Title:

A change in *SHATTERPROOF* protein lies at the origin of a fruit morphological novelty and a new strategy for seed dispersal in *Medicago* genus

Chloé Fourquin[a], Carolina del Cerro[a], Filipe C. Victoria[b,c], Aurélie Vialette-Guiraud[a], Antonio C. de Oliveira[b] and Cristina Ferrándiz[a]

[a] *Instituto de Biología Molecular y Celular de Plantas, UPV-CSIC, 46022 Valencia, Spain*

[b] *Plant Genomics and Breeding Center, Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas, RS, Brasil*

[c]*Antarctic Studies Plant Core, National Institute of Antarctic Science and Technology for Environmental Research, Universidade Federal do Pampa, São Gabriel, RS, Brasil.*

Summary: A study on the evolutionary origin of a novel fruit morphology in legumes and the importance of changes in coding regions of master genes to generate diversity

Financial source:

The work was funded by grants BIO2009-09920 from Spanish Ministerio de Ciencia e Innovación. to C.Ferr. and FP7-PEOPLE-PIRSES-2009-247589 from EU to C. Ferr. and A.C.O. C. Four. was the recipient of a Fellowship for Foreign Young Postdocs from Spanish Ministerio de Ciencia e Innovación.

Aurélie Vialette-Guiraud's Present Address: Reproduction et Developpement des Plantes, UMR 5667, CNRS-INRA-Universite de Lyon, Ecole Normale Superieure de Lyon, Lyon Cedex 07, France

Corresponding author :

Cristina Ferrándiz: cferrandiz@ibmcp.upv.es

Abstract

Angiosperms are the most diverse and numerous group of plants and it is generally accepted that this evolutionary success owes in part to the diversity found in fruits, key for protecting the developing seeds and ensuring seed dispersal. Although studies on the molecular basis of morphological innovations are few, they all illustrate the central role played by transcription factors acting as developmental regulators. Here, we show that a small change in the protein sequence of a MADS-box transcription factor correlates with the origin of a highly modified fruit morphology and the change in seed dispersal strategies that occurred in *Medicago*, a genus belonging to the large legume family. This protein sequence modification alters the functional properties of the protein, affecting the affinities for other protein partners involved in high-order complexes. Our work illustrates that variation in coding regions can generate evolutionary novelties not based on gene duplication/subfunctionalization but by interactions in complex networks, contributing also to the current debate on the relative importance of changes in regulatory or coding regions of master regulators in generating morphological novelties.

Introduction

Fruits are a major evolutionary innovation of flowering plants, essential for reproductive success and adaptation. Morphological and functional diversity in fruits found in nature is impressive and often allows for offspring dispersal, in the form of adaptations of fruits as vectors for seed dissemination. Thus, fruit morphology and associated dispersal strategies are evolutionarily labile characters of ecological importance under strong selective pressure (Bremer and Eriksson, 1992; Mummenhoff et al., 1997; Clausing et al., 2000), suggesting that modest genetic changes in key genes could be responsible for large phenotypic changes affecting fruit function observed within closely related taxa.

Most of the knowledge on the genetic networks that control fruit patterning has been obtained in *Arabidopsis thaliana*, a species with a dry dehiscent fruit formed from a bicarpellate pistil that opens through dehiscence zones formed at the valve margins and releases seeds upon maturity (Scutt et al., 2006; Ferrándiz et al., 2010). In *Arabidopsis*, the MADS-box genes *SHATTERPROOF1* and *2* (*SHP1*, *SHP2*) and the bHLH gene *INDEHISCENT* (*IND*) are the major regulators directing dehiscence zone formation, which includes the lignification of a precise subset of cells along the dehiscence zone, the lignified layer, and also the differentiation of an adjacent domain of cells with restricted growth, the separation layer. In *shp1 shp2* or *ind* mutants, the dehiscence zone does not form, and therefore neither the lignified nor the separation layers develop at the valve margins (Liljegren et al., 2000; Liljegren et al., 2004). *FRUITFULL* (*FUL*), another MADS-box gene, is required for ovary growth and expansion by restricting the expression domain of *SHP* and *IND* to the valve margins, and thus, in *ful* mutants,

misregulation of SHP/IND causes extensive ectopic lignification in the valves (Gu et al., 1998; Ferrándiz et al., 2000; Liljegren et al., 2004).

In spite of our increasing knowledge of genetic networks directing fruit patterning in *Arabidopsis*, little is known about the conservation of these networks in distantly related species and virtually nothing on how variations in these networks account for fruit morphological evolution in other higher plant taxa. Legumes are the third largest family of angiosperms, comprising more than 19,000 species and second only to grasses in agricultural importance (Lewis et al., 2005). The fruit (a legume or more generally a pod) is the defining characteristic of this group and typically consists of a dry dehiscent pod similar to the *Arabidopsis* silique, but is only derived from a monocarpellate pistil (Polhill, 1994). While the dry dehiscent pod is the most common fruit morphology in the legume family, fruit morphological innovations are also found (Lewis et al., 2005). Especially interesting are the fruits found in the *Medicago* genus. *Medicago* comprises about 87 species and is phylogenetically very close to economically important crop legumes such as pea, faba beans and lentils (Small and Jomphe, 1989; Small, 2011). Many of the species in the genus show a striking example of a morphological novelty, forming spiral pods, frequently spiny, that likely have an adaptive value. *Medicago* coiled pods represent a new mechanism of collective seed dispersal that upon development of spines can be more adapted to epizoochory, which allows long-range dissemination of the spiny ball-shaped fruits that attach to passing animals. In addition, some authors have also associated *Medicago* coiled pod morphology to increased pest resistance by reducing accesibility to insects (Small and Brookes, 1984).

In this work, we have addressed the molecular basis underlying the acquisition of coiled fruit morphology in *Medicago* by studying morphological traits associated with gene functions previously characterized for their role in fruit patterning in other angiosperm

model systems. Our study correlates coiled pod morphology in *Medicago* species with strong lignification at valve margins. Moreover, we show that the likely cause of this altered lignification pattern is a change in the protein coding sequence of orthologs of the *SHP* MADS-box gene, whose conserved function in fruit lignin deposition has already been demonstrated in distantly related angiosperm species (Liljegren et al., 2000; Tani et al., 2007; Fourquin and Ferrándiz, 2012).

Results

Monophyletic origin of coiled pod morphology in the *Medicago* genus.

We chose for our study 17 annual species of the *Medicago* genus that have a diploid genome and that flowered and formed fertile fruits under our greenhouse conditions. Fig. 1 shows the phylogenetic relationships among these 17 species on a tree based on the plastid *matK* gene sequence distances (Steele et al., 2010), which is in good agreement with other phylogenetic analyses of the genus (Bena, 2001; Maureira-Butler et al., 2008; Yoder et al., 2013). In the tree, species with coiled fruits, defined by pods that complete at least a 360° turn, and uncoiled fruits, defined by straight or slightly curved pods, never completing a full 360° turn, are indicated (for pictures of fruits see Fig. S1 and Fig. 2). Species with coiled fruits grouped together, suggesting a common evolutionary origin within the genus for this novel trait, which would be derived from the ancestral uncoiled condition.

Coiled pod morphology in *Medicago* genus correlates with strong lignification of the carpel margins.

Medicago fruits, like the vast majority of legumes, are derived from a monocarpellate pistil. By anthesis the carpel is fused along the margin forming a tubular structure with marginal placentation, which becomes the pod after fertilization (Benlloch et al., 2002;

Wang and Grusak, 2005). To identify possible mechanisms underlying fruit coiling, we compared pod growth patterns and morphology in coiled and uncoiled fruits from four different species, *M. truncatula, M. noeana, M. ruthenica* and *M. polyceratia* (Fig. 2A-T). At anthesis, when fertilization takes place and fruit development starts, ovary morphology looked strikingly similar in all four species (Fig 2B,G,L,Q). However, differences in ovary growth patterns were evident from early post-anthesis stages. In uncoiled pod species, the ovary grew mainly through symmetrical elongation in the longitudinal axis and in width, followed by expansion in pod diameter, similarly to what has been described for other legume species such as pea, soybean or *Lotus japonicus* (Ozga et al., 2002; Nautrup-Pedersen et al., 2010). Mature pods in these species had a similar shape to the anthesis ovaries, only larger (Fig. 2K,M,P,R). In contrast, for species with coiled pods, ovary elongation proceeded asymmetrically. The fused valve margins showed restricted elongation from early stages of fruit development while the rest of the ovary grew at a much higher rate, resulting in the formation of a spiral structure where the fused valve margins remained at interior positions (Fig. 2A,C,F,H).

Whole mount phloroglucinol staining of fruits from 12 of the species included in this study suggested differences in lignification patterns (Fig. 2D,I,N,S and Fig.S1). Transverse sections of elongated fruits from four species stained with toluidine blue to reveal ovary tissue organization confirmed that differences were observed in lignin deposition patterns between coiled and uncoiled pods. In coiled pods, the valve margins showed extensive lignification with large patches of lignified tissue on the interior of the spiral structure, while a small domain of lignification associated with the medial vascular bundle was observed on the exterior side of the spiral (Fig. 2E,J). In uncoiled pods, the lignified tissues were much less evident both at the medial bundle and the valve margins (Fig. 2O,T). In all cases, a positive correlation was found

between strong lignification at valve margins and pod coiling. Likewise, uncoiled pods showed weak symmetrical lignification at medial and marginal positions.

Are different pod morphologies defined by changes in the regulatory networks directing dehiscence zone formation?

