Genome Biology

Comparative transcriptomics among floral organs of the basal eudicot Eschscholzia californica as reference for floral evolutionary developmental studies

Laura M Zahn (lzahn@aaas.org)
Xuan Ma (xxm108@psu.edu)
Naomi S Altman (naomi@stat.psu.edu)
Qing Zhang (qzxu@hotmail.com)
P Kerr Wall (pkerrwall@gmail.com)
Donglan Tian (dut3@psu.edu)
Cynthia J Gibas (cgibas@uncc.edu)
Raad Gharaibeh (rgharaib@uncc.edu)
James H Leebens-Mack (jleebensmack@plantbio.uga.edu)
Claude W dePamphilis (cwd3@psu.edu)
Hong Ma (hxm16@psu.edu)

Comparative transcriptomics among floral organs of the basal eudicot *Eschscholzia californica* as reference for floral evolutionary developmental studies

Laura M Zahn[1,2,8]*, Xuan Ma[1,2,3]*, Naomi S Altman[2,4], Qing Zhang[2,4,9], P Kerr Wall[1,2,10],

Donglan Tian[1,11], Cynthia J Gibas[5], Raad Gharaibeh[5], James H Leebens-Mack[1,2,12], Claude W

dePamphilis[1,2] and Hong Ma[1,2,3,6,7,§]

[1]Department of Biology, The Pennsylvania State University, University Park, PA 16802 USA

[2]The Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA 16802, USA

[3]The Intercollege Graduate Program in Cell and Developmental Biology, The Pennsylvania State University, University Park, PA 16802, USA

[4]Department of Statistics, The Pennsylvania State University, University Park, PA 16802, USA

[5]Department of Bioinformatics and Genomics, The University of North Carolina at Charlotte, 9201 University City Boulevard, Charlotte, NC 28223, USA

[6]State Key Laboratory of Genetic Engineering and School of Life Sciences, Fudan University, 220 Handan Road, Shanghai 200433, China

[7]Institutes of Biomedical Sciences, Fudan University, 138 Yixueyuan Road, Shanghai 200032, China

[8]Current address: American Association for the Advancement of Science, 1200 New York Avenue NW, Washington DC 20005, USA

[9]Current address: 2367 Setter Run Lane, State College, PA 16802, USA

[10] Current address: BASF Plant Science, 26 Davis Drive, Research Triangle Park, NC 27709, USA

[11] Current address: Department of Entomology, The Pennsylvania State University, University Park, PA 16802, USA

[12] Current address: Department of Plant Biology, University of Georgia, 120 Carlton Street, Athens, GA 30602, USA

§Correspondence, Email: hxm16@psu.edu

* These authors contributed equally to this work

Abstract

Background

Molecular genetic studies of floral development have concentrated in several core eudicots and grasses (monocots), which have canalized floral forms. Basal eudicots possess a wider range of floral morphologies than the core eudicots and grasses and can serve as an evolutionary link between core eudicots and monocots, and provide a reference for studies of other basal angiosperms. Recent advances in genomics have enabled researchers to profile gene activities during floral development, primarily in the eudicot *Arabidopsis thaliana* and the monocots rice and maize. However, our understanding of floral developmental processes among the basal eudicots remains limited.

Results

Using a recently generated EST (expressed sequence tag) set, we have designed an oligonucleotide microarray for the basal eudicot *Eschscholzia californica* (California poppy). We performed microarray experiments with an interwoven-loop design in order to characterize the *E. californica* floral transcriptome and to identify differentially expressed genes in flower buds with pre-meiotic and meiotic cells, four floral organs at pre-anthesis stages (sepals, petals, stamens and carpels), developing fruits, and leaves.

Conclusions

Our results provide a foundation for comparative gene expression studies between eudicots and basal angiosperms. We identified whorl-specific gene expression patterns in *E. californica* and examined the floral expression of several gene families. Interestingly, most *E. californica* homologs of *Arabidopsis* genes important for flower development, except for genes encoding MADS-box transcription factors, show different expression patterns between the two species. Our comparative transcriptomics study highlights the unique evolutionary position of *E. californica* compared with basal angiosperms and core eudicots.

Background

The eudicots are believed to have originated approximately 130 million years ago [1]. They include about 70% of all flowering plant species and are comprised of core eudicots [2-4], which include the groups containing *Arabidopsis thaliana* and *Antirrhinum majus*, respectively, and species that branched earlier from these groups and are at basal positions within the eudicot clade. The earliest branching lineage of the eudicots, the Ranunculales, contains the Papaveraceae (poppy) family, of which *Eschscholzia californica* (California poppy) is a member [5]. The core eudicots commonly have stable (i.e., canalized) flower architecture (Figure 1A); by contrast, the basal eudicots exhibit a wider range of floral patterns [6] (see examples in Figure 1A). Comparing the morphology and the underlying mechanisms of flower development between the core and basal eudicots may help us better understand the evolution of flower structures and development.

Molecular genetic studies in *Arabidopsis*, *Antirrhinum* and other core eudicots have uncovered the functions of many genes involved in regulating flowering time and floral organ identity and development [7-9]. In particular, it is known that a number of MADS-box genes are required to control flowering time and floral organ identities, as well as anther, ovule and fruit development. These include the well-known ABC genes *APETALA1* (A function), *APETALA3* and *PISTILLATA* (B function), and *AGAMOUS* (C function) from *Arabidopsis*, and their respective orthologs from *Antirrhinum* (*SQUAMOSA*, *DEFICIENS*, *GLOBOSA*, and *PLENA*) [10-12]. Comparative studies of core eudicots suggest that homologs of B and C function genes have relatively conserved functions, although some divergences have also been observed. Putative orthologs of these MADS-box genes may have diverged expression patterns in different species and the expression difference between recent duplicates is often associated with subfunctionalization [11, 12]. In addition, several MADS-box genes have been found to be important for floral organ identities in the monocots [13-16]. However, both the long evolutionary distance and the highly diverged flower architectures between monocots and core eudicots have made it difficult to study the evolution of floral gene function.

The investigation of floral gene function in the basal eudicots serves to bridge the gap between core eudicots and monocots. Molecular and expression studies of floral genes have been reported

for some basal eudicots, providing informative initial knowledge of the conservation and divergence of floral gene activities among eudicots [17-19]. Molecular evolutionary studies of several MADS-box subfamilies, complemented by expression analyses, support that some of the MADS-box genes have maintained conserved functions throughout angiosperm evolution [11, 20-23]. For example, expression studies of floral MADS-box genes in *E. californica* demonstrated that genes in the *AGAMOUS*, *GLOBOSA* and *SEPALLATA* subfamilies are highly conserved between basal and core eudicots [11, 12, 21]. Additionally, in other ranunculids, expression divergences have also been observed between recently duplicated MADS-box genes [11, 12].

High-throughput technologies, including microarrays, can be used to analyze transcriptomes of individual floral organs at specific developmental stages. Transcriptome studies have been performed extensively for *Arabidopsis* and, to a lesser extent, several other highly derived core eudicots [19, 24-29]. Among basal eudicots, such studies have only been carried out recently in the basal eudicot *Aquilegia*, which represents a different ranunculid lineage than *E. californica* [30]. *E. californica* is a potential model organism because (1) it has a relatively small plant size, many seeds per fruit and a short generation time, which facilitate genetic studies; (2) it does not have determinate flowering and produces multiple flowers over its lifespan, providing easy access to floral materials [31]; (3) it has a relatively small genome; and (4) it both has an efficient system for virally induced gene silencing and is transformable [21, 32-35]. Previous gene expression studies in *E. californica* showed that there is very good correlation between regions of gene expression and domains of gene function [19, 34, 36, 37]. An *E. californica* EST collection of over 6000 unigenes was constructed from a pre-meiotic floral cDNA library [21], which provides gene sequence information for microarray analysis of *E. californica* leaf and floral transcriptomes. A transcriptome-level analysis facilitates our understanding of floral development in basal eudicots and sheds light on potential floral regulatory genes in *E. californica*.

In this study, we used microarray technology to investigate transcriptomes in *E. californica* and to identify differentially expressed genes in developing leaves and floral buds at pre-meiotic (small buds) and meiotic (medium buds) stages. Additionally, we examined the transcriptomes of developing fruits and four types of floral organs (sepals, petals, stamens, and carpels) at the pre-

anthesis stage. We identified genes that are significantly differentially expressed in different floral organs or at different floral stages, in comparison with developing fruit and leaf tissues. We also analyzed the expression of genes in several regulatory gene families, some of which contain homologs of known floral genes from other organisms. Finally, we compared our results with similar studies in *Arabidopsis* and recent studies [30, 38] in *Aquilegia* and *Persea* (avocado), a basal angiosperm related to magnolia, to assess conservation and divergence in gene expression and discuss their implications for evolution of floral development in the eudicots.

Results and discussion

Construction and use of a microarray chip for *E. californica*

To investigate the leaf and reproductive transcriptomes of *E. californica*, we generated a custom Agilent microarray chip with features for 6446 unigenes from the *E. californica* EST collection [21] (see Materials and methods for additional information). The oligonucleotide sequences for the probes were selected using available sequence information from *E. californica* ESTs, as well as other public sequence information, avoiding non-specific hybridization as much as possible. Additional criteria were used to consider potential secondary structure and hybridization temperature (see Materials and methods).

A primary objective was to obtain expression profiles with the power to detect differential expression between vegetative (leaves) and reproductive organs, between different floral stages, and between different floral organs. Therefore, we sampled the *E. californica* plants for the following eight representative organs and stages (for convenience, referred to generally as tissues hereafter): leaves, early floral buds, medium floral buds, four floral organs (sepals, petals, stamens, and carpels) at pre-anthesis, and young fruits. Four sets of plants were sampled at the same time daily (8:30-10:30AM) to minimize variation due to circadian rhythms, yielding four biological replicates. RNAs from these 32 samples were used to generate cDNAs and labeled with Cy3 and Cy5 dyes, for two-channel microarray experiments. Finally, we used an interwoven loop design (Figure 1B) to maximize the comparative statistical power using a limited number of hybridizations [39].

In an interwoven loop design, differences in gene expression can be estimated for all pairs of tissues with a relatively small number of hybridizations [40]. Each of the 8 tissues was directly compared on the same slide with one of four other tissues, with one biological replicate for each comparison, resulting in a total of 16 hybridizations. The comparison of the two tissues on the same arrays allowed more precise results than those compared indirectly via other tissues. The specific pairings on the same array were chosen to optimize precision of comparisons for biologically important comparisons, while keeping the precision of different comparisons as similar as possible. Because our EST library was constructed with floral bud mRNAs, we compared developing floral buds at different stages with each of the four floral organs, and compared each of these tissues with leaves, the only vegetative organ in this study, and developing fruits. The comparison between small buds and leaves was aimed at identifying differentially expressed genes at early reproductive stages. We hypothesized that the sepal should be the most leaf-like tissue among all floral organs; whereas previous studies [25] suggest that the stamens might have the most complex transcriptome among the four major floral organs [27]. In this study, the fruit tissue represents the only post-anthesis tissue. We also considered the ABC model, which posits that sepals and petals both require A function genes, petals and stamens both need B function genes, and stamens and carpels both depend on C function genes. In addition, carpels and fruits were developmentally related tissues, with small and medium buds representing two consecutive stages in floral development.

After microarray hybridizations, we tested the quality of the microarray experiments. We assessed the reproducibility of the microarray hybridizations by determining the Pearson's correlation coefficients between the biological replicates for each of the eight tissues (see Figure 2 for an example; the plots for the remaining seven tissues can be found in Additional file 1, Figure S1). As shown in Figure 2, the Pearson's correlation coefficients between any pair of the four biological replicates of small buds, one of the most complex tissues in this study, ranged from 0.94 to 0.97. The high correlation values indicate that our results were highly reproducible.

