube or pistil mutant affecting pollen-pistil interactions. Moreover, this method lends itself readily to analysis of the impact of environmental change on the reproductive transcriptome of crop plants and holds the potential to uncover how pollen tube-pistil interactions change in response to abiotic and biotic stresses that limit crop productivity.

METHODS

Plant Material

Seeds were grown on 0.7% agar Murashige and Skoog medium (Sigma-Aldrich). Arabidopsis thaliana plants were grown in chambers at 21°C under illumination (100 μmol m$^{-2}$ s$^{-1}$ with a 16-h photoperiod). Seedlings were transplanted to sterile 2MX potting media (Fafard) between 7 and 10 d after plating. All plants were grown at 20°C, 50 to 60% humidity, on a 16-h/8-h day/night cycle in growth chambers (Environmental growth chambers). SIV pollinations were performed as previously described (Palanivelu and Preuss, 2006; Qin et al., 2009).

mRNA-Seq Library Construction

RNA was isolated from excised pistil tissues either unpollinated or 8 h after pollination using the NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs). At least 10 μg of total RNA was purified with oligo(dT) beads and RNA-seq libraries were prepared according to the NEBNext protocol. Primers 4, 6, and 12 were used from NEBNext Multiplex Oligos for Illumina (New England Biolabs). Amplification was for 15 cycles. Libraries were tested by qPCR after amplification to validate amplification.

mRNA-Seq Data Analysis

Pistils were collected with and without pollination in triplicate (15 samples). Single-end sequencing was performed with mRNA-seq libraries on an Illumina HiSeq machine. Read length was 50 bp. Sequencing quality was analyzed using fastqc. Reads were aligned to the A. thaliana genome (TAIR10) using TopHat v2.0.13 (Kim et al., 2013). TopHat options were segment-length 22–segment-mismatches 1–max-segment-intron 11,000. Reducing segment length or mismatch number yielded no additional mapped reads, indicating that this segment length, mismatch number was optimal for mapping our libraries to the TAIR10 genome. RNA-seq data have been deposited at the Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra; NCBI BioProject, PRJNA329634; SRA accession, SRP078926). Aligned reads were parsed using a PERL script released by the Ghering Lab (Pignatta et al., 2014). Read counts per gene were quantified using HTSeq (Anders et al., 2015) and EasyRNASeq (Delhomme et al., 2012). Multiply-mapping reads were counted because many pollen expressed genes were observed to be in redundant gene families due to gene duplication.

GO Analysis

GO analysis was performed using the DAVID bioinformatics resource version 6.7 (Huang et al., 2009). Reported P values were corrected using the Benjamini method.

Calling Variants

RNA sequencing reads from unpollinated Cvi-0 and Cvi-0 8 h after self-pollination were analyzed for novel SNP variants following the GATK Best Practices workflow (McKenna et al., 2010). Briefly, mapped reads (TopHat2) were prepared for variant calling: read group addition, mark duplicates, SplitNTrim, ReassignMappingQuality, indel realignment, base recalibration. Variant calling was performed using the GATK HaplotypeCaller. Variant filtration was performed (FS > 30, QD < 2.0). SNPs supported with <10 reads or that were not homozygous were discarded. Any SNPs in common with previously established Cvi-0 SNPs (Pignatta et al., 2014) were considered validated and set aside from the list of novel SNPs. Unpollinated Cvi-0 and Cvi-0 self-pollinated RNA-seq library reads were then run against this novel SNP list using a previously published Perl pipeline (Pignatta et al., 2014). Any SNP locations that returned a Col-0 SNP were discarded from the BED file. This eliminated 2000 SNPs and created a final additional 15,995 SNPs (in 237 genes previously without documented SNPs and 5217 genes that had one or more previously described SNP).

Variant Analysis

The effect of SNPs was predicted using SnpEff and SnpSift software (Cingolani et al., 2012a, 2012b). Master SNP lists were correlated to the TAIR10 gene annotation file to identify the distribution of SNPs in genomic elements using a range overlap function via the iRanges package (Lawrence et al., 2013) and a custom R script (R Core Team, 2013; http://www.R-project.org/). The SNP rate was computed for each gene (rate = SNPs/nucleotide), and the average SNP rate of individual gene lists was compared with background frequencies simulated using randomly generated lists of the same size as input lists. Each comparison background frequency was the average of 100 simulations. Similarly, the rate of missense mutations was calculated for each gene and averaged for the individual gene lists. These frequencies were compared with the rate of missense mutations across the entire genome. R scripts and data for analysis are part of the package SexSpecificExpression, which is available on GitHub at https://github.com/evenable/SexSpecificExpression.git.

Transcriptional Profile Correlation Analysis

Pairwise correlations were generated between our RNA-seq data and microarray data from previous studies. Data on uniloculate microspores, tricellular pollen, and mature pollen grains are from Honys and Twell (2004).
Data on dry pollen, 0.5 h, 4 h, and SIV-grown pollen tubes are from Qin et al. (2009). Data from unpollinated and pollinated pistils and ovules are from Boavida et al. (2011). Data from stigmas and ovaries are from Swanson et al. (2005). Data from roots, seedlings, and rosette leaves are from Schmid et al. (2005). For each pair, Pearson-r was computed over all genes with a unique SNP, using mean reads from three biological replicates for RNA-seq and mean absolute intensity for multiple microarray replicates.

