The role of *BRANCHLESS TRICHOME* in the evolution of a novel trichome shape in the genus *Physaria* (Brassicaceae)

By

Abigail R. Mazie

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Botany)

at the

UNIVERSITY OF WISCONSIN-MADISON

2016

Date of final oral examination: 4 December 2015

The dissertation is approved by the following members of the Final Oral Committee:
David A. Baum, Professor, Botany
Cécile Ané, Associate Professor, Botany and Statistics
Donna E., Fernandez, Professor, Botany
Marisa S. Otegui, Professor, Botany
Kenneth J. Sytsma, Professor, Botany

Table of Contents

Abstract

Cell shape evolution is an important driver of cell specialization within organisms and diversification between organisms. I studied the unicellular trichomes of the mustard family (Brassicaceae) as a model to better understand the molecular and developmental mechanisms that underlie the evolution of cell shape. Although unicellular, the three-dimensional shape of Brassicaceae trichomes is quite variable. Most Brassicaceae trichomes consist of a stalk with multiple branching outgrowths arranged in a dendritic ("tree-like") pattern. In contrast, members of the genus *Physaria* display trichomes with branches that radiate from a short stalk and lie parallel to the epidermis in a stellate ("star-like") branching pattern. The stellate shape is a synapomorphy of *Physaria*, and all close relatives have retained the dendritic shape.

My research focused on understanding the contribution of *BRANCHLESS TRICHOME* (*BLT*), which plays a role in determining the number and position of trichome branches in *Arabidopsis thaliana*. Using sequence-based statistical analyses, I established that selection acted differently on *BLT* in *Physaria* than in other members of the Physarieae clade. Specifically, I found evidence of positive selection favoring changes in the amino acid sequence at a subset of sites in *Physaria* BLT.

To further explore the role of *BLT* in trichome shape change, I conducted interspecies transformation experiments in which the *P. fendleri BLT* promoter and coding regions were introduced into *A. thaliana blt* mutant plants. The *Physaria BLT* promoter appeared to be interchangeable with its *Arabidopsis* homolog, but the *Physaria* coding region resulted in trichomes that were less branched and had narrower, shorter stalks than those rescued with the *Arabidopsis* coding region. Together, the molecular evolutionary and interspecies transformation experiments indicate that changes to *Physaria BLT* did not cause the switch from dendritic to

stellate trichomes. Instead, these results suggest that selection on *Physaria BLT* acted to modify its interactions with other branch promoting proteins, perhaps to stabilize trichome development in response to protein changes that occurred during the transition from dendritic to stellate trichomes or, alternatively, during the subsequent diversification of the stellate form in different *Physaria* species.

Chapter I: *Physaria* trichomes as a model for cell shape evolution

A central goal of evolutionary biology is to understand the causes of organismal diversity. Cell shape diversity is especially interesting because cells are the most basic building blocks of all organisms. To better understand the molecular and developmental mechanisms that underlie the rich cell shape diversity in the natural world, my research investigates the unicellular trichomes of the mustard family (Brassicaceae).

Plant trichomes are epidermal outgrowths that most commonly play roles in defense from herbivores and abiotic stressors (Werker 2000; Wagner et al. 2004; Hanley et al. 2007). Although unicellular, Brassicaceae trichomes exhibit a stunning diversity of forms, especially in their branching pattern (Appel and Al-Shehbaz 2003; Al-Shehbaz et al. 2006; Beilstein et al. 2006; Beilstein et al. 2008). Though most branched trichomes have a dendritic, or tree-like, appearance, a few taxa deviate from this pattern (Beilstein et al. 2006; Beilstein et al. 2008). One striking departure from the typical dendritic pattern is seen in the genus *Physaria* (tribe Physarieae), where trichomes display a stellate, or star-like, branching pattern (Rollins and Shaw 1973; Rollins and Banerjee 1975; Rollins and Banerjee 1976). The stellate shape is a synapomorphy of the genus, as this morphology is shared by all members of *Physaria* but not close relatives, which still retain the dendritic shape (Beilstein et al. 2006; Beilstein et al. 2008; Fuentes-Soriano and Al-Shehbaz 2013).

Physaria stellate trichomes are interesting not only for their unique morphology, but also for their potential adaptive significance. The dense cover of overlapping trichomes seen in many species of *Physaria* has been proposed as an adaptation to the dry, open habitats of the western United States and northern Mexico where *Physaria* is widespread (Rollins and Shaw 1973; Rollins and Banerjee 1976). The short stalks and parallel branches of stellate trichomes seem

particularly suitable for reflecting light away from the leaf surface to protect the plants from solar damage. Research into stellate trichome evolution may therefore increase our understanding of an important structural adaptation to high light environments. Additionally, if we can identify the molecular and developmental changes that resulted in the transition from producing dendritic to producing stellate trichomes, this could shed light on the more general problem of how cell shape evolves. A key factor that makes *Physaria* trichomes such an exciting model system for answering this problem is it phylogenetic proximity to *Arabidopsis thaliana*, the premier model system for studies of trichome developmental.

Arabidopsis thaliana dendritic trichome development

Trichome development has been extensively researched in *A. thaliana*, which, like *Physaria*, is a member of Brassicaceae Lineage I. *A. thaliana* trichomes are dendritic, and they appear to have retained the ancestral trichome shape for Brassicaceae Lineage I (Beilstein et al. 2006; Beilstein et al. 2008). The basic sequence of *A. thaliana* dendritic trichome development has been carefully characterized and divided into six developmental stages (Hülskamp et al. 1994; Marks 1994; Szymanski et al. 1998; Szymanski et al. 1999). During stages one and two, trichome precursors enlarge by radial expansion, then undergo diffuse extension growth away from the leaf surface to form the main trichome stalk. During stages three and four, secondary growth foci form, typically establishing 3-4 trichome branches. The final stages of development include diffuse expansion of the entire trichome cell, sharpening of the branch tips, thickening of the cell wall, and formation of cell-wall papillae.

Mature trichomes are large, polyploid cells, reaching a final height of 200-300μm (Marks 1998; Oppenheimer 1998) and a typical DNA content of 32C (Hülskamp et al. 1994; Schnittger

and Hülskamp 2002), with a range of 4-64C (Melaragno et al. 1993). The increased DNA content is achieved through repeated endoreplication cycles, in which nuclear DNA is synthesized in the absence of mitosis and cell division (Hülskamp et al. 1994; Schnittger and Hülskamp 2002). Endoreplication cycles are postulated to contribute to cell size regulation and cell identity maintenance (De Veylder et al. 2011; Grebe 2012; Yang and Ye 2013). Mutations that affect the number of endoreplication cycles can affect trichome morphology - mutants with increased endoreplication often having more branches than normal, and those with decreased endoreplication having fewer (Hülskamp et al. 1994; Perazza et al. 1999; Downes et al. 2003; del Mar Castellano et al. 2004; Sonoda et al. 2009; Roodbarkelari et al. 2010; Breuer et al. 2012). The importance of endoreplication cycles to growth and development is widespread feature of both plant and animal cells (Edgar et al. 2014).

The genetic basis of trichome identity establishment

Extensive research has led to a greater understanding of the genetic underpinnings of trichome cell identity establishment and morphogenesis (Hülskamp 2004; Ishida et al. 2008; Tominaga-Wada et al. 2011; Grebe 2012; Yang and Ye 2013; Pattanaik et al. 2014; Marks 2015). Trichome identity is established in cells where the R2R3 MYB transcription factor *GLABRA1* (*GL1*) and the basic helix-loop-helix (bHLH) transcription factor *GLABRA3* (*GL3*) form an activation complex with the WD40 repeat domain containing protein *TRANSPARENT TESTA GLABRA1* (*TTG1*) that initiates trichome formation by controlling the transcription of key regulatory genes required for trichome development (Oppenheimer et al. 1991; Walker et al. 1999; Payne et al. 2000; Morohashi et al. 2007; Zhao et al. 2008; Morohashi and Grotewold 2009). Although *GL1* and *GL3* were the first members of the trichome activator complex to be

discovered, other, typically redundant, copies of R2R3 MYBs or bHLH proteins that can serve as activators of trichome development have been identified (R2R3 MYBs: Kirik et al. 2001; Kirik 2005; Liang et al. 2014; bHLH: Zhang et al. 2003; Maes et al. 2008; Symonds et al. 2011; Zhao et al. 2012). Phytohormones, including gibberellic acids, cytokinins, and jasmonates, promote trichome initiation through interactions with members of the activator complex (Pattanaik et al. 2014).

Immediate, downstream targets of the MYB-bHLH-WD40 complex include *SIAMESE* (*SIM*), a cell cycle inhibitor that mediates the transition from mitosis to endoreplication (Walker et al. 2000; Churchman et al. 2006; Morohashi and Grotewold 2009), along with *GLABRA2* (Rerie et al. 1994; Szymanski et al. 1998; Morohashi et al. 2007; Wang and Chen 2008; Zhao et al. 2008; Morohashi and Grotewold 2009; Wang et al. 2010) and *TRANSPARENT TESTA GLABRA2* (Johnson et al. 2002; Ishida et al. 2007; Zhao et al. 2008; Morohashi and Grotewold 2009), transcription factors that are thought to regulate the expression of downstream differentiation genes. The MYB-bHLH-WD40 complex also activates expression of inhibitors, including *TRIPTYCHON* (*TRY*) and *CAPRICE* (*CPC*), that travel to neighboring cells where they bind to the bHLH transcription factors to repress trichome development in the vicinity of newly established trichomes (Wang and Chen 2014). This results in a nonrandom distribution of trichomes across the leaf surface, with initiating trichomes typically separated by 3-4 cells (Hülskamp et al. 1994; Larkin et al. 1996; Bramsiepe et al. 2010). Other trichome patterning mechanisms are thought to work in concert with the activator-inhibitor model to regulate trichome spacing (Pesch and Hülskamp 2009; Balkunde et al. 2010; Ryu et al. 2013).