In addition, we examined signal intensities. Because the EST library used for the probe design was constructed from mRNAs of flower buds, we assumed that expression of most genes should be detected in our microarray experiments from mostly flower-related tissues. The value of 5.41 for log2 of hybridization intensity (10% quantile of all genes on the chip) was selected as a cut-

off to identify "present" signal (see Table 1; for alternative cutoffs, see Additional file 2 for gene numbers with 5% or 15% quantiles) similar to previous microarray experiments in *Arabidopsis*[41]. For the 10% quantile, we identified the number of genes detected in leaves (5905), small buds (5906), medium buds (5876), sepals (5876), petals (5870), stamens (5877), carpels (5851) and fruits (5881). These results were not surprising because the unigenes were derived from EST data, which tend to favor genes that are expressed at relatively high levels. Therefore, our microarray chip and hybridization experiments were able to detect the expression of several thousand genes in eight major tissues of *E. californica*. Of the genes examined, the majority of genes present in leaf were also observed in small buds and medium buds (Figure 3A). In addition, most genes expressed in sepal were also expressed in petal (Figure 3B), suggesting similar gene expression levels between these two tissues. There was significant overlap of genes expressed in petal and/or sepal with genes expressed in carpel and stamen (Figure 3C). Similarly, there was considerable overlap of expressed genes between the carpel and fruit (Figure 3D); this is not surprising since fruit is derived from the ovary containing large carpel tissues. Using the same cut-off for detection of expression, 5554 genes were expressed in all 8 tissues (Table S1 in Additional file 2). We then examined GO categorization of all 5554 genes and found that the "unknown" genes (homolog of genes annotated as unknown in *Arabidopsis*) were under-represented while some specific functional categories were slightly over-represented, including transferase and protein binding group (Additional file 3 and Additional file 1, Figure S2). The observation that most of the genes in this study were expressed in all tissues might be because our EST collection represented relatively abundant genes, including most house-keeping genes. This might also explain why the "unknown" category was under-represented because widely expressed genes tend to have known annotations.

To verify our microarray results, real-time reverse-transcription polymerase chain reaction (RT-PCR) was performed using RNAs from the same 8 tissues as those in microarray experiments. Nine representative genes were examined relative to our reference gene (Additional file 1, Figure S3), including three MADS-box genes, *EScaAGL2* (87251), *EScaAGL6* (86583), *EScaDEF1* (83744) [11]. The other genes were homologs of a transcription factor MYB35 (86850), a gamma-tip protein (84392), a putative ferrodoxin (85140), a transducin family/ WD-40 repeat family protein (84618), and homologs (86386 & 88941) of two *Arabidopsis* genes encoding different "expressed proteins" without a known function. The real time RT-PCR results indicate

that the gene expression patterns were generally supportive of the microarray results, and were also consistent with previous in RNA situ hybridization experiments [11, 12, 42, 43].

An overview of differential expression profiling of floral development

Although the *E. californica* ESTs were obtained from a cDNA library that was constructed with mRNAs from multiple stages of floral development [21], many of the corresponding genes were also expressed in leaves, different stages and various organs of the flower, as well as fruits. To determine additional transcriptome characteristics, we investigated whether specific genes were expressed similarly or differentially in the tissues tested. Of the 6446 unigenes examined, most genes (4513/6446) were not significantly differentially expressed with more than a two-fold change between any two of the eight tissues (with P-value < 0.05).

Nevertheless, 1933 genes were found to be differentially expressed between at least two tissues (Table S2 in Additional file 4); however, most of these 1933 genes showed similar expression levels in the other tissues (Figure 4A). Not surprisingly, carpel and fruit, as well as small and medium buds, showed the most similar expression patterns at sequential development stages. Leaf, the only vegetative organ in our study, had similar expression patterns to those of the green organs (carpel and fruit), which may be due to shared high expression of photosynthesis-related genes (see below). Interestingly, stamen had the most different expression profile, suggesting a distinct developmental process relative to the other floral organs.

To obtain additional insights into functions of those differentially expressed genes, we examined the GO categorization for the most similar *Arabidopsis* homologs of each poppy gene using functions within TAIR website [44] (Additional file 3). Genes encoding proteins categorized as "other enzyme activity" (chi-square test with p-value < 0.01) and "structural molecule" (p-value < 0.001) were enriched among those genes differentially expressed between at least two tissues (Figure 4C) relative to the control group of all genes on the microarray chip (Figure 4B). These results suggested that variation in the expression of metabolic genes across those tissues might be in part responsible for their morphological and/or physiological differences in *E. californica*.

Similar expression pattern of vegetative preferential genes in *E. californica* and in *Arabidopsis*

To identify genes with greater expression in either vegetative or reproductive tissues, we performed pairwise comparisons among all tissues as well as groups of floral organs and/or stages. Only one gene, 90036 (with no significant BLASTX hits to *Arabidopsis* predicted proteome, nor the NCBI NR database), was significantly two fold greater in all reproductive tissues and through all stages including fruit, compared to leaf tissue. However, 65 genes were expressed significantly higher in leaves compared to all floral tissues and stages (Table S2 in Additional file 4). To obtain overall expression patterns of vegetative genes, we constructed a heat-map (Figure 5A) resulting in two main clusters. In the first cluster, most genes that were highly expressed in leaves were also highly expressed in floral tissues except stamens. In the second cluster, most genes were highly expressed in leaves but not in the other tissues.

To compare gene expression pattern of leaf-preferential genes in *E. californica* and their homologs in *Arabidopsis*, we used BLAST to search the *E. californica* EST sequences against the *Arabidopsis* genome. Our BLAST results (with $10E^{-10}$ as cut-off) indicate that 58 out of the 65 leaf-preferential genes have identifiable homologs in *Arabidopsis*. On the basis of previous microarray data, of these 58 genes all but one (*RBCS1A*) of their *Arabidopsis* homologs were also preferentially expressed in leaves (Table S4 in Additional file 5) [45]. According to TAIR9 annotation, most of these genes encode for proteins that are localized in the chloroplast. GO categorization on the basis of gene function (methods) indicate that most of these genes are likely to be involved in photosynthesis, encoding homologs of protochlorophyllide reductases, photosystem I reaction center subunits and oxygen-evolving enhancer proteins.

Comparing transcriptome profiles at crucial stages of floral development in *E. californica* and in *Arabidopsis*

To identify developmental stage-specific genes in *E. californica* flowers, we examined the expression patterns of genes in the pre-meiotic (small buds), meiotic (medium buds) and pre-anthesis stages (four floral organs: sepals, petals, stamens and carpels). Pre-meiotic buds (small

buds < 5mm) had 49 differentially expressed genes in comparison with any other tissues examined (*p*-value <0.05 and two-fold cut-off, Table S2 in Additional file 4). Among these genes, 30 had identifiable *Arabidopsis* homologs, 24 of which have expression data available (Table S4 in Additional file 5). Unlike leaf-preferential genes, only seven of these 24 genes showed expression peaks in early *Arabidopsis* flower buds while the rest were predominately expressed in specific floral organs at higher levels than in leaves. The proteins encoded by these seven genes include two transcription factors, one oxidoreductase, two peroxidases, one electron carrier and one gene of unknown function (Table 1, genes and annotation with peak expression in small floral buds, information obtained from Markus' results [45]). The *Arabidopsis* homologs for two transcription factors, *MYB35,* which regulates anther cell layer formation at early stages and a bHLH gene that has not been fully studied [46, 47], were also preferentially expressed in anthers (Xuan Ma and Baomin Feng, unpublished data). However, the corresponding *E. californica* genes were expressed at low levels in the pre-anthesis stamens, possibly because either these genes are not highly expressed in *E. californica* stamens or that our stamen expression data from pre-anthesis stamens was too late relative to the stages of highest expression in *Arabidopsis*, which may be during earlier anther developmental stages.

In medium buds (which span the meiotic stage), we found eight genes that were expressed two fold significantly higher and none that were significantly down regulated compared with any of the other tissues examined (Table 1). All of these genes have homologs in *Arabidopsis* and most encode proteins that may have enzymatic activities (Table 1). However, none of the *Arabidopsis* homologs of those genes show expression peaks in the equivalent stages as our medium buds in *Arabidopsis* [45] (Table 1, Table S4 in Additional file 5). Interestingly, the homolog of *E. californica* gene 88096 in *Arabidopsis* (AT3G11450) encodes a DnaJ heat shock protein proposed to be involved in either mitosis or meiosis. The expression pattern of these homologs differs in that in *Arabidopsis* it is highly expressed in both vegetative and reproductive tissues. It is possible that the gene function might have diverged after the separation of basal eudicots from core eudicots.

In fruits, 9 genes were expressed significantly two fold higher than the other tissues in *E. californica* (Table 1). None of their homologs showed an expression peak in the *Arabidopsis* fruit. Among the genes of particular interest, the *Arabidopsis* homolog of 86118 (At5g62200,

MMI9) plays an important role in embryo development [48], and given its high expression in the fruits, suggesting that its *E. californica* homolog might have a similar function.

Identification of putative genes under control of certain genes in ABC model

According to the ABC model, A-function genes are transcription factors that are required to properly specify the sepal (alone) and petal (along with B-function genes) identities, with B-function genes specifying the stamens (along with C function genes), and C function specifying the carpels. Thus, genes expressed in the sepals and petals (regions encompassing the A-domain) are called A-domain genes, genes expressed in the petals and stamens are called B-domain genes, and genes expressed in the stamens and carpels are called C-domain genes. Although the homologs of *Arabidopsis* A function genes (such as AP1 and AP2) might not have conserved functions in other eudicots [47-49], because of the distinct sepals and petals in *E. californica*, we tried to identify putative A-function genes on the basis of regulatory genes expressed in the A-domain, hypothesizing that they may function in specifying the sepal and petal identities in *E. californica*.

From our hypothesis that A-domain genes should be more highly expressed in the sepals and possibly in the petals, than in the other floral organs we compared them with three tissues: leaf, stamen and carpel collected approximately 1 day pre anthesis. We found significantly greater expression of 64 genes in sepals over each of the above three tissues and 49 genes in petals over each of the three tissues, respectively (Table S5 in Additional file 6). When compared with all seven other tissues, 34 genes in sepals and 29 genes in petals were significantly preferentially expressed (Table S2 in Additional file 4). Whereas genes highly expressed in sepals or petals tended to be expressed in all tissues at moderately high levels (Figure 5B and C), genes with lower expression in sepals and/or petals were scarcely expressed in other tissues. On the basis of comparisons of petals and sepals with leaves, stamens and carpels, only 5 genes were expressed two fold greater in tissues controlled by A-function genes (Table 2). Interestingly, two of these genes are members of the MADS-box family. However, the expression of their closest *Arabidopsis* homologs, *AGL2* /*SEP1* and *AGL6*, is not sepal, petal or even floral specific (Figure 6D and F). *SEP1* is an E-function gene [49, 50], and is involved in the development of all floral

organs in *Arabidopsis*. A homolog of *SEP1* in soybean (*GmSEP1*) is expressed in reproductive development especially in petals and seed coats [51]. *AGL6* and its homologs have been shown to function in flower development not only in eudicots, like *Arabidopsis* and *Petunia*, but also in orchid, rice, and other monocots. In the grasses, *AGL6* has high expression in paleas, lodicules, carpels and ovule integuments, as well as the receptacle [52-56]. We hypothesize that other *MADS* genes, possibly *SEP* homologs, may serve as A-function genes in *E. californica* instead of *AP1* and *AP2* in *Arabidopsis*, in part because the *AP1* subfamily is closely related to the *AGL6* and *SEP* subfamilies [57].

B-function genes, such as the *Arabidopsis APETALA3* and *PISTILLATA* genes, are required for the identities of petals and stamens [10, 12, 58]. In monocots like tulip, homologs of *AP3* and *PI* are expressed in the tepals (petal-like organs found in the outer two whorls). We searched for putative B-domain genes on the basis of their expression patterns in *E. californica* and found that 60 genes in petals and 180 genes in stamens were expressed significantly higher in these organs than sepals, carpels and leaves (Table S5 in Additional file 6). And 94 genes were 2 fold significantly greater in stamens than all the other organs (Table S2 in Additional file 4). The large number of genes with stamen-preferential expression patterns suggested that the development of stamen requires more specialized genes. Alternatively, the larger number of stamen-preferential genes identified here may be explained by the fact that stamens comprise much of the biomass of developing *E. californica* buds, relative to other developing floral organs (Figure 5D).

We combined the expression data from petals and stamens to represent the B-domain group and compared their expression levels with those of leaves, sepals, carpels and fruits (Table 2), identifying 13 genes as preferentially expressed in the B-domain organs. A homolog of *PI* (87167) and two homologs of *AP3* (83744 and 87005) were identified in this group [12] (Table S5 in Additional file 6). Since *PI* and *AP3* are B-function genes in *Arabidopsis* and other species, such as lily [59-61], it is possible that their homologs in *E. californica* function in a similar manner. It should also be noted that *in situ* analysis showed that the *AP3* homologs are also expressed in ovules in *E. californica* [12], suggesting that they may have roles outside of B-function.

Of the genes preferentially expressed in the B-domain, one is a homolog of the At*bZIP61* gene, which encodes a putative transcription factor and is expressed in *Arabidopsis* flowers, with especially high expression in petals. It is not known whether At*bZIP61* regulates floral development in *Arabidopsis*. However, on the basis of its expression pattern and that of its homolog in *E. californica* we speculate that its function is downstream of the B function genes to regulate petal development.

