


RESEARCH

Open Access



Evolution of the *SPATULA/ALCATRAZ* gene lineage and expression analyses in the basal eudicot, *Bocconia frutescens* L. (Papaveraceae)

Cecilia Zumajo-Cardona^{1,2,3}, Barbara Ann Ambrose² and Natalia Pabón-Mora^{1*} 

Abstract

Background: *SPATULA* (*SPT*) and *ALCATRAZ* (*ALC*) are recent paralogs that belong to the large bHLH transcription factor family. Orthologs of these genes have been found in all core eudicots, whereas pre-duplication genes, named *paleoSPATULA/ALCATRAZ*, have been found in basal eudicots, monocots, basal angiosperms and gymnosperms. Nevertheless, functional studies have only been performed in *Arabidopsis thaliana*, where *SPT* and *ALC* are partially redundant in carpel and valve margin development and *ALC* has a unique role in the dehiscence zone. Further analyses of pre-duplication genes are necessary to assess the functional evolution of this gene lineage.

Results: We isolated additional *paleoSPT/ALC* genes from *Aristolochia fimbriata*, *Bocconia frutescens*, *Cattleya trianae* and *Hypoxis decumbens* from our transcriptome libraries and performed phylogenetic analyses. We identified the previously described bHLH domain in all analyzed sequences and also new conserved motifs using the MEME suite. Finally, we analyzed the expression of three *paleoSPT/ALC* genes (*BofrSPT1/2/3*) from *Bocconia frutescens*, a basal eudicot in the Papaveraceae. To determine the developmental stages at which these genes were expressed, pre- and post-anthesis carpels and fruits of *B. frutescens* were collected, sectioned, stained, and examined using light microscopy. Using in situ hybridization we detected that *BofrSPT1/2/3* genes are expressed in floral buds, early sepal initiation, stamens and carpel primordia and later during fruit development in the dehiscence zone of the opercular fruit.

Conclusions: Our expression results, in comparison with those available for core eudicots, suggest conserved roles of members of the *SPT/ALC* gene lineage across eudicots in the specification of carpel margins and the dehiscence zone of the mature fruits. Although there is some redundancy between *ALC* and *SPT*, these gene clades seem to have undergone some degree of sub-functionalization in the core eudicots, likely by changes in cis regulatory regions and to some extent in coding sequences, at least in Brassicaceae. Our results also indicate that in *Bocconia frutescens*, *paleoSPT/ALC* genes may play a role in early floral organ specification that was subsequently lost in core eudicot lineages.

Keywords: *ALCATRAZ*, Basal eudicots, *Bocconia frutescens*, Fruit development, *paleoSPT/ALC*, Papaveraceae, *SPATULA*

Background

Plants provide a great experimental model to assess functional evolution after gene duplication as most plant genomes have been shaped by ancient whole genome

duplication (WGD) events [1, 2]. Gene and genome duplications can result in mutational loss of duplicates, redundancy, the acquisition of new roles due to changes in regulatory or protein interactions (neofunctionalization) or the redistribution of functions among paralogs by uncoupling of regulatory elements (sub-functionalization) [2–5]. Functional diversification after gene and genome duplication has received considerable attention in polyploid crops, but can also be assessed by comparing

*Correspondence: lucia.pabon@udea.edu.co

¹ Instituto de Biología, Universidad de Antioquia, Medellín, 1226, Colombia

Full list of author information is available at the end of the article

gene functions at different time points in the phylogeny prior to and after major WGD in angiosperms [2, 6]. Paleopolyploidy has been traced back to WGD occurring before the diversification of angiosperms, prior to the origin of core eudicots, within the Brassicales and Solanales, and concomitant with monocot diversification [7–9]. In this framework, basal eudicots have become a unique reference for assessing gene functional evolution in core eudicots, the latter include 75% of flowering plant species with unique paleopolyploidy events, whereas the former include species with pre-duplication genes, often single copy, most of the time exemplifying the ancestral role prior to the WGD events.

The basic/helix-loop-helix (bHLH) proteins are a superfamily of transcription factors that has been better characterized in animals than in plants [10, 11]. However, the *Arabidopsis* genome possesses at least 147 bHLH protein-encoding genes, making this, one of the largest transcription factor families in this model species, likely having important unexplored functions [12]. The key developmental transcription factors *ALCATRAZ* (*ALC*) and *SPATULA* (*SPT*) are two closely related factors that belong to this transcription factor family [12–15]. *SPATULA* (*SPT*) is expressed in developing carpel margins, leaves and petals, as well as the dehiscence zone of fruits and anthers [16]. In *Arabidopsis thaliana*, *SPT* is important for carpel margin development which impacts proper carpel fusion, transmitting tract development, as well as style and stigma development [13, 16–19]. Loss of function phenotypes in *spt* mutants exhibit defective carpel margin fusion, particularly at the distal-most portion of the congenitally fused, bicarpellate gynoecium [17, 19]. *ALCATRAZ* is turned on in the petal margins, in the stamens, stigmas and carpellary margins and later in development, at the layer of non-lignified cells in the silique or silicle [20, 21]. *ALC* is required for proper fruit dehiscence zone development, in particular as it specifies the identity of the separation layer, rich in cell-wall degrading enzymes [20, 22]. In addition, *SPT* and *ALC* proteins are able to form heterodimers and have redundant roles in gynoecium development. More specifically, they are key factors during carpel and valve (fruit wall) margin development, as shown by the *spt/alc* double mutant, which exhibits increased severity in the carpel separation as well as defects in style and stigma patterning and the histogenesis of the valve margin, and the dehiscence zone [21]. However, while *SPT* overexpression can fully compensate *alc* defects during fruit development and dehiscence, *ALC* overexpression can only partially compensate for *spt* defects by increasing the short fruit size and restoring, to some extent, apical fusion between the carpels [21]. *SPATULA* also interacts with *HECATE1* (*HEC1*), *HEC2*, and *HEC3*, all bHLH transcription

factors involved in septum, transmitting tract and stigma development [23, 24]. On the other hand, *ALCATRAZ* interacts with *ALC* interacting protein1 (*ALC1*), a protein expressed in vascular and mesocarp cells in *Arabidopsis* [25]. Thus, it is likely that despite the fact that they act partly redundantly in early gynoecium patterning and late fruit development, there is some specialization due to changes in expression patterns and protein interactions.

The current working model for a gene regulatory network in fruit development includes both *SPT* and *ALC* functioning downstream of two MADS-box genes, *FRUITFULL* and *SHATTERPROOF* which establish valve and valve margin identity, thus delineating the dehiscence zone [20, 26, 27]. *FRUITFULL* (*FUL*) controls valve identity, as shown by the prematurely exploded fruits with bursting seeds of the *ful* mutants [26]. *SHATTERPROOF1/2* (*SHP1/2*) controls the dehiscence zone differentiation by promoting adjacent lignified and unlignified cells in the valve margin [27]. In the *ful* mutant, *SHP* expands its expression to the fruit valves while, in the *shp1/2* mutant, *FUL* does not expand its expression to the valve margin [27]. Consistent with the proposed interactions, in the *alc* mutant, the expression of both upstream transcription factors *SHP1/2* and *FUL* is not altered [27]. Both, *SPT* and *ALC*, are activated by *SHP1/2* [20], that in turn are negatively regulated by *FUL* [28]. However, this genetic network is restricted to *Arabidopsis*, as a number of these genes is not conserved in all angiosperms due to lineage specific duplications mostly occurring in Brassicaceae and core eudicots [29].

Phylogenetic analyses across seed plants have shown that *SPATULA* and *ALCATRAZ* belong to paralogous clades resulting from a duplication event that occurred prior to the diversification of the core eudicots; in non-core eudicots, pre-duplication genes are referred to as the *paleoSPT/ALC* genes [29]. Comparisons between *SPT* and *ALC* sequences across flowering plants reveal that the bHLH domain is highly conserved. Nevertheless, *SPT* proteins have a conserved acidic domain and amphipathic helix N terminal to the bHLH domain. The amphipathic helix but not the acidic domain has been identified in *ALC* proteins [13, 21, 29, 30]. *paleoSPT/ALC* orthologs have the acidic domain and exhibit conserved key functional residues in the bHLH domain, suggesting that *paleoSPT/ALC* may have similar downstream targets as the *Arabidopsis* *SPT* and *ALC* [29].

Functional and expression analyses of *SPT* and *ALC* have only been performed in a few core eudicots including *Arabidopsis thaliana*, *Lepidium campestre*, *Fragaria vesca* and *Prunus persica* [13, 20, 30–32]. In *Lepidium campestre* (Brassicaceae), *ALC* is expressed in the dehiscence zone of the fruit, suggesting that the function of *ALC* is

conserved in several members of Brassicaceae exhibiting the distinct silique or silicle fruit types [32]. In *Fragaria vesca*, functional studies have shown that *FaSPT*, the *SPT* ortholog, is involved in strawberry development, as *faspt* fruits in early stages exhibited reduced size [31] similar to a role for *Arabidopsis SPT* in regulating organ size [33]. In *Prunus persica* (peach), expression analysis of the *SPT* ortholog, *PPERALCATRAZ/SPATULA* (*PPERALC/SPT*), showed that it is expressed in the perianth, ovary and later in the endocarp margins as well as in leaves [30].

As expression and functional analyses in this gene lineage are restricted to only a few core eudicot species, it is difficult to predict the functional evolution of this gene lineage. Thus, to better understand the evolution of this gene lineage we here report: (1) expanded sampling of *SPT/ALC* homologs across flowering plants; (2) the analysis of coding sequences of *SPT* and *ALC* genes prior to and after the core eudicot duplication to identify conserved regions between the pre- and post-duplication homologs that may help predict putative shifts in

protein function; (3) the comparisons of expression patterns of pre- and post-duplication homologs using the online available tools for model species (i.e., eFP browser in *Arabidopsis thaliana*, *Medicago truncatula*, *Solanum lycopersicum*, *S. tuberosum* and *Oryza sativa*; and (4) the expression analyses of selected *paleoSPT/ALC* genes in *Bocconia frutescens* (Papaveraceae, Basal eudicot). We chose *B. frutescens* because: 1) it is a member of the basal eudicots; thus, it possesses *paleoSPT/ALC* genes predating the core eudicot duplication; 2) there is a floral/fruit transcriptome available from this species and fresh tissue as it is widely cultivated in the tropics; and 3) it exhibits a dry dehiscent fruit with complete valve separation from a ring-like persistent septum (opercular dehiscence) that resembles the silique or silicle of *Arabidopsis* (Fig. 1) [34].