The molecular basis of trichome morphogenesis

In *Arabidopsis*, the microtubule and actin cytoskeletons play an important role in the establishment and shaping of trichome branches (Smith and Oppenheimer 2005; Mathur 2006; Szymanski 2009; Marks 2015). Early pharmacological studies showed that trichomes treated with microtubule destabilizing drugs became bloated and did not form branches, whereas those treated with actin destabilizing drugs initiated branches that displayed abnormal twisting or reduced expansion or both, indicating a failure of the branches to properly mature (Mathur et al. 1999; Szymanski et al. 1999; Mathur and Chua 2000). Trichomes on actin-destabilized plants looked identical to those of the *distorted* class of trichome mutants (Mathur et al. 1999; Szymanski et al. 1999), which were later identified as members of the ACTIN-RELATED PROTEIN 2/3 (ARP2/3) complex, that nucleates actin microfilaments, and the WAVE (WASP family verprolin homologous) / SCAR (suppressor of cAMP receptor) complex (W/SRC) that regulates ARP2/3 activity (Hülskamp 2004; Mathur 2004; Szymanski 2005; Yanagisawa et al. 2013). These analyses indicate that microtubules are crucial in establishing polarized growth axes, while actin functions primarily in the expansion and refinement of the microtubule-established shape.

The necessity of microtubules for branch initiation is further supported by studies of mutations to genes important for microtubule development and dynamics. Microtubules with dysfunctional or immature α- and β-tubulin subunits have trichomes with fewer than normal branches (Kirik, Grini, et al. 2002; Kirik, Mathur, et al. 2002; Abe et al. 2004; Gu et al. 2008). Trichome branching is also reduced in plants with mutations to genes that affect the angle of microtubule growth from existing cortical arrays (Nakamura and Hashimoto 2009; Kirik et al. 2012) and mutants with reduced microtubule severing (Burk et al. 2001).

One model for trichome branch initiation proposes that the development of new branch foci is driven by the local reorientation of cortical microtubules (Mathur and Chua 2000; Marks 2015). During cell development, turgor pressure drives expansion uniformly across the cell surface. Microtubules are postulated to determine the directionality of microfibril synthesis (Paredez et al. 2006; Lindeboom et al. 2008; Gutierrez et al. 2009), which directs cell elongation by restricting turgor driven cell expansion parallel to their orientation, thereby redirecting growth perpendicular to their axes. Although microtubules of mature trichomes have an oblique or longitudinal orientation, during early, pre-branching trichome growth, cortical microtubules are randomly distributed across the trichome (Tian et al. 2015). Cortical microtubules then become aligned in a transverse, circumferential array around the elongating trichome, with a microtubule-depleted zone at the apex and a decreasing microtubule gradient from the cell tip to the cell base, leading to anisotropic growth of the trichome perpendicular to the plane of the leaf (Mathur and Chua 2000; Beilstein and Szymanski 2004; Sambade et al. 2014; Tian et al. 2015). Loss of the microtubule density gradient in *an1* mutants is thought to contribute to their reduced branching phenotype, possibly through a membrane trafficking defect (Folkers et al. 2002; Kim et al. 2002; Minamisawa et al. 2011). A transverse ring of microtubules, similar to the microtubule ring seen in unbranched trichomes, also encircles elongating branches (Beilstein and Szymanski 2004; Sambade et al. 2014; Tian et al. 2015; Yanagisawa et al. 2015). Therefore, reorientation of cortical microtubules to create new microtubule rings may allow for the development of additional growth foci – the branches. Dynamic rearrangements of cortical microtubules are believed to more generally underlie cell shape formation (Shaw 2013; Oda 2015).

Once trichome branches have been established, actin is thought to play a crucial role in polarized branch elongation by regulating the intracellular transport of secretory organelles to control the deposition of membrane and cell wall materials. However, the precise mechanisms through which actin regulates anisotropic cell growth remain unclear (Smith and Oppenheimer 2005; Hussey et al. 2006; Ojangu et al. 2012; Yanagisawa et al. 2013; Yanagisawa et al. 2015). Disruption of actin-related genes and their upstream regulators typically results in trichomes of normal branch number but defective branch elongation (Hülskamp 2004; Mathur 2004; Szymanski 2005; Yanagisawa et al. 2013; Yanagisawa et al. 2015). ARP2/3, an actin filament nucleator, is positively regulated by WAVE/SCAR (W/SRC) (Szymanski 2005; J. Zhang et al. 2005; Djakovic et al. 2006; Le et al. 2006; Zhang et al. 2008; Yanagisawa et al. 2013). The W/SRC-ARP2/3 complex is thought to maintain polarized cell elongation by accumulating at developing branch apices to nucleate and organize actin (Dyachok et al. 2008; Yanagisawa et al. 2015).

In prebranching trichomes, cytoplasmic actin cables align with the long axis of the cell (Tian et al. 2015). During branch development, these cytoplasmic cables continue to extend from the cell's base to the branch tips, while a fine mesh of mostly transverse actin forms a cap at the elongating branch tips (Zhang et al. 2005; Sambade et al. 2014; Tian et al. 2015; Yanagisawa et al. 2015). The organization of the actin cytoskeleton maintains an increasing gradient of cell wall thickness from the branch tip to the branch base that contributes to the directional outgrowth of branches (Yanagisawa et al. 2015).

In addition to the separate roles of the microtubule and actin cytoskeletons in branch initiation and elongation, coordination between these two cytoskeletal systems is thought to be important in regulating cell shape changes (Saedler et al. 2004; Zhang et al. 2005; Petrášek and

Schwarzerová 2009; Sampathkumar et al. 2011; Yanagisawa et al. 2013; Sambade et al. 2014; Havelková et al. 2015; Yanagisawa et al. 2015). One protein, ZWICHEL (ZWI, also known as KCBP, kinesin-like calmodulin-binding protein), has been proposed to integrate microtubule and actin assembly to regulate branch growth (Tian et al. 2015). *A. thaliana zwi* mutants have trichomes with fewer branches, often with impaired growth and blunt tips (Hülskamp et al. 1994; Oppenheimer et al. 1997), and the ZWI protein localizes to initiating and elongating branch foci (Tian et al. 2015). *ZWI* is thought to be a kinesin-myosin chimera, with a domain that allows the protein to bind to both microtubules and actin filaments (Reddy and Reddy 1999; Tian et al. 2015). *ZWI* also has a calmodulin binding domain that can be bound by calmodulin or KCBP-interacting Ca^{2+} binding protein (KIC) to negatively regulate the microtubule-binding ability of ZWI in a calcium-dependent manner (Reddy et al. 1996; Deavours et al. 1998; Narasimhulu and Reddy 1998; Reddy and Reddy 2002; Reddy et al. 2004). However, this model is contradicted by the fact that trichome branching is normal in mutant plants transformed with a construct in which the calmodulin binding domain was deleted (Tian et al. 2015).

Rho of Plant (ROP) GTPases are regulators of plant cell polarity (Yang 2008; Bloch and Yalovsky 2013) that are known to modulate microtubule dynamics and to interact with W/SRC to positively regulate ARP2/3 (Fu et al. 2002; Basu et al. 2004; Fu et al. 2005; Uhrig et al. 2007; Yalovsky et al. 2008; Fu et al. 2009; Xu et al. 2010; Oda and Fukuda 2012; Lin et al. 2013). The loss of branching in mutants of *SPIKE1*, a guanine nucleotide exchange factor thought to activate ROP and W/SRC, is postulated to be a consequence of decreased ROP activation leading to misregulation of microtubule assembly and actin polymerization (Qiu 2002; Basu et al. 2008). Plants that express a protein that interferes with ROP activity look similar to *spike1* mutants (Singh et al. 2013), and mutants of the ROP-activated receptor-like cytoplasmic kinase (AtRLCK

VI_A3) have increased trichome branching (Reiner et al. 2015), suggesting a role for ROPs in trichome morphogenesis.

Developmental differences between the dendritic and stellate forms

As documented more fully in Chapter III, comparisons of stellate trichome development in *Physaria* to dendritic trichome development in its sister genus, *Paysonia,* show that the key differences between their mature trichomes result from differences in the pattern of primary branch initiation. In dendritic taxa, the primary growth focus continues to expand with branches initiated subterminally in a sequential pattern. In contrast, the primary growth focus of stellate trichomes ceases to expand once branches are initiated, and all further growth is directed to the resulting branches. Dendritic and stellate trichomes also differed in the number of primary branches initiated: while dendritic trichomes initiated only a single primary branch, stellate trichomes initiated at least three (and up to seventeen) branches simultaneously around the primary growth apex. Given these considerations, the simplest explanation for the transition from dendritic to stellate trichomes is a change in the number of branch initiation points and their position in relation to the trichome stalk (see Chapter III). Given this model, I hypothesized that good candidates for the evolution of changes in branching patterns would be proteins that regulate the reorganization of microtubules to create new branch foci, but are not directly involved in microtubule assembly or dynamics.

***STICHEL* and *BLT* are good candidates for the evolution of stellate trichomes**

Two genes, *STICHEL* (*STI*) and *BRANCHLESS TRICHOME* (*BLT*), seemed to be especially good candidates for contributing to the evolution of trichome shape. Both genes are

positive regulators of trichome branching, with loss-of-function mutants exhibiting completely unbranched trichomes (Hülskamp et al. 1994; Ilgenfritz et al. 2003; Marks et al. 2009; Kasili et al. 2011). *STI* contains a domain with sequence similarity to eubacterial DNA-polymerase III γ-subunits (Ilgenfritz et al. 2003) and *BLT* plays a role in linking endoreplication levels with branch number (Kasili et al. 2011), which suggests a link between cell cycle regulation and cell morphogenesis. *STI* and *BLT* expression was detected in the trichomes of *gl3-shapeshifter sim* mutants, which are arrested at an early stage of trichome development, but not in mature, wild-type trichomes (Marks et al. 2009). These results point to a role for *STI* and *BLT* during early stages of trichome development as primary regulators of trichome branch initiation.