In *Arabidopsis* C-function is controlled by *AGAMOUS,* which specifies the stamens and carpels. When compared with leaves, sepals and petals, 26 genes were preferentially expressed in carpels (compared to 168 genes in stamens) (Table S5 in Additional file 6). We searched for C-domain genes and found that 7 genes (Table S5 in Additional file 6) were expressed 2-fold significantly greater in stamens and carpels than in leaves, sepals and petals. Among them was a homolog of the *Arabidopsis* C-function gene *AG* [61]. Since both monocots (rice) and other eudicots have *AG* homologs functioning in stamen and carpel development, we hypothesize that the *AG* homolog in *E. californica* has similar functions [11, 62, 63]. It has been proposed that D-domain genes are required for ovule development, but only one *E. californica* gene (88769) was expressed in carpels two fold significantly higher over all other tissues. This EST did not have an identifiable *Arabidopsis* homolog.

To uncover additional candidates of A-, B- or C-domain genes, we used less stringent criteria and selected genes with expression levels at least two fold higher in each pre-anthesis reproductive tissue than in leaves (with FDR <0.05) (Figure 3E and F, Table S6 in Additional file 7). We found that most of these genes were expressed in a whorl-specific manner and only a small numbers of genes were co-upregulated in sepals and petals, in petals and stamens, or in stamens and carpels. Furthermore, the overlap of A/B-domain and that of B/C-domain genes were even smaller (Figure 3E and F). Unlike studies in *Persea* and *Aquilegia*, whose floral transcriptomes were interpreted as support for a "fading borders" model of floral organ identity [30, 64], the *E. californica* floral transcriptomes were rather distinctive, providing a molecular explanation for the morphologically different sepals and petals. Therefore, *E. californica* might have adopted an ABC model with relatively sharp borders, similar to those found in core eudicots. Because *E. californica* is basal to *Aquilegia* within the Ranunculales, as determined by phylogenetic analyses [38], it may be that sharply defined floral organ borders represent an

ancestral state for all eudicots, but has been lost in some more derived lineages.

Expression profiles of members of regulatory gene families

To gain further insights into the transcriptional activities of putative regulatory genes in floral development, we focused on gene families that are homologous to known regulators of plant development, particularly those encoding known or putative transcription factors. For convenience, we will refer to their predicted functions without using the words putative or predicted.

MADS: Genes containing a MADS-box DNA binding domain represent the best-studied floral gene family of which multiple members are crucial for floral development. In *E. californica* the expression of *EscaAG1* (84248), *EscaAG2* (86612), *EScaAGL2* (87251), *EScaAGL9* (87125), *EScaAGL11* (89484), *EScaGLO* (87167), *EScaDEF1* (83744) and *EScaDEF2* (87005) have been studied with *in situ* hybridization [11, 12, 42, 43]. Additionally, MADS-box genes homologous to those lacking characterized functions in *Arabidopsis* were included on our array, such as *EScaAGL54* (87912). Expression of *EScaAGL54* was highest in small buds, but showed similar levels in all the other tissues, suggesting a putative function in early floral stages.

To further understand the expression of the *E. californica* MADS-box genes, we plotted *E. californica* unigene expression profiles in comparison to the closest *Arabidopsis* homologs [12]. Expression patterns were largely similar between the two species, but there were some interesting differences (Figure 6). Both of the *E. californica AP3* homologs showed similar expression patterns to *AP3*, differing only in that 87005 (*EscaDEF2*) showed lower expression in all tissues relative to 83744 (*EscaDEF1*) or *AP3* in *Arabidopsis* (Figure 6A). At the same time, 87167 (*EscaGLO*), a homolog of *PI,* showed similar expression to *PI* in *Arabidopsis* (Figure 6B). Additionally, the *E. californica* homologs of the *Arabidopsis* C-function gene *AG* both showed similar expression to that of *AG*, again with one expressed lower than the other (Figure 6C). Besides those key *MADS* box genes regulating floral development, we found that *E. californica* homologs of E-function genes also have similar expression pattern as E-function genes in *Arabidopsis* (Figure 6D, E).

Homologs of other MADS box genes demonstrated different expression patterns. Unigene 84248 (*EscaAG1*, an *AG* homolog [65]) was highly expressed in stamens and carpels as expected, while 86612 (*EscaAG2*, a second *AG* homolog [65]) exhibited similar levels of expression in all floral tissues, suggesting a divergent function for this gene in *E. californica* flower development (Figure 6C). Also, the homolog of *AGL6* (86583) also showed a higher expression in sepals and petals (Figure 6F), in contrast to the low expression of the *Arabidopsis AGL6* gene in sepals on the basis of microarray expression [66]. Since a homolog of A-function gene has not been found in *E. californica*, it is possible that 86583 may function in the outer two whorls as an A- function gene (Figure 6F).

AGO: The ARGONAUTE family is involved in RNA post-transcriptional regulation [67]. In *Arabidopsis*, members of the AGO family are involved in floral development, most likely through miRNA and siRNA silencing. Our microarray included 10 members of the AGO family, all of which were differentially expressed in at least one tissue (expression data of family members listed in Additional file 8; Figure 7A and B). Among those genes, there was an interesting pattern, which identified three genes that were generally highly expressed in all organs while the remaining seven genes were expressed at a moderate/low level.

Among the genes examined in this study, three *AGO1* homologs (1 in the high expression group and two in the low expression group) shared similar expression patterns: two fold higher expression in petals, pre-meiotic and meiotic buds than in sepals. The AGO genes in *Arabidopsis* encode proteins with a PAZ domain (with nucleic acid binding activity [68]) and are expressed at similar levels in different tissues except PAZ-1, which was preferentially expressed in carpel, pre-meiotic and meiotic buds compared with sepal with more than two fold changes.

MYB: MYB transcription factors contain DNA binding domains and some have been identified as flower developmental regulators [69, 70]. Eleven *E. californica MYB* genes were included on our microarray. Most *MYB* genes showed dramatic differential expression among tissues, but two of them were not differentially expressed among any of the tissues tested (Additional file 8, Figure 7C). One homolog of At4g32730 (*MYB1*) was expressed at higher levels in mature petals and stamens, suggesting that this gene may have a role in B-function. A homolog of At4g32730 (*AtMYB3R1*) was significantly preferentially expressed (> two fold higher) in the pre-meiotic

bud in comparison with sepals, petals, and stamens and carpels. A homolog of At3g28470 (*AtMYB35*) was also preferentially expressed in pre-meiotic buds compared with all seven other tissues. A homolog of At4g01680 (*AtMYB55*) was significantly preferentially expressed in fruit in comparison with leaves, pre-meiotic and meiotic buds, petals, sepals and stamens. An At2g37630 (*AtMYB91/AS1*) homolog was more varied in expression but generally showed lower expression in stamens than in carpels, fruits, leaves, pre-meiotic and meiotic buds and lesser down regulation in petals relative to carpels, leaves and pre-meiotic buds. Last but not the least, a homolog of At3g61250 (*AtMYB17*) was expressed two fold significantly higher in meiotic buds compared with fruits.

ZF-HD: Zinc finger homeodomains are expressed during floral development in *Arabidopsis* [71]. Our microarray contained four genes in this family. Two homologs of At1g75240 (ATHB33) were expressed without significant difference across all tissues. Of these two genes, one (88691) was expressed highly in both vegetative and reproductive organs while the other was barely expressed in all tissues, suggesting a functional divergence between these two paralogs (Figure 7D).

ARF: Auxin-response factors (ARFs) are believed to regulate auxin responsive genes [72, 73]. This family contains *ETTIN* (At2G33860)*, a developmental regulatory gene that acts on regional identity in the perianth, stamens and carpels [74]. Most of the poppy ARF genes that were included on our microarray showed no differential expression among the tissues examined (Figure 7E). Only one gene, a homolog of the At5g62000 (ARF2, 84471), showed two fold significantly differential expression: two fold lower in stamens when compared with all tissues but sepal; and two fold lower in sepals compared with carpels, fruits and pre-meiotic buds.

bZIP: The bZIP protein family contains the *Arabidopsis FD* (At4G35900, FD-1) and *PERIANTHIA* (At1G68640) genes, which are involved in flower development and the *HY5* (At5G11260) gene involved in root development. Our array contained 12 members of this family, one of which was not differentially expressed among all tissues examined (Additional file 8, Figure 7F). From our microarray results, most of these genes showed only slightly different expression levels except the homologs of *bZIP7* (83748) and *bZIP8* (87035), both of which were expressed highly in stamens, with *bZIP8* also highly expressed in petals. Previous

studies of genes in bZIP family suggested that some of them may act downstream of B-function genes to regulate floral development [75-77]. Because the homolog of *bZIP8* was co-expressed with B function genes, we speculate that this gene might have a function similar to that of the *Arabidopsis* homolog. In addition, a homolog of At4g38900 is expressed at a level two fold higher in sepals than in stamens.

bHLH: The basic helix-loop-helix family contains several *Arabidopsis* genes regulating flower development including *SPATULA*, which controls the development of the carpel margins [78]. Eleven members of this family were included on our microarray, seven of which showed no significant differential expression (Additional file 8, Figure 7G). The other four genes demonstrated two fold differential expressions among tissues examined. A homolog of At2g31210 (bHLH91, 89282) was most highly expressed in pre-meiotic buds and the expression level was at least two fold higher than in all the other tissues; and its expression level in meiotic buds was at least two fold higher than any other floral organs. Since At2g31210 has an important role in anther development in *Arabidopsis* [47], its homolog in *E. californica* may function in a similar manner. Another gene, a homolog of At5g09460 (bHLH143), was also expressed at a higher level in the pre-meiotic buds than in sepals, petals and stamens and in meiotic buds. Additionally, this gene was expressed two fold higher in carpels and fruits than in stamens. A homolog of At1g26260 (bHLH76, CIB5) was expressed in pre-meiotic buds significantly two fold higher than in fruits and stamens. A homolog of At3g26744 (bHLH116/ICE1) was significantly down regulated by two fold in stamens relative to carpels, fruits, leaves, meiotic buds and petals. This gene was also significantly more highly expressed by two fold in petals over sepals. The expression patterns of bHLH genes suggest that they might regulate several aspects of floral development and/or physiology, but not necessarily associated with ABC functions. Further study of bHLH genes, and indeed many of the floral gene families examined here, in *Arabidopsis* and other species, including *E. californica*, may uncover their functions and reveal possible functional conservation among the eudicots.

Conclusions

We examined transcriptome landscapes from eight tissues of the basal eudicot *E. californica* and identified preferentially expressed genes within and among floral developmental tissues, fruits and leaves. By comparing genes showing tissue-preferential expression patterns in *E. californica*, we found that genes preferentially expressed in specific reproductive organs or at certain stages tended to have less conserved expression levels compared with *Arabidopsis* than those preferentially expressed in leaves (Table 1 and 2, Table S4 in Additional file 5). One possible explanation is that most of the leaf-preferential genes encode highly conserved chloroplast proteins.

We also identified the co-expressed and tissue-specific floral genes and characterized the signature of ABC domain genes. Our comparison of the gene expression patterns in *E. californica*, *Aquilegia*, *Persea* and *Arabidopsis* showed that the *E. californica* results support a "sharp border" model, similar to that for core eudicot such as *Arabidopsis*, rather than the "fading border" model in other basal angiosperms [30, 64]. This is consistent with the clear morphological distinction of sepals and petals, and the lack of intermediate floral organs such as staminodes in *E. californica* flowers. In contrast, *Aquilegia* flowers have similar outer perianth organs and a distinct type of floral organ between stamens and the carpels, which is in good agreement with the microarray results of the floral organs [30]. Therefore, although both *E. californica* and *Aquilegia* are basal eudicots, the morphological and expression characteristics strongly suggested that they have divergent developmental programs, with *E. californica* more similar to core eudicots and *Aquilegia* resembling basal angiosperms. Our analysis of *E. californica* further suggested that flowers with distinct perianth organs might have originated at an earlier time than the ancestor of core eudicots. This study along with other works [22, 30] highlight the importance of careful analysis of basal eudicots as an intermediate group of flowering plants to provide crucial information to bridge the gap between highly canalized core eudicots and morphological flexible basal angiosperms.

Out data also provided an overview of divergence and conservation between different species. The highly similar expression patterns of B- and C-function genes compared with the varied expression levels of other MADS box genes in *Arabidopsis* and *E. californica* suggested that the conserved expression of only a few key genes may result in the high similarity of flower morphology between *Arabidopsis* and *E. californica* . The transcriptome analysis of other

families with known functions in floral development indicates their possible roles in *E. californica*. Recent study of protein-protein interaction in basal eudicots (*Euptelea pleiospermum*, *Akebia trifoliata* and *Pachysandra terminalis*) suggested that *MADS* box genes that interact with each other have co-evolved. This is most likely due to the fact that the majority of the protein-protein interactions are expected to be conserved to some extent to orchestrate floral architecture [79]. However, Zhao *et al.* [73] showed the AP1 lineage had a distinct interaction pattern; this together with our results that *AGL6* and *SEP* homologs are expressed in the A-domain support that A-function genes show less conservation [58]. In *Arabidopsis*, *AP1* not only regulates the development of sepal and petal, but also integrates growth, patterning and hormonal pathways [80]. This dual function of *AP1* observed in the core eudicots might be a more recent innovation that evolved since the divergence of the core from the basal eudicots.