Methods

Transcriptome analysis

We prepared transcriptomes from *Aristolochia fimbriata* (Aristolochiaceae), *Bocconia frutescens* (Papaveraceae),

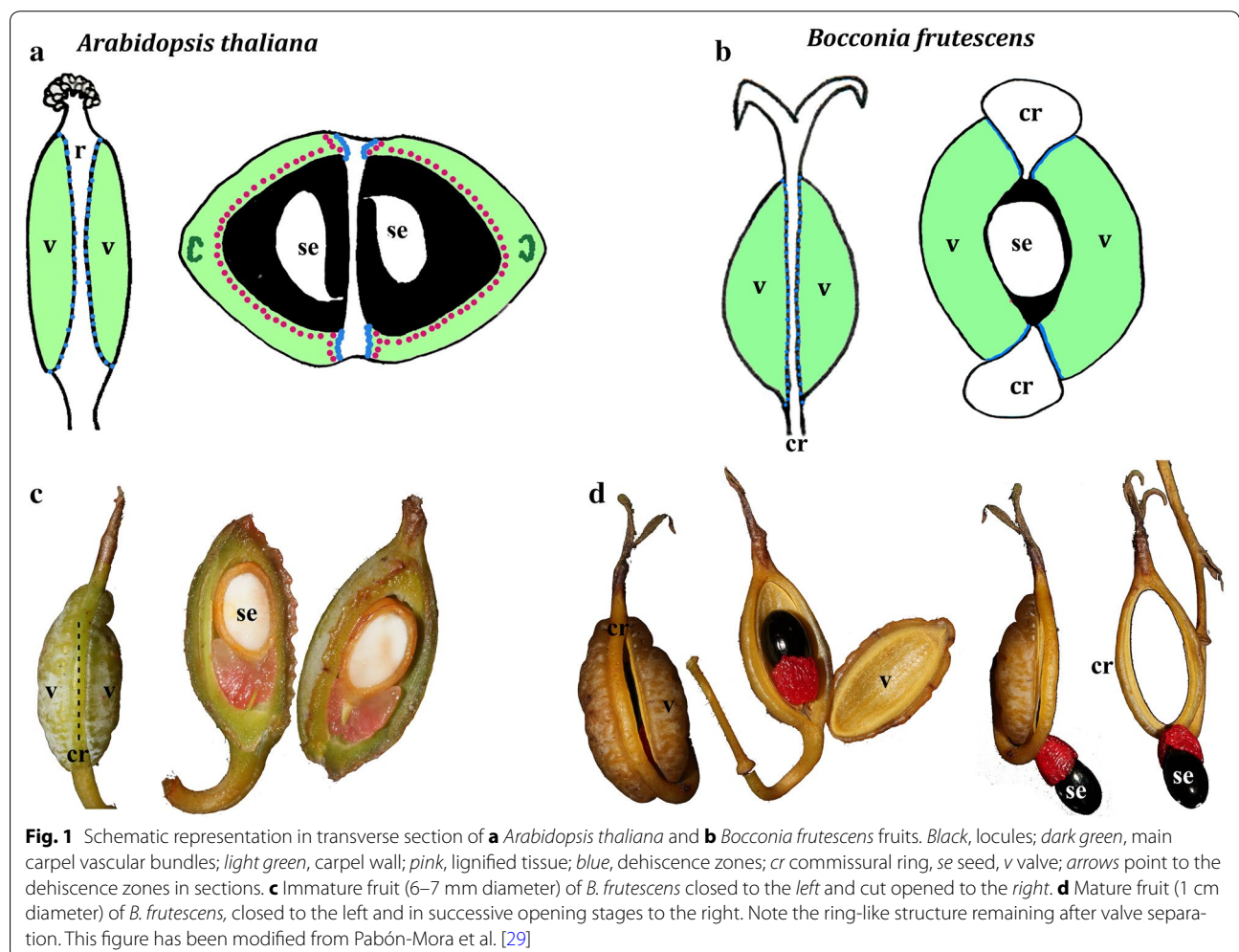


Fig. 1 Schematic representation in transverse section of **a** *Arabidopsis thaliana* and **b** *Bocconia frutescens* fruits. Black, locules; dark green, main carpel vascular bundles; light green, carpel wall; pink, lignified tissue; blue, dehiscence zones; cr commissural ring, se seed, v valve; arrows point to the dehiscence zones in sections. **c** Immature fruit (6–7 mm diameter) of *B. frutescens* closed to the left and cut opened to the right. **d** Mature fruit (1 cm diameter) of *B. frutescens*, closed to the left and in successive opening stages to the right. Note the ring-like structure remaining after valve separation. This figure has been modified from Pabón-Mora et al. [29]

Cattleya trianae (Orchidaceae), and *Hypoxis decumbens* (Hypoxidaceae), as previously described [35–37]. Fresh inflorescence, floral tissue, and vegetative tissue from all mentioned species were ground using liquid nitrogen, and further total RNA extraction was carried out using Trizol Reagent (Invitrogen). RNAseq experiments for each species were conducted using truseq mRNA library construction kit (Illumina) and sequenced in a HiSeq 2000 instrument reading 100 bases paired end reads. Read cleaning was performed with PRINSEQ-LITE with a quality threshold of Q35, and contig assembly was computed using Trinity package following default settings. For *Aristolochia fimbriata*, contig metrics are as follows: total assembled bases: 85,608,833; total number of contigs (>101 bp): 118,941; average contig length: 719 bp; largest contig: 16,972 bp; contig N50: 1823 bp; contig GC%: 42.71% [35]. For *Bocconia frutescens*, contig metrics are as follows: total assembled bases: 149,710,500; total number of contigs (>101 bp): 211,821; average contig length: 706 b; largest contig: 17,004 b; contig N50: 1877 bp; contig GC%: 40.09 [36]. For *Cattleya trianae*, contig metrics are as follows: total assembled bases: 63,287,862 bp; total number of contigs (>101 bp): 109,708; average contig length: 576 bp; largest contig: 9,321 bp; contig N50: 1,401 bp; contig GC%: 42.73 [37]. For *Hypoxis decumbens*, contig metrics are as follows: total assembled bases: 73,787,751 bp; total number of contigs (>101 bp): 157,153; average contig length: 469 bp; largest contig: 15,554 bp; contig N50: 1,075 bp; contig GC%: 46.42 [37].

Phylogenetic analyses

To expand sampling of *SPT/ALC* homologs, we isolated sequences using BLAST from our own generated transcriptomes from *Aristolochia fimbriata*, *Bocconia frutescens*, *Cattleya trianae* and *Hypoxis decumbens*. To show the phylogenetic position of the *Bocconia frutescens* homologs, we included *BofrSPT1*, 2 and 3 in a matrix consisting of selected *ALC* and *SPT* from all major plant groups, expanding the sampling done by Pabón-Mora et al. [29]. Most sequences were obtained from the plant transcriptome repositories of the OneKP database (<https://sites.google.com/a/ualberta.ca/onekp/>) and the genome repository Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>). Sequences were compiled with Bioedit (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>) and manually edited to exclusively keep the open reading frame for all transcripts, as many sequences from transcriptomic databases include the 5' and 3' untranslated reading frames (UTR's). Nucleotide sequences were subsequently aligned using the online version of MAFFT (<http://mafft.cbrc.jp/alignment/software/>) [38] with a gap open penalty of 3.0, offset value of 1.0, and all other

default settings. The alignment was then refined by hand using Bioedit around the bHLH domain. Maximum likelihood (ML) phylogenetic analyses using the nucleotide sequences were performed with RaxML-HPC2 BlackBox [39], through the CIPRES Science Gateway [40]. Bootstrapping was performed according to the default criteria in RaxML where the bootstrapping stopped after 200–600 replicates. Trees were observed and edited using FigTree v 1.4.3. (<http://tree.bio.ed.ac.uk/software/figtree/>). This analysis included all 54 sequences used for motif search (see above) from angiosperms and used the 6 gymnosperm paleoSPT/*ALC* outgroup sequences (Additional file 1: Table S1). Newly isolated sequences from our own generated transcriptomes from *Aristolochia fimbriata*, *Bocconia frutescens*, *Cattleya trianae* and *Hypoxis decumbens* are available under Genbank numbers KY421362–KY421369.

Identification of protein motifs across flowering plants

To detect reported as well as new conserved motifs, 60 complete sequences of *SPT/ALC* homologs were selected representing major seed plant lineages (25 from core eudicots, 9 from basal eudicots, 14 from monocots, 6 from basal angiosperms and 6 from gymnosperms). Sequences were permanently translated and uploaded as amino acids to the online MEME server (<http://meme.nbcr.net>) [41] and run with all the default options. The motifs retrieved by MEME are reported according to their statistical significance. The suite MEME finds in the given sequences the most statistically significant (low E-value) motifs first. We did the search for 20 motifs arbitrarily to search beyond the already identified motifs for *SPT* and *ALC* homologs. The E-value of a motif is based on its log likelihood ratio, width, sites, and the size of the set. The motifs identified with our dataset range from 9.3 e to 3095 (motif 1) found in all 60 input sequences, to 2.7 e–0.25 (motif 20) found in only 5 sequences (Additional file 2: Figure S2). Protein motifs provide important information for better assessing shifts in *SPT/ALC* proteins across seed plants, and as all motifs found are statistically significant (none of them is larger than 0.05) we decided to report them here even though more detail functional analyses are required in order to better understand how meaningful are the shifts here identified.

Developmental series of flowers and fruits of *B. frutescens*

Inflorescences, young buds, flowers and fruits were collected in the field (voucher: Colombia, Antioquia, Medellín, Las Palmas, Envigado, sobre la via principal, Km 12 retorno No 10. May 2015, C. Zumajo-Cardona and N. Pabón-Mora 03, HUA) and immediately fixed in formaldehyde-acetic acid–ethanol (FAA; 3.7% formaldehyde: 5% glacial acetic acid: 50% ethanol). For light

microscopy, fixed material was manually dehydrated through an alcohol-histochoice series, and embedded in Paraplast X-tra (Fisher Healthcare, Houston, Texas, USA). The samples were sectioned at 10–20 μm with an AO Spencer 820 (GMI Inc. Minnesota, USA) rotary microtome. Sections were stained with Johansen's safranin, to identify lignification and presence of cuticle, and 0.5% Astra Blue [42] and mounted in Permount (Fisher Scientific, Pittsburgh, Pennsylvania, USA). Sections were viewed and digitally photographed with a Zeiss Axioplan compound microscope equipped with a Nikon DXM1200C digital camera with ACT – 1 software. Different stages in flower and fruit development of *B. frutescens* were described using stages already identified in this species and other Papaveraceae as a Ref. [36, 44, 45]. Late fruit developmental stages were photographed and included in Fig. 1. In addition, a comparative drawing with respect to the Arabidopsis fruit was done based on fresh material and it corresponds to Fig. 1b.

Expression analyses by *In Situ* Hybridization

Inflorescences with flowers at different stages as well as mature carpels and immature fruits of *B. frutescens* were collected in the field (voucher: Colombia, Antioquia, Medellín, Las Palmas, Envigado, sobre la via principal, Km 12 retorno No 10. May 2015, C. Zumajo-Cardona and N. Pabón-Mora 03, HUA) and fixed in freshly prepared, cold FAA. After a 4-h incubation, samples were dehydrated in an ethanol series and then transferred to fresh Paraplast and stored at 4 °C until use. Samples were sectioned with a microtome at 10 μm . Samples were sectioned on a Microm HM3555 rotary microtome. DNA templates for RNA probe synthesis were obtained by PCR amplification of 312–412 bp fragments. To ensure specificity, the probe templates were designed to amplify the 3' sequence flanking the bHLH domain (Additional file 3: Figure S1; Additional file 4: Table S2). Because of the high percentage of similarity in the sequences, we were not able to design specific probes that recognized *BofrSPT1* from *BofrSPT2*; thus, there is a single *BofrSPT1/2* probe, and a different probe for *BofrSPT3* (Additional file 3: Figure S1). Fragments were cleaned using QIAquick PCR purification Kit (Qiagen, Valencia, CA, USA). Digoxigenin labeled RNA probes were prepared using T7 polymerase (Roche, Switzerland), murine RNase inhibitor (New England Biolabs, Ipswich, MA, USA), and RNA labeling-mix (Roche, Switzerland) according to each manufacturers protocol. RNA in situ hybridization was performed according to Ambrose et al. [45] and Ferrándiz et al. [46], optimized to hybridize overnight at 55 °C. *In situ* hybridized sections were subsequently dehydrated and permanently mounted in Permount (Fisher, Waltham, MA, USA). All sections were