In *Arabidopsis*, the FUL/SHP/IND pathway controls critical aspects of dehiscence zone formation and lignification of the fruit (Ferrándiz, 2002; Balanzá et al., 2006). In *shp1 shp2* or *ind* mutants, valve margins are not lignified and the typical small cells of the separation layer adjacent to the lignified patches are missing (Liljegren et al., 2000; Liljegren et al., 2004). Conversely, in *ful* mutants, SHP and IND expression expands to the valves, causing ectopic lignification and restricted cell elongation in the valves (Ferrándiz et al., 2000) (Fig. 3A-C). Additional evidence obtained through functional studies in other distantly related species indicate that the role of the FUL/SHP network in fruit dehiscence and lignification is widely conserved at least in dicot plants (Smykal et al., 2007; Fourquin and Ferrándiz, 2012; Pabon-Mora et al., 2012).

Because the *FUL/SHP/IND* genes are at the top of the genetic hierarchy controlling both lignification and cell differentiation and expansion in the fruit, we hypothesized that changes in this pathway could lead to the morphological variation of pod shape and lignification patterns occurring in the *Medicago* genus. This idea was supported by the phenotypes of the *ful-3* mutant line. This line bears an insertion of an *Spm*-transposon element in the *FUL* gene region that causes strong loss-of-function *ful* phenotypes (Ferrándiz et al., 2000). Spontaneous somatic excision of the *Spm*-element restores FUL activity, and in the *ful-3* line we occasionally observed fruits where *FUL* wildtype sectors underwent cell growth and expansion while cells in *ful* mutant sectors remained small and lignified, resulting in differential growth rates of FUL^{+} and ful^{-} sectors in the

ovary. These different patterns of cell elongation and lignification within the ovary resulted in coiling of the silique (Fig 3 D-F).

Given the *ful-3* phenotype, we hypothesized that changes in expression patterns or protein activities in the *Medicago* orthologs of *FUL, SHP* or *IND* could cause changes in ovary growth and lignification patterns in coiled pod species. Thus, for example, changes in the expression patterns could lead to different patterns of lignin deposition and cell expansion. Alternatively, a change in protein activity, for example a higher SHP or IND activity leading to a stronger activation of the lignin biosynthetic pathway, could also result in increased lignification of valve margins if these genes were only expressed at carpel/valve margins.

The expression patterns of SHP orthologs are similar in *Medicago* species with coiled and uncoiled pod morphology.

To test these hypotheses, we searched for orthologs of *FUL*, *SHP* and *IND* in the fully sequenced *Medicago truncatula* genome. We found two *FUL* genes, *MtFULa* and *MtFULb* (Berbel et al., 2012), while only one *SHP* gene, that we named *MtruSHP*. No clear ortholog was found for *IND*, in agreement with the fact that *IND* is a paralogue of *HECATE3 (HEC3)* that is only present in the Brassicaceae (Kay et al., 2012). Therefore, we focused our study on *Medicago* orthologs of *FUL* and *SHP*. First, we cloned these genes in most of the species included in our study (for accession numbers see Table S1). To check whether there was a correlation between the expression pattern of these genes and pod morphology, we performed a qRT-PCR experiment for the three genes in six species, three having uncoiled pods and three with coiled pods, on young fruits at the time of anthesis, when morphological differences are just beginning to originate. No correlation was found between levels of expression and pod shape in these two groups (Fig 4A), suggesting that it was unlikely that differences in

pod coiling were based on differences in expression level of these genes. However, since in *Arabidopsis* the *SHP* expression domain at anthesis marks the subsequent differentiation of the dehiscence zone during fruit development, it still could be possible that more subtle changes in the spatial pattern of *SHP* genes (either caused by different elements in *SHP* regulatory sequences or by gain/loss of activity of FUL, a *SHP* negative regulator) had a significant effect in the extension of the lignified patches at carpel margins. Hence, we studied the expression pattern of *SHP* orthologs in eight species, four with uncoiled pods and four with coiled pods. In anthesis fruits of *M. truncatula*, a coiled pod species, *SHP* was strongly expressed in ovules, at a lower level in the inner epidermal layer of the ovary, and in the carpel margins (Fig. 4B-C), an expression pattern resembling that of the Arabidopsis *SHP* genes (Ferrándiz et al., 2000). Highly similar patterns of *SHP* expression were observed for *M. orbicularis, M. littoralis* and *M. tenoreana*, coiled pod species, and *M. polyceratia, M. platycarpa, M. monspeliaca* and *M. ruthenica*, uncoiled pod species (Fig 4D-K). These results suggested that differences in cell growth and lignification patterns of coiled and uncoiled species were not likely caused by differences in regulatory elements of *SHP* genes or in the function of *SHP* upstream regulators like FUL. A second important conclusion from these experiments was that *SHP* expression in the ovary of both coiled and uncoiled pods was not symmetrical, but higher in the valve margins, the domain that is extensively lignified in coiled pods.

A sequence polymorphism in SHP orthologs correlates with coiled/uncoiled pod morphology.

Our results led us to investigate the alternative hypothesis, that differences in protein activity resulted in differences in the extent of lignified tissues associated with valve

margins. For this reason, we compared the deduced protein sequences of FUL and SHP from different *Medicago* species. For *FULa* and *FULb* orthologs, no correlation was found between any sequence polymorphism and fruit shape (Fig S2). However, for *SHP* orthologs, a sequence signature was found that perfectly correlated with coiled and uncoiled pod morphology in 17 species (Fig. 5A and Fig. S3). Moreover, we found the uncoiled signature in other legume SHP orthologs found in public databases (pea, soybean, *Lotus japonicus*), all of which are from species with similar uncoiled pods (Fig 5A).

If the correlation between the sequence variation and the coiled/uncoiled morphology had functional relevance, we would expect that the corresponding sequence polymorphism would be under selective pressure. In order to analyze the selective pressure on a given site, the ratio of replacement and silent mutations in this site is normalized to the number of replacement and silent sites in the gene (dN/dS), where the number of replacement mutations per replacement site is called dN, and the number of silent mutations per silent site is called dS (Lemey et al., 2009). Through a sequence alignment, we compared SHP coding regions from species possessing coiled fruits against those with uncoiled fruits from *Medicago* species and other legume plants in various codon positions. Each codon position was tested for the selective pressure acting on that site, using all homologues from coiled and uncoiled pod species. The evolutionary fingerprint mean observed for all codons in the ortholog SHP genes indicates a negative selection (Fig. S4). However, using the SLAC method (Single Likelihood Ancestor Counting) available in the HyPhy Package (Pond et al., 2005), we found evidence for diversifying positive selection, where the non-synonymous changes outweigh the synonymous changes (dN/dS>1), in the coiled/uncoiled signature sequence found in the SHP gene (ACA-to-GCA, or threonine-to-alanine in coiled pod

species, Fig. 5A,B,C; Table 1). These results suggested that the codon transition was under positive Darwinian selection. The selection rates for the ACA-GCA transition, measured for the three possible conditions, i.e., coiled, uncoiled and all species, are shown on Table 2. The results indicate that the evolutionary rates at ACA codon are higher than at the GCA, or that natural/domestication forces are higher towards the formation of coiled phenotypes, being more conservative on the genotypes presenting the GCA codon.

The two sequence variants of SHP have different protein activities.

To test whether this change in sequence could have an effect on protein activity, we generated two constructs expressing the coding sequence of the *SHP* ortholog from a coiled (*M. truncatula*) and an uncoiled pod species (*M. polyceratia*) under the control of the constitutive CaMV 35S promoter (Benfey and Chua, 1989). We transformed *Arabidopsis* wildtype plants with these constructs, recovered more than 150 independent primary transgenic lines (T1) for each and compared the phenotypes of the individual plants. Both transgenes caused a range of phenotypes in T1 lines related to those observed in *Arabidopsis* lines that overexpressed the endogenous *AtSHP* genes, with different degrees of leaves curled upwards, early flowering, conversion of sepals into carpelloid structures and shorter wrinkled fruits (Fig. 6A) (Liljegren et al., 2000). However, we were able to observe significant differences in the relative effects of both constructs: overexpression of *MpolSHP*, the *SHP* gene from *M. polyceratia* (uncoiled pods), only produced short wrinkled fruit phenotypes in 25% of the T1 lines, while the overexpression of *MtruSHP*, the *SHP* gene from *M. truncatula* (coiled pods) affected a significantly higher proportion of lines, 41% (Fig. 6B). Moreover, the strongest phenotypes appeared in 11% of the transgenic lines overexpressing the gene

from the coiled pod species (*MtruSHP*) while only in 4% of the lines overexpressing the gene from the uncoiled pod species (*MpolSHP*), therefore indicating that the MtruSHP protein was more active than MpolSHP (Fig. 6B).

Because protein-protein interactions among MADS-box proteins are crucial for their function (Honma and Goto, 2001), we speculated that the sequence differences between the two SHP variants from coiled and uncoiled pod species could affect the range of these interactions. MADS-box proteins from the AG/SHP subfamily have been shown to interact with SEPALLATA (SEP) MADS-box proteins to direct carpel development (de Folter et al., 2005; Immink et al., 2009). We used yeast two-hybrid experiments to test the ability of MtruSHP and MpolSHP to interact with SEP orthologs from the corresponding species. We found both MtruSHP and MpolSHP interacted with their corresponding SEP1 and SEP3 orthologs, but not with SEP4, in agreement with previously reported interactions for AG/SHP factors from other species (de Folter et al., 2005; Immink et al., 2009; Airoldi et al., 2010). However, we also found that the SHP-SEP3 dimers interacted with different affinities: the dimer from the coiled pod species *M.truncatula* interacting more strongly than the dimer from the uncoiled pod species *M. polyceratia*. The interaction of the SHP-SEP3 dimer from *M. truncatula* was maintained even when 80 mM 3AT was added to the growth medium, while the SHP-SEP3 dimer from *M. polyceratia* was not able to form at 3AT concentrations higher than 40 mM (Fig 6C).