Many of the genes showing tissue specific expression noted in this study have homologs in *Arabidopsis* that are currently lacking in functional analyses. This study, when compared with similar studies in *Arabidopsis* and other species should help us identify genes of interest that may play important, conserved, roles in floral development [27, 29, 30]. We have identified a number of candidate genes that share similar expression patterns between *E. californica* and *Arabidopsis*, but have not been functionally characterized. Our results suggest that *E. californica* has a similar floral program to the core eudicots, despite a mostly divergent set of genes outside of the MADS-box family. These results not only indicate that different regulatory machinery may operate among basal eudicots, but that canalized floral development might have originated prior to the core eudicots. Our findings also allow for informative comparisons with other species, allowing hypotheses formulation and stimulating further experimentation in model organisms, which now includes *E. californica*.

Materials and methods

Tissue collection and RNA isolation

Sixteen *E. californica* cv. 'Aurantica Orange' (J.L. Hudson Seedsman) plants were grown from seeds in a controlled greenhouse environment at the Pennsylvania State University (University Park, PA) under 16 hours light and watered and fertilized as needed. To avoid potential expression differences among collections due to circadian rhythms leaves, floral tissues were only collected from individual plants between 8:30-10:30 AM. Developing leaves of less than 5mm length, developing fruits, pre-meiotic (small) buds less than 5mm long, meiotic (medium) buds of 5-10mm length and pre-anthesis sepals, petals, stamens and carpels were collected from 16 plants, immediately placed in liquid nitrogen and stored in a -80°C freezer until RNA extraction. Tissues from a group of four plants were then pooled to create one biological replicate, for a total of four replicates.

Probe design for the *E. californica* transcriptome

To design oligonucleotide probes for *E. californica*, a two-stage pipeline for oligonucleotide probe design, Microarray Oligonucleotide Design and Integration Tool (MODIT) was used (probe information provided in Additional file 9). Briefly, MODIT integrates two existing programs: Array Oligo Selector (6) (AOS) and OligoArray (8) (OA), with subsequent independent evaluation and optimization steps. The pipeline enables one to design a set of probes having well-defined sequence and thermodynamic properties by first taking advantage of the strict thermodynamic criteria of OA, to produce a partial set of optimized probes, and then fills in the set from among the large number of probes selected by AOS, after screening them for thermodynamic compatibility.

The MODIT pipeline screens candidate probes based on three parameters: high sequence specificity, appropriate melting temperature T_m, and lack of stable secondary structure. The first criterion, sequence specificity, was determined using BLAST and Smith-Waterman local alignment tools to eliminate probes having a match to any non-target sequences of more than 15 consecutive nucleotides, or an overall match of more than 30 nucleotides [81-83]. The second criterion was that the probe set should have very similar T_m. The MODIT user is informed of probes with T_m outside a recommended range by flagging in the database, and she/he can decide whether to use such probes. A third criterion was the lack of stable secondary structure. MODIT

allows values of probe ΔG_{SS} above -0.5 kcal.mol^{-1}, less than the energy of one hydrogen bond between bases [84]. We use melting temperature to independently recalculate a consistent set of thermodynamic properties for the probes and check for consistency [85]. The pipeline stores comprehensive information about probe thermodynamic properties and potential cross reactions in a MySQL database, so that they can subsequently be used in array data analysis.

The MODIT pipeline was used to generate one 60 base probe for each gene in the 6,846 *E. californica* Unigene set [86, 87], after masking regions that were conserved in multigene families in *Arabidopsis,* rice (*Oryza*) and *Populus*. Unigenes were sorted into gene families using PlantTribes [88] and conserved sites in the multiple sequence alignment were identified using the column score metric calculated by CLUSTAL [89]. Sodium concentration of 0.5 M was used in modeling of thermodynamic properties, following hybridization conditions recommended by Agilent for their 60-mer *Arabidopsis* Oligo Microarray Kit, and the conditions modeled by Lee *et al*. [90]. The probe concentration range that was used in the thermodynamics calculations is 2.44 mM following the calculations of Riccelli *et al*. and assuming the default 1 nM target recommended in [91, 92]. In the OA run, duplex melting temperature T_m was constrained above 70 °C, and the duplex T_m for predicted cross-reactions and stable secondary structures was constrained below 60 °C. For the AOS run, the constraint on GC content was maintained around 52%. Duplex melting temperature was constrained to keep 20 °C separation between the upper and lower T_m limits, to allow for selection of more candidate probes**.** The probe maximum and minimum match for non-target sequences were maintained at 15 and 10 nucleotides, respectively. When the two sets of probes were merged, the constraints applied to the merged set were: 80 °C $\leq T_m \leq$ 90 °C, overall match with non-target as well as with consensus sequences should be less than 30 nucleotides and ΔG_{SS} above -0.5 kcal.mol^{-1}. Since one goal of this design was to obtain complete coverage of all target sequences, a selection of known suboptimal probes was added back to the final design (Column 5, Additional file 9), and their sequence and thermodynamic properties tracked in the MODIT database. The design results obtained using MODIT for the target sequences from *E. californica* are summarized in Additional file 9. No application, including MODIT, could provide 100% target coverage while satisfying all of the design criteria for each probe. However, MODIT improved on target coverage and significantly limited potential cross reactions relative to OA, while nearly eliminating probes which were predicted to form stable secondary structure.

Oligonucleotides of 60 base pair length were designed from 6446 *E. californica* unigenes obtained from a floral EST library [21] and cell culture suspension library [93]. Unigene builds were performed as described by Carlson *et al*. (2006) and the sorted into putative gene families using the PlantTribes database [21, 94]. Because the complete genome of *E. californica* is not yet sequenced, oligos were designed to specifically exclude conserved regions, when identified, so that expression analyses putatively represent single genes (see above). Oligonucleotide probes were arrayed on glass slides by Agilent (La Jolla, CA, USA).

RNA extraction, microarray hybridization and scanning

RNA was isolated from eight tissues examined each with four biological replicate pools and cleaned using the RNeasy plant mini Kit (Qiagen, USA) following Agilent's instructions. RNA concentrations were quantified using an Agilent 2100 Bioanalyzer and stored at -80°C before use, with yields of 20-35 micrograms of total RNAs from ~100 mg of tissues. Approximately 400 ng of total RNAs were used for cRNA synthesis with Cyanine 3-dCTP and Cyanine5-dCTP (Perkin–Elmer Life Sciences, Inc., USA) incorporation, using the Agilent Low RNA Input Kit (Agilent, USA), according to the manufacturers' protocol. Qiagen's RNeasy mini spin columns (Qiagen, USA) were used to purify amplified cRNA samples. Sample concentrations and were quantified using a NanoDrop spectrometer (NanoDrop Technologies, USA). Hybridization was performed using the In situ Hybridization Kit (Agilent, USA) with 35ng of Cy3 and Cy5 labeled cRNA followed the instructions at 65°C for 17h. Prior to scanning, each slide was washed, rinsed and dried in Agilent's Stabilization and Drying Solution, as directed. Scanning was performed using a Gene Pix 4000A scanner and the Gene Pix Pro 3.0.6 Software (Axon Instruments (now Molecular Devices, USA) to produce two TIFF images at 532nm and 635nm. The microarray data were submitted to the GEO database, with accession number [GEO:GSE24237].

Statistical analyses of genes differentially expressed among tissues and developmental stages

Analyses were performed with the R programming language [95] and the limma package Bioconductor [96]. Arrays were background corrected and loess normalized within arrays and

Aq normalized between arrays [97]. Agilent controls and other control probes were removed from the data. For the 93 *E. californica* oligos with multiple probes, we chose the probe with the highest 75% quantile value among the normalized "A" intensities of all 16 arrays. A one-way single-channel empirical Bayes ANOVA was used to identify those genes [98, 99] that were significantly differentially expressed among the seven floral RNAs and one leaf RNA examined, with an FDR of 0.05. Additionally, significant differences between combinations of more than one floral organ and leaf were also identified under the same parameters.

In order to identify those genes that were most likely to be organ/stage specific in *E. californica*, we examined those genes with a significantly (FDR=0.05) two fold greater expression in a single organ/stage relative to all other tissues stages examined. The expression of these genes was then compared to the expression, as determined by Affymetrix arrays [29], to their closest identified *Arabidopsis* homolog based on a tribe-MCL analysis, when available, to determine which genes may have conserved expression profiles. We were able to directly compare expression in pre-meiotic and meiotic buds in *E. californica* vs. inflorescences containing stage 1-9 flowers in *Arabidopsis* (developing inflorescences), the *E. californica* fruit, capsules, vs. the *Arabidopsis* fruit, siliques, *Arabidopsis* flowers at stage twelve nearing pre-anthesis vs. sepals, petals stamens and carpels at anthesis in *E. californica* and genes preferentially expressed in leaves in both organisms.

Real-time PCR experiments

To test the reliability of our microarray hybridizations, nine genes and one reference were investigated using Quantitative Real-Time PCR. RNA (1ug) of each tissue was treated with DNase (Invitrogen, USA), followed by reverse transcription using the Superscript III reverse transcriptase (Invitrogen, USA). We then performed real time PCR using DyNAmo SYBR Green qPCR Kit from New England Biolabs (NEB, USA) under the following parameters: 95°C 10min, 40 cycle for 95°C 30 sec, 60°C for 1min. Fluorescence intensity was measured using ABI's 7300 Sequence Detection System. Eca_2514 (Unigene84142) was chosen as the reference gene as it was not significantly differentially expressed among any of our examined tissues in the microarray experiments and it was expressed at a moderate level in all our tissues compared to

all other genes. The relative amounts of cRNA converted from a messenger RNA was calculated using intensities corresponding to "experimental" genes relative to the reference gene. We performed triplicate reactions for all tissues with samples containing no reverse transcriptase and no RNA as negative controls. All primer information is provided in Additional file 9.

Abbreviations

ARF, Auxin-response factors; AG, AGAMOUS; AGL, AGAMOUS-like gene; AGO, ARGONAUTE; AP, APETALA; bHLH, basic Helix-Loop-Helix; bZIP, basic-leucine Zipper; DEF, DEFICIENS; ESca, *Eschscholzia californica*; EST, Expressed Sequence Tag; GO, Gene Ontology Consortium; GLO, GLOBOSA; MYB, Myeloblastosis-like gene; PI, PISTILLATA; RACE, Rapid Amplification of cDNA Ends; TAIR, The Arabidopsis Information Resource web site; ZF-HD, Zinc finger homeodomains; L, leaves; C, carpels; F, fruits; P, petals; SB, small buds; MB, medium buds; S, sepals; ST, stamens; STD, staminodia; OT, outer tepals; IT, inner tepals.

Authors' contributions

LMZ, HM, NSA, JLM and CWD designed the study; DT and LMZ performed tissue collection and RNA isolation; CJG and RG designed the oligonucleotides for the probes used in the microarray chip; XM, LMZ and DT performed RT-PCR experiments; LMZ, XM, NSA, QZ, and PKW performed data analysis; LMZ, XM, and HM wrote the manuscript drafts; LMZ, XM, HM, NSA, JLM, CWD, CJG, RG, and PKW edited the manuscript; all authors approved the manuscript.