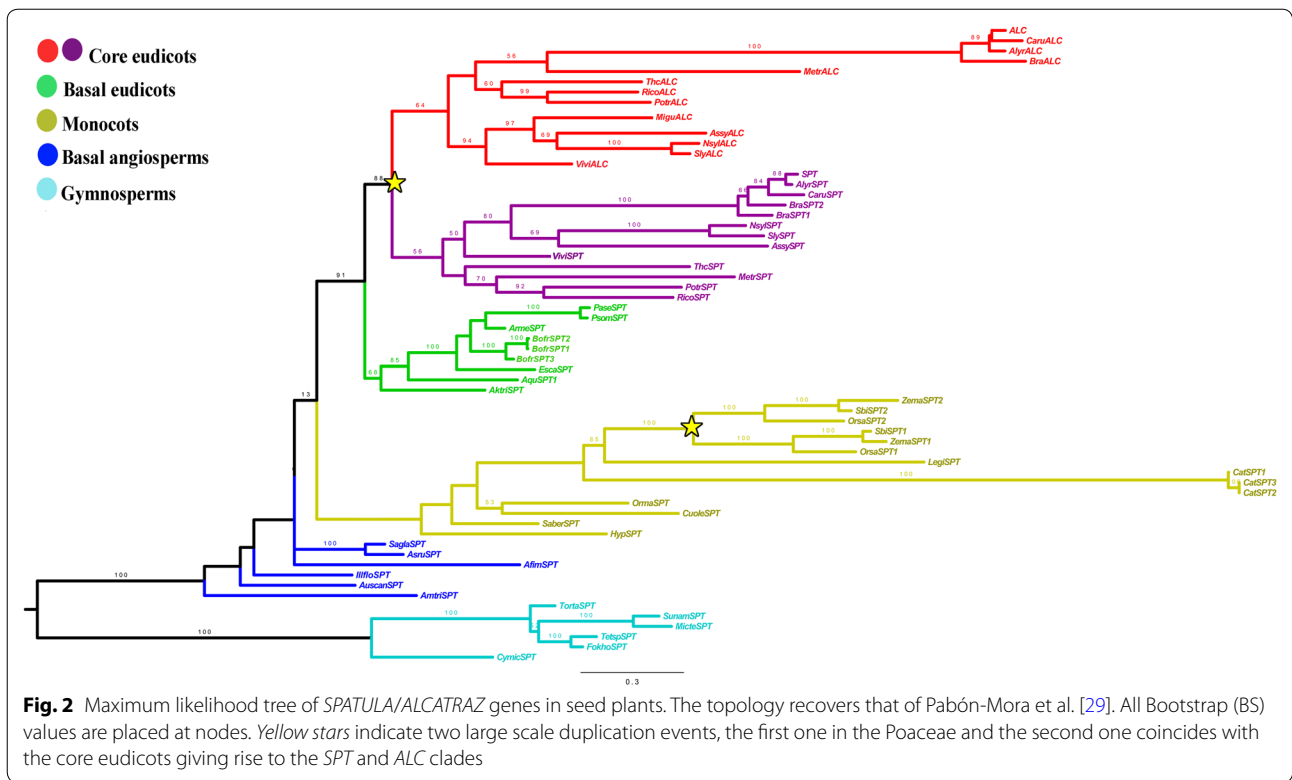
digitally photographed using a Zeiss Axioplan microscope equipped with a Nikon DXM1200C digital camera.

Results

ALCATRAZ/SPATULA gene evolution

To reconstruct the *SPT/ALC* gene lineage evolution, we included 60 sequences from all major seed plant groups. Unlike our previous analysis that used exclusively the bHLH domain [29], here we have included the complete coding sequences of all homologs. The resulting topology, however, shows the same two duplication events we previously reported (Fig. 2). One duplication event correlates with the diversification of the Poaceae, with a very high support (100 BS bootstrap value) and another duplication (BS 88), results in the SPT (BS 56) and ALC (BS 64) clades in core eudicots.

Our MEME analysis resulted in the identification of conserved protein motifs across homologous sequences in seed plants as well as more specific motifs only present in some taxa (Figs. 3, 4). We found that the bHLH domain is highly conserved in all seed plant sequences, corresponding to previously identified motifs 1, 2 and 3 (Figs. 3, 4) [21, 29]. In the 5' flank of the bHLH domain, we identified motif 4, corresponding to the acidic domain [21, 29] DDY/FDCESEEGVE, present in 42 sequences out of the 60 sampled. Motif 4 is lacking from all Brassicaceae ALC homologs, as well as in many monocot paleoSPT/ALC. Motif 5, D/EE/DI/MSL/FRSSSSSSSS, corresponds partly to the amphipathic helix (DELSSF/IRQI/VL) reported by Groszmann et al. [21]. It is rich in serine and is present in all eudicot SPT and ALC protein sequences, whereas it is lacking in monocots, basal angiosperms or gymnosperms. Motif 6 is the β strand [sensu Groszmann et al. 2011]; (LPGxLQPxxQLPQ); it is found right after the 3' end of the bHLH domain and is conserved in most angiosperm homologs sampled except in ALC, some monocot proteins (ZemaSPT2, SbiSPT2) and some gymnosperm sequences (MictSPT, SunamSPT and CymicSPT). Motif 7 (HxGP/SFQLS/LT/ASSEE-ICRED) is located toward the end of almost all protein sequences, except in the Brassicaceae ALC copies, *Medicago truncatula* ALC (MtrALC) and the gymnosperm homologs. Motif 8 seems to be present only in angiosperms SPT/ALC homologs, but is apparently lost in Brassicales as neither ALC nor SPT Brassicaceae proteins have it. Motifs 9 and 10 are exclusive to ALC Solanaceae homologs. Motif 11 G/VT/MLPV/LNQE/DSST/AxxxF is present in most non-core eudicot proteins, but it is absent in Brassicaceae SPT and ALC homologs. Motif 12 is at the C-terminus of most proteins, except for two gymnosperm sequences (FokhoSPT and TetspSPT) that have motif 12 at the N-terminus. Motifs 13 and 20 are exclusive of gymnosperm homologs. Motif 14 (DAVTV/



ASVKRRKV/F) is present in basal eudicot copies, the Chloranthales homologs, and the *Vitis vinifera SPT* copy (ViviSPT). Similarly, motif 17 is characteristic of most basal eudicot SPT/ALC homologs, including two of the three *Bocconia frutescens* SPT/ALC proteins. Motif 10 is exclusive of the Solanaceae ALC orthologs, whereas motif 15 is exclusive of the Brassicaceae SPT orthologs. Finally, motifs 16 and 18 are exclusive to the C-terminus of the three *Cattleya trianae* (Orchidaceae) SPT/ALC homologs (Figs. 3, 4).

Expression of paleoSPT/ALC *B. frutescens* orthologs (BofrSPT1/2/3)

Bocconia frutescens L., has determinate inflorescences with numerous flowers, where partial inflorescences are formed by 2–3 apetalous flowers formed by two sepals, a single whorl of homeotic stamens replacing petals, two–three whorls of true stamens and finally, a bicarpellate gynoecium with papillose exerted stigmas [36]. Flowers develop basipetally with the terminal flower always bigger. In order to properly describe the molecular genetics of fruit and flower development in *B. frutescens*, we performed developmental analyses by scanning electron and light microscopy. Similar to other floral developmental studies, particularly in the Papaveraceae [43, 44], we have defined 11 stages of floral and fruit development based on the following landmarks (Table 1; Fig. 5). Stage 1—the

floral meristem can first be distinguished (Fig. 5a). Stage 2—the two sepal primordia initiate (Fig. 5a). Stage 3—usually corresponding to petal initiation, hereby replaced by the first whorl of homeotic stamens (Fig. 5b, c). Stage 4—the next two to three staminal whorls are formed (Fig. 5d, e). Stage 5—the initiation of the bicarpellate gynoecium closing the floral meristem (Fig. 5f, g). Stage 6- the two carpels overtop the single ovule (Fig. 5h, i) and Stage 7—the apical most distal regions of the gynoecium including the style and the stigma differentiate (Fig. 5i, k). During Stage 8 the gynoecium in *B. frutescens* differentiates a medial–lateral plane of two valves derived from the two carpels, each with a central midvein, separated by a persistent commissural ring-like tissue, also irrigated by two massive vascular traces (Fig. 5l–n). It also differentiates three main proximo-distal zones including a gynophore, an ovary and a short style with two massive vascularized stigmas (Fig. 5l–n). The ovary wall is formed by 12 layers including both the outer and inner epidermis in pre-anthesis (Fig. 5n) and remains the same in anthesis (Stage 9).

During fruit development, after anthesis, two additional stages were identified: Stage 10 is defined by young fruits, when the valves expand by both anticlinal and limited periclinal cell division reaching up to 15 layers in the fruit wall (Fig. 5o, p). The commissural ring also expands outwards and develops a larger central vascular

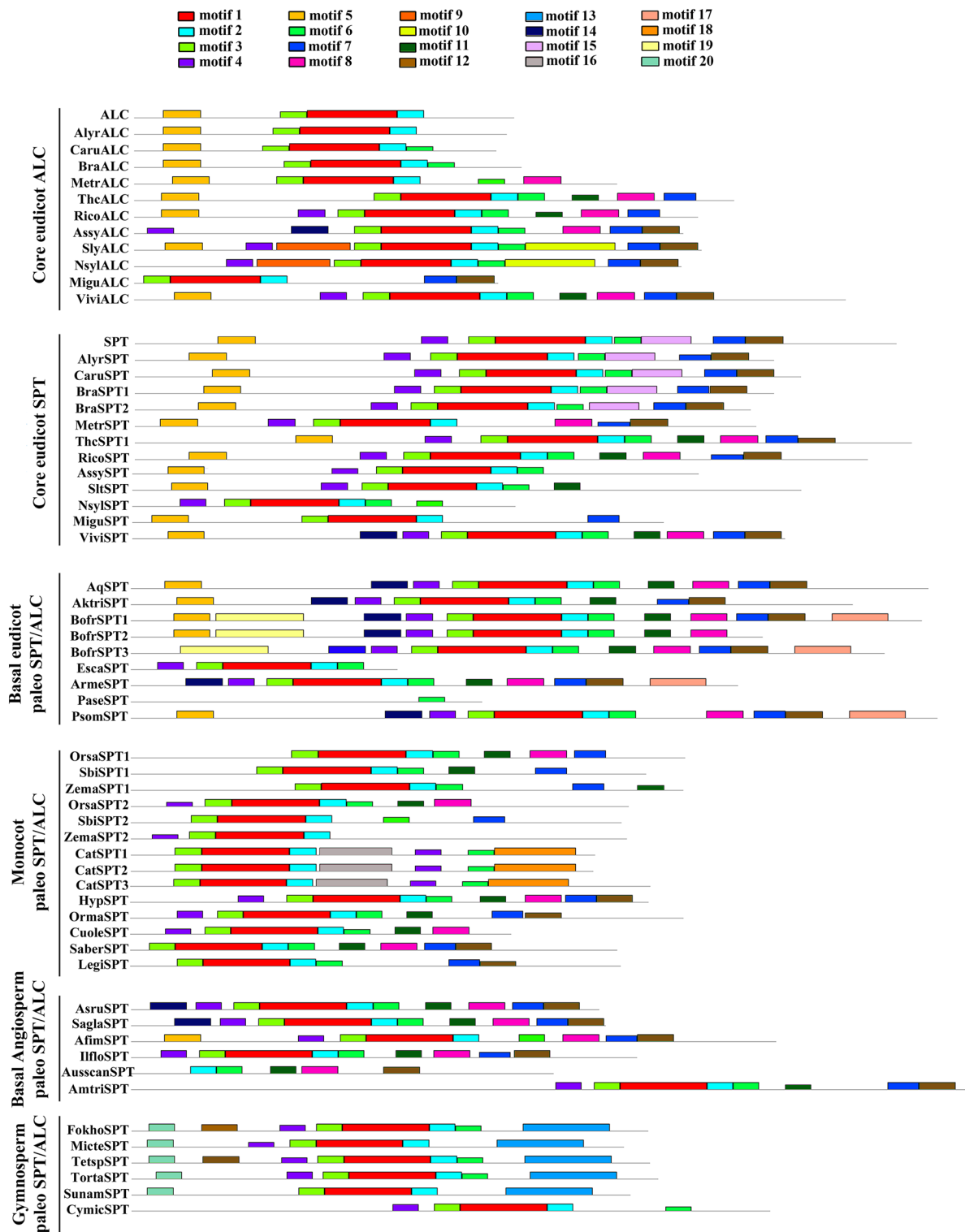


Fig. 3 Conserved motifs of SPT/ALC proteins across seed plants identified through a MEME analysis. Each motif is represented by a colored box numbered at the top. The black lines represent unique sequences. The bHLH domain (here represented by motifs 1, 2 and 3) is highly conserved across seed plant sequences. Other conserved motifs include the acidic domain (motif 4), the amphipathic helix (motif 5) and the β strand (motif 6). Scale bar indicates number of amino acids (AA). Names to the left indicate the clades to which the sequences belong to according to Fig. 2