In summary, these results suggested that differential SHP activity could be at the origin of coiled fruit morphology and that this differential activity did not seem to be based on differential expression, but on different protein functional properties.

Discussion

Our work shows a tight correlation between pod coiled morphology and increased valve margin lignification in *Medicago* fruits. The lignified areas in coiled pods fail to elongate during fruit development, similarly to what is observed in partially revertant fruits of *ful*-3 mutants that also adopt a coiled morphology. We also show that *SHP* genes in *Medicago* fruits are expressed in the valve margins, the domains that lignify extensively in coiled pod species, and we have found a correlation of coiled pod morphologies with a sequence polymorphism under positive Darwinian selection in SHP sequences, strongly suggesting that the modification of SHP protein underlies the evolutionary origin of pod coiling in *Medicago*.

Several pieces of evidence support this hypothesis. First, there is a strong correlation between pod coiling, extent of valve margin lignification and a functionally relevant polymorphism in the SHP sequence. The role of *SHP* genes in the control of lignin deposition during late fruit development has been described for other distantly related dicot species. Thus, in *Arabidopsis* or *Nicotiana benthamiana*, SHP directs lignification of dehiscence zones (Liljegren et al., 2000; Fourquin and Ferrándiz, 2012) and in peach, increased levels of *PpSHP* are found in varieties prone to split-pit formation and overlignification of the fruit (Tani et al., 2007). These conserved roles across dicot species support the idea that modifications in SHP sequence, causing a change in protein activity, would be likely to induce changes in lignification patterns in valve margins such as those observed in the coiled pods of *Medicago* species. Second, we have proven that the variation in SHP sequence observed in coiled pod species causes a change in SHP functional properties, as seen in the relative effects of *SHP* overexpression in *Arabidopsis* and in protein interaction studies. A naturally occurring variation in a single amino acid of the related PLENA (PLE) and FARINELLI (FAR) MADS-box proteins from *Antirrhinum majus* has been shown to have profound effects on their ability to specify male and female reproductive organs by altering its range of

protein–protein interactions with SEP homologs (Airoldi et al., 2010). Thus, we have uncovered a similar scenario, where a small sequence modification in SHP amino acid sequence increases the SHP affinity for SEP3 protein partners that could lead to hyperactivation of the downstream targets. Finally, positive selective pressure in the ACA-GCA transition indicates a selective sweep fixing the new allele. Interestingly, the remaining codons are under a negative selection, indicating a strong conserving force that suggests an essential role of SHP for plant survival. Coiled pod morphology in the *Medicago* genus appears to be evolutionarily fixed, suggesting an adaptive value of this novel trait. In fact, pod coiling in *Medicago* has been associated with an increase in reproductive success due to enhanced resistance to certain phytophagous insects, which encounter increased physical constraints to feed on the seeds of coiled pods (Small and Brookes, 1984). Moreover, the coiled pod morphology represents a change in seed dispersal strategies with a likely impact on population structures and fitness (Heyn, 1963; Lesins and Lesins, 1979; Polhill, 1994; Levin et al., 2003; Yan et al., 2009). The possible evolutionary advantages of the novel pod morphology in *Medicago* species could explain the low variability in the coiled-SHP sequence signature and supports the hypothesis of this signature as the origin of the trait.

We are aware that additional functional data would be desirable for the validation of our hypothesis. For this purpose, we identified an insertion in the *MtruSHP* gene by reverse genetics in the collection of lines tagged with the transposable element of *Nicotiana tabacum* cell type1 (Tnt1) at the Noble Foundation (Tadege et al., 2008), but the original primary line from which genomic DNA was isolated did not produce seeds, and therefore the corresponding mutants were lost. Likewise, when we generated transgenic lines harboring a hairpin RNAi construct designed to silence *MtruSHP* we observed unusually reduced transformation efficiency and, in the small number of lines

recovered, almost no reduction in *MtruSHP* levels was obtained. Thus, these experiments did not help to characterize the role of *SHP* in pod development, but might suggest an essential role of *SHP* for reproductive success in *Medicago*, as already indicated by the analyses on sequence evolution that we have performed.

Our work on *Medicago* SHP orthologs provides an example of a modest change in the protein sequence of a key regulatory factor as the likely origin of a developmental innovation, in this case a dramatic change in fruit morphology likely affecting fitness and seed dispersal strategies. A widely accepted postulate of evolutionary developmental biology science (evo-devo) is that many developmental innovations affecting morphology are more likely linked to changes in the *cis*-regulatory regions than in the protein-coding regions of key developmental genes. This idea is mainly based on the modular nature of *cis*-regulatory elements, which allows local modifications of the activity of key regulators of developmental pathways without other deleterious pleiotropic effects (Carroll, 2000; Carroll, 2008). However, there is a lively debate on the relative importance of changes in protein activity to create new developmental functions, which some authors consider also to be highly significant (Hoekstra and Coyne, 2007; Haag and Lenski, 2011). In addition to our study and the PLE/FAR case already mentioned, we can find other examples of amino acid substitutions underlying the evolution of novel traits, for instance, an asparagine to lysine substitution in the coding sequence of the *SH4* gene in rice causing a reduction in seed shattering that was selected by humans in the process of rice domestication (Li et al., 2006), a lysine to glutamate substitution in the ENHANCER OF TRY AND CPC 2 factor, which determines trichome patterning in natural *Arabidopsis* populations (Hilscher et al., 2009) or a lysine to a asparagine substitution in the *tga1* gene of maize that appears to be responsible for the origin of the naked grains selected in domesticated maize (Wang et

al., 2005). Taken together, these studies highlight the importance of modified protein activity acting as a driving force in evolution. This expands the general postulates of evo-devo which state that form evolves mainly through *cis*-regulatory changes that alter expression patterns of functionally conserved proteins with highly pleiotropic functions.

Methods

Plant material and growth conditions *Medicago* plants were grown in the greenhouse, at 22°C (day) and 18°C (night) with a 16-h light/8-h dark photoperiod, in soil irrigated with Hoagland no. 1 solution supplemented with oligoelements (Hewitt, 1966). Arabidopsis plants were grown in cabinets at 21ºC under long day conditions (16h light), illuminated by cool-white fluorescent lamps (150µE m^{-2} s^{-1}) in a 1:1:1 mixture of sphagnum:perlite:vermiculite.

For *Arabidopsis* transformation, *MtruSHP* and *MpolSHP* CDS were amplified with primers Sal1MSHPFor and BamH1MSHPRev and cloned in the pBIN-JIT vector (Ferrándiz et al., 2000). Each vector was introduced into *Agrobacterium* PMP90 for *Arabidopsis* transformation using the floral dip protocol (Clough and Bent, 1998). T1 plants were selected based on Kanamycin selection.

Phenotypic categories for 35S::MtruSHP/MpolSHP fruits were scored as follows. Silique length was measured for 10 fully developed fruits in the main inflorescence of 5 WT plants and each of the T1 plants obtained for each construct showing morphological alterations such as modified sepals or curled leaves. Average silique length of each transgenic plant was compared to the WT value and categorized as “strong” if it was less than 75% of WT and the ovary showed an evident bumpy or wrinkled surface, while “intermediate” phenotypes corresponded to plants where average silique length was between 75 and 95% of WT and ovary surface was not smooth.

Cloning and sequence analysis. All sequences were isolated by RT-PCR on cDNA of young flowers. The full-length coding sequence of *MtruSHP* gene was isolated using primers PeaSHPFor and PeaSHPRev designed from the sequence of the *Pisum sativum SHP* gene (*PM8*, AY884292). From the identified *MtruSHP* gene, the primers MSHPFor and MSHPRev were designed and used to isolate the *SHP* genes from the rest of the Medicago species. The primers MFULaFor/MFULaRev and MFULbFor/MFULbRev were designed from the *M. truncatula FUL* available sequences (*MtFULa*, TC84496 and *MtFULb*, TC82227) and used to amplify the *FULa* and *FULb* genes from the different *Medicago* species. The deduced amino acid sequences alignments were analyzed using the Macvector 12.5 software. See table S2 for primers sequences and table S1 for the accession numbers of the identified sequences.

For the analyses on patterns of sequence polymorphism, an evolutionary fingerprint analysis was made to infer about the main selection pressure in the overall codon sites within the SHP orthologs from *Medicago* and other legume species. To verify the putative selection, the transitional ACA-GCA site was analyzed individually. The analyses were run using the Fast Unbiased AppRoximate Bayesian approach (FUBAR) for selection pressure evidences. To verify the existence of mixed effects in the selected codon site a MEME analysis was carried out. All analyses were made using the HyPhy package (Pond et al., 2005) in situ or the webserver available www.datamonkey.org.

In situ hybridization. RNA *in situ* hybridization with digoxigenin-labeled probes was performed on 8-μm paraffin sections of *Medicago* flowers as described (Ferrándiz et al., 2000). The RNA antisense and sense probes were generated from a 478-bp fragment of the *MtruSHP* cDNA (positions 258 to 735 from start codon). The same probe was

used for the 8 *Medicago* species studied as the sequences present more than 95% of identity.