Acknowledgements

References

1. Friis EM, Pedersen KR, Crane PR: **Cretaceous anglosperin flowers: Innovation and evolution in plant reproduction**. *Palaeogeogr Palaeocl* 2006, **232**:251-293.
2. Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE: **Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots**. *Proc Natl Acad Sci USA* 2010, **107**:4623-4628.
3. Anderson CL, Bremer K, Friis EM: **Dating phylogenetically basal eudicots using rbcL sequences and multiple fossil reference points**. *Am J Bot* 2005, **92**:1737-1748.
4. Leebens-Mack J, Raubeson LA, Cui L, Kuehl JV, Fourcade MH, Chumley TW, Boore JL, Jansen RK, depamphilis CW: **Identifying the basal angiosperm node in chloroplast genome phylogenies: sampling one's way out of the Felsenstein zone**. *Mol Biol Evol* 2005, **22**:1948-1963.
5. Cajsa Lisa Anderson KB, Else Marie Friis **Dating phylogenetically basal eudicots using rbcL sequences and multiple fossil reference points**. *Annu Bot* 2005, **92**:12.
6. Endress PK, Doyle JA: **Floral phyllotaxis in basal angiosperms: development and evolution**. *Curr Opin Plant Biol* 2007, **10**:52-57.
7. Putterill J, Laurie R, Macknight R: **It's time to flower: the genetic control of flowering time**. *Bioessays* 2004, **26**:363-373.
8. Baek IS, Park HY, You MK, Lee JH, Kim JK: **Functional conservation and divergence of *FVE* genes that control flowering time and cold response in rice and *Arabidopsis***. *Mol Cells* 2008, **26**:368-372.
9. Liu C, Thong Z, Yu H: **Coming into bloom: the specification of floral meristems**. *Development* 2009, **136**:3379-3391.
10. Soltis DE, Ma H, Frohlich MW, Soltis PS, Albert VA, Oppenheimer DG, Altman NS, dePamphilis C, Leebens-Mack J: **The floral genome: an evolutionary history of gene duplication and shifting patterns of gene expression**. *Trends Plant Sci* 2007, **12**:358-367.
11. Zahn LM, Kong H, Leebens-Mack JH, Kim S, Soltis PS, Landherr LL, Soltis DE, dePamphilis CW, Ma H: **The evolution of the *SEPALLATA* subfamily of MADS-box genes: a preangiosperm origin with multiple duplications throughout angiosperm history**. *Genetics* 2005, **169**:2209-2223.
12. Zahn LM, Leebens-Mack J, DePamphilis CW, Ma H, Theissen G: **To B or not to B a flower: the role of *DEFICIENS* and *GLOBOSA* orthologs in the evolution of the angiosperms**. *J Hered* 2005, **96**:225-240.
13. Paolacci AR, Tanzarella OA, Porceddu E, Varotto S, Ciaffi M: **Molecular and phylogenetic analysis of MADS-box genes of MIKC type and chromosome location**

of *SEP*-like genes in wheat (*Triticum aestivum* **L.**). *Mol Genet Genomics* 2007, **278**:689-708.

14. Yamaguchi T, Hirano HY: **Function and diversification of MADS-box genes in rice**. *Sci World J* 2006, **6**:1923-1932.
15. Ma H, dePamphilis C: **The ABCs of floral evolution**. *Cell* 2000, **101**:5-8.
16. Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Munster T, Winter KU, Saedler H: **A short history of MADS-box genes in plants**. *Plant Mol Biol* 2000, **42**:115-149.
17. Soltis DE, Bell CD, Kim S, Soltis PS: **Origin and early evolution of angiosperms**. *Annu N Y Acad Sci* 2008, **1133**:3-25.
18. Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A, Albert VA, Ma H, dePamphilis CW: **Widespread genome duplications throughout the history of flowering plants**. *Genome Res* 2006, **16**:738-749.
19. Orashakova S, Lange M, Lange S, Wege S, Becker A: **The *CRABS CLAW* ortholog from California poppy (*Eschscholzia californica*, Papaveraceae), *EcCRC*, is involved in floral meristem termination, gynoecium differentiation and ovule initiation**. *Plant J* 2009, **58**:682-693.
20. Jaramillo MA, Kramer EM: **Molecular evolution of the petal and stamen identity genes, *APETALA3* and *PISTILLATA*, after petal loss in the Piperales**. *Mol Phylogenet Evol* 2007, **44**:598-609.
21. Carlson JE, Leebens-Mack JH, Wall PK, Zahn LM, Mueller LA, Landherr LL, Hu Y, Ilut DC, Arrington JM, Choirean S, Becker A, Field D, Tanksley SD, Ma H, dePamphilis CW: **EST database for early flower development in California poppy (*Eschscholzia californica* Cham., Papaveraceae) tags over 6,000 genes from a basal eudicot**. *Plant Mol Biol* 2006, **62**:351-369.
22. Nam J, Kim J, Lee S, An G, Ma H, Nei M: **Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms**. *P Natl Acad Sci USA* 2004, **101**:1910-1915.
23. Kramer EM, Dorit RL, Irish VF: **Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the *APETALA3* and *PISTILLATA* MADS-box gene lineages**. *Genetics* 1998, **149**:765-783.
24. Galbraith DW: **DNA microarray analyses in higher plants**. *OMICS* 2006, **10**:455-473.
25. Ma H: **Molecular genetic analyses of microsporogenesis and microgametogenesis in flowering plants**. *Annu Rev Plant Biol* 2005, **56**:393-434.
26. Tung CW, Dwyer KG, Nasrallah ME, Nasrallah JB: **Genome-wide identification of genes expressed in *Arabidopsis* pistils specifically along the path of pollen tube growth**. *Plant Physiol* 2005, **138**:977-989.
27. Wellmer A, von Mering M, Spreer A, Diem R, Eiffert H, Noeske C, Bunkowski S, Gold R, Nau R: **Experimental pneumococcal meningitis: impaired clearance of bacteria from the blood due to increased apoptosis in the spleen in *Bcl-2*-deficient mice**. *Infect Immun* 2004, **72**:3113-3119.
28. Yang WY, Yu Y, Zhang Y, Hu XR, Wang Y, Zhou YC, Lu BR: **Inheritance and expression of stripe rust resistance in common wheat (*Triticum aestivum*) transferred from *Aegilops tauschii* and its utilization**. *Hereditas* 2003, **139**:49-55.

29. Zhang X, Feng B, Zhang Q, Zhang D, Altman N, Ma H: **Genome-wide expression profiling and identification of gene activities during early flower development in Arabidopsis**. *Plant Mol Biol* 2005, **58**:401-419.

30. Voelckel C, Borevitz JO, Kramer EM, Hodges SA: **Within and between whorls: comparative transcriptional profiling of *Aquilegia* and *Arabidopsis***. *PLoS One* 2010, **5**:e9735.

31. Becker A, Gleissberg S, Smyth DR: **Floral and vegetative morphogenesis in California poppy (*Eschscholzia californica* Cham.)**. *Int J Plant Sci* 2005, **166**:537-555.

32. Bennett MD, Leitch IJ: **Nuclear DNA amounts in angiosperms: progress, problems and prospects**. *Annu Bot* 2005, **95**:45-90.

33. Park SU, Facchini PJ: ***Agrobacterium* rhizogenes-mediated transformation of opium poppy, *Papaver somniferum* l., and California poppy, *Eschscholzia californica* Cham., root cultures**. *J Exp Bot* 2000, **51**:1005-1016.

34. Becker A, Lange M: **VIGS--genomics goes functional**. *Trends Plant Sci* 2010, **15**:1-4.

35. Wege S, Scholz A, Gleissberg S, Becker A: **Highly efficient virus-induced gene silencing (VIGS) in California poppy (*Eschscholzia californica*): an evaluation of VIGS as a strategy to obtain functional data from non-model plants**. *Annu Bot* 2007, **100**:641-649.

36. Liscombe DK, Ziegler J, Schmidt J, Ammer C, Facchini PJ: **Targeted metabolite and transcript profiling for elucidating enzyme function: isolation of novel N-methyltransferases from three benzylisoquinoline alkaloid-producing species**. *Plant J* 2009, **60**:729-743.

37. Wall PK, Leebens-Mack J, Chanderbali AS, Barakat A, Wolcott E, Liang H, Landherr L, Tomsho LP, Hu Y, Carlson JE, Ma H, Schuster SC, Soltis DE, Soltis PS, Altman N, dePamphilis CW: **Comparison of next generation sequencing technologies for transcriptome characterization**. *BMC Genomics* 2009, **10**:347.

38. Moore MJ, Bell CD, Soltis PS, Soltis DE: **Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms**. *Proc Natl Acad Sci USA* 2007, **104**:19363-19368.

39. Sacan A, Ferhatosmanoglu N, Ferhatosmanoglu H: **MicroarrayDesigner: an online search tool and repository for near-optimal microarray experimental designs**. *BMC Bioinformatics* 2009, **10**:304.

40. Altman NS, Hua J: **Extending the loop design for two-channel microarray experiments**. *Genet Res* 2006, **88**:153-163.

41. Zhang X, Feng B, Zhang Q, Zhang D, Altman N, Ma H: **Genome-wide expression profiling and identification of gene activities during early flower development in *Arabidopsis***. *Plant Mol Biol* 2005, **58**:401-419.

42. Kim S, Koh J, Yoo MJ, Kong H, Hu Y, Ma H, Soltis PS, Soltis DE: **Expression of floral MADS-box genes in basal angiosperms: implications for the evolution of floral regulators**. *Plant J* 2005, **43**:724-744.

43. Drea S, Hileman LC, de Martino G, Irish VF: **Functional analyses of genetic pathways controlling petal specification in poppy**. *Development* 2007, **134**:4157-4166.

44. **The Arabidopsis Information Resource** [www.arabidopsis.org]

45. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU: **A gene expression map of *Arabidopsis thaliana* development**. *Nat Genet* 2005, **37**:501-506.

46. Zhu J, Chen H, Li H, Gao JF, Jiang H, Wang C, Guan YF, Yang ZN: **Defective in** *Tapetal Development and Function 1* **is essential for anther development and tapetal function for microspore maturation in** *Arabidopsis*. *Plant J* 2008, **55**:266-277.

47. Zhang W, Sun Y, Timofejeva L, Chen C, Grossniklaus U, Ma H: **Regulation of** *Arabidopsis* **tapetum development and function by** *DYSFUNCTIONAL TAPETUM1* (*DYT1*) **encoding a putative bHLH transcription factor**. *Development* 2006, **133**:3085-3095.

48. Ascencio-Ibanez JT, Sozzani R, Lee TJ, Chu TM, Wolfinger RD, Cella R, Hanley-Bowdoin L: **Global analysis of Arabidopsis gene expression uncovers a complex array of changes impacting pathogen response and cell cycle during geminivirus infection**. *Plant Physiol* 2008, **148**:436-454.

49. Cui R, Han J, Zhao S, Su K, Wu F, Du X, Xu Q, Chong K, Theissen G, Meng Z: **Functional conservation and diversification of class E floral homeotic genes in rice** (*Oryza sativa*). *Plant J* 2010, **61**:767-781.

50. Ma H, Yanofsky MF, Meyerowitz EM: *AGL1-AGL6,* **an** *Arabidopsis* **gene family with similarity to floral homeotic and transcription factor genes**. *Genes Dev* 1991, **5**:484-495.

51. Huang F, Chi Y, Gai J, Yu D: **Identification of transcription factors predominantly expressed in soybean flowers and characterization of** *GmSEP1* **encoding a SEPALLATA1-like protein**. *Gene* 2009, **438**:40-48.

52. Li H, Liang W, Jia R, Yin C, Zong J, Kong H, Zhang D: **The** *AGL6-like gene* *OsMADS6* **regulates floral organ and meristem identities in rice**. *Cell Res* 2010, **20**:299-313.

53. Ohmori S, Kimizu M, Sugita M, Miyao A, Hirochika H, Uchida E, Nagato Y, Yoshida H: *MOSAIC FLORAL ORGANS1,* **an** *AGL6-***like MADS box gene, regulates floral organ identity and meristem fate in rice**. *Plant Cell* 2009, **21**:3008-3025.

54. Reinheimer R, Kellogg EA: **Evolution of** *AGL6-***like MADS box genes in grasses (Poaceae): ovule expression is ancient and palea expression is new**. *Plant Cell* 2009, **21**:2591-2605.

55. Rijpkema AS, Zethof J, Gerats T, Vandenbussche M: **The petunia** *AGL6* **gene has a** *SEPALLATA-***like function in floral patterning**. *Plant J* 2009, **60**:1-9.

56. Hsu HF, Huang CH, Chou LT, Yang CH: **Ectopic expression of an orchid (Oncidium Gower Ramsey)** *AGL6-***like gene promotes flowering by activating flowering time genes in** *Arabidopsis thaliana*. *Plant Cell Physiol* 2003, **44**:783-794.

57. Litt A, Irish VF: **Duplication and diversification in the** *APETALA1/FRUITFULL* **floral homeotic gene lineage: implications for the evolution of floral development**. *Genetics* 2003, **165**:821-833.

58. Kanno A, Nakada M, Akita Y, Hirai M: **Class B gene expression and the modified ABC model in nongrass monocots**. *Sci World J* 2007, **7**:268-279.

59. Schultz EA, Pickett FB, Haughn GW: **The** *FLO10* **gene product regulates the expression domain of homeotic genes** *AP3* **and** *PI* **in** *Arabidopsis* **flowers**. *Plant Cell* 1991, **3**:1221-1237.

60. Tzeng TY, Kong LR, Chen CH, Shaw CC, Yang CH: **Overexpression of the lily** *p70(s6k)* **gene in Arabidopsis affects elongation of flower organs and indicates TOR-dependent regulation of AP3, PI and SUP translation**. *Plant Cell Physiol* 2009, **50**:1695-1709.