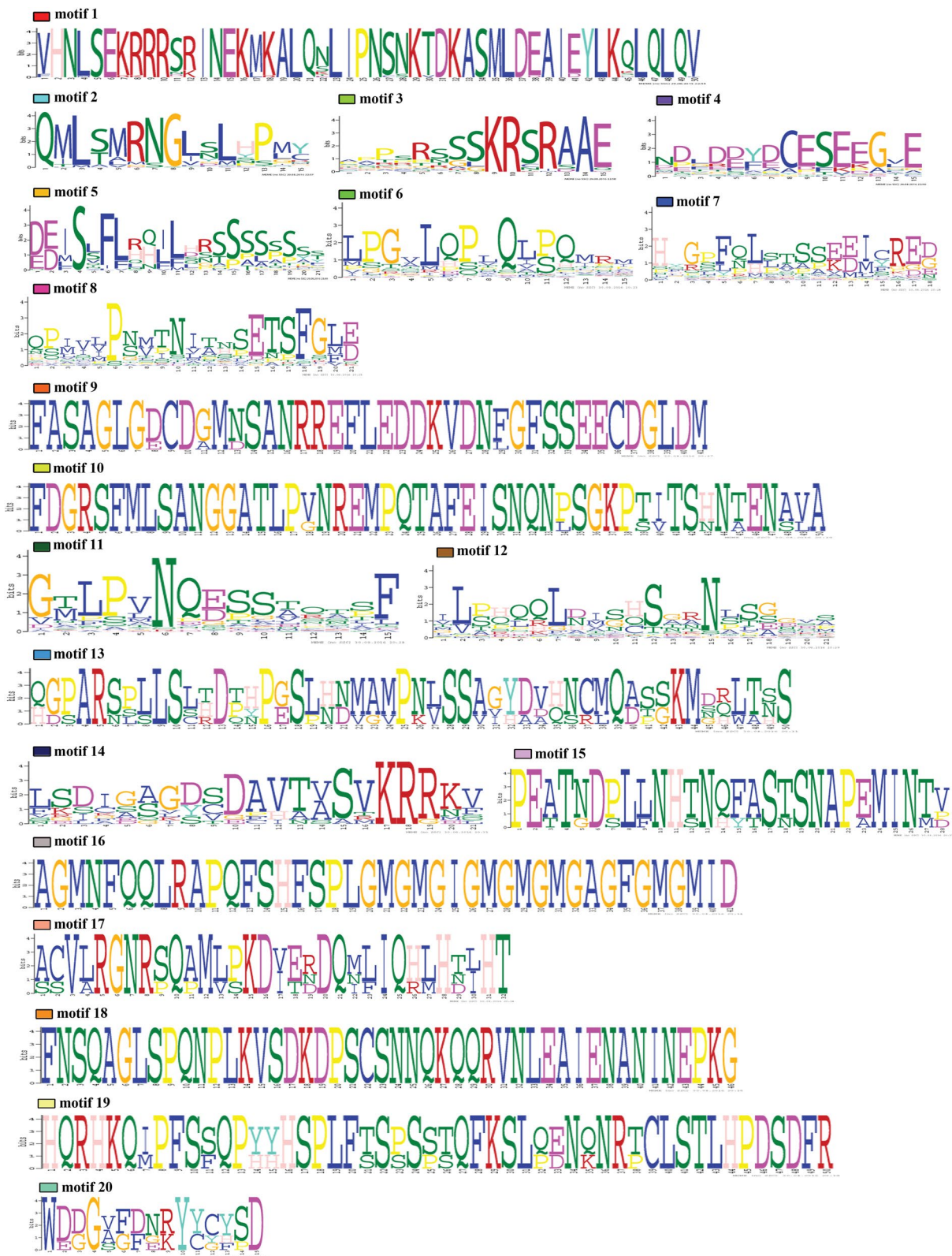


Fig. 4 Sequences of the conserved motifs detected by the MEME analysis on the SPT/ALC homologs across seed plants. Letter size denotes the degree of conservation of each amino acid

Table 1 Developmental landmarks for each stage identified during flower and fruit development

Stage	Developmental landmarks
Stage 1	The floral meristem can first be distinguished (Fig. 5a)
Stage 2	The two sepal primordia initiate (Fig. 5a)
Stage 3	Initiation of the first whorl of homeotic stamens (Fig. 5b, c)
Stage 4	Formation of the next two to three staminal whorls (Fig. 5d, e)
Stage 5	Initiation of the bicarpellate gynoecium closing the floral meristem (Fig. 5f, g)
Stage 6	Overtopping of the two carpels around the single ovule (Fig. 5h, i)
Stage 7	Differentiation of the style and the stigma (Fig. 5i, k)
Stage 8	Medial–lateral plane differentiation in the carpel. Two valves are distinguished, each with a central midvein, separated by a persistent commissural ring-like tissue, also irrigated by two massive vascular traces (Fig. 5l–n) Differentiation of the proximo-distal zones including a gynophore, an ovary and a short style with two massive vascularized stigmas (Fig. 5l–n)
Stage 9	Anthesis. Formation of up to 12 layers in the ovary wall including both the outer and inner epidermis
Stage 10	Young fruits. Expansion of the valves by both anticlinal and limited periclinal cell division reaching up to 15 layers in the fruit wall (Fig. 5o, p) Expansion of the commissural ring outwards developing a larger central vascular bundle surrounded abaxially and adaxially by collenchyma (Fig. 5o, p)
Stage 11	Mature fruits. Radial elongation of the outer epidermis accompanied by tangential elongation of the hypodermal cell layers in the mesoderm (Fig. 5q–s) Flattening of the two inner-most cell layers in the endoderm in the periphery of the dehiscence zone where they expand and become sclerenchymatic (Fig. 5q–s) Formation of the dehiscence zone by 2–3 layers of smaller cells in the limits between the commissural ring and the fruit valves (Fig. 5q–s)
Stage 12	Opercular dehiscence, which occurs between the valves and the persistent ring-like tissue and the single seed remains attached to the base of the ring through the funicle, exposing a fleshy red aril that may be a bird attractant tissue for seed dispersal (Fig. 1c, d)

bundle surrounded abaxially and adaxially by collenchyma (Fig. 5o, p). Stage 11 is defined by mature fruits, when the outer epidermis continues to enlarge into rectangular radially elongated cells and while the hypodermal cell layers in the mesoderm start to tangentially elongate the cells toward the inside of the vascular bundles begin cell expansion (Fig. 5q–s). The two inner-most cell layers in the endoderm are flattened, except in the periphery of the dehiscence zone where they expand and become sclerenchymatic (Fig. 5q–s). The dehiscence zone is formed by 2–3 layers of smaller cells in the limits between the commissural ring and the fruit valves. No adjacent layers of lignified cells can be observed at any developmental stage (Fig. 5q–s). The opercular dehiscence marks Stage 12, which occurs between the valves and the persistent ring-like tissue and the single seed remains attached to the base of the ring through the funicle, exposing a fleshy red aril that may be a bird attractant tissue for seed dispersal (Fig. 1c, d).

As there are superficial similarities between the dry dehiscent fruit with complete valve separation from a ring-like persistent replum (commissural ring) by opercular dehiscence in *Bocconia* and that of the silique of *Arabidopsis*, the comparison between the genetic mechanisms controlling fruit dehiscence in these distantly related taxa becomes even more relevant (Fig. 1). In this context, the study of expression of *paleoSPT/*

ALC genes in *B. frutescens* serves as a starting point to better assess the role of *SPT/ALC* genes prior to the core eudicot duplication and thus the functional evolution of the *SPT/ALC* gene lineage. Using in situ hybridization, we evaluated the expression patterns of *paleoSPT/ALC* copies found in the transcriptome of *B. frutescens* during flower and fruit development. Although we found three *paleoSPT/ALC* genes due to sequence similarity, we were not able to design different probes for *BofrSPT1* and *BofrSPT2*, in consequence, we used a single probe to detect *BofrSPT1/BofrSPT2* and a second gene specific probe for *BofrSPT3* (Additional file 3: Figure S1).

Our results show that *BofrSPT1/SPT2* and *BofrSPT3* have similar expression patterns early in flower development. *BofrSPT1/2/3* are turned on in the youngest floral meristem in stages 1 and 2, particularly strong in the adaxial surface of the sepals and in the limits between the sepals and the rest of the developing floral primordia (Figs. 6a, b; 7a, b). All three copies are expressed during stages 3 and 4 in the growing tips of the sepals, in the developing stamen primordia, and in the center of the floral meristem prior to carpel initiation (Figs. 6d, e; 7c). During stages 5 and 6, all copies expand their expression to the elongating carpel primordia and to the ovule and are continuously expressed in the distal-most portion of the stamens and restricted only to the tip of the sepals (Figs. 6f–i; 7e–g). During stage 7, *BofrSPT1/2/3* are