Quantitative RT-PCR. Total RNA was extracted from flowers in anthesis from the different *Medicago* species with the RNeasy Plant Mini kit (Qiagen). 4 µg of total RNA were used for cDNA synthesis with the First-Strand cDNA Synthesis kit (Invitrogen) and the qPCR master mix was prepared using the iQTM SYBR Green Supermix (Bio-rad). Results were normalized to the expression of the *M. truncatula OGT* mRNA (XM_003590858) amplified with primers qMConstiFor/qMConstiRev. The primers used to amplify the *SHP*, *FULa* and *FULb* genes of *Medicago* generated products of 51 bp and did not show any cross-amplification. The efficiency in the amplification of *MedicagoSHP*, *MedicagoFULa*, *MedicagoFULb* and the reference gene was similar. The PCR reactions were run and analyzed using the ABI PRISM 7700 Sequence detection system (Applied Biosystems). See Table S2 for primer sequences.

Scanning electron microscopy (SEM). Samples were vacuum infiltrated with FAE (3,7% formaldehyde, 5% acetic acid, 50% ethanol) for 10 min and fixed with fresh solution for 16h at 4ºC. Samples were dehydrated in an ethanol series and critical point dried in liquid CO2 (Polaron E300 apparatus). Dried samples were mounted on stubs and coated with gold palladium (4:1) in a Sputter Coater SCD005 (BALTEC). Scanning electron microscopy was performed with a JEOL JSM-5410 microcope (10 kV).

Lignin staining. Fruits were fixed in FAE overnight and then embedded into paraffin. 12-µm sections were stained in a 0,2% toluidine blue solution for 2 min and then washed in water. Alternatively, sections were stained in phloroglucinol 2,5% for 30 min, and then soaked 30 sec in HCl 50% before being photographed under the microscope. For

whole-mount lignin observation, fruits were fixed in FAE overnight, stained for 30 min in phloroglucinol 2,5% and soaked 30 sec in HCl 50%.

Yeast Two-Hybrid Assays (Y2H). The Y2H assays were performed in the yeast strain PJ69-4α using the vectors pGADT7 (pAD) and pGBKT7 (pBD) (kindly provided by M. Colombo and B. Causier). The coding sequences were cloned using Topo cloning and Gateway recombination (Invitrogen). *MtruSHP* and *MpolSHP* were cloned into the pBD vector using primers MSHPFor/MSHPRev. *MtruSEP1*, *MtruSEP3*, *MpolSEP1* and *MpolSEP3* were cloned into the pAD vector using primers MSEP1For/MSEP1Rev and MSEP3For/MSEPRev. Two-hybrid interactions were assayed on selective yeast synthetic dropout medium lacking Leu, Trp, and His, supplemented with different concentrations of 3-aminotriazole (3-AT; 40, 60 and 80 mM). Selections were performed at 28°C. The experiments were repeated four times, and growth of the colonies on control and selective plates was assessed.

Acknowledgements

We thank Rafael Martínez-Pardo and Eugenio Grau for technical support, and Barbara Ambrose (NYBG, USA), Mario Fares and Francisco Madueño (IBMCP, Spain) for helpful discussions and critical reading of the manuscript. Germplasm used in this study was obtained from the National Genetic Resources Program (USA).

Literature Cited

Airoldi CA, Bergonzi S, Davies B (2010) Single amino acid change alters the ability to specify male or female organ identity. Proc Natl Acad Sci USA **107:** 18898-18902

Balanzá V, Navarrete M, Trigueros M, Ferrándiz C (2006) Patterning the female side of Arabidopsis: the importance of hormones. J Exp Bot **57:** 3457-3469

Bena G (2001) Molecular phylogeny supports the morphologically based taxonomic transfer of the "medicagoid" Trigonella species to the genus Medicago L. Plant Syst. Evol **229:** 217-236

Benfey PN, Chua N-H (1989) The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. EMBO J. **8:** 2195-2202

Benlloch R, Navarro C, Beltran JP, Canas LA (2002) Floral development of the model legume Medicago truncatula: ontogeny studies as a tool to better characterize homeotic mutations. Sex Plant Reprod **15:** 231-241

Berbel A, Ferrandiz C, Hecht V, Dalmais M, Lund OS, Sussmilch FC, Taylor SA, Bendahmane A, Ellis TH, Beltran JP, Weller JL, Madueno F (2012) VEGETATIVE1 is essential for development of the compound inflorescence in pea. Nature Commun **3:** 797

Bremer B, Eriksson O (1992) Evolution of fruit characters and dispersal modes in the tropical family Rubiaceae. Biol J Linn Soc **47:** 75-95

Carroll SB (2000) Endless forms: the evolution of gene regulation and morphological diversity. Cell **101:** 577-580

Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell **134:** 25-36

Clausing G, Meyer K, Renner S (2000) Correlations among fruit traits and evolution of different fruits within Melastomataceae. Bot J Linn Soc **133:** 303-326

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J **16:** 735-743

de Folter S, Immink RGH, Kieffer M, Parenicová L, Henz SR, Weigel D, Busscher M, Kooiker M, Colombo L, Kater MM, Davies B, Angenent GC (2005) Comprehensive interaction map of the Arabidopsis MADS Box transcription factors. Plant Cell **17:** 1424-1433

Ferrándiz C (2002) Regulation of fruit dehiscence in Arabidopsis. J Exp Bot **53:** 2031-2038

Ferrándiz C, Fourquin C, Prunet N, Scutt C, Sundberg E, Trehin C, Vialette-Guiraud A (2010) Carpel Development. Adv Bot Res **55:** 1-74

Ferrándiz C, Gu Q, Martienssen R, Yanofsky MF (2000) Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. Development **127:** 725-734

Ferrándiz C, Liljegren SJ, Yanofsky MF (2000) Negative regulation of the SHATTERPROOF genes by FRUITFULL during Arabidopsis fruit development. Science **289:** 436-438

Fourquin C, Ferrandiz C (2012) Functional analyses of AGAMOUS family members in Nicotiana benthamiana clarify the evolution of early and late roles of C-function genes in eudicots. Plant J**:** 990-1001

Gu Q, Ferrandiz C, Yanofsky MF, Martienssen R (1998) The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development. Development **125:** 1509-1517

Haag ES, Lenski RE (2011) L'enfant terrible at 30: the maturation of evolutionary developmental biology. Development **138:** 2633-2637

Hewitt Y (1966) Sand and Water Culture Methods used in theStudy of Plant Nutrition, Ed 2nd ed. Farnham: Commonwealth Agricultural Bureau.

Heyn C (1963) The annual species of Medicago. Magnes Press, The Hebrew University, Jerusalem

Hilscher J, Schlötterer C, Hauser M-T (2009) A single amino acid replacement in ETC2 shapes trichome patterning in natural Arabidopsis populations. Curr Biol **19:** 1747-1751

Hoekstra HE, Coyne JA (2007) The locus of evolution: evo devo and the genetics of adaptation. Evolution **61:** 995-1016

Honma T, Goto K (2001) Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature **409:** 525-529

Immink RG, Tonaco IA, de Folter S, Shchennikova A, van Dijk AD, Busscher-Lange J, Borst JW, Angenent GC (2009) SEPALLATA3: the 'glue' for MADS box transcription factor complex formation. Genome Biol **10:** R24

Kay P, Groszmann M, Ross JJ, Parish RW, Swain SM (2012) Modifications of a conserved regulatory network involving INDEHISCENT controls multiple aspects of reproductive tissue development in Arabidopsis. New Phytol **197:** 73-87

Lemey P, Salemi M, Vandamme AM (2009) The Phylogenetic Handbook. A Practical Approach to Phylogenetic Analysis and Hypothesis Testing, Ed 2nd. Cambridge University Press.

Lesins KA, Lesins I (1979) Genus Medicago (Leguminosae): a taxogenetic study. W. Junk., The Hague

Levin SA, Muller-Landau HC, Nathan R, Chave J (2003) The Ecology and Evolution of Seed Dispersal: A Theoretical Perspective. Ann Rev Ecol Evol System **34:** 575-604

Lewis GP, Schrire B, Mackinder B, Lock M (2005) Legumes of the World. Kew Publishing, Richmond, UK

Li C, Zhou A, Sang T (2006) Rice domestication by reducing shattering. Science **311:** 1936-1939

Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF (2000) SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature **404:** 766-770

Liljegren SJ, Roeder AHK, Kempin SA, Gremski K, Østergaard L, Guimil S, Reyes DK, Yanofsky MF (2004) Control of fruit patterning in Arabidopsis by INDEHISCENT. Cell **116:** 843-853

Maureira-Butler IJ, Pfeil BE, Muangprom A, Osborn TC, Doyle JJ (2008) The reticulate history of Medicago (Fabaceae) Syst Biol **57**:466-82.