61. Riechmann JL, Meyerowitz EM: **Determination of floral organ identity by** *Arabidopsis* **MADS domain homeotic proteins AP1, AP3, PI, and AG is independent of their DNA-binding specificity**. *Mol Biol Cell* 1997, **8**:1243-1259.

62. Lee S, Kim J, Han JJ, Han MJ, An G: **Functional analyses of the flowering time gene** *OsMADS50*, **the putative** *SUPPRESSOR OF OVEREXPRESSION OF CO 1/AGAMOUS-LIKE 20* (*SOC1/AGL20*) **ortholog in rice**. *Plant J* 2004, **38**:754-764.

63. Kyozuka J, Shimamoto K: **Ectopic expression of** *OsMADS3*, **a rice ortholog of** *AGAMOUS*, **caused a homeotic transformation of lodicules to stamens in transgenic rice plants**. *Plant Cell Physiol* 2002, **43**:130-135.

64. Chanderbali AS, Albert VA, Leebens-Mack J, Altman NS, Soltis DE, Soltis PS: **Transcriptional signatures of ancient floral developmental genetics in avocado (Persea americana; Lauraceae)**. *Proc Natl Acad Sci U S A* 2009, **106**:8929-8934.

65. Zahn LM, Leebens-Mack JH, Arrington JM, Hu Y, Landherr LL, dePamphilis CW, Becker A, Theissen G, Ma H: **Conservation and divergence in the AGAMOUS subfamily of MADS-box genes: evidence of independent sub- and neofunctionalization events**. *Evol Dev* 2006, **8**:30-45.

66. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU: **A gene expression map of Arabidopsis thaliana development**. *Nature Genetics* 2005, **37**:501-506.

67. Chen X: **Small RNAs and their roles in plant development**. *Annu Rev Cell Dev Biol* 2009, **25**:21-44.

68. Lingel A, Simon B, Izaurralde E, Sattler M: **Structure and nucleic-acid binding of the** *Drosophila* **Argonaute 2 PAZ domain**. *Nature* 2003, **426**:465-469.

69. Peng J: **Gibberellin and jasmonate crosstalk during stamen development**. *J Integr Plant Biol* 2009, **51**:1064-1070.

70. Martin C, Bhatt K, Baumann K, Jin H, Zachgo S, Roberts K, Schwarz-Sommer Z, Glover B, Perez-Rodrigues M: **The mechanics of cell fate determination in petals**. *Philos Trans R Soc Lond B Biol Sci* 2002, **357**:809-813.

71. Tan QK, Irish VF: **The Arabidopsis zinc finger-homeodomain genes encode proteins with unique biochemical properties that are coordinately expressed during floral development**. *Plant Physiol* 2006, **140**:1095-1108.

72. Ellis CM, Nagpal P, Young JC, Hagen G, Guilfoyle TJ, Reed JW: **AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in** *Arabidopsis thaliana*. *Development* 2005, **132**:4563-4574.

73. Goetz M, Hooper LC, Johnson SD, Rodrigues JCM, Vivian-Smith A, Koltunow AM: **Expression of aberrant forms of AUXIN RESPONSE FACTOR8 stimulates parthenocarpy in Arabidopsis and tomato**. *Plant Physiol* 2007, **145**:351-366.

74. Sessions A, Nemhauser JL, McColl A, Roe JL, Feldmann KA, Zambryski PC: **ETTIN patterns the** *Arabidopsis* **floral meristem and reproductive organs**. *Development* 1997, **124**:4481-4491.

75. Jakoby M, Weisshaar B, Droge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F: **bZIP transcription factors in** *Arabidopsis*. *Trends Plant Sci* 2002, **7**:106-111.

76. Gibalova A, Renak D, Matczuk K, Dupl'akova N, Chab D, Twell D, Honys D: **AtbZIP34 is required for Arabidopsis pollen wall patterning and the control of several metabolic pathways in developing pollen**. *Plant Mol Biol* 2009, **70**:581-601.

77. Das P, Ito T, Wellmer F, Vernoux T, Dedieu A, Traas J, Meyerowitz EM: **Floral stem cell termination involves the direct regulation of *AGAMOUS* by *PERIANTHIA***. *Development* 2009, **136**:1605-1611.

78. Heisler MG, Atkinson A, Bylstra YH, Walsh R, Smyth DR: ***SPATULA*, a gene that controls development of carpel margin tissues in *Arabidopsis*, encodes a bHLH protein**. *Development* 2001, **128**:1089-1098.

79. Davies B, EgeaCortines M, Silva ED, Saedler H, Sommer H: **Multiple interactions amongst floral homeotic MADS box proteins**. *EMBO J* 1996, **15**:4330-4343.

80. Kaufmann K, Wellmer F, Muiño JM, Ferrier T, Wuest SE, Kumar V, Serrano-Mislata A, Madueño F, Krajewski P, Meyerowitz EM, Angenent GC, Riechmann JL: **Orchestration of floral initiation by *APETALA1***. *Science* 2010, **328**:85-89.

81. Kane MD, Jatkoe TA, Stumpf CR, Lu J, Thomas JD, Madore SJ: **Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays**. *Nucleic Acids Res* 2000, **28**:4552-4557.

82. Li F, Stormo GD: **Selection of optimal DNA oligos for gene expression arrays**. *Bioinformatics* 2001, **17**:1067-1076.

83. Relogio A, Schwager C, Richter A, Ansorge W, Valcarcel J: **Optimization of oligonucleotide-based DNA microarrays**. *Nucleic Acids Res* 2002, **30**:e51.

84. Bloomfield VA, Crothers DM, Tinoco I Jr.: ***Nucleic acids: structures, properties, and functions***. Sausalito, Calif: University Science Books; 2000.

85. Le Novere N: **MELTING, computing the melting temperature of nucleic acid duplex**. *Bioinformatics* 2001, **17**:1226-1227.

86. Wang Y, Kong F, Yang Y, Gilbert GL: **A multiplex PCR-based reverse line blot hybridization (mPCR/RLB) assay for detection of bacterial respiratory pathogens in children with pneumonia**. *Pediatr Pulmonol* 2008, **43**:150-159.

87. Wang JP, Lindsay BG, Leebens-Mack J, Cui L, Wall K, Miller WC, dePamphilis CW: **EST clustering error evaluation and correction**. *Bioinformatics* 2004, **20**:2973-2984.

88. Wall PK, Leebens-Mack J, Muller KF, Field D, Altman NS, dePamphilis CW: **PlantTribes: a gene and gene family resource for comparative genomics in plants**. *Nucleic Acids Res* 2008, **36**:D970-D976.

89. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools**. *Nucleic Acids Res* 1997, **25**:4876-4882.

90. Lee I, Dombkowski AA, Athey BD: **Guidelines for incorporating non-perfectly matched oligonucleotides into target-specific hybridization probes for a DNA microarray**. *Nucleic Acids Res* 2004, **32**:681-690.

91. Chou HH, Hsia AP, Mooney DL, Schnable PS: **Picky: oligo microarray design for large genomes**. *Bioinformatics* 2004, **20**:2893-2902.

92. Riccelli PV, Merante F, Leung KT, Bortolin S, Zastawny RL, Janeczko R, Benight AS: **Hybridization of single-stranded DNA targets to immobilized complementary DNA probes: comparison of hairpin versus linear capture probes**. *Nucleic Acids Res* 2001, **29**:996-1004.

93. Lin HC, Morcillo F, Dussert S, Tranchant-Dubreuil C, Tregear JW, Tranbarger TJ: **Transcriptome analysis during somatic embryogenesis of the tropical monocot Elaeis guineensis: evidence for conserved gene functions in early development**. *Plant Mol Biol* 2009, **70**:173-192.

94.   Wall PK, Leebens-Mack J, Muller KF, Field D, Altman NS, dePamphilis CW:
        PlantTribes: a gene and gene family resource for comparative genomics in plants.
        Nucleic Acids Res 2008, **36**:D970-976.
95.   **The R Project for Statistical Computing** [www.r-project.org]
96.   **Bioconductor** [www.bioconductor.org]
97.   Yang YH, Thorne NP: **Normalization for two-color cDNA microarray data.** In
        Statistics and Science: A Festschrift for Terry Speed. Volume 40. Edited by Goldstein
        DR. Beachwood, OH: Institute of Mathematical Statistics; 2003:403-418.
98.   Smyth GK: **Linear models and empirical bayes methods for assessing differential
        expression in microarray experiments**. *Stat Appl Genet Mol Biol* 2004, **3**:Article3.
99.   **Limma: Linear Models for Microarray Data - User's Guide**
        [http://bioconductor.org/packages/1.9/bioc/vignettes/limma/inst/doc/usersguide.pdf]

Figure legends

Figure 1. An angiosperm phylogram with illustration of flower structures and the loop design of the *E. californica* microarray experiments. (A) A phylogram of angiosperms with flower architectures for several representative species. S represents sepal; P, petal; St, stamen; Std, staminodia; C, carpel; Ot, outer tepals; It, inner tepals. **(B)** We sampled from eight different tissues, including leaves, small floral buds, medium floral buds, four floral organs (sepals, petals, stamens, and pistils) at anthesis, and young fruits (four replicates for each tissue, 32 in total). Each line connects samples from two tissues in one microarray hybridization reaction, and four different colors represent four replicates of each tissue. The arrows point to the samples labeled with Cy5 dyes while base point to the samples labeled with Cy3 dyes.

Figure 2. Correlation coefficients between signal intensities from four biological replicates of the small floral buds. Pearson's correlation coefficients were between 0.94 and 0.97 between any pair of the four biological replicates, indicating that the results were highly reproducible.

Figure 3. Venn Diagrams of Genes expressed in reproductive tissues (the same abbreviations were used as in tables). (A) - **(D)** genes expressed in different tissues and their intersections. **(E)** - **(F)** genes significantly preferentially expressed compared with leaf with more than two-fold differences and their intersections.

Figure 4. Heat maps and GO annotation pie chart of genes differentially expressed between any two tissues. (A) Heat map for the mRNA profiles of 1921 genes differentially expressed between any two tissues. Red color represents high expression while green color represents low expression. HCL clustering was performed on transcripts ratios of all tissues across tissues and genes. Two major clusters had been identified as C1 and C2. L represents leaf; C represents carpel; F represents fruit; P represents petal; SB represents small bud; MB represents medium bud; S represents sepal; ST represents stamen. **(B)** GO categorization of all *Arabidopsis* homologs of Poppy genes included in our chip as control. **(C)** GO categorization of all *Arabidopsis* homologs of poppy genes that were statistically significantly differentially expressed.

Figure 5. Heat maps of genes preferentially expressed in different tissues. Red color represents high expression while green color represents low expression. Heat map of genes preferentially expressed in **(A)** leaf compared with all the other tissues, **(B)** sepal compared with all the other tissues, and **(C)** petal compared with all the other tissues. **(D)** stamen compared with all the other tissues. See definition of abbreviation in figure 4A.

Figure 6. The expression levels of MADS transcription factors families. All the expression values are log2 ratio. L represents leaf; C represents carpel; F represents fruit; P represents petal; SB represents small bud; MB represents medium bud; S represents sepal; ST represents stamen. The y-axis is the log2 ratio of gene expression levels.

Figure 7. The expression levels of ARGONAUTE, MYB, Zinc-finger, Homeodomain, ARF, bZIP and bHLH families. All the expression values are log2 ratio. The same abbreviations of different tissues were used as in figure 5.