Over-expression of *STI* (Ilgenfritz et al. 2003) or *BLT* (Kasili et al. 2011) increases trichome branching, and weak *sti* alleles develop two branches (Folkers et al. 1997; Ilgenfritz et al. 2003), suggesting a possible dosage dependent effect on branch number regulation. The STI and BLT proteins co-localize to the expanding tips of emerging trichome branches (Marks et al. 2009), and physically interact (Kasili et al. 2011). STI and BLT are thought to form a protein complex that regulates dynamic microtubule rearrangements to direct the polarity of branch initiation sites (Marks et al. 2009; Kasili et al. 2011; Marks 2015). Although *STI* and *BLT* play a critical role in trichome branch initiation, mutants were normal in all other aspects of growth and development (Ilgenfritz et al. 2003; Kasili et al. 2011). Based on their strong and specific effects on the regulation of trichome branching, coupled with their lack of pleiotropic phenotypes, I hypothesized that *STI* and *BLT* played a role in the evolution of trichome branching patterns that occurred during the transition from the dendritic to stellate forms.

The role of *BLT* in stellate trichome evolution

In the research described in this dissertation, I focused on the possible role of *BLT* in the evolution of the stellate form. I started by looking for evidence of a change in selective pressure on *BLT* during the transition from dendritic to stellate trichomes (Chapter II). This entailed using maximum likelihood models of codon evolution to evaluate whether *BLT* experienced a selective shift during the initial evolution of stellate trichomes or during subsequent evolution in the *Physaria* clade. My results support the latter scenario: I found evidence for a shift in selective pressure on *BLT* in the *Physaria* crown lineage, but not on the stem lineage where stellate trichomes evolved. In addition, I found evidence for positive selection on a subset of *Physaria BLT* codons, and discovered that most sites in the protein with the highest probability of having evolved under positive selection were situated in a coiled-coil domain. These results point to changes in protein interactions between BLT and other trichome branch promoters in the *Physaria* clade relative to that seen in *Paysonia* and other outgroups.

To complement my molecular evolutionary analyses of *Physaria BLT*, Chapter III details developmental studies of trichome development and interspecies transformation experiments that introduced the promoter and coding region of *P. fendleri BLT* into *A. thaliana blt* mutants and compared the transformant plants to those rescued with the homologous *A. thaliana* gene regions. I found that the *BLT* promoter is interchangeable between the two taxa, but that the *P. fendleri* coding region is less effective than the *A. thaliana* coding region at rescuing the mutant phenotype. Plants transformed with the *P. fendleri BLT* coding region had fewer branches and narrower, shorter trichome stalks than those transformed with *A. thaliana BLT*.

Overall my research suggests that changes to *BLT* in the *Physaria* clade were not the sole cause of the evolution of the stellate form. Rather, given that *BLT* is under positive selection in

the *Physaria* crown lineage and that *A. thaliana blt* mutants transformed with the *P. fendleri BLT* coding region are less effective at rescuing the mutant phenotype, it is likely that selection acted on the amino acid sequence of BLT to modify its interactions with other proteins subsequent to the evolution of the stellate form. One possibility is that sequence changes that acted to stabilize protein interactions in the context of stellate trichome development were favored by positive selection. Alternatively, positive selection could reflect changes in protein interactions associated with the diversification of stellate trichome form within *Physaria* following the abrupt switch from dendritic to stellate trichomes.

Whichever selective hypothesis holds, my research succeeded in identifying an altered pattern of evolution in an important trichome development gene that coincides with a dramatic change in trichome shape. Although my data do not suggest that *BLT* itself drove the transition from dendritic to stellate trichomes in *Physaria*, they did suggest compelling hypotheses about how selection can shape genetic variation in developmentally important genes across species. As a result, this research provides a first step towards a greater understanding of the types of cell-level genetic and developmental changes that create morphological diversity. Through such analyses of *BLT* and other trichome shape genes in Brassicaceae, I believe that Brassicaceae trichomes will emerge as a premier model system for exploring the evolution of cell shape control in multicellular eukaryotes.

Literature cited

Abe T, Thitamadee S, Hashimoto T. 2004. Microtubule defects and cell morphogenesis in the *lefty1lefty2* tubulin mutant of *Arabidopsis thaliana*. Plant and Cell Physiology 45:211–220.

Al-Shehbaz IA, Beilstein MA, Kellogg EA. 2006. Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Plant Syst. Evol. 259:89–120.

Appel O, Al-Shehbaz IA. 2003. Cruciferae. In: Kubitzki K, editor. Families and genera of vascular plants. Vol. 5. Berlin (Germany): Springer Berlin Heidelberg. pp. 75–174.

Balkunde R, Pesch M, Hülskamp M. 2010. Trichome patterning in *Arabidopsis thaliana*: from genetic to molecular models. Current topics in developmental biology. 91:299-321.

Basu D, El-Assal SE-D, Le J, Mallery EL, Szymanski DB. 2004. Interchangeable functions of Arabidopsis PIROGI and the human WAVE complex subunit SRA1 during leaf epidermal development. Development 131:4345–4355.

Basu D, Le J, Zakharova T, Mallery EL, Szymanski DB. 2008. A SPIKE1 signaling complex controls actin-dependent cell morphogenesis through the heteromeric WAVE and ARP2/3 complexes. Proc. Natl. Acad. Sci. U.S.A. 105:4044–4049.

Beilstein M, Szymanski D. 2004. Cytoskeletal requirements during Arabidopsis trichome development. In: Hussey PJ, editor. The Plant Cytoskeleton in Cell Differentiation and Development. Annual Plant Reviews. pp. 265–289.

Beilstein MA, Al-Shehbaz IA, Kellogg EA. 2006. Brassicaceae phylogeny and trichome evolution. American Journal of Botany 93:607–619.

Beilstein MA, Al-Shehbaz IA, Mathews S, Kellogg EA. 2008. Brassicaceae phylogeny inferred from phytochrome A and *ndhF* sequence data: tribes and trichomes revisited. American Journal of Botany 95:1307–1327.

Bloch D, Yalovsky S. 2013. Cell polarity signaling. Current Opinion in Plant Biology 16:734–742.

Bramsiepe J, Wester K, Weinl C, Roodbarkelari F, Kasili R, Larkin JC, Hülskamp M, Schnittger A. 2010. Endoreplication Controls Cell Fate Maintenance.Qu L-J, editor. PLoS Genet 6:e1000996.

Breuer C, Morohashi K, Kawamura A, Takahashi N, Ishida T, Umeda M, Grotewold E, Sugimoto K. 2012. Transcriptional repression of the APC/C activator CCS52A1 promotes active termination of cell growth. The EMBO Journal 31:4488–4501.

Burk DH, Liu B, Zhong R, Morrison WH, Ye Z-H. 2001. A Katanin-like protein regulates normal cell wall biosynthesis and cell elongation. The Plant Cell 13:807–827.

Churchman ML, Brown ML, Kato N, Kirik V, Hülskamp M, Inzé D, De Veylder L, Walker JD,

Zheng Z, Oppenheimer DG, et al. 2006. SIAMESE, a plant-specific cell cycle regulator, controls endoreplication onset in *Arabidopsis thaliana*. THE PLANT CELL ONLINE 18:3145–3157.

De Veylder L, Larkin JC, Schnittger A. 2011. Molecular control and function of endoreplication in development and physiology. Trends in Plant Science 16:624–634.

Deavours BE, Reddy A, Walker RA. 1998. Ca^{2+}/calmodulin regulation of the *Arabidopsis* kinesin-like calmodulin-binding protein. Cell motility and the cytoskeleton 40:408–416.

del Mar Castellano M, Boniotti MB, Caro E, Schnittger A, Gutierrez C. 2004. DNA Replication Licensing Affects Cell Proliferation or Endoreplication in a Cell Type–Specific Manner. The Plant Cell 16:2380–2393.

Djakovic S, Dyachok J, Burke M, Frank MJ, Smith LG. 2006. BRICK1/HSPC300 functions with SCAR and the ARP2/3 complex to regulate epidermal cell shape in Arabidopsis. Development 133:1091–1100.

Downes BP, Stupar RM, Gingerich DJ, Vierstra RD. 2003. The HECT ubiquitin-protein ligase (UPL) family in *Arabidopsis*: UPL3 has a specific role in trichome development. The Plant Journal 35:729–742.

Dyachok J, Shao M-R, Vaughn K, Bowling A, Facette M, Djakovic S, Clark L, Smith L. 2008. Plasma membrane-associated SCAR complex subunits promote cortical f-actin accumulation and normal growth characteristics in *Arabidopsis* roots. Molecular Plant 1:990–1006.

Edgar BA, Zielke N, Gutierrez C. 2014. Endocycles: a recurrent evolutionary innovation for post-mitotic cell growth. Nat Rev Mol Cell Biol 15:197–210.

Folkers U, Berger J, Hülskamp M. 1997. Cell morphogenesis of trichomes in *Arabidopsis*: differential control of primary and secondary branching by branch initiation regulators and cell growth. Development:3779–3786.

Folkers U, Kirik V, Schobinger U, Falk S, Krishnakumar S, Pollock MA, Oppenheimer DG, Day I, Reddy AR, Jürgens G, et al. 2002. The cell morphogenesis gene *ANGUSTIFOLIA* encodes a CtBP/BARS-like protein and is involved in the control of the microtubule cytoskeleton. The EMBO Journal 21:1280–1288.

Fu Y, Gu Y, Zheng Z, Wasteneys G, Yang Z. 2005. *Arabidopsis* Interdigitating cell growth requires two antagonistic pathways with opposing action on cell morphogenesis. Cell 120:687–700.

Fu Y, Li H, Yang Z. 2002. The ROP2 GTPase Controls the Formation of Cortical Fine F-Actin and the Early Phase of Directional Cell Expansion during Arabidopsis Organogenesis. The Plant Cell 14:777–794.

Fu Y, Xu T, Zhu L, Wen M, Yang Z. 2009. A ROP GTPase signaling pathway controls cortical microtubule ordering and cell expansion in *Arabidopsis*. Current Biology 19:1827–1832.

Fuentes-Soriano S, Al-Shehbaz I. 2013. Phylogenetic relationships of mustards with multiaperturate pollen (Physarieae, Brassicaceae) based on the plastid *ndhF* gene: implications for morphological diversification. Systematic Botany 38:178–191.