Mummenhoff K, Franzke A, Koch M (1997) Molecular data reveal convergence in fruit characters used in classification of Thlaspi sl (Brassicaceae). Bot J Linn Soc **125:** 183-199

Nautrup-Pedersen G, Dam S, Laursen BS, Siegumfeldt AL, Nielsen K, Goffard N, Staerfeldt HH, Friis C, Sato S, Tabata S, Lorentzen A, Roepstorff P, Stougaard J (2010) Proteome analysis of pod and seed development in the model legume Lotus japonicus. J Proteome Res **9:** 5715-5726

Ozga JA, van Huizen R, Reinecke DM (2002) Hormone and seed-specific regulation of pea fruit growth. Plant Physiol **128:** 1379-1389

Pabon-Mora N, Ambrose BA, Litt A (2012) Poppy APETALA1/FRUITFULL orthologs control flowering time, branching, perianth identity, and fruit development. Plant Physiol **158:** 1685-1704

Polhill RM (1994) Classification of the Leguminoseae and complete synopsis of legume genera. *In* FA Bisby, J Buckingham, JB Harborne, eds, Phytochemical dictionary of the Leguminoseae, Vol 1. Chapman & Hall, London

Pond SL, Frost SD, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. Bioinformatics **21:** 676-679

Scutt CP, Vinauger-Douard M, Fourquin C, Finet C, Dumas C (2006) An evolutionary perspective on the regulation of carpel development. J Exp Bot **57:** 2143-2152

Small E (2011) Alfalfa and relatives: Evolution and classification of Medicago. NRC Research Press, Otawa, Ontario, Canada

Small E, Brookes B (1984) Coiling of alfalfa pods in relation to resistance against seed chalcids: additional observations. Can J Plant Sci **64:** 659-665

Small E, Jomphe M (1989) A synopsis of the genus Medicago (Leguminosae). Can J. Bot **67:** 3260-3294

Smykal P, Gennen J, De Bodt S, Ranganath V, Melzer S (2007) Flowering of strict photoperiodic Nicotiana varieties in non-inductive conditions by transgenic approaches. Plant Mol Biol **65:** 233-242

Steele KP, Ickert-Bond SM, Zarre S, Wojciechowski MF (2010) Phylogeny and character evolution in Medicago (Leguminosae): Evidence from analyses of plastid trnK/matK and nuclear GA3ox1 sequences. Am J Bot **97:** 1142-1155

Tadege M, Wen J, He J, Tu H, Kwak Y, Eschstruth A, Cayrel A, Endre G, Zhao PX, Chabaud M, Ratet P, Mysore KS (2008) Large-scale insertional mutagenesis

using the Tnt1 retrotransposon in the model legume Medicago truncatula. Plant J **54:** 335-347

Tani E, Polidoros AN, Tsaftaris AS (2007) Characterization and expression analysis of FRUITFULL- and SHATTERPROOF-like genes from peach (Prunus persica) and their role in split-pit formation. Tree Physiol **27:** 649-659

Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Faller M, Bomblies K, Lukens L, Doebley JF (2005) The origin of the naked grains of maize. Nature **436:** 714-719

Wang HL, Grusak MA (2005) Structure and development of Medicago truncatula pod wall and seed coat. Ann Bot **95:** 737-747

Yan J, Chu HJ, Wang HC, Li JQ, Sang T (2009) Population genetic structure of two Medicago species shaped by distinct life form, mating system and seed dispersal. Ann Bot **103:** 825-834

Yoder JB, Briskine R, Mudge J, Farmer A, Paape T, Steele K, Weiblen GD, Bharti AK, Zhou P, May GD, Young ND, Tiffin P (2013) Phylogenetic signal variation in the genomes of Medicago (Fabaceae). Syst Biol **doi:** 10.1093/sysbio/syt009

Figure legends

Figure 1. Simplified species tree of the *Medicago* genus.

The phylogenetic relationships among the species included in this study are shown. The tree has been redrawn based on the phylogeny based on the gene sequence distances of the plastid *matk* genes published in (Steele et al., 2010), eliminating the species not used in our study. Species with coiled pod morphology are shaded in orange and identified with an orange solid circle. Species with uncoiled pod morphology are indicated with green bars. The related genera *Trigonella* and *Melilotus* are used to root the tree.

Figure 2. Comparative morphological study of fruit development in *Medicago* species with coiled and uncoiled pod morphology

(**A)** to **(E**) *M. truncatula* fruits. **(F)** to **(J)**. *M. noeana* fruits. **(K)** to **(O)**. *M. ruthenica* fruits. **(P)** to **(T).** *M. polyceratia* fruits. **A, F, K, P**. Seven developmental stages of fruit development from anthesis to mature fruit. Scale bar = 0,5 cm. **(B, C, G, H, L, M, Q, R)** Scanning electron micrographs of fruits at anthesis (B, G, L and Q) and at stage 3 according to the developmental series in upper panels (C, H, M and R). Scale bar = 1 mm. (**D, I, N, S**) Whole mount phloroglucinol staining of *Medicago* fruits at stage 5. Lignin appears in pink/red. (**E, J, O, T**). Lignification patterns observed in histological sections of *Medicago* fruits at stage 5. The sections have been stained with toluidine blue and the lignin, only evident in coiled pods, appears in turquoise blue (arrow). The orientation of the sections of coiled fruits corresponds to a plane paralel to the background in C or H/ perpendicular to the background in D or I. Scale bar = 0,5 mm.

Figure 3. Arabidopsis phenotypes related to altered lignification patterns.

(A) to (C). Lignification patterns in transverse sections of mature fruits at valve (carpel) margins as revealed by phloroglucinol staining in pink/red. (**A**) Wild type. Arrows mark the lignified patches at valve margins. The inner cell layer of the valve also lignifies. (**B**) *ful* loss-of-function mutant. Arrow marks ectopic lignification in the valve (carpel wall). (**C**) *shp1 shp2* double mutant. Note the absence of lignin at the valve margin (asterisks). (**D) to (F**). Scanning electron micrographs of two fruits from the same *ful-3* mutant plant, which bears an active transposon element inserted in the coding sequence of the *FUL* gene. (**D**) Fruit with typical *ful* phenotype. (**E**) Fruit where the excision of the transposon element has created a *FUL*+ revertant sector, therefore restoring FUL activity and the concomitant repression of *SHP*. The *FUL*+ sector grows normally, while growth in the *ful* sector is restricted by the ectopic expression of genes promoting dehiscence zone formation and lignification (SHP, IND). **(F)** Close up of the revertant sector of the *ful-3* fruit shown in (E). Arrows indicate developing stomata, typical of wild type valve tissue. For comparison of differences in cell expansion, one cell of the revertant sector has been shaded in green, while one cell of the mutant sector is colored in pink. Scale bar: 1 mm.

Figure 4. Expression analyses of *FUL* and *SHP* genes in *Medicago* species with different pod morphology.

(**A**) Expression levels of *FULa*, *FULb* and *SHP* in anthesis fruits of three *Medicago* species with coiled pod morphology (*M. noeana, M. truncatula* and *M. littoralis*) and of three *Medicago* species with uncoiled pod morphology (*M. polyceratia, M. ruthenica* and *M. monspeliaca*). No correlation of pod shape and overall levels of expression are observed for any of the genes.

(**B**) to **(K)**. Expression patterns of *MedicagoSHP* in anthesis flowers of *Medicago* species with coiled (B to F) or uncoiled (G to K) pod morphologies. (**B**) Longitudinal section of a *M. truncatula* flower. Expression is mainly detected in ovules (ov) and the inner epidermal layer of the carpel (arrow). (**C**) Transversal section of a *M. truncatula* fruit. Expression is observed in ovules (ov), inner epidermal layer and at the carpel margin (arrow). (**D** to **F**). Transversal sections equivalent to that shown in panel C for *M. orbicularis (*D), *M. littoralis* (E), and *M. tenoreana* (F), all showing very similar expression patterns. (**G**) Longitudinal section of a *M. polyceratia* flower. Expression is detected in ovules (ov) and the inner epidermal layer of the carpel (arrow). (**H**) Transversal section of a *M. polyceratia* fruit. Again, expression is observed in ovules, inner epidermal layer and at the carpel margin (arrow). (**I** to **K**) Transversal sections equivalent to that shown in panels (C) or (H) for *M. platycarpa* (I), *M. monspeliaca* (J), and *M. ruthenica* (K), all showing very similar expression patterns. Scale bar = 50 μm.