Table 1. Genes preferentially expressed at pre-meiotic, meiotic stage and fruit in poppy.

gene	BestATHit	L	SB	MB	S	P	ST	C	F	annotation
					Preferentially expressed in pre-meiotic buds					
89282	AT2G31210.1	5.3	9.0	7.1	5.6	5.3	5.3	5.3	5.1	bHLH
83967	AT5G16920.1	7.1	9.9	8.5	7.3	7.0	6.9	6.9	6.9	
84082	AT1G68540.1	6.8	10.2	8.9	6.8	6.9	6.7	6.5	6.2	oxidoreductase
87393	AT1G44970.1	5.1	7.9	5.9	5.1	5.0	5.0	5.6	5.2	peroxidase
86946	AT4G33870.1	7.8	9.5	8.1	8.0	7.8	7.9	7.8	7.8	peroxidase
86850	AT3G28470.1	6.2	7.5	6.4	6.1	6.1	6.1	6.1	6.0	ATMYB35
85123	AT5G09970.1	5.9	9.5	7.6	5.4	5.4	5.1	6.5	7.3	CYP78A7
					Preferentially expressed in meiotic buds					
84975	AT5G35630.2	6.9	6.7	8.5	6.8	6.6	6.7	6.6	6.9	GS2
85233	AT1G11910.1	5.6	7.4	10.2	9.1	6.1	8.5	6.1	8.4	aspartyl protease
86094	AT1G54220.1	6.8	7.8	9.9	7.5	7.3	8.6	7.0	7.2	dihydrolipoamide S-acetyltransferase
88004	AT4G16260.1	5.7	7.5	9.7	6.0	5.9	6.1	5.4	5.8	hydrolase
88092	AT4G12910.1	9.1	9.3	10.9	8.9	8.5	8.4	9.0	9.4	scpl20
88096	AT3G11450.1	7.8	8.2	9.9	7.8	7.8	8.2	7.9	7.9	cell division protein-related
88675	AT4G35160.1	6.3	6.6	7.9	6.6	6.3	6.2	6.1	6.2	O-methyltransferase
89901	AT5G03880.1	7.6	7.6	8.7	7.7	7.4	7.6	7.3	7.5	electron carrier
					Preferentially expressed in fruits					
83998		6.4	5.8	5.7	6.3	6.5	5.8	6.2	8.5	
84097	AT5G54160.1	9.4	9.1	10.0	9.1	8.6	8.1	9.0	11.1	ATOMT1
86118	AT5G62200.1	7.6	7.0	7.4	7.6	7.6	7.7	7.3	9.3	embryo-specific protein
86486	AT1G07080.1	6.5	6.6	6.9	6.8	6.3	6.8	6.6	10.1	GILT
87027		5.8	5.5	5.5	5.7	5.6	6.0	5.8	7.3	
87195	AT5G12380.1	6.6	6.2	6.5	6.7	6.4	6.5	7.2	9.6	annexin
87830	AT5G08260.1	6.0	5.9	6.2	6.0	6.1	5.9	6.1	7.4	scpl35
88106	AT1G20030.2	6.6	6.3	6.8	7.3	5.9	6.4	6.5	9.0	pathogenesis-related thaumatin
89333		8.8	6.5	8.0	8.4	5.9	7.6	7.1	10.5	

The first column is the gene number for gene represented by poppy ESTs. The second column is the closest *Arabidopsis* homolog of poppy genes. All expression values are log2 ratio. L represents leaf; C represents carpel; F represents fruit; P represents petal; SB represents small bud; MB represents medium bud; S represents sepal; ST represents stamen. Annotations are from TAIR version 9.

Table 2. Expression levels of putative ABC genes in poppy.

gene	BestATHit	L	SB	MB	S	P	ST	C	F	annotation
					A function genes					
84392	AT2G36830.1	14.4	13.9	14.3	16.2	16.3	15.1	14.5	14.4	GAMMA-TIP
86583	AT2G45650.1	6.4	10.1	10.4	11.9	10.7	7.0	8.9	8.9	AGL6
87043	AT3G05490.1	8.9	9.1	9.8	10.6	10.9	9.6	9.2	9.3	RALFL22
87251	AT5G15800.1	6.3	9.2	9.4	10.5	9.5	6.9	8.4	8.3	SEP1, AGL2
85671		7.3	7.0	6.9	10.5	11.1	8.4	7.3	7.7	
					B function genes					
83744	AT3G54340.1	8.2	11.4	12	9.1	11.9	13.1	9.9	9.1	AP3
83763	AT1G69500.1	5.2	5.7	6.1	5.9	7.7	7.3	6.1	5.7	electron carrier
83991	AT5G19770.1	10.0	10.1	10.3	9.0	11	11.2	9.8	10.1	TUA3
84789	AT5G64250.2	11.9	11.4	13.2	13.8	15.6	15.0	13.9	13.2	2-nitropropane dioxygenase
85140	AT2G27510.1	9.2	11.1	12.2	10.8	13.8	11.9	10.5	10.0	ferredoxin 3
85166	AT5G62690.1	9.2	10.0	10.2	8.3	10.4	10.9	9.3	9.8	TUB2
85610	AT4G36250.1	5.4	6.5	8.5	6.4	7.8	7.6	6.1	5.5	Aldehyde Dehydrogenase 3F1
87005	AT3G54340.1	4.6	7.8	8.3	6.0	10.2	8.1	5.3	5.1	AP3
87035	AT3G58120.1	5.6	5.5	5.7	5.5	7.9	8.0	5.1	5.2	ATBZIP61
87167	AT5G20240.1	7.3	11.2	12.0	9.2	12.8	11.9	8.4	8.1	PI
87294	AT5G03690.2	8.0	9.3	10.0	7.8	9.8	10.0	8.7	9.2	fructose-bisphosphate aldolase
89750	AT4G37990.1	8.2	8.8	9.6	8.5	11.4	10.2	7.7	7.9	mannitol dehydrogenase
89805	AT5G66310.1	5.9	6.5	6.8	5.3	7.2	7.8	6.1	6.4	kinesin motor
					C function genes					
84248	AT4G18960.1	6.7	10.6	11.1	7.2	6.6	11.5	11.6	11.6	AG
84252	AT4G26220.1	7.3	10.9	10.9	7.0	7.2	10.9	10.1	6.6	caffeoyl-CoA 3-O-methyltransferase
84340	AT3G44260.1	7.9	8.4	8.3	7.8	8	9.9	9.2	8.6	CCR4-NOT transcription complex protein
84512	AT1G11910.1	7.2	9.5	10.1	7.0	7.2	8.8	9.1	9.1	aspartyl protease
84691	AT2G44480.1	9.1	12.4	12.8	8.7	8.6	12.6	12	12.9	BETA GLUCOSIDASE 17
89115	AT3G20240.1	6.4	7.4	7.3	6.3	6.2	8.2	7.6	7.0	mitochondrial substrate carrier
89980	AT1G35720.1	7.1	8.8	9.5	7.2	7.9	10.1	9.2	8.5	ANNEXIN ARABIDOPSIS 1

The first column is the gene number for gene represented by poppy ESTs. The second column is the closest *Arabidopsis* homolog of poppy genes. All expression values are log2 ratio. L represents leaf; C represents carpel; F represents fruit; P represents petal; SB represents small bud; MB represents medium bud; S represents sepal; ST represents stamen. Annotations are from TAIR version 9.

Additional files

Additional file 1. Supplemental figures. Supplemental Figure 1: Correlation coefficients between signal intensities from four biological replicates of seven tissues. Pearson's correlation coefficients were between 0.88 and 0.97 between any pair of the four biological replicates, indicating that the results were highly reproducible. Supplemental Figure 2: GO annotation pie chart of genes present across all tissues. GO categorization of all *Arabidopsis* homologs of poppy genes which were expressed across all the eight tissues with log2 values of signal intensity larger than 5.41 (10% percentile, control provided in figure 4). Supplemental Figure 3: Real-time PCR results consistent with microarray data. Nine genes were verified using real-time PCR. The lines in blue represent the real-time RT-PCR results while red the microarray results. All the numbers shown in this figure are the fold changes of expression intensities in reproductive tissues compared with leaf. The left Y-axis is for microarray results and right Y-axis for real-time PCR results.

Additional file 2. Numbers of genes expressed in eight tissues respectively using different cut-off and gene lists. This additional file contains supplemental tables, including: Table S1, A summary of numbers of genes expressed in eight tissues respectively using different cut-off percentiles (5%, 10%, 15%); Table S2, a list of genes expressed in both leaves and medium buds; Table S3, a list of genes expressed in leaves, small buds and medium buds; Table S4, a list of genes expressed in both leaves and small buds; Table S5, a list of genes expressed in both small buds and medium buds; Table S6, a list of genes expressed in leaves, sepals and petals; Table S6, a list of genes expressed in both carpels and stamens; Table S7, a list of genes expressed in either sepals and/or petals; Table S8, a list of genes expressed in carpels and either sepals and/or petals; Table S9, a list of genes expressed in carpels, stamens and either sepals and/or petals; Table S10, a list of genes expressed in stamens and either sepals and/or petals.

Additional file 3. GO comparison between all genes on the chip and differentially expressed genes. Gene numbers comparing all genes on the chip, genes expressed across different tissues and those differentially expressed between any two tissues in each GO category.

Additional file 4. Genes preferentially expressed in eight tissues respectively and all the genes differentially expressed. This additional file contains information about lists of all the genes preferentially and differentially expressed between any two tissues and lists of genes preferentially expressed in each tissue over all the other tissues. Column sequence, abbreviation and the version of annotation are as those used as in table 1 and all the other supplemental tables. All the expression values are log2 ratio.

Additional file 5. Expression levels of *Arabidopsis* homologs of selected poppy genes. This additional file contains information about the expression levels of *Arabidopsis* homologs of selected poppy genes of interests listed in the tables in our study.

Additional file 6. Lists of genes identified A-, B- and C-domain genes.
This additional file contain information about gene lists of identified putative A-, B- and C-domain genes.

Additional file 7. Lists of genes preferentially expressed in reproductive tissues compared with leaf. This additional file contains information about list of genes preferentially expressed in reproductive tissues compared with leaves.

Additional file 8. Gene expression of different gene families. This additional file contains information about the expression levels of different gene families.

Additional file 9. Probe design in the microarray and properties of *E. californica* probe sets designed by MODIT and other methods and primers used for Real-time PCR experiments. This additional file contains probe design and orientation of the custom microarray, properties of probe sets and primers for real-time PCR.
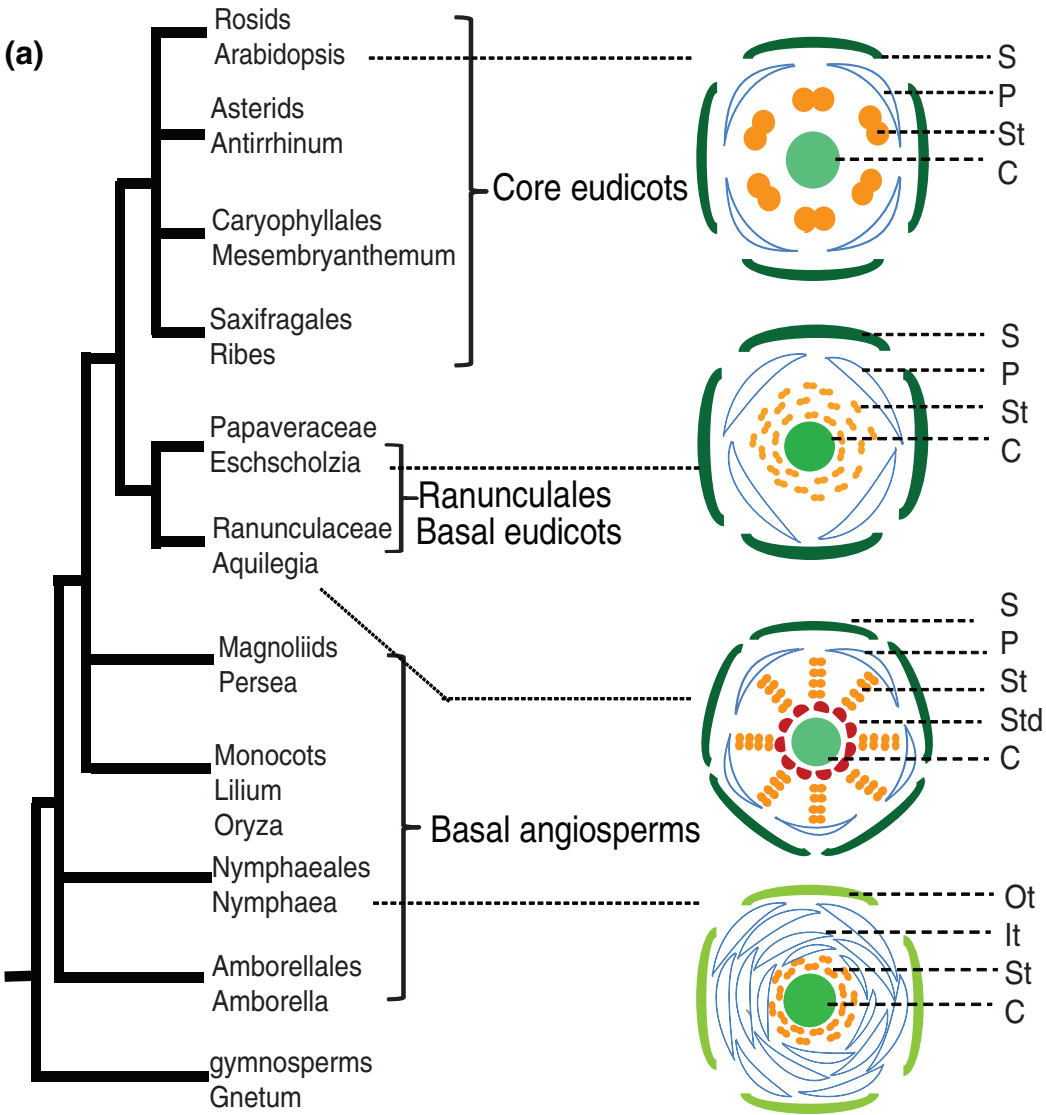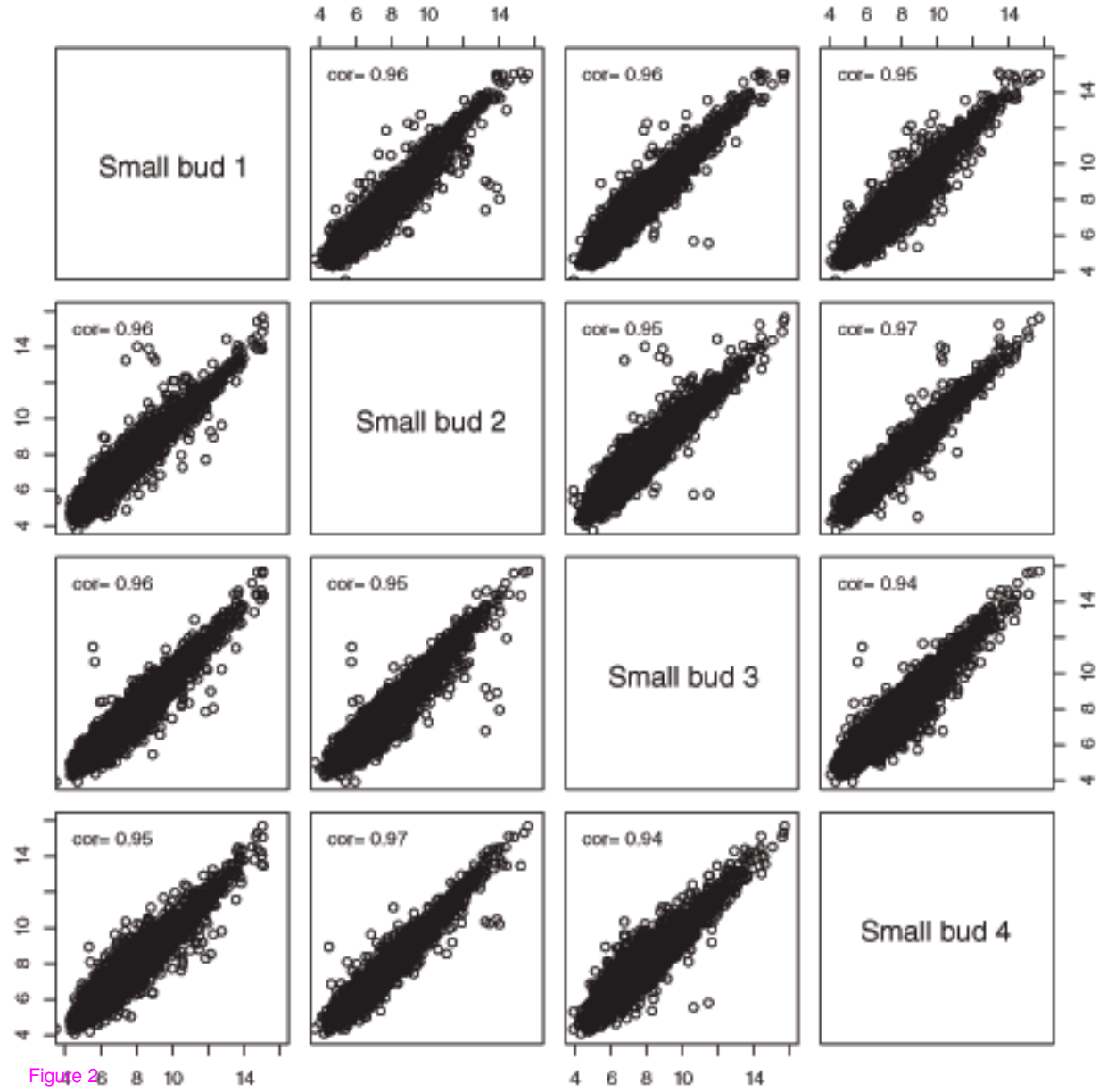
(a)