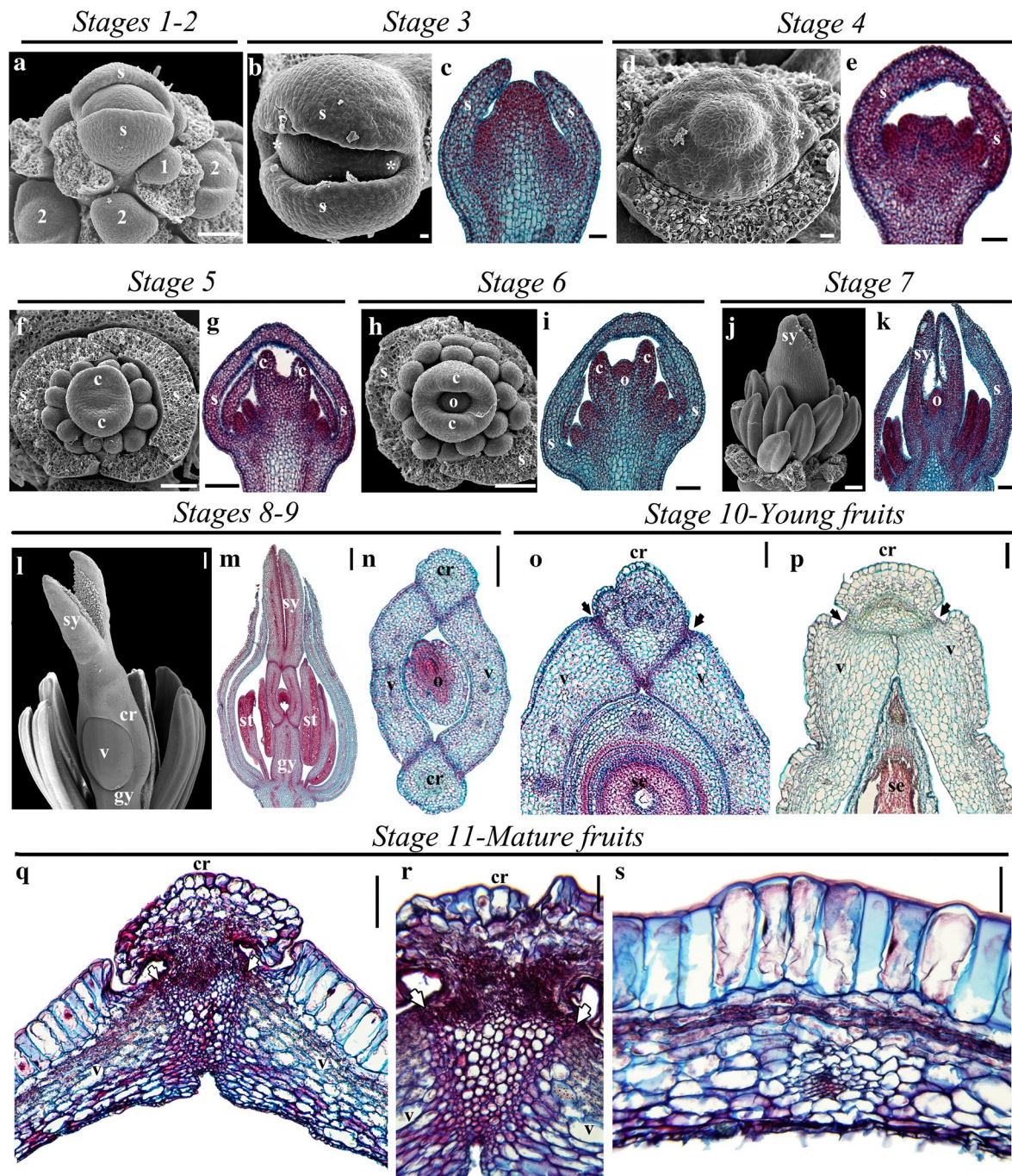


Fig. 5 Flower and fruit development stages of *Bocconia frutescens* shown in SEM and anatomical sections. **a** Inflorescence apex with floral primordia in stage 1 and floral buds during sepal initiation in stage 2. **b, c** Floral buds in the stage 3 during stamen initiation in the second whorl. **d, e** Floral buds in the stage 4 during stamen initiation in the next three inner whorls. **f, g** Floral buds in stage 5 during carpel initiation. **h, i** Floral buds in stage 6 during when the ovule is formed. **j, k** Floral buds in stage 7 during stigma elongation. **l-n** Floral buds in stages 8 and 9 during carpel proximodistal differentiation. Note the formation of the gynophore, the valves and the commissural ring. **o-p** Cross sections of *B. frutescens* fruits. **o-p** Young fruits (stage 10) with close up to the dehiscence zone in a young fruit (3 mm diameter), right after anthesis in the center (**o**) and the tip (**p**). **q-s** Mature fruits at stage 11 with close up to the dehiscence zone, the commissural ring and its main vascular bundle (**q, r**) and fruit wall (**s**) of a more mature fruit (6 mm diameter) of *B. frutescens*. 1: stage 1, 2: stage 2, c carpel, cr commissural ring, gy gynoeceum, o ovule, s sepal, sf stamen, sy stigma, v valve, *: first pair of stamens that appear in the place of petals. Arrows indicate the dehiscence zone. Scale bars: 10 μ m (**b, d**), 40 μ m (**r, s**) 50 μ m (**c, e, q**), 100 μ m (**a, f, g, h, i, j, k, o, p**), 200 μ m (**l**), 250 μ m (**n, o**), 500 μ m (**m**)

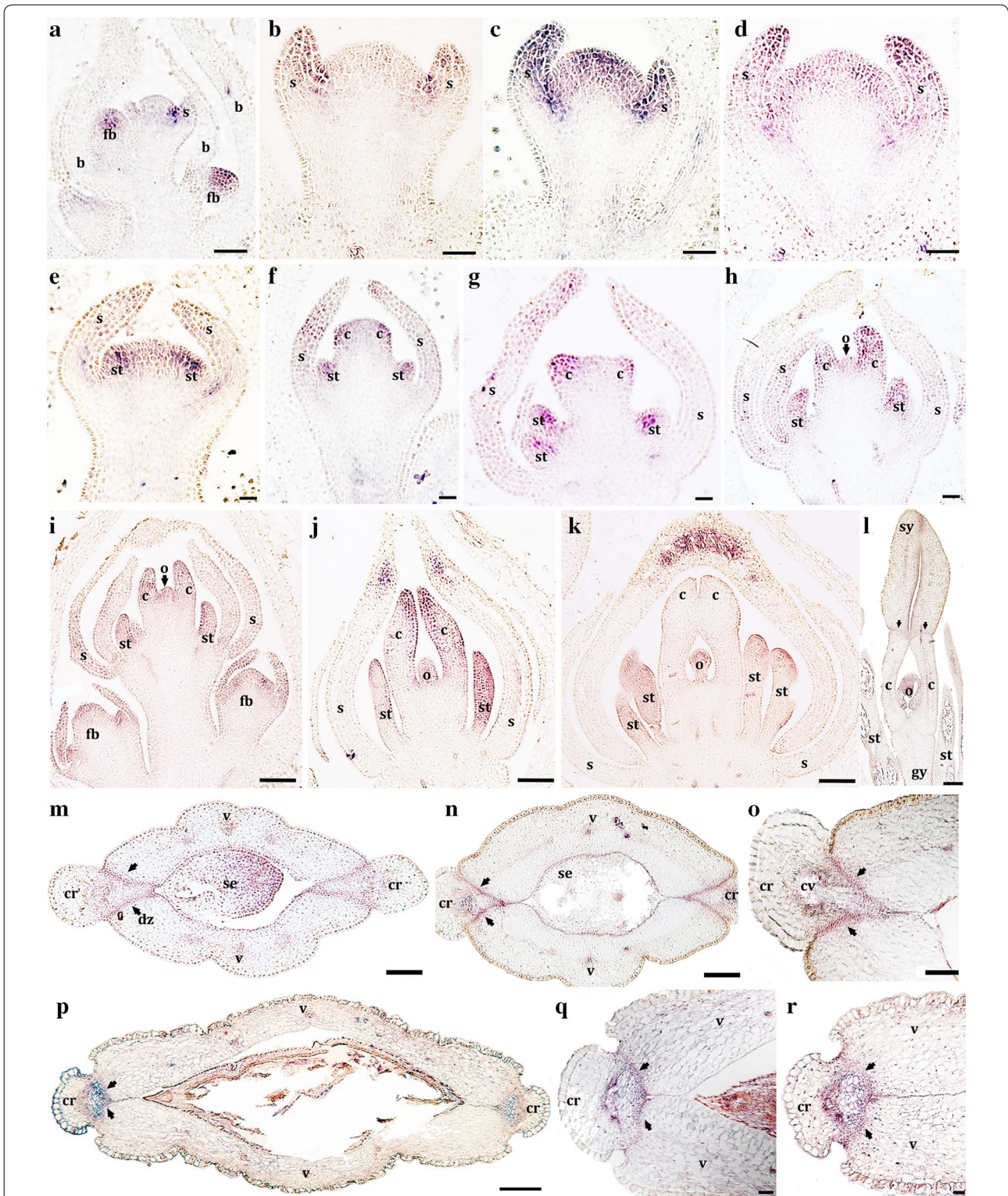


Fig. 6 Expression of *BofrSPT1/2* using *In situ* mRNA hybridization. **a** Inflorescence apex with floral buds in stages 1 and 2. **b–i** Floral stages 2–6. **J, k** Floral stage 7. **l** Floral stage 8. **m–r** Young fruits in cross section at the mid-level (**m–o**) and at the tip (**p–r**). Black arrowheads point to the dehiscence zones. *b* bract, *fb* floral bud, *s* sepal, *st* stamen, *c* carpel, *o* ovule, *sy* stigmata, *gy* gynoecium, *v* valve, *cr* commissural ring, *cv* central vascular bundle in the commissural ring, *dz* dehiscence zone, *se* seed. Scale bars: 100 μ m (**a–g, h**), 0.2 mm (**i, k, m, n**), 0.1 mm (**l, o, q, r**), 0.5 mm (**p**)

expressed in the growing tips of the two carpels that fuse to each other enclosing the single ovule (Figs. 6j, k; 7h). Later during carpel development until anthesis (stages 8 and 9), the expression patterns of *BofrSPT1/2* and *BofrSPT3* begin to diverge. During carpel elongation and differentiation, *BofrSPT1/2* are expressed in the elongating stigmas, particularly toward the adaxial epidermis of each massive stigma (Fig. 6l) but *BofrSPT3* is not. Some expression of *BofrSPT1/2* is detected in the junction between the ring-like structure derived from the commissural tissue in between the carpels and the upper stigmas (Fig. 6l), but *BofrSPT3* expression is not detected in these same areas. During the transition to fruit development (stages 10–11), *BofrSPT1/2/3* are expressed in the 3–4 cell layers between the valves and the commissural ring that will form the dehiscence zone in the mature fruits (Figs. 6m–o; 7i, j). *BofrSPT3* is also expressed in the phloem cells in the vascular bundle directly in contact with the xylem cells forming the commissural vascular bundle (Fig. 7i, j). During late fruit development prior to the dehiscence (stage 11), all three paralogs are expressed in the dehiscence zone that becomes compressed as both the valves and the commissural ring thicken (Figs. 6p–r; 7k, l).