Grebe M. 2012. The patterning of epidermal hairs in *Arabidopsis*—updated. Current Opinion in Plant Biology 15:31–37.

Gu Y, Deng Z, Paredez AR, DeBolt S, Wang ZY, Somerville C. 2008. Prefoldin 6 is required for normal microtubule dynamics and organization in *Arabidopsis*. Proc. Natl. Acad. Sci. U.S.A. 105:18064–18069.

Gutierrez R, Lindeboom JJ, Paredez AR, Emons AMC, Ehrhardt DW. 2009. *Arabidopsis* cortical microtubules position cellulose synthase delivery to the plasma membrane and interact with cellulose synthase trafficking compartments. Nature Cell Biology 11:797–806.

Hanley ME, Lamont BB, Fairbanks MM, Rafferty CM. 2007. Plant structural traits and their role in anti-herbivore defence. Perspectives in Plant Ecology, Evolution and Systematics 8:157–178.

Havelková L, Nanda G, Martinek J, Bellinvia E, Sikorová L, Šlajcherová K, Seifertová D, Fischer L, Fišerová J, Petrášek J, et al. 2015. Arp2/3 complex subunit ARPC2 binds to microtubules. Plant Science 241:96–108.

Hussey PJ, Ketelaar T, Deeks MJ. 2006. Control of the actin cytoskeleton in plant cell growth. Annu. Rev. Plant Biol. 57:109–125.

Hülskamp M, Miséra S, Jürgens G. 1994. Genetic dissection of trichome cell development in Arabidopsis. Cell 76:555–566.

Hülskamp M. 2004. Plant trichomes: a model for cell differentiation. Nat Rev Mol Cell Biol 5:471–480.

Ilgenfritz H, Bouyer D, Schnittger A, Mathur J, Kirik V, Schwab B, Chua N-H, Jürgens G, Hülskamp M. 2003. The Arabidopsis *STICHEL* gene is a regulator of trichome branch number and encodes a novel protein. Plant Physiology 131:643–655.

Ishida T, Hattori S, Sano R, Inoue K, Shirano Y, Hayashi H, Shibata D, Sato S, Kato T, Tabata S, et al. 2007. Arabidopsis TRANSPARENT TESTA GLABRA2 Is Directly Regulated by R2R3 MYB Transcription Factors and Is Involved in Regulation of GLABRA2 Transcription in Epidermal Differentiation. The Plant Cell 19:2531–2543.

Ishida T, Kurata T, Okada K, Wada T. 2008. A genetic regulatory network in the development of trichomes and root hairs. Annu. Rev. Plant Biol. 59:365-386.

Johnson CS, Ben Kolevski, Smyth DR. 2002. TRANSPARENT TESTA GLABRA2, a Trichome and Seed Coat Development Gene of Arabidopsis, Encodes a WRKY Transcription Factor. The Plant Cell 14:1359–1375.

Kasili R, Huang CC, Walker JD, Simmons LA, Zhou J, Faulk C, Hülskamp M, Larkin JC. 2011. *BRANCHLESS TRICHOMES* links cell shape and cell cycle control in *Arabidopsis* trichomes. Development 138:2379–2388.

Kim GT, Shoda K, Tsuge T, Cho KH, Uchimiya H, Yokoyama R, Nishitani K, Tsukaya H. 2002. The ANGUSTIFOLIA gene of Arabidopsis, a plant CtBP gene, regulates leaf-cell expansion, the arrangement of cortical microtubules in leaf cells and expression of a gene involved in cell-wall formation. The EMBO Journal 21:1267–1279.

Kirik A, Ehrhardt DW, Kirik V. 2012. TONNEAU2/FASS Regulates the Geometry of Microtubule Nucleation and Cortical Array Organization in Interphase Arabidopsis Cells. The Plant Cell 24:1158–1170.

Kirik V, Grini PE, Mathur J, Klinkhammer I, Adler K, Bechtold N, Herzog M, Bonneville J-M, Hülskamp M. 2002. The Arabidopsis TUBULIN-FOLDING COFACTOR A Gene Is Involved in the Control of the α/β-Tubulin Monomer Balance. The Plant Cell 14:2265–2276.

Kirik V, Mathur J, Grini PE, Klinkhammer I, Adler K, Bechtold N, Herzog M, Bonneville J-M, Hülskamp M. 2002. Functional Analysis of the Tubulin-Folding Cofactor C in Arabidopsis thaliana. Current Biology 12:1519–1523.

Kirik V, Schnittger A, Radchuk V, Adler K, Hülskamp M, Bäumlein H. 2001. Ectopic Expression of the Arabidopsis AtMYB23 Gene Induces Differentiation of Trichome Cells. Developmental Biology 235:366–377.

Kirik V. 2005. Functional diversification of MYB23 and GL1 genes in trichome morphogenesis and initiation. Development 132:1477–1485.

Larkin JC, Young N, Prigge M, Marks MD. 1996. The control of trichome spacing and number in Arabidopsis. Development 122:997–1005.

Le J, Mallery EL, Zhang C, Brankle S, Szymanski DB. 2006. Arabidopsis BRICK1/HSPC300 Is an Essential WAVE-Complex Subunit that Selectively Stabilizes the Arp2/3 Activator SCAR2. Current Biology 16:895–901.

Liang G, He H, Li Y, Ai Q, Yu D. 2014. MYB82 functions in regulation of trichome development in Arabidopsis. Journal of Experimental Botany 65:3215–3223.

Lin D, Cao L, Zhou Z, Zhu L, Ehrhardt D, Yang Z, Fu Y. 2013. Rho GTPase Signaling Activates Microtubule Severing to Promote Microtubule Ordering in Arabidopsis. Current Biology 23:290–297.

Lindeboom J, Mulder BM, Vos JW, Ketelaar T, EMONS AMC. 2008. Cellulose microfibril deposition: coordinated activity at the plant plasma membrane. Journal of Microscopy 231:192–200.

Maes L, Inzé D, Goossens A. 2008. Functional Specialization of the TRANSPARENT TESTA GLABRA1 Network Allows Differential Hormonal Control of Laminal and Marginal

Trichome Initiation in Arabidopsis Rosette Leaves. Plant Physiology 148:1453–1464.

Marks M. 1998. Molecular genetic analysis of trichome development in *Arabidopsis*. Annu. Rev. Plant Biol. 48:137–163.

Marks MD, Wenger JP, Gilding E, Jilk R, Dixon RA. 2009. Transcriptome analysis of *Arabidopsis* wild-type and *gl3-sst sim* trichomes identifies four additional genes required for trichome development. Molecular Plant 2:803–822.

Marks MD. 1994. Plant development: the making of a plant hair. Current Biology 4:621–623.

Marks MD. 2015. Arabidopsis trichome morphogenesis and the role of microtubules in controlling trichome branch formation. In: Fukuda H, editor. Plant cell wall patterning and cell shape. 1st ed. Hoboken, NJ, USA: John Wiley & Sons, Inc. pp. 269–285.

Mathur J, Chua N-H. 2000. Microtubule stabilization leads to growth reorientation in Arabidopsis trichomes. THE PLANT CELL ONLINE 12:465–477.

Mathur J, Spielhofer P, Kost B, Chua N. 1999. The actin cytoskeleton is required to elaborate and maintain spatial patterning during trichome cell morphogenesis in Arabidopsis thaliana. Development 126:5559–5568.

Mathur J. 2004. Cell shape development in plants. Trends in Plant Science 9:583–590.

Mathur J. 2006. Local interactions shape plant cells. Curr. Opin. Cell Biol. 18:40–46.

Melaragno JE, Mehrotra B, Coleman AW. 1993. Relationship between Endopolyploidy and Cell Size in Epidermal Tissue of Arabidopsis. THE PLANT CELL ONLINE 5:1661–1668.

Minamisawa N, Sato M, Cho KH, Ueno H, Takechi K, Kajikawa M, Yamato KT, Ohyama K, Toyooka K, Kim GT, et al. 2011. ANGUSTIFOLIA, a plant homolog of CtBP/BARS, functions outside the nucleus. Plant J. 68:788–799.

Morohashi K, Grotewold E. 2009. A Systems Approach Reveals Regulatory Circuitry for Arabidopsis Trichome Initiation by the GL3 and GL1 Selectors. PLoS Genet 5:e1000396.

Morohashi K, Zhao M, Yang M, Read B, Lloyd A, Lamb R, Grotewold E. 2007. Participation of the Arabidopsis bHLH Factor GL3 in Trichome Initiation Regulatory Events. Plant Physiology 145:736–746.

Nakamura M, Hashimoto T. 2009. A mutation in the Arabidopsis γ-tubulin-containing complex causes helical growth and abnormal microtubule branching. J. Cell. Sci. 122:2208–2217.

Narasimhulu SB, Reddy ASN. 1998. Characterization of Microtubule Binding Domains in the Arabidopsis Kinesin-like Calmodulin Binding Protein. The Plant Cell 10:957–965.

Oda Y, Fukuda H. 2012. Initiation of Cell Wall Pattern by a Rho- and Microtubule-Driven Symmetry Breaking. Science 337:1333–1336.

Oda Y. 2015. Cortical microtubule rearrangements and cell wall patterning. Frontiers in Plant Science 6:236.

Ojangu E-L, Tanner K, Pata P, Järve K, Holweg CL, Truve E, Paves H. 2012. Myosins XI-K, XI-1, and XI-2 are required for development of pavement cells, trichomes, and stigmatic papillae in Arabidopsis. BMC Plant Biol 12:81.

Oppenheimer DG, Herman PL, Sivakumaran S, Esch J, Marks MD. 1991. A myb gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules. Cell 67:483–493.

Oppenheimer DG, Pollock MA, Vacik J, Szymanski DB, Ericson B, Feldmann K, Marks MD. 1997. Essential role of a kinesin-like protein in *Arabidopsis* trichome morphogenesis. Proc. Natl. Acad. Sci. USA 94:6261–6266.

Oppenheimer DG. 1998. Genetics of plant cell shape. Current Opinion in Plant Biology 1:520–524.