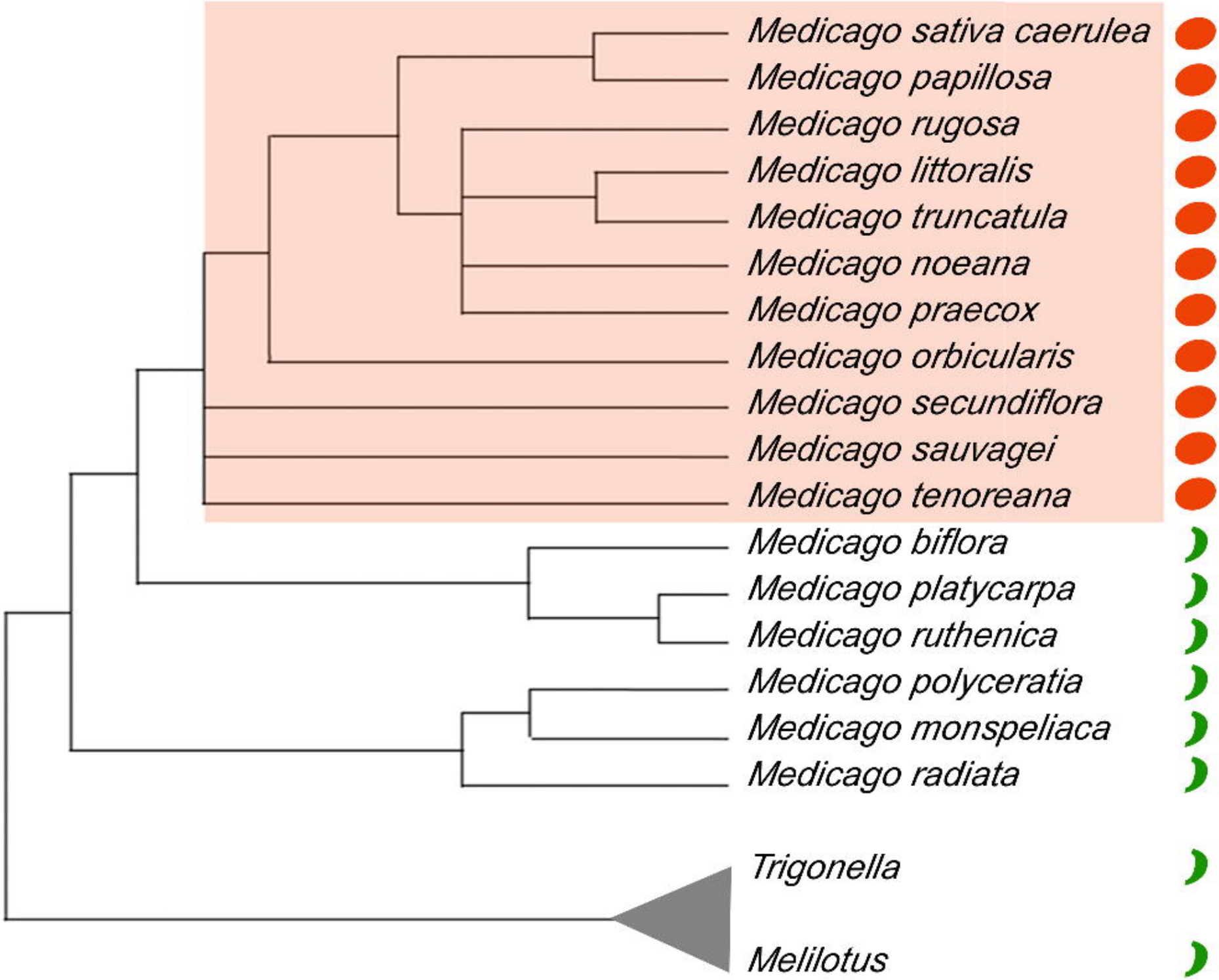
Figure 5. Sequence analyses on *SHP* genes from *Medicago* species with coiled or uncoiled pod morphology.

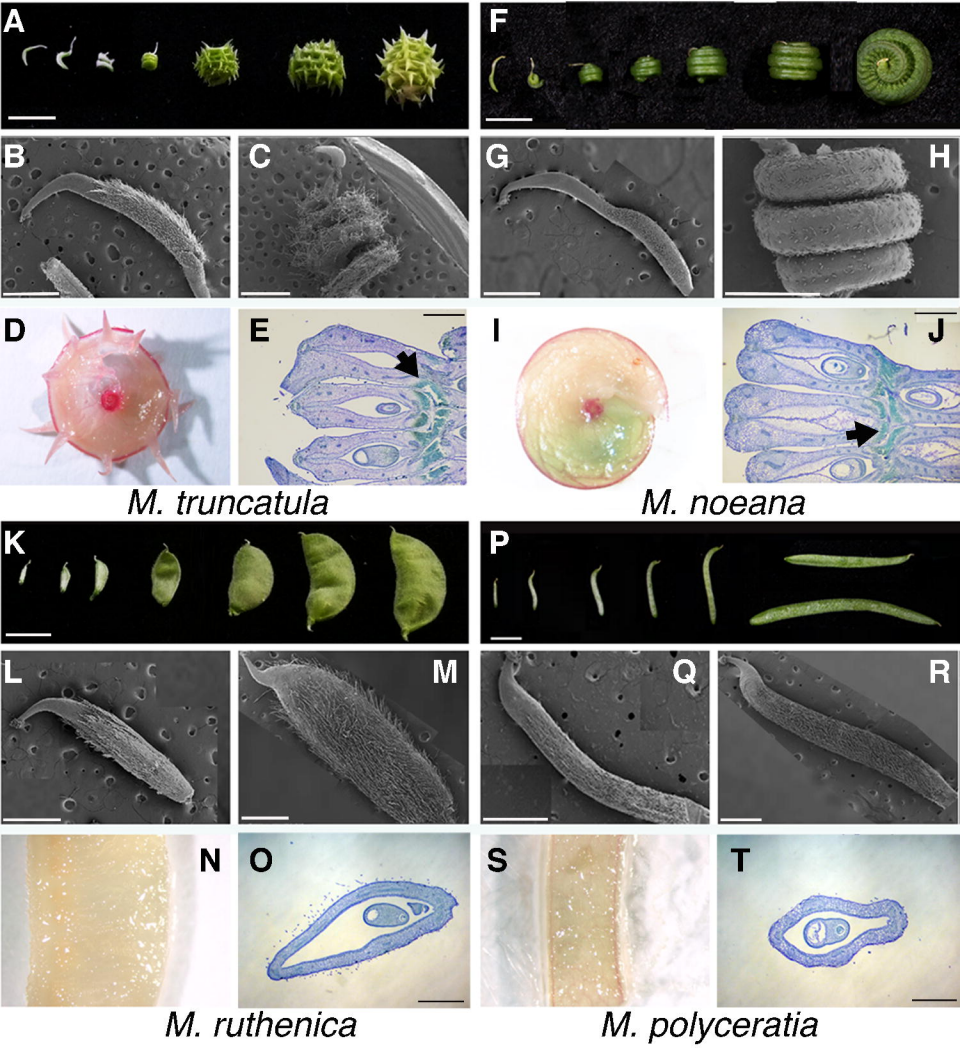
(**A**) Alignment of a fragment of SHP protein sequences from *Medicago* coiled (orange full circles) and uncoiled pod species (green bars). Sequences from SHP proteins from other legumes species with uncoiled pods (*Pisum sativum,* acc. number AY884292*; Lotus japonicus,* BT135371; *Glycine max,* XM_003551863) have also been added. The polymorphism is framed in red and the aminoacid corresponding to the transition codon used for the evolutionary fingerprint analyses is marked by a red asterisk. (**B**) Alignment of the corresponding nucleotide sequences of *SHP* genes around the polymorphism analyzed. The transition codon is shaded in orange for species with coiled fruits and in green for species with uncoiled fruits. (**C**) Selection pressure evidences in the transition ACA-GCA site in *Medicago SHATTERPROOF* orthologs suggested by the Bayesian

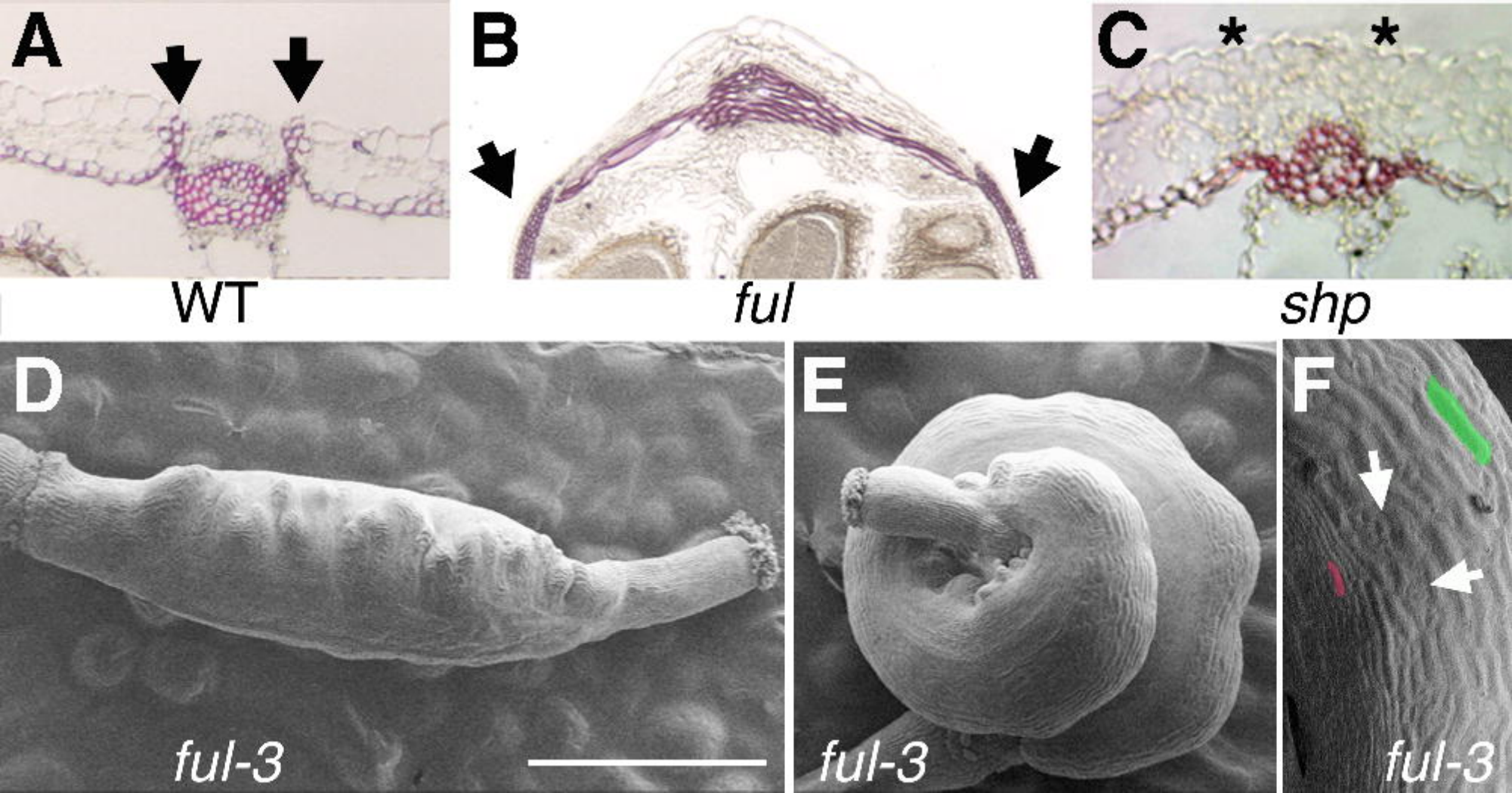
analysis. Posterior mean shown on the vertical axis indicate mean values of sites. Rate class weights (individual sites) around the transition site are shown as a red peak rising above the neutral selection level (blue plane).