Discussion

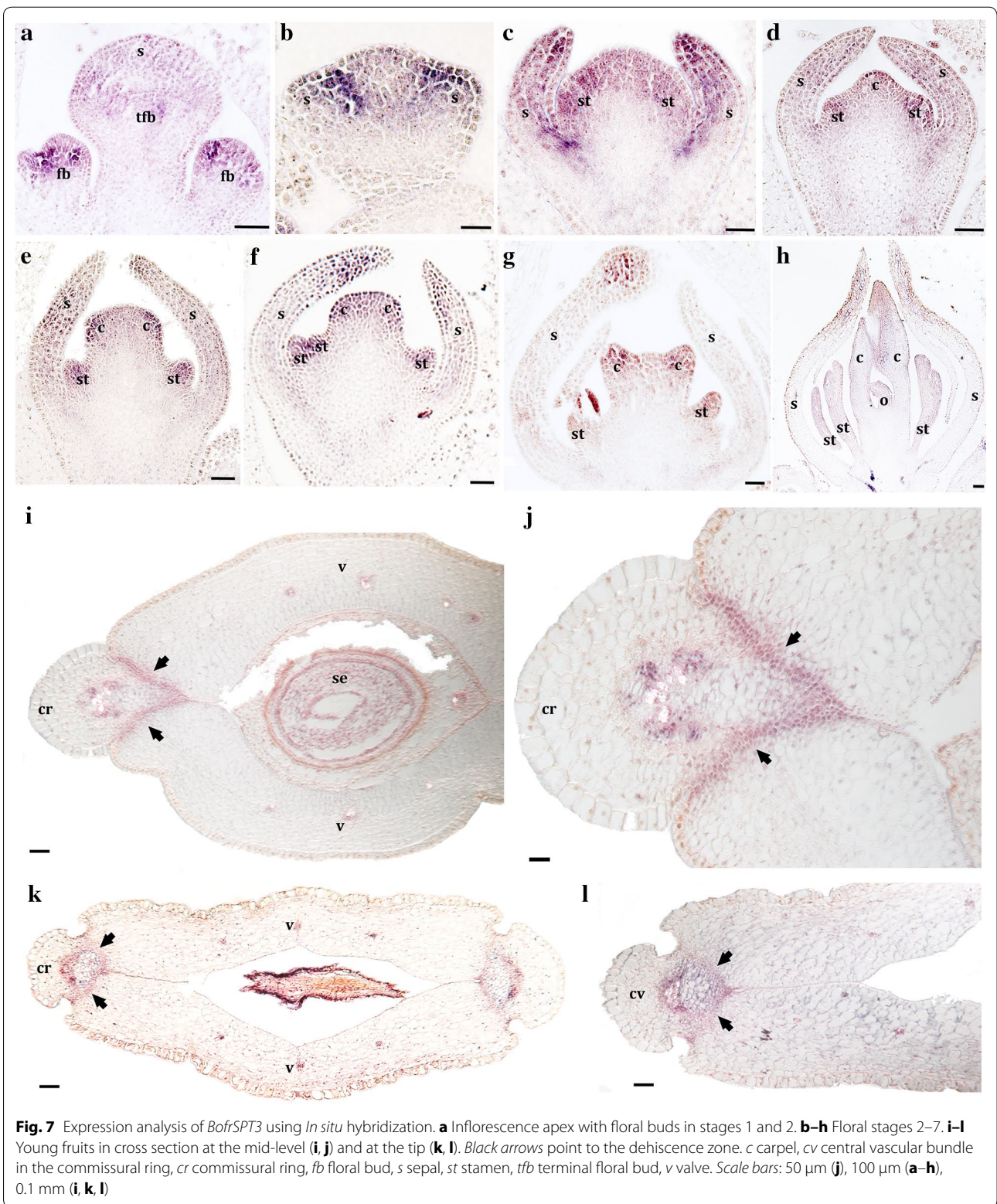
Our phylogenetic analysis confirms that a major duplication event in the *SPT/ALC* gene lineage, occurred at the base of the core eudicots and coincides with the γ whole genome duplication (WGD) 130 Mya [9, 29]. This finding contradicts previous reports that point to a Brassicaceae-specific duplication of this gene lineage in the Brassicaceae, at the β WGD event, 40–70 Mya [21, 47]. Understanding how sequences evolve and how their changes underpin diversity is one of the major questions in evolutionary developmental studies. In the absence of functional data is not possible to tease apart differences in function due to protein interactions, the expression domains, and interactions in trans [27, 48–51, 52]. However, in the absence of functional data, we can provide a framework for future studies based on sequence analyses and expression analyses.

One of our goals was to determine how conserved are the protein domains in these sequences across seed plants as this can indicate conservation or changes in protein–protein interactions of *SPT*, *ALC*, and/or paleo*SPT/ALC*, with the caveat that presence/absence of motifs is often not sufficient to explain functional specificity [50]. Our expanded dataset has allowed us to confirm that the bHLH domain and the 3' β -strand (LQLQVQ; motif 6) are highly conserved in all analyzed sequences (Figs. 3, 4). This suggests that the function in conferring specificity typically assigned to the basic

region, as well as the ability to form homo- or heterodimers of the bHLH region, are both maintained in *ALC*, *SPT* as well as in paleo*SPT/ALC* proteins (Fig. 3) [13, 29, 53, 54]. Because both the bHLH domain and the flanking regions are known to be necessary and sufficient for *SPT* homodimerization and the heterodimerization with other bHLH proteins, it is possible that some *SPT/ALC* proteins in other flowering plants could share some common partners when compared to those identified for *SPT* and *ALC*. A persistent motif in *SPT* and *ALC* clades [13, 21] and also in the basal eudicots proteins is the amphipathic helix in motif 5, located near the N-terminus of the protein (Figs. 3, 4). This amphipathic helix is required for full complementation of *spt* mutants in *Arabidopsis*, and it is thought to function in recruiting co-activators [13]. Our data show that motif 5 is only present in eudicots, and although core eudicot *SPT* functional data is scarce, at least *SlySPT* (the *Solanum lycopersicum* *SPT* homolog) can fully complement and restore *spt* mutants in *Arabidopsis*, although this does not necessarily mean their functions are conserved [13]. Because motif 5 is lacking outside eudicots it is possible that functions reported for the amphipathic helix are likely not plesiomorphic in the lineage.

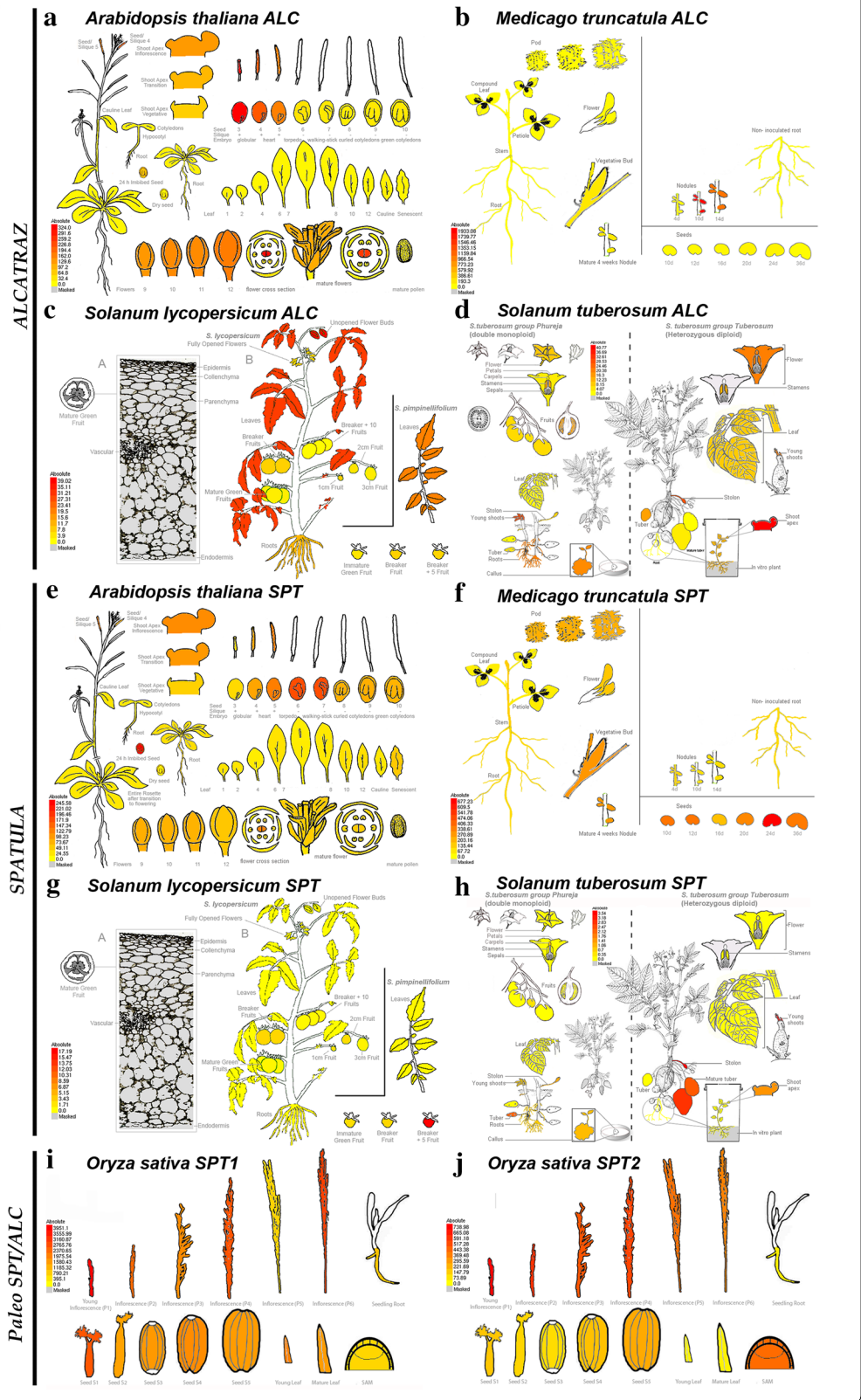
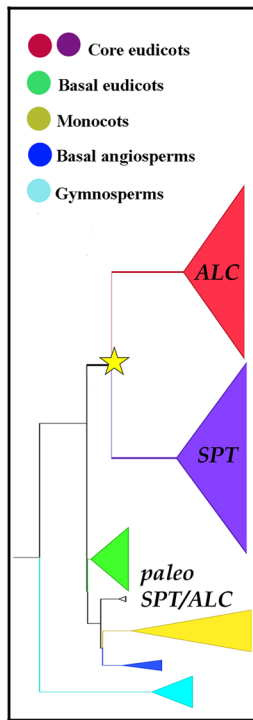
The acidic domain (motif 4) was found in *SPT*, *ALC* and paleo*SPT/ALC* but not in Brassicaceae *ALC* sequences and *SPT1* orthologs in monocots [21, 29]. Thus, the acidic domain is present in gymnosperms, basal angiosperms, *SPT2* orthologs in monocots, and basal eudicots. The acidic domain is key to restoring *spt Arabidopsis* mutants, and it is absolutely required for *SPT* function [13]. A close inspection of the rice paralogs indicates that *OrsaSPT1* lacks the acidic domain but possesses motif 7, which is rich in Phenylalanine (F), Leucine (L), Serine (S) and Glutamic Acid (E) similar to an acidic motif. Moreover, according to the eFP Browser, *OrsaSPT1* is co-expressed with *OrsaSPT2*, pointing to an interesting system to test functional evolution of close paralogs and shifts in protein folding and putative partners (Fig. 8). Altogether, this suggests that sequences with the acidic domain across seed plants could have functional resemblance and a closer set of interactions to *SPT* compared to *ALC*.

It had been hypothesized that the *SPT/ALC* genes, in comparison with their closely related bHLH *PIF 3/4/5* genes, lost the 5' APB domain, which is a negative regulator of shade responses [15]. Such evolutionary loss is a prerequisite for the recruitment of *SPT* homologs in carpel development, in a light-independent manner [15], and our analyses confirm that there are no remnants of an APB domain in any of the seed plant *SPT/ALC* sequences. Interestingly, motifs 8 and 14 are only present in angiosperms and lacking in gymnosperm homologs,



whereas motifs 13 and 20 are exclusive to gymnosperms. However, the *Amborella trichopoda* *AtriSPT* sequence lacks both motifs 8 and 14, suggesting that

there are no synapomorphic motifs for all angiosperm sequences likely responsible for new capabilities in carpel development.



(See figure on previous page.)