Paredez AR, Somerville CR, Ehrhardt DW. 2006. Visualization of Cellulose Synthase Demonstrates Functional Association with Microtubules. Science 312:1491–1495.

Pattanaik S, Patra B, Singh SK, Yuan L. 2014. An overview of the gene regulatory network controlling trichome development in the model plant, Arabidopsis. Frontiers in Plant Science 5:259.

Payne CT, Zhang F, Lloyd AM. 2000. GL3 encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. Genetics 156:1349–1362.

Perazza D, Herzog M, Hülskamp M, Brown S, Dorne A-M, Bonneville J-M. 1999. Trichome Cell Growth in Arabidopsis thaliana Can Be Derepressed by Mutations in at Least Five Genes. Genetics 152:461–476.

Pesch M, Hülskamp M. 2009. One, two, three…models for trichome patterning in Arabidopsis? Current Opinion in Plant Biology 12:587–592.

Petrášek J, Schwarzerová K. 2009. Actin and microtubule cytoskeleton interactions. Current Opinion in Plant Biology 12:728–734.

Qiu JL. 2002. The Arabidopsis SPIKE1 Gene Is Required for Normal Cell Shape Control and Tissue Development. THE PLANT CELL ONLINE 14:101–118.

Reddy ASN, Safadi F, Narasimhulu SB, Golovkin M, Hu X. 1996. A Novel Plant Calmodulin-binding Protein with a Kinesin Heavy Chain Motor Domain. J. Biol. Chem. 271:7052–7060.

Reddy VS, Day IS, Thomas T, Reddy ASN. 2004. KIC, a Novel Ca2+ Binding Protein with One EF-Hand Motif, Interacts with a Microtubule Motor Protein and Regulates Trichome Morphogenesis. The Plant Cell 16:185–200.

Reddy VS, Reddy A. 1999. A plant calmodulin-binding motor is part kinesin and part myosin. Bioinformatics 15:1055–1057.

Reddy VS, Reddy ASN. 2002. The Calmodulin-binding Domain from a Plant Kinesin Functions as a Modular Domain in Conferring Ca2+-Calmodulin Regulation to Animal Plus- and Minus-end Kinesins. J. Biol. Chem. 277:48058–48065.

Reiner T, Hoefle C, Huesmann C, Ménesi D, Fehér A, Hückelhoven R. 2015. The Arabidopsis ROP-activated receptor-like cytoplasmic kinase RLCK VI_A3 is involved in control of basal resistance to powdery mildew and trichome branching. Plant Cell Rep 34:457–468.

Rerie WG, Feldmann KA, Marks MD. 1994. The GLABRA2 gene encodes a homeo domain protein required for normal trichome development in Arabidopsis. Genes Dev. 8:1388–1399.

Rollins R, Shaw E. 1973. The genus *Lesquerella* (Cruciferae) in North America. Cambridge, Mass.: Harvard University Press

Rollins RC, Banerjee UC. 1975. Atlas of the trichomes of *Lesquerella* (Cruciferae). Cambridge: Mass.

Rollins RC, Banerjee UC. 1976. Trichomes in studies of the Cruciferae. In: Vaughan JG, Macleod AJ, Jones BMG, editors. The biology and chemistry of the Cruciferae. London: the Biology and Chemistry of the Cruciferae. pp. 145–166.

Roodbarkelari F, Bramsiepe J, Weinl C, Marquardt S, Novák B, Jakoby MJ, Lechner E, Genschik P, Schnittger A. 2010. CULLIN 4-RING FINGER-LIGASE plays a key role in the control of endoreplication cycles in Arabidopsis trichomes. Proc. Natl. Acad. Sci. U.S.A. 107:15275–15280.

Ryu KH, Zheng X, Huang L, Schiefelbein J. 2013. Computational modeling of epidermal cell fate determination systems. Current Opinion in Plant Biology 16:5–10.

Saedler R, Mathur N, Srinivas BP, Kernebeck B, Hülskamp M, Mathur J. 2004. Actin Control Over Microtubules Suggested by DISTORTED2 Encoding the Arabidopsis ARPC2 Subunit Homolog. Plant and Cell Physiology 45:813–822.

Sambade A, Findlay K, Schaffner AR, Lloyd CW, Buschmann H. 2014. Actin-Dependent and -Independent Functions of Cortical Microtubules in the Differentiation of Arabidopsis Leaf Trichomes. THE PLANT CELL ONLINE 26:1629–1644.

Sampathkumar A, Lindeboom JJ, Debolt S, Gutierrez R, Ehrhardt DW, Ketelaar T, Persson S. 2011. Live Cell Imaging Reveals Structural Associations between the Actin and Microtubule Cytoskeleton in Arabidopsis. The Plant Cell 23:2302–2313.

Schnittger A, Hülskamp M. 2002. Trichome morphogenesis: a cell-cycle perspective. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 357:823–826.

Shaw SL. 2013. Reorganization of the plant cortical microtubule array. Current Opinion in Plant

Biology 16:693–697.

Singh MK, Ren F, Giesemann T, Bosco CD, Pasternak TP, Blein T, Ruperti B, Schmidt G, Aktories K, Molendijk AJ, et al. 2013. Modification of plant Rac/Rop GTPase signalling using bacterial toxin transgenes. Plant J. 73:314–324.

Smith LG, Oppenheimer DG. 2005. Spatial control of cell expansion by the plant cytoskeleton. Annu. Rev. Cell Dev. Biol. 21:271–295.

Sonoda Y, Sako K, Maki Y, Yamazaki N, Yamamoto H, Ikeda A, Yamaguchi J. 2009. Regulation of leaf organ size by the Arabidopsis RPT2a 19S proteasome subunit. Plant J. 60:68–78.

Symonds VV, Hatlestad G, Lloyd AM. 2011. Natural Allelic Variation Defines a Role for ATMYC1 : Trichome Cell Fate Determination. PLoS Genet 7:e1002069.

Szymanski DB, Jilk RA, Pollock SM, Marks MD. 1998. Control of GL2 expression in Arabidopsis leaves and trichomes. Development 125:1161–1171.

Szymanski DB, Marks MD, Wick SM. 1999. Organized F-Actin Is Essential for Normal Trichome Morphogenesis in Arabidopsis. The Plant Cell 11:2331–2347.

Szymanski DB. 2005. Breaking the WAVE complex: the point of Arabidopsis trichomes. Current Opinion in Plant Biology 8:103–112.

Szymanski DB. 2009. Plant cells taking shape: new insights into cytoplasmic control. Current Opinion in Plant Biology 12:735–744.

Tian J, Han L, Feng Z, Wang G, Liu W, Ma Y, Yu Y, Kong Z, McCormick S. 2015. Orchestration of microtubules and the actin cytoskeleton in trichome cell shape determination by a plant-unique kinesin. McCormick S, editor. eLife Sciences 4:e09351.

Tominaga-Wada R, Ishida T, Wada T. 2011. New Insights into the Mechanism of Development of *Arabidopsis* Root Hairs and Trichomes. In: Jeon KW, editor. International Review of Cell and Molecular Biology. Vol. 286. Elsevier Inc. pp. 67–106.

Uhrig JF, Mutondo M, Zimmermann I, Deeks MJ, Machesky LM, Thomas P, Uhrig S, Rambke C, Hussey PJ, Hülskamp M. 2007. The role of Arabidopsis SCAR genes in ARP2-ARP3-dependent cell morphogenesis. Development 134:967–977.

Wagner GJ, Wang E, Shepherd RW. 2004. New approaches for studying and exploiting an old protuberance, the plant trichome. Annals of Botany 93:3–11.

Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD, Gray JC. 1999. The TRANSPARENT TESTA GLABRA1 Locus, Which Regulates Trichome Differentiation and Anthocyanin Biosynthesis in Arabidopsis, Encodes a WD40 Repeat Protein. The Plant Cell 11:1337–1349.

Walker JD, Oppenheimer DG, Concienne J, Larkin JC. 2000. SIAMESE, a gene controlling the endoreduplication cell cycle in Arabidopsis thaliana trichomes. Development 127:3931–3940.

Wang S, Barron C, Schiefelbein J, Chen J-G. 2010. Distinct relationships between GLABRA2 and single-repeat R3 MYB transcription factors in the regulation of trichome and root hair patterning in Arabidopsis. New Phytol. 185:387–400.

Wang S, Chen J-G. 2008. Arabidopsis Transient Expression Analysis Reveals that Activation of GLABRA2 May Require Concurrent Binding of GLABRA1 and GLABRA3 to the Promoter of GLABRA2. Plant and Cell Physiology 49:1792–1804.

Wang S, Chen J-G. 2014. Regulation of cell fate determination by single-repeat R3 MYB transcription factors in Arabidopsis. Frontiers in Plant Science 5:133.

Werker E. 2000. Trichome diversity and development. Advances in botanical research 31:1–35.

Xu T, Wen M, Nagawa S, Fu Y, Chen J-G, Wu M-J, Perrot-Rechenmann C, Friml J, Jones AM, Yang Z. 2010. Cell Surface- and Rho GTPase-Based Auxin Signaling Controls Cellular Interdigitation in Arabidopsis. Cell 143:99–110.

Yalovsky S, Bloch D, Sorek N, Kost B. 2008. Regulation of Membrane Trafficking, Cytoskeleton Dynamics, and Cell Polarity by ROP/RAC GTPases. Plant Physiology 147:1527–1543.

Yanagisawa M, Desyatova AS, Belteton SA, Mallery EL, Turner JA, Szymanski DB. 2015. Patterning mechanisms of cytoskeletal and cell wall systems during leaf trichome morphogenesis. Nature Plants 1:1–8.

Yanagisawa M, Zhang C, Szymanski D. 2013. ARP2/3-dependent growth in the plant kingdom: SCARs for life. Frontiers in plant science 4:1–12.

Yang C, Ye Z. 2013. Trichomes as models for studying plant cell differentiation. Cellular and Molecular Life Sciences 70:1937–1948.

Yang Z. 2008. Cell Polarity Signaling in Arabidopsis. Annu. Rev. Cell Dev. Biol. 24:551–575.