Rosids
Arabidopsis — S, P, St, C

Asterids
Antirrhinum

Caryophyllales
Mesembryanthemum

Saxifragales
Ribes

Core eudicots

Papaveraceae
Eschscholzia — S, P, St, C

Ranunculaceae
Aquilegia

Ranunculales
Basal eudicots

Magnoliids
Persea — S, P, St, Std, C

Monocots
Lilium
Oryza

Nymphaeales
Nymphaea — Ot, It, St, C

Amborellales
Amborella

Basal angiosperms

gymnosperms
Gnetum

(b)

leaf

small bud

sepal

medium bud

petal

fruit

stamen

carpel

Figure 1

Figure 2

(a)
SB 5906
49
53
69
5735
75
30
42
L 5905
MB 5876

(b)
114
5761
109
S 5876
P 5870

(c)
St 5877
47
91
16
5723
89
30
82
P+S 5985
C 5851

(d)
97
5754
127
C 5851
F 5851

(e)
St 372
280
36
32
24
86
95
17
P 163
S 168

(f)
St 372
302
43
10
17
90
13
46
P 163
C 86

Figure 3

(a) MB SB L C F P S ST

4.0 8.0 12.0

(b)

9% 18% 9% 7% 4% 4% 3% 6% 2% 7% 7% 7% 2% 0% 15%

(c)

11% 14% 9% 7% 3% 3% 3% 6% 3% 6% 10% 3% 3% 0% 19%

■ Other binding
■ Other enzyme activity
■ Hydrolase activity
■ Transferase activity
■ Transporter activity
■ Kinase activity
■ Transcription factor activity
■ DNA or RNA binding
■ Other molecular functions
■ Protein binding
■ Nucleotide binding
■ Structural molecule activity
■ Nucleic acid binding
■ Receptor binding or activity
■ Unknown molecular functions

Figure 4

(a)
0.0   8.8   10.6
L S F MB SB C P ST

(b)
0.0   8.8   10.4
S L C F MB SB P ST

(c)
0.0   8.0   9.6
P C F L S MB SB ST

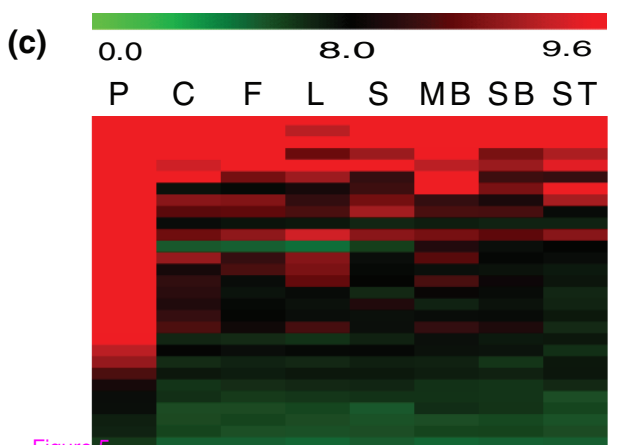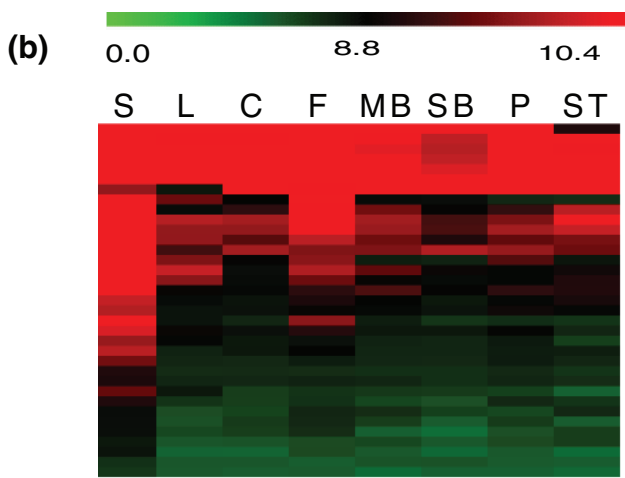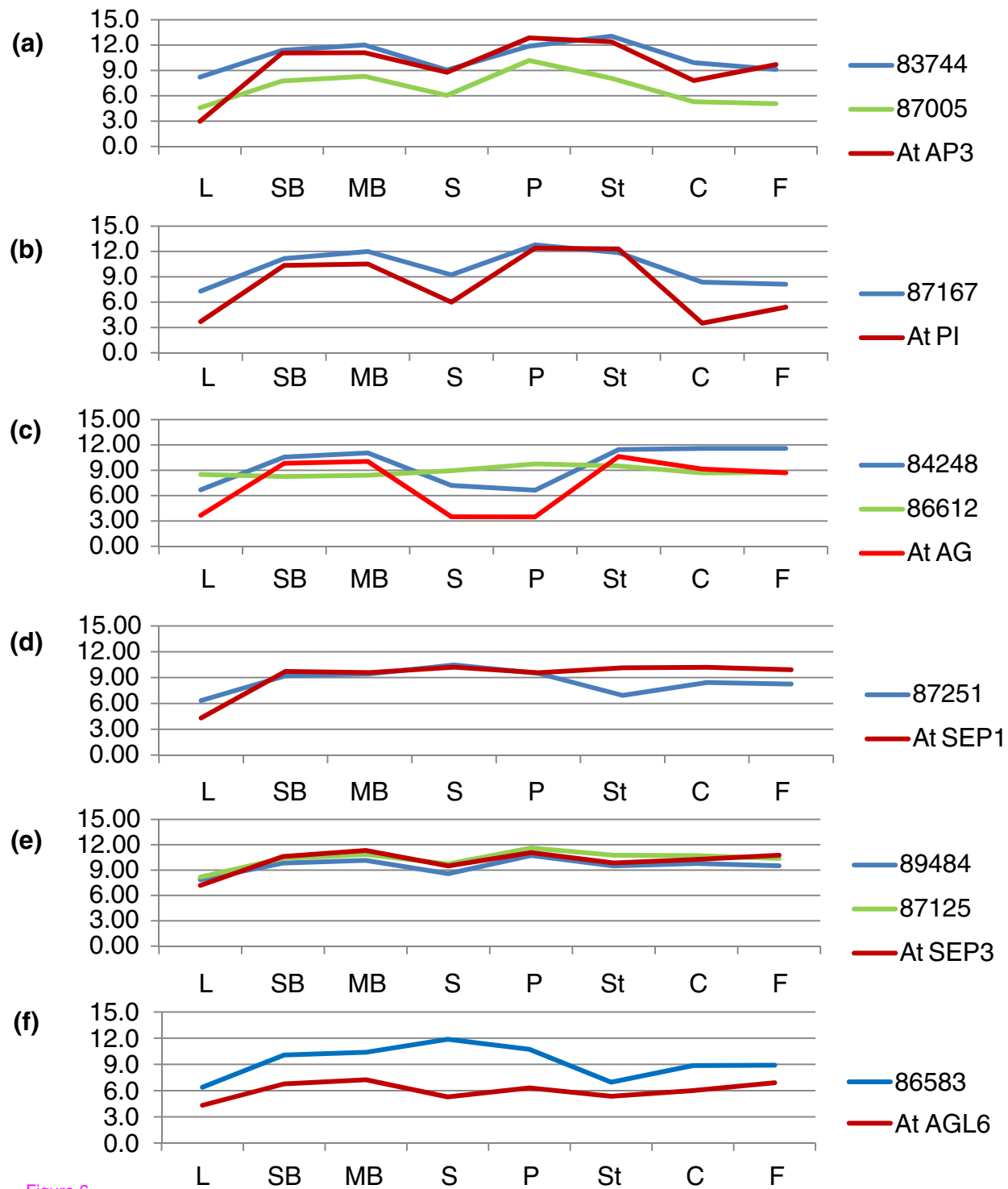(d)
0.0   7.9   9.8
ST MB SB C F P L S

Figure 5

Figure 6

Figure 7

Additional files provided with this submission:

Additional file 1: suppl9.ppt, 483K
http://genomebiology.com/imedia/1438749908466445/supp1.ppt
Additional file 2: suppl1.xlsx, 8512K
http://genomebiology.com/imedia/8856412594664458/supp2.xlsx
Additional file 3: suppl3.xls, 20K
http://genomebiology.com/imedia/1747823604466445/supp3.xls
Additional file 4: suppl2.xls, 904K
http://genomebiology.com/imedia/2665450334664462/supp4.xls
Additional file 5: suppl4.xls, 34K
http://genomebiology.com/imedia/1931746531466446/supp5.xls
Additional file 6: suppl5.xls, 326K
http://genomebiology.com/imedia/1404344799466446/supp6.xls
Additional file 7: suppl6.xls, 312K
http://genomebiology.com/imedia/2057167452466446/supp7.xls
Additional file 8: suppl7.xls, 49K
http://genomebiology.com/imedia/6885426574664471/supp8.xls
Additional file 9: suppl8.xls, 3981K
http://genomebiology.com/imedia/3364606844664464/supp9.xls