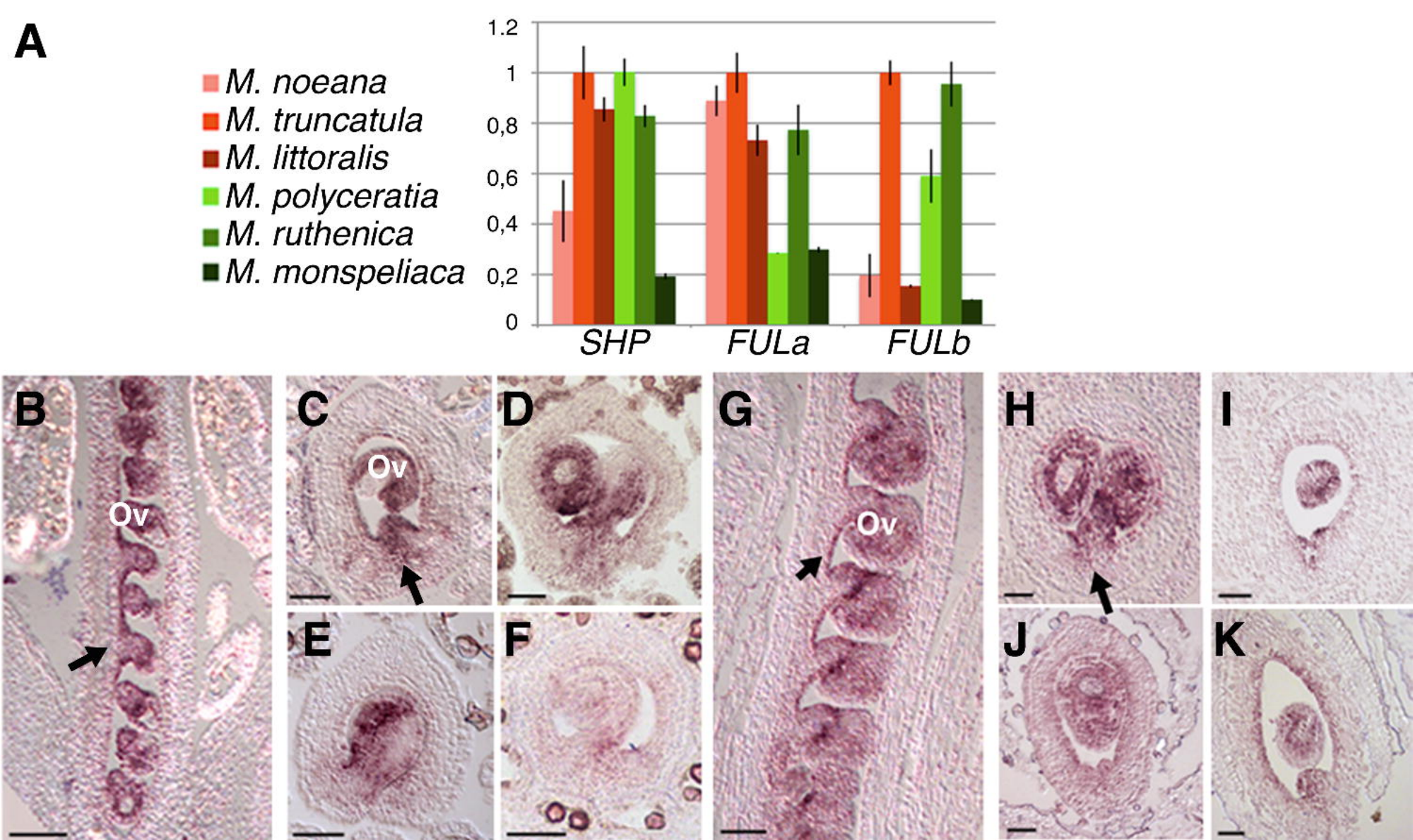
Figure 6. Different activities of SHP proteins from *Medicago* species with coiled or uncoiled pod morphology.

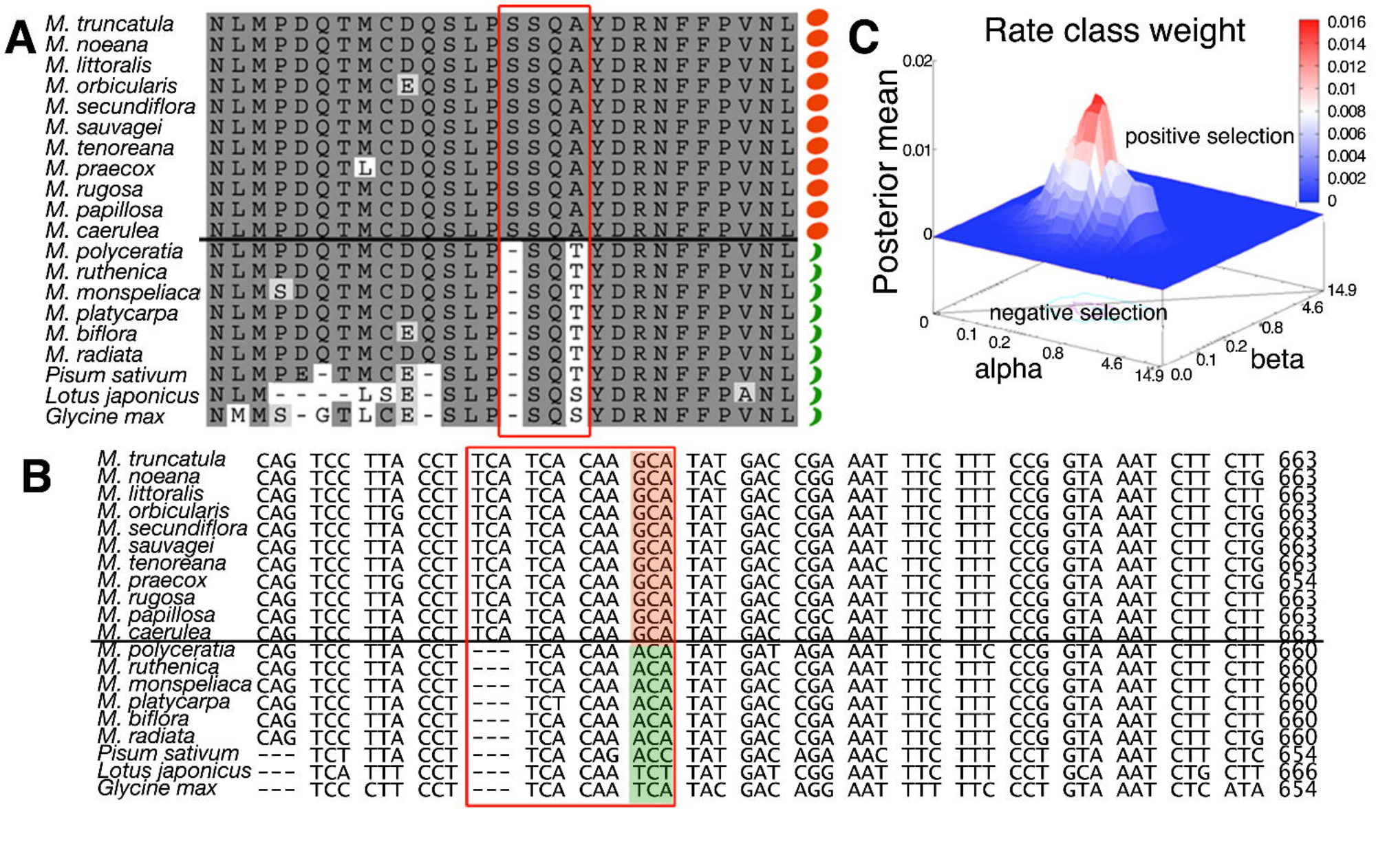
(**A, B**) Effect of the overexpression of *Medicago* SHP genes in Arabidopsis. (**A**) Wild type Arabidopsis fruit (left) and examples of intermediate and strong fruit phenotypes observed in *35S::MpolSHP* or 35S::*MtruSHP* T1 lines (center). The right panel also shows the strongest fruit phenotypes observed when the Arabidopsis *SHP1* gene is overexpressed. (**B**) Relative proportions of phenotypes observed in the different T1 plants transformed with 35S::*MtruSHP* (from *M. truncatula*, coiled pods) and 35S::*MpolSHP* (from *M. polyceratia*, uncoiled pods). More that 150 T1 plants were generated for each construct. (**C**) Yeast-two-hybrid experiment to reveal dimer formation between SEP and SHP proteins from *M. truncatula* (coiled pods) and *M. polyceratia* (uncoiled pods). 3-AT concentration (mM) for each experiment is indicated. *MpolSEP1* and *MpolSEP3* were cloned for these analyses (see acc. Numbers in Table S1). Acc. Numbers: *MtSEP1*, AC146650a; *MtSEP3*, AC144644.


