## Supplementary material

**Table S1:** Microarray data and differentially expressed genes. The first spreadsheet ('Data') lists processed and normalized microarray data for genes that were differentially expressed in at least one of the experiments described in this study. Gene identifiers, names and descriptions are listed together with fold change and intensity data, as well as (Benjamini & Hochberg) adjusted *p* values for all experiments. In the remaining spreadsheets, genes that were differentially expressed in the individual experiments (as indicated) are listed. In the last spreadsheet ('ms1 samples'), the assignment of genes to one of the twelve clusters shown in Figure 3 is indicated.

**Table S2:** Datasets used for comparison and analysis of gene lists (as shown in Figure 1). The first spreadsheet ('Stamen') lists genes with predicted specific or predominant expression in stamens after reannotation of microarray probes (see Material and Methods for details). In spreadsheet 'ap3 inflorescences', genes are listed with differentially expression in whole inflorescences of *ap3* mutant plants compared to the wild type. The gene lists derived from the overlap analyses (Venn diagrams of Figure 1B and 1C) are provided in spreadsheets 'Venn Diagram Figure 1B' and 'Venn Diagram Figure 1C'. In the last spreadsheet ('Figure 1D'), the genes included in the cluster diagram shown in Figure 1D are listed and their assignment to one of the clusters is indicated. In all spreadsheets, gene identifiers, names and descriptions are shown.

**Table S3:** Analysis of the dataset derived from the *ms1* developmental series. In the first spreadsheet, genes with differential expression in the *ms1* experiments were analyzed with respect to their presence in the datasets from several previous studies. Genes with specific or predominant expression in stamens ('Stamen') were taken from (Wellmer et al., 2004); genes predicted as pollen-selective or pollen-enriched were taken from (Pina et al., 2005); JA-induced genes were derived from (Mandaokar et al., 2006); pollen-specific transcripts were predicted by (Honys and Twell, 2004). For all genes, identifiers, names and descriptions, as well as the assignment of genes to one of the twelve clusters shown in Figure 3 are shown.

The second spreadsheet ('DNA binding proteins') lists genes found among the differentially expressed genes identified in the *ms1* experiments that code for transcription factors or other DNA binding proteins. The general direction of the expression change (up or down in the mutant compared to the wild type) is indicated.

**Table S4:** Genes with known expression in anthers. Data from the *ms1* developmental series for genes with known expression in anthers are shown. Their published expression patterns are briefly described and literature is referenced.

**Table S5:** T-DNA insertion lines. The T-DNA lines that were analyzed for this study and the corresponding target genes are listed. The position of the T-DNA insertions is briefly described. The sequences of primers used for genotyping of the T-DNA insertions (see Material and Methods for details) are shown.

**Table S6:** Primers used for reverse-transcriptase PCR and quantitative real-time PCR. The target genes and the corresponding primer sequences are shown. The last column indicates whether or not a gene was repressed in *ap3* inflorescences compared to the wild type.

**Table S7:** Primers used for generating probes for *in situ* hybridizations. The target genes and the corresponding primer sequences are shown.

## Supplemental Figures

**Figure S1:** Comparison of microarray datasets. **(A)** A Venn diagram depicts the overlap between genes identified as differentially expressed in *ms1* mutant flowers compared to the wild type and transcripts with previously predicted pollen-enriched and pollen-selective expression (see Supplementary Table 1 in (Pina et al., 2005). **(B)** A Venn diagram shows the overlap between pollen-selective transcripts and genes present in selected clusters (as indicated) of the diagram shown in Figure 3.

**Figure S2**: Expression pattern of *At2g42940* in stamens of *ms1* and wild-type plants. Arrowheads point to regions of expression. In *ms1* mutant flowers, expression was observed in tapetal cells of three adjacent anthers from two medial and one lateral stamen. In contrast, in wild-type plants (inset), expression was detected in anthers of medial position only. This result suggests a prolonged expression of the gene in *ms1* mutant flowers. Abbreviations: ca, carpel; td, tetrads; tp, tapetum; vr, vascular region. Scale bars: 25 µm.

**Figure S3:** Distribution of selected Gene Ontology (GO) terms in the *ms1* dataset. **(A to D)** Genes that were identified as downregulated **(A and B)** or as upregulated **(C and D)** in the different *ms1* flower samples or cluster **(E and F)** were analyzed using the program GOToolBox as outlined in Material and Methods. Solid lines denote the background frequencies of a given term in the genome and columns the frequencies of that term in the different gene groups. Bold numbers mark gene groups in which a term was either significantly over- or underrepresented relative to its genome-wide distribution.
**(A)** Distribution of GO term "Lipid Metabolism". Genes assigned to this term were statistically over-represented in *ms1* flower samples 3 to 5 (**B**) Distribution of GO term "Transcription Factor Activity". Genes assigned to this term were statistically under-represented in samples 1 to 4. **(C)** Distribution of GO term "Catalytic Activity". Genes assigned to this term were statistically over-represented in samples 2 to 6. **(D)** Distribution of GO term "Secondary Metabolism". Genes assigned to this term were statistically over-represented in samples 2 to 5. **(E)** and **(F)** Distribution of selected GO terms in clusters 1 to 8 of the diagram shown in Figure 3. **(E)** Distribution of GO term "Metabolism". Genes assigned to this term were statistically under-represented in clusters 2, 6 and 7 and over-represented in clusters 3 and 5. **(F)** Distribution of GO term "External Encapsulating Structures". Genes assigned to this term were statistically over-represented in clusters 1, 2 and 7.

**Figure S4:** Stage-dependent expression of transcription factor coding genes in developing anthers as derived from the *ms1* experiments. Identifiers, as well as gene names or gene

families, for genes upregulated **(A)** or downregulated **(B)** in *ms1* flowers are shown. Genes that are downregulated in one of the *ms1* tissue samples compared to the wild type are depicted in blue and upregulated genes in yellow. The intensities of the colors increase with increasing expression differences. Diagrams were generated using an agglomerative hierarchical clustering algorithm.

**Figure S5:** Characterization of an RNAi line for gene *At2g42940*. **(A)** and **(B)** Toluidine-blue-stained transverse sections through anthers of the RNAi line **(A)** and the wild type **(B)** at stage 11 of anther development. Several microspores of the RNAi line are morphologically abnormal. **(C)** and **(D)** Silique and seed formation in the RNAi line **(C)** and in the wild type **(D)**. Compared to the wild type, the RNAi line generated small siliques with few seeds. Abbreviations: mp, microspore; vr, vascular region. Scale bars: 20 µm (panels A and B); 0.25 cm (panels C and D).