Fig. 8 Expression patterns of selected *ALCATRAZ/SPATULA* homologs from selected model species. **a–d** Expression of *ALCATRAZ* homologs from core eudicots including *Arabidopsis thaliana* (At5g67110), *Medicago truncatula* (Medtr1g019240), *Solanum lycopersicum* (Solyc04g078690), *Solanum tuberosum* (PGSC003DMT400020534). **e–h** Expression of *SPATULA* homologs from the same core eudicots, *A. thaliana* (At4g36930), *M. truncatula* (Medtr5g017040), *S. lycopersicum* (Solyc02g093280), and *S. tuberosum* (PGSC003DMP400007151). **i–j** Expression of *paleoALC/SPT* from *Oryza sativa* (monocot), OrsaSPT1 (LOC_Os06g06900) and OrsaSPT2 (LOC_Os02g56140). Taken from the eFP browser (<http://bar.utoronto.ca/efp/cgi-bin/efp-Web.cgi>)

Our analysis corroborates previous findings that all model core eudicot plants including *Arabidopsis*, tomato, potato and *Medicago* have orthologs to *SPT* and to *ALC* as the duplication event occurred prior to their diversification [29]. This observation poses a new question regarding conservation in their expression patterns after the duplication as well as the expression patterns and putative roles of pre-duplication genes. In *Arabidopsis*, *ALC* and *SPT* are co-expressed in the leaf margin, medial ridge of the gynoecium by stage 8, in the valve margin from stage 9 and in the septum outgrowth and transmitting tract as well as in the stigma during stage 11 and in the ovules [16, 19–21, 48]. During late fruit development, expression of both *SPT* and *ALC* also overlaps in the dehiscence zone of the fruit [13, 16, 19–21]. However, in addition to these organs, *ALC* is expressed in the petal margins, the connective of developing anthers, nectaries, and the pedicel-stem abscission zone [20, 21]. On the other hand, *SPT* is expressed during early embryogenesis, particularly in the root meristem and the procambium [16]. Coexpression of both paralogs is consistent with their conserved role during carpel development and in the specification of the non-lignified layer of the dehiscence zone [20, 21]. Nevertheless, only *SPT* can fully complement *alc* mutants, whereas *ALC* can only partly complement *spt* mutants, likely because of both the shortening, and the loss of putative ancestral motifs of *ALC* orthologs in Brassicaceae (Figs. 2, 3) [21]. Sub-functionalization has been posed as the functional evolutionary scenario after duplication, as *SPT* has a more fundamental role during early gynoecium fusion, elongation and development, with direct *HECATE1,2,3* partners, whereas *ALC* has a more prominent role in the differentiation of the separation (non-lignified) layer in the dehiscence zone [20, 21, 55, 56]. *SPT* also plays important roles in specifying valve margin differentiation and dehiscence zone formation later in development, as it is a direct target of *INDEHISCENT* (*IND*) another bHLH gene closely related with *HEC3* which promotes the differentiation of both the lignified and the non-lignified layers [22].

Broad expression patterns for both paralogs are uncommon in other core eudicots (Fig. 8). In *Medicago truncatula* (another rosid like *Arabidopsis*), *Metralc* has only been detected in the root nodules, whereas *Metrspt* is expressed in the vegetative meristem and during fruit

and seed development. In *M. truncatula*, for instance, it is well known that *Metrspt*, the most important positive regulator upstream of *ALC* and *IND* according to the model, has conserved roles in establishing the dehiscence zone and controls the coiled fruit shape [56]. According to the expression patterns found in the public databases (eFP Browser), and considering upstream regulators are conserved, it is likely that it is the *Metrspt* and not the *Metralc* paralog playing important roles during flower and fruit development. Similar cases of sub- or neofunctionalization are likely to be occurring in Asterids. Also in the eFP Browser, in tomato (*Solanum lycopersicum*), while *Slyalc* is expressed in leaves and to some extent in flowers, *Slyspt* is mostly restricted to later developmental stages of fruit maturation (Fig. 8). In potato (*S. tuberosum*), while *Stualc* is expressed in roots, young shoots and tubers, sepals, petals, stamens and fruits, *Stuspt* is mostly restricted to tuber development (Fig. 8). Nonetheless, a functional scenario has so far been incomplete since there is no expression data available for pre-duplication genes that allow the assessment of the plesiomorphic role of the gene lineage.

Here, we present the first expression analysis of *paleoSPT/ALC* genes in the basal eudicot *Bocconia frutescens* (Papaveraceae). The study of expression patterns during flower and fruit development in a basal eudicot, placed before the duplication event, allows us to propose hypotheses in terms of the functional evolution of the *SPT/ALC* gene lineage. In *B. frutescens*, *paleoSPT/ALC* copies are expressed throughout the early floral meristem (stage 1), in the sepal primordia and quickly after sepal initiation, their expression delimits the boundaries between sepals and the rest of the floral primordia where all remaining organs will form (stage 2) (Figs. 6a, b; 7a, b). During early and mid-flower developmental stages (stages 3–6), *paleoSPT/ALC* genes are expressed in the growing tips of stamens and carpels (Figs. 6c–i, 7c–g). Importantly, *Bofrspt1,2,3* genes are expressed in the adaxial domain of the postgenitally fused carpels and in the growing ovule, during stages 7 and 8 (Figs. 6j–l; 7h). Our results show broader expression patterns during early floral organ specification, compared to that reported for *SPT* and *ALC* in *Arabidopsis*, but point to ancestral roles in eudicot *paleoSPT/ALC* homologs in early gynoecium patterning including the formation of the septum, the stigma and the style.

In addition to the expression recorded during flower development, all three *BofrSPT1,2,3* copies remain turned on during the gynoecium to fruit transition to the smaller cells marking the dehiscence zone present between the commissural ring and the valves (Figs. 6, 7). As dehiscence is opercular, the dehiscence zone as well as *BofrSPT1,2,3* expression is present from the proximal to the distal-most portions of the ovary (Figs. 1, 6, 7). This is exactly the same as the expression patterns detected for both *SPT* and *ALC* in the separation layer of the dehiscence zone during fruit maturation in *Arabidopsis* [20, 21, 56], suggesting that a role in specifying the dehiscence zone was already present in the pre-duplication *paleoSPT/ALC* genes and it is likely conserved in all early diverging angiosperms. It is important to highlight that the *B. frutescens* opercular capsule does not form a sclerenchymatic cell layer adjacent to the smaller celled separation layer, between the commissural ring and the valves. This suggests that the role of *paleoSPT/ALC* genes in delimiting the separation layer occurs independently of the formation of a lignified layer or the maintenance of *INDEHISCENT/HECATE3* functions in the dehiscence zone [22].

Teasing apart the evolution of the fruit developmental network will only be possible with the knockouts of each gene in the network from diverse basal eudicot species. However, our expression data supports the idea of a conserved fruit developmental genetic network in basal eudicots like *B. frutescens*, even in the absence of strict orthologs of *FUL* and *SHP* [29]. Both the *APETALA1/FRUITFULL* and the *AGAMOUS/SHATTERPROOF* gene lineage have duplicated extensively in core eudicots, basal eudicots and monocots independently, making it extremely difficult to tease out new plesiomorphic roles in plant development during plant evolution [57–59]. Gene down-regulation of *FRUITFULL*-like homologs in other Papaveraceae like *Eschscholzia californica* and *Papaver somniferum* results in shorter fruits with premature fruit wall rupture and pericarp defects suggesting conserved roles in fruit development [60]. In addition, *Bocconia frutescens* has three *FUL*-like genes *BofrFL1*, *FL2* and *FL3*, from which *BofrFL2* and *BofrFL3* show expression in early carpel patterning as well as during fruit development suggesting conserved roles of *FUL*-like genes in fruit development across Papaveraceae [36]. On the other hand, down-regulation of *AGAMOUS*-like homologs in Papaveraceae results in homeotic shifts from stamen and carpel identity to petal identity accompanied by the acquisition of indeterminacy in the floral meristem [61, 62]. As gene silencing of *AG*-like genes block carpel development, it is unclear whether they may play roles during fruit development, and particularly if they could have conserved roles in the specification of the dehiscence zone. The fact

that *BofrAG* is expressed during carpel development and maintained in early fruit development [36], together with the observation that *BofrSPT1*, 2 and 3 have conserved expression patterns in comparison with their core eudicot counterparts, allow us to propose that it is likely that the upstream gene regulatory network specifying *SPT* and *ALC* expression is also maintained in basal eudicots [63].

Conclusions

Basal eudicots are useful models to better understand the evolution of the fruit developmental network as they often carry the pre-duplication genes with respect to core eudicot model species. Although there is some redundancy between *ALC* and *SPT*, these gene clades have undergone some degree of sub-functionalization in the core eudicots. Our results also indicate that in *Bocconia frutescens*, *paleoSPT/ALC* may play a role in early floral organ specification, particularly in sepal and stamen morphogenesis that was subsequently lost in core eudicot lineages. It will be necessary to investigate the function of *paleoSPT/ALC* by knockouts or knockdowns in the Papaveraceae and do complementation assays in *Arabidopsis* to address conserved roles prior to the split of *ALC* and *SPT* clades. There is a large amount of dry fruit diversity in the Papaveraceae with diverse dehiscence mechanisms and functional studies in a range of species will help to elucidate the pre-duplication role of this gene lineage and its subsequent sub-functionalization.

Additional files

Additional file 1: Table S1. List of all genes included in the phylogenetic analyses of *SPT/ALC* gene lineage with their respective accession number.

Additional file 2: Figure S2. e-values for each motif as reported in the MEME search (<http://meme-suite.org/tools/meme>).

Additional file 3: Figure S1. Alignment of the *B. frutescens* *paleoSPT/ALC* protein sequences. Yellow boxes highlighting the position where the probe was designed for each sequence.

Additional file 4: Table S2. List of the primers used for the in situ hybridization.

Abbreviations

ALC: ALCATRAZ; bHLH: basic/helix-loop-helix; FUL: FRUITFULL; HEC: HECATE; IND: INDEHISCENT; ML: maximum likelihood; SPT: SPATULA; SHP: SHATTERPROOF; WGD: whole genome duplication.

Authors' contributions

All authors planned and designed the research, performed experiments, analyzed the data and wrote and approved the final version of the manuscript. All authors read and approved the final manuscript.

Author details

¹ Instituto de Biología, Universidad de Antioquia, Medellín, 1226, Colombia.

² New York Botanical Garden, Bronx, NY 10458, USA. ³ City University of New York, New York, NY 10016, USA.

Acknowledgements

We thank J.F. Alzate (Centro Nacional de Secuenciación de Genómica, SIU, Universidad de Antioquia, Medellín, Antioquia) for the assembly and storage of our own generated transcriptomes. We thank Favio Gonzalez (Universidad Nacional de Colombia) for his extensive help during field work and the processing of the samples. All samples were collected following the local and national guidelines and legislation for cultivated plants. Finally, we thank Rainer Melzer and an anonymous reviewer for their helpful suggestions and comments.

Competing interests

The authors declare that they have no competing interests.

Availability of supporting data

All supporting data are available in the supplementary files.

Consent for publication

All authors have provided consent for publication.

Ethical approval and consent to participate

The work was done with plants so animal ethical approval and consent to participate are not applicable. As the plants are cultivated, they are exempt of legal access to genetic resources as stated by the Andean and Colombian law.

Funding

This work was funded by COLCIENCIAS (111565842812) and the Committee for Research development (CODI), Convocatoria de Internacionalización 2015 at the Universidad de Antioquia. C.Z. was partially supported by an NSF-EDEN summer fellowship, and B.A.A. gratefully acknowledges funding from The Eppley Foundation for Research, Inc.