Medicago sativa caerulea
Medicago papillosa
Medicago rugosa
Medicago littoralis
Medicago truncatula
Medicago noeana
Medicago praecox
Medicago orbicularis
Medicago secundiflora
Medicago sauvagei
Medicago tenoreana
Medicago biflora
Medicago platycarpa
Medicago ruthenica
Medicago polyceratia
Medicago monspeliaca
Medicago radiata
Trigonella
Melilotus


A
F
B
C
G
H
D
E
I
J
M. truncatula
M. noeana
K
P
L
M
Q
R
N
O
S
T
M. ruthenica
M. polyceratia


A
B
C
*
*
WT
ful
shp
D
E
F
ful-3
ful-3
ful-3


A
M. noeana
M. truncatula
M. littoralis
M. polyceratia
M. ruthenica
M. monspeliaca
1.2
1
0,8
0,6
0,4
0,2
0
SHP
FULa
FULb
B
Ov
C
Ov
D
E
F
G
Ov
H
I
J
K


A
M. truncatula NLMPDQTMCDQSLPSSQAYDRNFFPVNL
M. noeana NLMPDQTMCDQSLPSSQAYDRNFFPVNL
M. littoralis NLMPDQTMCDQSLPSSQAYDRNFFPVNL
M. orbicularis NLMPDQTMCEQSLPSSQAYDRNFFPVNL
M. secundiflora NLMPDQTMCDQSLPSSQAYDRNFFPVNL
M. sauvagei NLMPDQTMCDQSLPSSQAYDRNFFPVNL
M. tenoreana NLMPDQTMCDQSLPSSQAYDRNFFPVNL
M. praecox NLMPDQTLCDQSLPSSQAYDRNFFPVNL
M. rugosa NLMPDQTMCDQSLPSSQAYDRNFFPVNL
M. papillosa NLMPDQTMCDQSLPSSQAYDRNFFPVNL
M. caerulea NLMPDQTMCDQSLPSSQAYDRNFFPVNL
M. polyceratia NLMPDQTMCDQSLP-SQTYDRNFFPVNL
M. ruthenica NLMPDQTMCDQSLP-SQTYDRNFFPVNL
M. monspeliaca NLMSDQTMCDQSLP-SQTYDRNFFPVNL
M. platycarpa NLMPDQTMCDQSLP-SQTYDRNFFPVNL
M. biflora NLMPDQTMCEQSLP-SQTYDRNFFPVNL
M. radiata NLMPDQTMCDQSLP-SQTYDRNFFPVNL
Pisum sativum NLMPE-TMCE-SLP-SQTYDRNFFPVNL
Lotus japonicus NLM-----LSE-SLP-SQSYDRNFFPANL
Glycine max NMMS-GTLCE-SLP-SQSYDRNFFPVNL
C
Rate class weight
Posterior mean
0.02
0.01
0
positive selection
negative selection
0.016
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
alpha
0
0.1
0.2
0.8
4.6
14.9
beta
0.0
0.1
0.2
0.8
4.6
14.9
B
M. truncatula CAG TCC TTA CCT TCA TCA CAA GCA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTT 663
M. noeana CAG TCC TTA CCT TCA TCA CAA GCA TAC GAC CGG AAT TTC TTT CCG GTA AAT CTT CTG 663
M. littoralis CAG TCC TTA CCT TCA TCA CAA GCA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTT 663
M. orbicularis CAG TCC TTG CCT TCA TCA CAA GCA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTG 663
M. secundiflora CAG TCC TTA CCT TCA TCA CAA GCA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTG 663
M. sauvagei CAG TCC TTA CCT TCA TCA CAA GCA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTG 663
M. tenoreana CAG TCC TTA CCT TCA TCA CAA GCA TAT GAC CGA AAC TTC TTT CCG GTA AAT CTT CTG 663
M. praecox CAG TCC TTG CCT TCA TCA CAA GCA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTG 654
M. rugosa CAG TCC TTA CCT TCA TCA CAA GCA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTT 663
M. papillosa CAG TCC TTA CCT TCA TCA CAA GCA TAT GAC CGC AAT TTC TTT CCG GTA AAT CTT CTT 663
M. caerulea CAG TCC TTA CCT TCA TCA CAA GCA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTT 663
M. polyceratia CAG TCC TTA CCT --- TCA CAA ACA TAT GAT AGA AAT TTC TTC CCG GTA AAT CTT CTT 660
M. ruthenica CAG TCC TTA CCT --- TCA CAA ACA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTT 660
M. monspeliaca CAG TCC TTA CCT --- TCA CAA ACA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTT 660
M. platycarpa CAG TCC TTA CCT --- TCT CAA ACA TAT GAC CGG AAT TTC TTT CCG GTA AAT CTT CTT 660
M. biflora CAG TCC TTA CCT --- TCA CAA ACA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTT 660
M. radiata CAG TCC TTA CCT --- TCA CAA ACA TAT GAC CGA AAT TTC TTT CCG GTA AAT CTT CTG 660
Pisum sativum --- TCT TTA CCT --- TCA CAG ACC TAT GAC AGA AAC TTC TTT CCT GTA AAT CTT CTC 654
Lotus japonicus --- TCA TTT CCT --- TCA CAA TCT TAT GAT CGG AAT TTC TTT CCT GCA AAT CTG CTT 666
Glycine max --- TCC CTT CCT --- TCA CAA TCA TAC GAC AGG AAT TTT TTC CCT GTA AAT CTC ATA 654

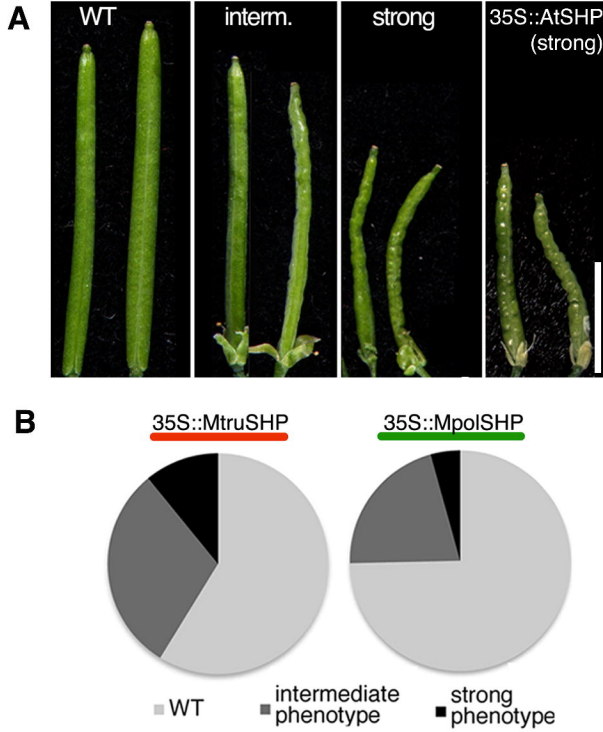

A
WT
interm.
strong
35S::AtSHP
(strong)
B
35S::MtruSHP
35S::MpolSHP
WT
intermediate
phenotype
strong
phenotype
C
MtruSHP
MtruSEP1
MpolSHP
MpolSEP1
MtruSHP
MtruSEP3
MpolSHP
MpolSEP3
[3AT]
40 mM
60 mM
80mM

Table 1. Single Likelihood Ancestor Counting (SLAC) results[#].

Genes	Posit. Sites (SLAC)*	Neg. Sites (SLAC)*	Mean dN-dS	Sites with evidence of directional selection**
Coiled	1	52	-0.050	0
Uncoiled	0	0	-0.010	0
All homologs	0	50	-0.095	2

[#]Positively and negatively selected sites (nucleotide changes), as well as mean dN-dS (difference between synonymous and non-synonymous sites) and Sites with evidence of directional selection, i.e., those having a concerted substitution towards a particular residue. *Significance 0.05, ** Mixed Effects Model of Evolution. For each codon, estimates of the numbers of inferred synonymous (s) and nonsynoymous (n) substitutions were calculated along with the numbers of sites that are estimated to be synonymous (S) and nonsynonymous (N). These estimates were produced using the joint Maximum Likelihood reconstructions of ancestral states under a Muse-Gaut model of codon substitution and Tamura-Nei model of nucleotide substitution for *coiled-like* and *uncoiled-like* homologues and General Time Reversible model for all genes. For estimating ML values, a tree topology was automatically computed. The statistical test dN-dS was used for detecting codons that have undergone positive selection, where dS is the number of synonymous substitutions per site (s/S) and dN is the number of nonsynonymous substitutions per site (n/N).

Table 2. Selection rate in the ACA-AGA transition site[†] using the coiled, uncoiled homologues and all genes clustered[#].

Triplet	Selection* Coiled	Seq. evol. rate**	Selection Uncoiled	Seq. evol. rate**	Selection* all genes	Seq. evol. rate**
ACA	n/a	n/a	0	0.496	0.644	0.496
GCA	0	0.226	n/a	n/a	0.500	0.461

[†]Codon identified as associated with the *Coiled* condition. *Selection acting on the site. **Mean (relative) sequence evolution rates are shown for each site next to the aimed site. These are normalized values such as sequence evolution rates across all sites is 1. Relative rates were estimated under the Tamura model (+gamma) for coiled-like and uncoiled-like homologues and under the General Time reversible model (+gamma+invar) for all genes. There were a total of 609 positions in the final dataset.