Zhang C, Mallery EL, Schlueter J, Huang S, Fan Y, Brankle S, Staiger CJ, Szymanski DB. 2008. Arabidopsis SCARs Function Interchangeably to Meet Actin-Related Protein 2/3 Activation Thresholds during Morphogenesis. The Plant Cell 20:995–1011.

Zhang F, Gonzalez A, Zhao M, Payne CT, Lloyd A. 2003. A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. Development 130:4859–4869.

Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an Improved Branch-Site Likelihood Method for Detecting Positive Selection at the Molecular Level. Molecular Biology and Evolution 22:2472–2479.

Zhang X, Dyachok J, Krishnakumar S, Smith LG, Oppenheimer DG. 2005. IRREGULAR TRICHOME BRANCH1 in Arabidopsis Encodes a Plant Homolog of the Actin-Related Protein2/3 Complex Activator Scar/WAVE That Regulates Actin and Microtubule Organization. THE PLANT CELL ONLINE 17:2314–2326.

Zhao H, Wang X, Zhu D, Cui S, Li X, Cao Y, Ma L. 2012. A single amino acid substitution in IIIf subfamily of basic helix-loop-helix transcription factor AtMYC1 leads to trichome and root hair patterning defects by abolishing its interaction with partner proteins in Arabidopsis. J. Biol. Chem. 287:14109–14121.

Zhao M, Morohashi K, Hatlestad G, Grotewold E, Lloyd A. 2008. The TTG1-bHLH-MYB complex controls trichome cell fate and patterning through direct targeting of regulatory loci. Development 135:1991–1999.

Chapter II: Changes in selection acting on *BRANCHLESS TRICHOME* in relation to the evolution of stellate trichomes in *Physaria* (Brassicaceae)

Abigail R. Mazie[a], mazie@wisc.edu

David A. Baum[a], dbaum@wisc.edu

[a]Department of Botany, University of Wisconsin-Madison, 132 Birge Hall, 430 Lincoln Drive, Madison, WI 53706

Corresponding author: Abigail R. Mazie, mazie@wisc.edu

ABSTRACT

Cell shape is an important phenotype that can have a direct effect on organismal fitness, yet we know little about the molecular basis of cell shape evolution. Single-celled epidermal trichomes vary greatly among species of Brassicaceae, providing an excellent opportunity to uncover general principles of cell shape evolution. We explored the molecular evolution of *BRANCHLESS TRICHOME* (*BLT*), a candidate gene for the transition from dendritic trichomes, as seen in most species of Brassicaceae, to the distinctive, stellate trichomes of the genus *Physaria*. We used maximum likelihood models of codon evolution to evaluate whether *BLT* experienced positive selection during the initial evolution of stellate trichomes and/or during subsequent evolution in the *Physaria* clade. Our results point to a long-term shift in selective pressure on *Physaria BLT*, and we found clear evidence for positive selection on a subset of *Physaria BLT* codons. Almost all of the 11 sites in the protein with the highest probability of having evolved under positive selection are clustered in a coiled-coil domain, pointing to changes in protein-protein interactions. Thus, our findings suggest that selection acted on BLT to modify its interactions with other proteins during the process of stabilizing development following the abrupt switch from the dendritic form to the stellate form and/or during subsequent diversification of the stellate form. These results point to the need for evolutionary developmental studies of BLT and interacting proteins in *Physaria*.

Keywords: molecular evolution; positive selection; trichome; cell shape; Brassicaceae; *Physaria*

1. INTRODUCTION

One of the great challenges facing biology at the start of the 21st century is identifying the genetic changes responsible for the evolution of derived morphologies. Methods that estimate relative rates of nonsynonymous to synonymous substitutions in protein-coding genes are commonly used to identify an evolutionary history of positive selection (recently reviewed in Anisimova and Kosiol 2009; Delport et al. 2009). Such a molecular evolutionary approach has been used to scan genomes for new candidate genes for species differences (e.g., Kim et al. 2011; Aguileta et al. 2012; Nery et al. 2013) and to explore prior candidate gene hypotheses (e.g., Jarvis et al. 2014; Parker et al. 2014; Tao et al. 2015; Vargas-Pinilla et al. 2015). A gene found to have experienced positive selection in a group that has a derived phenotype, but not in its sister group that lacks the phenotype, is likely to have played a role in the origin or developmental stabilization of the derived phenotype.

Cell shape varies greatly across cell types within an organism and between homologous cell types in different species. These differences are often functionally significant (e.g., Young 2006; Whitney et al. 2011; Chanut-Delalande et al. 2012; Hachet et al. 2012), yet we know relatively little about the genetic changes that have led to the evolution of cell shape in nature. In this study we used a molecular evolutionary approach to explore a potential candidate gene for the evolution of the shape of unicellular trichomes of the mustard family (Brassicaceae). Brassicaceae unicellular trichomes offer an excellent system for studying cell shape evolution because they are externally visible and vary greatly in form among species (Appel and Al-Shehbaz 2003; Al-Shehbaz et al. 2006; Beilstein et al. 2006; Beilstein et al. 2008). Extensive research on trichome development in the model Brassicaceae species *Arabidopsis thaliana* (recently reviewed in Tominaga-Wada et al. 2011; Grebe 2012; Marks 2015) facilitates the

identification of candidate genes for differences in trichome shape. Furthermore, the availability of *A. thaliana* mutants for most genes and the relative ease of transforming these mutants with homologous genes from other species open the possibility of directly testing for interspecies divergence in gene function. However, before embarking upon such experiments it seems prudent to first identify those candidate genes that appear to have experienced altered patterns of molecular evolution coincident with a change in trichome shape. In this study we focus on a dramatic change in trichome shape that occurred in the genus *Physaria* and look for evidence of positive selection on a candidate gene.

Unicellular, branching trichomes in Brassicaceae most commonly have a dendritic (i.e., tree-like) shape, with a pronounced stalk and two or more branches that angle away from the leaf surface (fig. 1A) (Beilstein et al. 2006; Beilstein et al. 2008). A striking departure from the dendritic pattern is found in the genus *Physaria* (tribe Physarieae), whose species all have stellate (i.e., star-like) trichomes, with numerous branches that radiate from a short stalk and lie parallel to the epidermis (fig. 1B) (Rollins and Shaw 1973; Rollins and Banerjee 1975; Rollins and Banerjee 1976). The stellate form is a synapomorphy of *Physaria*, as all other Physarieae genera have dendritic trichomes (Beilstein et al. 2006; Beilstein et al. 2008; Fuentes-Soriano and Al-Shehbaz 2013). Trichomes are known to protect plants from desiccation and solar damage (Wagner et al. 2004; Hauser 2014), and the dense cover of overlapping trichomes seen in many species of *Physaria* has been proposed as an adaptation to the dry, open habitats of the western United States and northern Mexico where *Physaria* is widespread (Rollins and Shaw 1973; Rollins and Banerjee 1976).

The main differences between *Physaria* stellate trichomes and the dendritic trichomes of other members of Physarieae are the length of the trichome stalk and the number and orientation

of primary branches that emerge from the stalk. A promising genetic candidate for the evolution of these differences is *BRANCHLESS TRICHOME* (*BLT*). In *A. thaliana*, *BLT* is a positive regulator of trichome branching, with loss-of-function mutants exhibiting completely unbranched trichomes (Marks et al. 2009; Kasili et al. 2011) and over-expression increasing trichome branching in a dosage dependent manner (Kasili et al. 2011). Although mutations in *BLT* dramatically change trichome shape, other aspects of plant growth and development are unaffected (Kasili et al. 2011).

The majority of the BLT protein folds into a coiled-coil (Marks et al. 2009), a secondary structure associated with mediating protein-protein interactions (Burkhard et al. 2001; Parry et al. 2008). Further, BLT is known to interact with another positive regulator of trichome branching, STICHEL (STI) (Kasili et al. 2011). Like *BLT*, strong *STI* mutants have completely unbranched trichomes, and *STI* over-expression increases trichome branching (Ilgenfritz et al. 2003). Both STI and BLT localize to the tips of emerging and elongating branches (Marks et al. 2009), where they are postulated to form a protein complex that directs the polarity of cell growth (Kasili et al. 2011).

Based on the strong and specific effects of *BLT* on *A. thaliana* trichome branch morphology, we hypothesized that changes at this locus in *Physaria* contributed to the transition from dendritic to stellate trichomes by increasing the number and modifying the orientation of branch points during trichome development. To test this hypothesis we looked for evidence of positive selection on *Physaria BLT*. We found that a subset of *BLT* codons show evidence of positive selection in the *Physaria* clade, with most positively selected sites concentrated in the predicted coiled-coil domain. Our findings are consistent with a role for *BLT* in cell shape

evolution and suggest the need for further functional studies, including interspecies transformation experiments.

2. MATERIALS AND METHODS

2.1 Sequence generation

We identified publicly available *Arabidopsis thaliana BLT* homologs with BLASTN searches of the nucleotide collection, expressed sequence tag, whole genome shotgun, and transcriptome shotgun assembly databases using *A. thaliana BLT* (NM_105144) as query sequence. We used the sequences found in BLASTN searches to design the degenerate primers BLTF1 and BLTR3 (supplementary table 1), which were used to amplify and sequence *Physaria subumbellata*, *P. tenella*, *P. pycnantha*, and *P. nelsonii BLT* homologs. These sequences were then used to design new degenerate primers with improved amplification efficiency (BLT190F and BLT740R, supplementary table 1).

DNA was extracted from fresh or dried leaf tissue with the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). BLT190F/BLT740R primers were used to generate new sequences of *BLT* from members of the three clades of Physarieae identified in Fuentes-Soriano and Al-Shehbaz (2013): the "DDNLS" clade, which includes the genera *Dimorphocarpa*, *Dithyrea*, *Nerisyrenia*, *Lyrocarpa*, and *Synthlipsis*, all with dendritic trichomes; *Paysonia* also with dendritic trichomes; and *Physaria*, with stellate trichomes. In addition to members of Physarieae, we generated new *BLT* sequences of *Berteroa incana*, *Bivonaea lutea*, and *Malcolmia maritima*. We also generated sequences of the internal transcribed spacer (ITS) region of nuclear ribosomal DNA to provide a point of comparison with *BLT* and to confirm species identities. We used

ITS5/ITS4 (White et al. 1990) or ITS.LEU/ITS4 (Baum et al. 1998) primer pairs (supplementary table 1) to amplify this locus.