Received: 21 January 2017 Accepted: 8 March 2017

Published online: 15 March 2017

References

- Wang K, Wang Z, Li F, Ye W, Wang J, Song G, Yue Z, Cong L, Shang H, Zhu S, Zou C, Li Q, Yuan Y, Lu Wei H, Gou C, Zheng Z, Yin Y, Zhang X, Liu K, Wang B, Song C, Shi N, Kohel RJ, Percy RG, Yu JZ, Zhu YX, Wang J, Yu S. The draft genome of a diploid cotton *Gossypium raimondii*. *Nat Genet*. 2012. doi:10.1038/ng.2371.
- Wendel JF, Jackson SA, Meyers BC, Wing RA. Evolution of plant genome architecture. *Genome Biol*. 2016. doi:10.1186/s13059-016-0908-1.
- Soltis PS, Soltis DE. Polyploidy and genome evolution. 1st ed. Berlin: Springer; 2012.
- Chen ZJ, Ni Z. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. *BioEssays*. 2006;28:240–52.
- Chen ZJ. Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu Rev Plant Biol*. 2007;58:377–406.
- Wendel JF. Genome evolution in polyploids. *Plant Mol Biol*. 2000;42:225–49.
- Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A, et al. Widespread genome duplications throughout the history of flowering plants. *Genome Res*. 2006;16:738–69.
- Flagel LE, Wendel JF. Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation. *New Phytol*. 2010;186:184–93.
- Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, de Pamphilis CW. Ancestral polyploidy in seed plants and angiosperms. *Nature*. 2011. doi:10.1038/nature09916.
- Atchley WR, Fitch WM. A natural classification of the basic helix–loop–helix class of transcription factors. *Proc Natl Acad Sci*. 1997;94:5172–6.
- Lendent V, Vervoort M. The basic helix–loop–helix protein family: comparative genomics and phylogenetic analysis. *Genome Res*. 2001;11:754–70.
- Toledo-Ortiz G, Huq E, Quail PH. The Arabidopsis basic/helix–loop–helix transcription factor family. *Plant Cell*. 2003;5:1749–70.
- Groszmann M, Paicu T, Smyth DR. Functional domains of SPATULA, a bHLH transcription factor involved in carpel and fruit development in Arabidopsis. *Plant J*. 2008;55:40–52.
- Pires N, Dolan L. Early evolution of bHLH proteins in plants. *Plant Signal Behav*. 2010;7:911–2.
- Reymond MC, Brunoud G, Chauvet A, Martínez-García JF, Martin-Magniette ML, Monéger F, Scutt CP. A light-regulated genetic module was recruited to carpel development in Arabidopsis following a structural change to SPATULA. *Plant Cell*. 2012;24:2812–25.
- Groszmann M, Bylstra Y, Lampugnani ER, Smyth DR. Regulation of tissue-specific expression of SPATULA, a bHLH gene involved in carpel development, seedling germination, and lateral organ growth in Arabidopsis. *J Exp Bot*. 2010;61:1495–508.
- Alvarez J, Smyth DR. CRABS CLAW and SPATULA, two Arabidopsis genes that control carpel development in parallel with AGAMOUS. *Development*. 1999;126:2377–86.
- Alvarez J, Smyth DR. CRABS CLAW and SPATULA genes regulate growth and pattern formation during gynoecium development in *Arabidopsis thaliana*. *Int J Plant Sci*. 2002;163:17–41.
- 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;128:1089–98.
- Rajani S, Sundaresan V. The Arabidopsis myc/bHLH gene ALCATRAZ enables cell separation in fruit dehiscence. *Curr Biol*. 2001;11:1914–22.
- Groszmann M, Paicu T, Alvarez JP, Swain SM, Smyth DR. SPATULA and ALCATRAZ, are partially redundant, functionally diverging bHLH genes required for Arabidopsis gynoecium and fruit development. *Plant J*. 2011;68:816–29.
- Girin T, Paicu T, Stephenson P, Fuentes S, Körner E, O'Brien M, Sorefan K, Wood TA, Balanzá V, Ferrandiz C, Smyth DR, Østergaard L. INDEHISCENT and SPATULA interact to specify carpel and valve margin tissue and thus promote seed dispersal in Arabidopsis. *Plant Cell*. 2011;23:3641–53.
- Gremski K, Ditta G, Yanofsky MF. The HECATE genes regulate female reproductive tract development in *Arabidopsis thaliana*. *Development*. 2007;134:3593–601.
- Geisler-Lee J, O'Toole N, Ammar R, Provart NJ, Millar AH, Geisler M. A predicted interactome for Arabidopsis. *Plant Physiol*. 2007;2:317–29.
- Wang F, Shi DQ, Liu J, Yang WC. Novel nuclear protein ALC-INTERACTING PROTEIN1 is expressed in vascular and mesocarp cells in Arabidopsis. *J Integr Plant Biol*. 2008. doi:10.1111/j.1744-7909.2008.00694.
- Gu Q, Ferrándiz C, Yanofsky MF, Martienssen R. The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development. *Development*. 1998;125:1509–17.
- Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF. SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. *Nature*. 2000;404:766–70.
- Ripoll JJ, Roeder AHK, Ditta GS, Yanofsky MF. A novel role for the floral homeotic gene APETALA2 during Arabidopsis fruit development. *Development*. 2011;138:5167–76.
- Pabón-Mora N, Wong KG, Ambrose B. Evolution of fruit development genes in flowering plants. *Front Plant Sci*. 2014. doi:10.3389/fpls.2014.00300.
- Tani E, Tsuballa A, Stedel C, Kalloniati C, Papaefthimiou D, Polidoros A, Darzentas N, Ganopoulos I, Flemetakis E, Katinakis P, Tsaftaris A. The study of a SPATULA-like bHLH transcription factor expressed during peach (*Prunus persica*) fruit development. *Plant Physiol Biochem*. 2011;49:654–63.
- Tisza V, Kovács L, Balogh A, Heszky L, Kiss E. Characterization of FaSPT, a SPATULA gene encoding a bHLH transcriptional factor from the non-climacteric strawberry fruit. *Plant Physiol Biochem*. 2010. doi:10.1016/j.plaphy.2010.08.001.
- Lenster T, Theißen G. Conservation of fruit dehiscence pathways between *Lepidium campestre* and *Arabidopsis thaliana* sheds light on the regulation of INDEHISCENT. *Plant J*. 2013;76:545–56.
- Makkena S, Lamb RS. The bHLH transcription factor SPATULA is a key regulator of organ size in Arabidopsis thaliana. 2013; doi:10.4161/psb.24140
- Kadereit JW. Papaveraceae. In: Kubitzki K, editor. The families and genera of vascular plants. Berlin: Springer; 1993. p. 494–506.
- Pabón-Mora N, Suárez-Baron H, Ambrose BA, González F. Flower development and perianth identity candidate genes in the basal angiosperm *Aristolochia fimbriata* (Piperaleae: Aristolochiaceae). *Front Plant Sci*. 2015. doi:10.3389/fpls.2015.01095.

36. Arango-Ocampo C, González F, Alzate JF, Pabón-Mora N. The developmental and genetic bases of apetalry in *Bocconia frutescens* (Chelidoniaceae: Papaveraceae). *EvoDevo*. 2016. doi:10.1186/s13227-016-0054-6.
37. Madrigal Y, Alzate JF, Pabón-Mora N. Evolution and expression patterns of TCP genes in Asparagales. *Front Plant Sci*. 2017; doi:10.3389/fpls.2017.00009
38. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucl Acids Res*. 2002;30:3059–66.
39. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol*. 2008;57:758–71.
40. Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES science gateway for interference of large phylogenetic trees. In: Proceedings of the gateway computing environments workshop (GCE). doi:10.1109/GCE.2010.5676129
41. Bailey TL, Williams N, Misleh C, Li WW. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucl Acids Res*. 2006;34:369–73.
42. Kraus JE, De Souza HC, Rezende MH, Castro NM, Vecchi C, Luque R. Astra blue and basic fuchsin double staining of plant material. *Biotech Histochem*. 1998;73:235–43.
43. Becker A, Gleissberg S, Smyth DR. Floral and vegetative morphogenesis in California Poppy (*Eschscholzia californica* Cham.). *Int J Plant Sci*. 2005;166:537–55.
44. Drea S, Hileman LC, de Martino G, Irish VF. Functional analyses of genetic pathways controlling petal specification in poppy. *Development*. 2007;134:4157–66.
45. Ambrose BA, Lerner DR, Ciceri P, Padilla CM, Yanofsky MF, Schmid RJ. Molecular and genetic analyses of the Silky1 gene reveal conservation in floral organ specification between eudicots and monocots. *Mol Cell*. 2000;5:569–79.
46. Ferrándiz C, Gu Q, Martienssen R, Yanofsky MF. Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. *Development*. 2000;127:725–34.
47. Franzke A, Lysak MA, Al-Shehbaz IA, Kochand MA, Mummenhoff K. Cabbage family affairs: the evolutionary history of Brassicaceae. *Trends Plant Sci*. 2011. doi:10.1016/j.tplants.2010.11.005.
48. Kramer EM, Jaramillo A, Di Stilio VS. Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS box genes in angiosperms. *Genetics*. 2004. doi:10.1534/genetics.166.2.1011.
49. Davies B, Motte P, Keck E, Saedler H, Sommer H, Schwarz-Sommer Z. PLENA and FARINELLI: redundancy and regulatory interactions between two Antirrhinum MADS-box factors controlling flower development. *EMBO J*. 1999;18:4023–34.
50. Moore RC, Purugganan MD. The evolutionary dynamics of plant duplicate gene. *Curr Opin Plant Biol*. 2005. doi:10.1016/j.pbi.2004.12.001.
51. Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, Yanofsky MF. Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature*. 2003. doi:10.1038/nature01741.
52. Slattery M, Zhou T, Yang L, Dantas Machado AC, Gordân R, Rohs R. Absence of a simple code: how transcription factors read the genome. *Trends Biochem Sci*. 2014;39:381–99.
53. Ma PCM, Rould MC, Weintraub H, Pabo CO. Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation. *Cell*. 1994;77:451–9.
54. Shimizu T, Toumoto A, Ihara K, Shimizu M, Kyogoku Y, Ogawa N, Oshima N, Hakoshima T. Crystal structure of PHO4 bHLH domain-DNA complex: flanking base recognition. *EMBO J*. 1997;16:4689–97.
55. Ichihashi Y, Horiguchi G, Gleissberg S, Tsukaya H. The bHLH transcription factor SPATULA controls final leaf size in *Arabidopsis thaliana*. *Plant Cell Physiol*. 2010. doi:10.1093/pcp/pcp184.
56. Girin T, Sorefan K, Østergaard L. Meristematic sculpting in fruit development. *J Exp Bot*. 2009;60:1493–502.
57. Fourquin C, del Cerro C, Victoria FC, Vialette-Guiraud A, de Oliveira AC, Ferrandiz C. A change in SHATTERPROOF protein lies at the origin of a fruit morphological novelty and a new strategy for seed dispersal in *Medicago* genus. *Plant Physiol*. 2013;12:907–17.
58. Litt A, Irish VF. Duplication and diversification in the APETALA1/FRUITFULL floral homeotic gene lineage: implications for the evolution of floral development. *Genetics*. 2003;2:821–33.
59. Dreni L, Kater MM. MADS reloaded: evolution of the AGAMOUS subfamily genes. *New Phytol*. 2014. doi:10.1111/nph.12555.
60. Pabon-Mora N, Ambrose BA, Litt A. Poppy APETALA1/FRUITFULL orthologs control flowering time, branching, perianth identity, and fruit development. *Plant Physiol*. 2012;158:1685–704.
61. Yellina A, Orashakova S, Lange S, Erdmann R, Leebens-Mack J, Becker A. Floral homeotic C function genes repress specific B function genes in the carpel whorl of the basal eudicot California poppy (*Eschscholzia californica*). *EvoDevo*. 2010. doi:10.1186/2041-9139-1-13.
62. Hands P, Vosnakis N, Betts D, Irish VF, Drea S. Alternate transcripts of a floral developmental regulator have both distinct and redundant functions in opium poppy. *Ann Bot Lond*. 2011;107:1557–66.
63. Ferrandiz C, Fourquin C. Role of the FUL-SHP network in the evolution of fruit morphology and function. *J Exp Bot*. 2014;65:4505–13.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