Graphing Sequencing Data
Scatterplots were generated from sequence data using Python’s Matplotlib module.

Analysis of Differential Gene Expression
Three differential gene expression callers, DESeq, DESeq2, and EdgeR (Anders and Huber, 2010; Robinson et al., 2010; Love et al., 2014), were tested to define differentially expressed genes using the Cvi x Col versus Cvi x myb triple mutant as a test. DESeq differential expression was performed using the default parameters as per the DESeq vignette (estimateSizeFactors, estimateDispersions, nbinomTest; Anders and Huber, 2010). Differential expression was performed using DESeq2 and a trimmed mean approach (replaceOutliersWithTrimmedMean; Love et al., 2014). Data were visualized according to the vignette for DESeq2. Differential expression with EdgeR was performed as per the EdgeR vignette with the user-defined estimateGLMCommonDisp - Estimate Common Dispersion for Negative Binomial GLMs, estimateGLMTrendedDisp - Estimate Trended Dispersion for Negative Binomial GLMs, estimateGLMTagwiseDisp - Empirical Bayes Tagwise Dispersions for Negative Binomial GLMs functions. All subsequent differential expression analysis was performed using DESeq2 as described above.

GFP Fusion Constructs
To generate ACA:GFP and PL:GFP constructs, genomic fragments containing the promoter (0.96 kb upstream of the ACA start codon; 1.18 kb upstream of the PL start codon), introns, and exons (excluding the stop codon) were amplified by PCR and inserted into the pENTR-D vector (Invitrogen). Genomic clones in pENTR-D were integrated into pMDC107 with LR Clonase II (Invitrogen). Genomic clones in pENTR-D were integrated into pMDC107 using Green qPCR Master Mix (Invitrogen) on an ABI 7000 PRISM sequence detection system (Applied Biosystems). All oligonucleotides used are given in Supplemental Table 1.

Supplemental Data
Supplemental Figure 1. SNP frequency in genomic features.
Supplemental Figure 2. Flowchart of data analysis utilized to discover sex-specific differential gene expression.
Supplemental Figure 3. Quantitative analysis of reads with SNPs suggests they are distributed across transcribed regions of genes similarly to reads without SNPs.
Supplemental Figure 4. Predicting optimal accessions to detect pollen tube expressed transcripts.
Supplemental Figure 5. Sex-specific gene expression profiles correspond to known reproductive transcriptional profiles.
Supplemental Figure 6. Comparison of differential expression callers’ ability to identify MYB-dependent genes.
Supplemental Figure 7. Expression of SUC9:GFP and ACA4:GFP in female gametophytic cells.
Supplemental Figure 8. Accession-specific gene expression of MYB-dependent genes.

Supplemental Figure 9. Rapid evolution of the CRP2460 thionin-related proteins by expansion and positive selection.

Supplemental Table 1. Oligonucleotides used in this study.

Supplemental Data Set 1. Final SNP database.

Supplemental Data Set 2. Table of mRNA-seq libraries generated in this study.

Supplemental Data Set 3. Table of read counts per gene for Cvi x Col interecotype pollinations.


Supplemental Data Set 5. Intersection and Venn diagram of SIV-expressed genes.


Supplemental Data Set 7. Pistil-specific genes that are differentially expressed after pollination.

Supplemental Data Set 8. Detecting differential expression of MYB-dependent genes.


Supplemental Data Set 10. Genes differentially expressed between Col x Col and Cvi x Cvi pollinations.

Supplemental Data Set 11. Pistil-expressed genes with elevated transcript levels in interecotype pollinations.

Supplemental Data Set 12. Alignment of thionin CRPs used to produce the phylogenetic tree (Figure 6).

Supplemental Data Set 13. Gene lists used in Table 1: SNP rates are elevated in MYB-dependent pollen-specific genes.

Supplemental Data Set 14. Annotation of Cvi SNPs to Col with strong effects on coding sequences and gene function.

Supplemental Data Set 15. Analysis of natural selection in MYB-dependent CRP2460 family.


Supplemental Movie Legends.

ACKNOWLEDGMENTS

We thank Mary Gehring (Whitehead Institute, MIT) for generous access to bioinformatic resources and Christoph Schorl (Brown University Genomics Core Facility) for RNA-sequencing. This work was funded by National Science Foundation Grant IOS-1353798.

AUTHOR CONTRIBUTIONS

A.R.L. performed the experiments. E.V. and C.W. performed the bioinformatics analyses. A.R. and J.M.W. provided input and reagents for the SUC9 experiments. M.A.J. and A.R.L. conceived the study, designed the experiments, and wrote the manuscript.

Received October 28, 2016; revised January 30, 2017; accepted April 10, 2017; published April 11, 2017.

REFERENCES