Most amplifications were performed with the following 25uL reaction mixture: 1X Green GoTaq Flexi Buffer, 2.5mM MgCl2, 0.2mM dNTPs, 0.2µM each primer, and either 0.625u (*BLT*) or 1u (ITS) GoTaq DNA polymerase (Promega Corp., Madison, WI). BSA (0.1mg/mL; New England BioLabs, Ipswich, MA) was added to some reactions to improve amplification efficiency. Typical *BLT* PCR conditions were as follows: initial denaturation of 95°C (2 min); 10 cycles with a 95°C (1 min) denaturation, an initial 60°C (1 min) annealing, which was decreased by 0.5°C per cycle, and a 72°C (1 min) extension; 25 cycles of a 95°C (1 min) denaturation, 50°C (1 min) annealing, and 72°C (1 min) extension; a final 72°C (5 min) extension. Annealing temperatures were decreased by as much as 5°C in a few taxa that were difficult to amplify. ITS PCR conditions were as follows: initial denaturation of 94°C (1 min); 32 cycles with a 94°C (30 sec) denaturation, 60°C (20 sec) annealing, and 72°C (20 sec) extension; a final 72°C (7 min) extension.

Physaria fendleri accession 596455 sequence information was extended by genome walking (GenomeWalker Universal Kit, Clontech, Mountain View, CA). PCR products were extracted with the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) and cloned with the pGEM-T vector system (Promega Corp., Madison, WI) and DH5α competent cells (Invitrogen, Life Technologies, Grand Island, NY). Insert-containing colonies were identified by blue-white screening and amplified with GoTaq DNA Polymerase (Promega Corp., Madison, WI).

PCR products were purified using either magnetic bead (Agencourt AMPure XP, Beckman Coulter, Brea, CA) or enzymatic cleanup (ExoSAP-IT, Affymetrix, Santa Clara, CA) and sequenced with the same primers used for amplification. ITS PCR products that were

amplified with ITS.LEU and ITS4 primers were additionally sequenced with the internal primers ITS2 and ITS3B (Baum et al. 1998). The BigDye Terminator Kit (Applied Biosystems, Life Technologies, Grand Island, NY) was used for sequencing reactions, excess dye was removed with Agencourt CleanSEQ (Beckman Coulter, Brea, CA), and sequences were analyzed with an ABI 3730XL DNA Analyzer at the UW-Madison Biotechnology Center. All sequences generated for this study are deposited in GenBank under accession numbers XX-XX. GenBank accession numbers for individual taxa are listed in supplementary tables 2 and 3.

2.2 Sequence alignment and phylogenetic inference

Newly generated sequences were combined with publicly available sequences to generate *BLT* and ITS datasets for the same set of taxa. Sequences were assembled, edited, and aligned using Geneious Pro (version 5.6.6, Biomatters, Auckland, New Zealand; http://www.geneious.com; Kearse et al. 2012). *BLT* sequences were manually aligned by codon, whereas ITS sequences were aligned with the MUSCLE algorithm (Edgar 2004) and then manually adjusted.

We sequenced several accessions of some species of *Paysonia* (*P. auriculata, P. densipila*, *P. grandiflora*, *P. lasiocarpa, P. stonensis*) and *Physaria* (*P. argyraea, P. fendleri, P. occidentalis*, *P. pycnantha*). In order to remove redundant accessions from our alignments and check taxon identifications, we conducted preliminary phylogenetic analyses of initial alignments. When sequences were similar for all accessions of a species, we only retained the accession with the longest *BLT* sequence length. If all *BLT* sequences were similar in length, we used the accession with the shortest terminal *BLT* branch length. Accessions of a particular species that did not form a clade with other accessions in *BLT* or ITS phylogenies were retained

in the alignments used for molecular evolutionary analyses. Supplementary table 2 lists all taxa used in molecular evolutionary analysis; supplementary table 3 lists the additional Physarieae accessions that were included in preliminary analyses, but not retained. *BLT* and ITS alignments used in molecular evolutionary analyses are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.xxxxx.

Phylogenies were estimated for *BLT*, ITS, and concatenated *BLT*-ITS alignments using maximum likelihood (RAxML-HPC BlackBox 8.1.11; Stamatakis 2014) on the CIPRES Science Gateway Version 3 (citation) and Bayesian methods (MrBayes plugin version 2.0.5 in Geneious Pro version 5.6.6, Biomatters, Auckland, New Zealand; Huelsenbeck and Ronquist 2001). Both analyses used the GTR+gamma model of nucleotide substitution. Maximum likelihood was performed with rapid bootstrapping, with the number of replicates determined automatically. Bayesian MCMC settings were as follows: nruns=2; chain length = 1,100,000; heated chains = 4; heated chain temp = 0.2; subsampling frequency = 200; burn-in length = 100,000; with *Tarenaya hassleriana* (Cleomaceae) identified as the outgroup. Convergence was confirmed by identifying the average standard deviation of split frequencies to be <0.02.

2.3 Analysis of BLT Molecular Evolution

We modeled codon evolution by analyzing the nonsynonymous (d_N) to synonymous (d_S) substitution ratio (ω) with the CODEML program in the Phylogenetic Analysis by Maximum Likelihood package (PAML, version 4.8; Yang 2007). Analyses were performed on the unaltered *BLT* sequence alignment (cleandata = 0). Exploratory analyses with ambiguities and alignment gaps removed (cleandata = 1) showed essentially identical, but less robust, results and will not be reported further. We used the *BLT* single-gene and concatenated *BLT*-ITS maximum likelihood

trees generated with RAxML as input trees. Branch lengths were optimized in PAML, starting from the RAxML-estimated branch lengths. We estimated codon frequencies as a product of nucleotide frequencies at the three codon positions (F3X4 model). The transition/transversion ratio (κ) and *d*N/*d*S ratios (ω) were estimated from the data.

We used branch models (Yang 1998) to assess whether selection acted differently on *BLT* in the *Physaria* clade in comparison to other Brassicaceae lineages. Likelihood ratio tests were used to compare the fit of the models to the data. We used the branch-site test of positive selection (Zhang et al. 2005) to determine whether a subset of *BLT* codons evolved under positive selection in *Physaria*. This test compares a model in which a subset of codon sites is allowed to have evolved under positive selection in a designated lineage to a null model where codon sites have only evolved under purifying selection or neutrally. The designated lineage is labeled the foreground, while all other lineages in the phylogeny are labeled as the background. The program estimates the proportion of sites that are under purifying selection in both background and foreground (site class 0), neutrally evolving in both background and foreground (site class 1), under purifying selection in the background and positive selection in the foreground (site class 2a), and neutrally evolving in the background and under positive selection in the foreground (site class 2b). A likelihood ratio test compared the full model to a null model where ω was capped at 1.0, meaning no positive selection in the foreground. We also reversed the foreground and background lineage designations to identify sites under positive selection in other members of Brassicaceae while under purifying selection or neutrally evolving in *Physaria*. A Bayes Empirical Bayes technique was used to identify the posterior probability that a particular site evolved under positive selection (Yang et al. 2005).

2.4 Protein structure analysis

We translated *BLT* sequences of *Physaria* (*P. johnstonii*, *P. arizonica*, *P. tenella*, *P. bellii*, *P. newberryi*, *P. eriocarpa*, *P. nelsonii*, *P. pallida*, *P. mexicana*, *P. fendleri* 596455), *Paysonia* (*P. grandiflora*, *P. densipila* 355035), *Lyrocarpa coulteri*, *Nerisyrenia johnstonii*, and *Arabidopsis thaliana* into their corresponding amino acid residues with Geneious Pro (version 5.6.6, Biomatters, Auckland, New Zealand; http://www.geneious.com; Kearse et al. 2012). Each sequence was queried for the presence of a coiled-coil domain with the web-interface of COILS (Lupas et al. 1991; Lupas 1996) on the Bioinformatics Toolkit Platform at the Max-Plank Institute for Developmental Biology (Biegert et al. 2006). COILS uses a database of known coiled-coils to calculate the probability that a sequence region will adopt a coiled-coil conformation. We used the MTIDK scoring option without weighting. All residues with >50% probability of belonging to a coiled-coil domain in the 28 residue scanning window (as recommended by Lupas 1996) were identified.

3. RESULTS

3.1 Sequence characteristics

To assess if positive selection occurred on *BLT* during the evolution of *Physaria* stellate trichomes, we estimated the phylogeny of putative *BLT* homologs. We generated or obtained published sequences from taxa that spanned the phylogenetic breadth of the Brassicaceae, but heavily concentrated in the tribe Physarieae. This included samples from Brassicaceae lineages 1-3 (Beilstein et al. 2006), as well as Aethionemeae, which is sister to the rest of the family (Al-Shehbaz 2012), and Cleomaceae, the family most closely related to the Brassicaceae (Hall et al. 2002; Hall et al. 2004). Beyond the published data, we generated 59 new *BLT* sequences

(supplementary tables 2,3). The final *BLT* alignment (after deletion of duplicate sequences from individual species clades) includes 58 sequences and is 954 bps in length. It covers the entire coding region of *BLT*, except approximately 250 bps at the 5' end and 100 bps at the 3' end, which are missing from most taxa because of the internal position of the primers. As a consequence of this missing sequence data and internal indels, Physarieae *BLT* sequences we analyzed are typically about 450 bp long (range = 375-771 bps). The entire matrix has 361 parsimony informative sites, a minimum pairwise nucleotide identity of 65.3%, and a minimum amino acid identity of 50.4%. Within Physarieae there are 111 informative sites, a minimum pairwise nucleotide identity of 87.3% and a minimum amino acid identity of 75.3%.

To provide a point of comparison with *BLT* and check species identification, we also conducted a phylogenetic analysis of ITS for the same set of taxa. We chose ITS because it is the most widely used locus in Brassicaceae phylogenetic studies (Warwick et al. 2010; Kiefer et al. 2014), which allowed us to compare publicly available sequences to the accessions used in our study. The final ITS alignment includes the ITS1, 5.8S, and ITS2 regions and is 709 bps in length. Due to missing data and indels, Physarieae ITS sequences are typically about 620 bp long (range = 556-624 bps). The entire ITS matrix has 281 parsimony informative sites and a minimum pairwise nucleotide identity of 67.1%. Within Physarieae there are 127 informative sites and a minimum pairwise nucleotide identity of 85.6%.

3.2 Phylogenetic trees

We used maximum likelihood and Bayesian methods to infer the evolutionary history of Brassicaceae *BLT* and ITS (figs. 2 and 3). There is substantial agreement between the two trees, with few well-supported differences. Physarieae is strongly supported as a monophyletic group,

with 100% maximum likelihood bootstrap support (MLBS) and a 1.0 posterior probability (PP) for both *BLT* and ITS. Members of other Lineage 1 tribes, Camelineae (*A. thaliana, A. lyrata, Capsella rubella*, *Camelina sativa*) and Boechereae (*Boechera gunnisoniana*), are the closest relatives of Physarieae. The clade that includes all three tribes has relatively high support with *BLT* (MLBS: 100%, PP: 1.0) and ITS (MLBS: 68%, PP: 0.99). *Malcolmia maritima* (Lineage 3) is also included in this clade for *BLT*, but this relationship is not supported with ITS. Within Physarieae, both genes strongly support three monophyletic clades: (1) a group of five small genera, *Dimorphocarpa*, *Dithyrea*, *Nerisyrenia*, *Lyrocarpa*, and *Synthlipsis*, the "DDNLS clade" (*BLT* MLBS: 99%, PP: 1.0; ITS MLBS: 100%, PP: 1.0), (2) *Physaria* (*BLT* MLBS: 100%, PP: 1.0; ITS MLBS: 100%, PP: 1.0), and (3) *Paysonia* (*BLT* MLBS: 95%, PP: 1.0; ITS MLBS: 99%, PP: 1.0). Both genes support, albeit weakly, a *Physaria-Paysonia* clade.

Within *Physaria* and *Paysonia*, minor differences are found in the relationships between species, depending on whether the phylogeny was inferred with *BLT* or ITS. However, only two disagreements are strongly supported: (1) the placement of *P. pycnantha* as sister to *P. nelsonii* with *BLT* (MLBS: 100%, PP: 1.0), but in a clade with *P. goodrichii* and *P. humilis* with ITS (MLBS: 52%, PP: 0.95) and (2) the placement of *P. arizonica* as sister to *P. argyraea* 643169 (MLBS: 100%, PP: 1.0) in a clade with *P. bellii*, *P. acutifolia*, and *P. newberryi* (MLBS: 68%, PP: 1.0) with *BLT*, but sister to *P. intermedia* (MLBS: 99%, PP: 1.0) in a clade with *P. tenella* (MLBS: 76%, PP: 0.98) with ITS. Despite these minor differences, we analyzed a concatenated *BLT*-ITS alignment. The concatenated phylogeny (fig. 4) provides stronger support for relationships within Physarieae, including the sister relationship between *Physaria* and *Paysonia* (MLBS: 86%, PP: 0.99). All molecular evolutionary analyses were therefore performed on both the *BLT* single-gene phylogeny and the combined *BLT*-ITS phylogeny.

A few accessions nominally assigned to the same species do not form clades in the combined analysis. Although correct identification of species is not important for our molecular evolutionary conclusions, several accessions were obtained through the USDA's National Plant Germplasm System, which is widely used, so we discuss these taxonomic issues in Supplementary Materials (supplementary information 1).

3.3 Branch Models

Is the presence of stellate trichomes in *Physaria* associated with changes in the direction and strength of selection acting on *BLT*, a gene known from *A. thaliana* to play a role in trichome branching? To answer this question, we analyzed the nonsynonymous (d_N) to synonymous (d_S) substitution ratio (ω) using the CODEML program in the PAML package (Yang 2007). The numerical value of ω indicates the direction of selection on the protein. Specifically, proteins under purifying selection have ω<1 because nonsynonymous substitutions often disrupt protein function and are disfavored by selection. Conversely, some nonsynonymous changes that alter protein function are favored by selection; proteins with favored nonsynonymous changes will have ω>1 and are considered to be under positive selection. Neutrally evolving proteins have ω approximately equal to 1, as neither type of substitution is favored.

We used the branch models of Yang (1998) to uncover selection occurring in the *Physaria* lineage. We first estimated ω across both the *BLT* and combined *BLT*-ITS phylogenies (Model A, fig. 5). This estimate effectively averages ω across all branches of the tree and all codon sites in the alignment. The one-ratio model shows that, in net, *BLT* is under purifying selection, ω=0.24, as expected for a gene that codes for a functionally important protein (table 1). To see if selection acted differently during or after the evolution of stellate trichomes, we

estimated ω separately for the entire *Physaria* clade (including its stem lineage) and all other lineages on the tree (Model B, fig. 5). Under Model B, *Physaria* has an estimated ω of 0.39 (*BLT*) or 0.37 (*BLT*-ITS), as contrasted with a ω of 0.22 (both trees) for the remaining taxa (table 1). Regardless of which tree was considered, the likelihood score of the two-ratio model was significantly higher than the one-ratio model, as judged with a likelihood ratio test (LRT: 1 degree of freedom, $p<0.001$). This shows that selection has acted differently on *BLT* in *Physaria* compared to outgroups and, specifically, that the proportion of non-synonymous changes is higher in *Physaria* than in other Brassicaceae.

To assess whether the elevated ω seen in the *Physaria BLT* clade could be due to an episodic burst of selection along the *Physaria* stem branch, we allowed this branch to have its own unique ω, which was estimated separately from all other branches on the tree (Model C, fig. 5). Although, the ω of the *Physaria* stem branch was estimated to be higher (ω=0.30, *BLT*, or 0.27, *BLT*-ITS) than other lineages (ω=0.24, both trees), a likelihood ratio test showed that Model C is not significantly better than Model A, where one ω was averaged across the entire tree (table 1). The complementary two-ratio test (Model D, fig. 5), which estimated ω separately for the *Physaria* crown lineage and all other branches on the tree, provided a significantly higher likelihood score than Model A (LRT: 1 degree of freedom, $p<0.001$). To explore this further, we estimated ω under model E, which allowed a different ω for the *Physaria* crown lineage, the *Physaria* stem branch, and the remaining lineages (fig. 5). As might be expected, the estimated ω of the *Physaria* crown (ω=0.40, *BLT*, or 0.38, *BLT*-ITS) and stem (ω=0.29, *BLT*, or 0.26, *BLT*-ITS) was increased in comparison to the remaining lineages (table 1). Nonetheless, Model E did not fit the data significantly better than the best fitting two-ratio model (Model D) suggesting that the change in selective pressure primarily occurred during the radiation of *Physaria*, after

stellate trichomes had already arisen, rather than coincident with the evolution of stellate trichomes. On the other hand, we cannot rule out the possibility that stellate trichomes arose very close to the crown node of *Physaria*, such that the proportion of the stem lineage subject to the altered selection was too little to be detected.

We wanted to confirm that the change in selective pressure was a feature of the *Physaria* clade, rather than a consequence of an elevated ω ratio for Physarieae as a whole. We therefore compared Model B to a three-ratio model (Model F, fig. 5) where ω was estimated separately for *Physaria* (ω=0.39, *BLT*, or 0.37, *BLT*-ITS), other members of Physarieae included in our analysis (*Paysonia, Lyrocarpa, Nerisyrenia, and Synthlipsis*, ω=0.19, both trees), and the remaining outgroup lineages (ω=0.22, *BLT*, or 0.23, *BLT*-ITS) (table 1). A likelihood-ratio test did not support the addition of a separate ratio for the latter two clades, thus suggesting that the elevated ω value seen in *Physaria* is specific to this clade, not to the broader Physarieae.

3.4 Branch-site Models

The elevated ω found for *Physaria* with branch models could be a result of relaxed selection on *Physaria BLT*, positive selection on a subset of sites within the coding region of the gene, or both. To distinguish between these possibilities, we used the branch-site test of positive selection, which compares a model that allows a subset of codon sites to have evolved under positive selection in a designated lineage to a null model where codon sites have only evolved under purifying selection or neutrally (Zhang et al. 2005). We designated the entire *Physaria* clade, but not its stem branch, as the foreground lineage, as this model, Model D, was the best-fitting branch model (table 1). We found evidence of positive selection on a subset of *Physaria BLT* codons, regardless of tree used. A likelihood ratio test judged the full model (LogL:

−8875.16, *BLT*; −8938.81, *BLT*-ITS) to fit the data significantly better than a null model (Log-L: −8891.54, *BLT*; −8958.63, *BLT*-ITS) that did not allow for positive selection (ω capped at 1.0; LRT: 1 degree of freedom, $p<0.0001$). Most codon sites were estimated to be under purifying selection (table 2, site class 0) or neutrally evolving (table 2, site class 1) across the phylogeny, but a subset (8-13%) was estimated to be under relatively strong positive selection in *Physaria* (table 2, site classes 2a, b).

In the previously described model, *Physaria* was designated as the foreground lineage, whereas the rest of the tree was designated as the background. We flipped these designations (*Physaria*, including its stem branch, as the background lineage; the rest of the taxa as foreground) to identify any sites under positive selection in other Brassicaceae taxa that are evolving under purifying or neutral selection in *Physaria*. We failed to find evidence of any sites evolving under positive selection, as ω was never estimated to be >1 in the foreground lineage even though its value was unconstrained (data not shown).

3.5 Protein evolution

In addition to estimating the proportion of sites under positive selection across *Physaria BLT*, we used a Bayes Empirical Bayes approach (Yang et al. 2005) to assign a posterior probability to the hypothesis that each site in the alignment is evolving under purifying selection, positive selection, or neutrally in *Physaria*. We identified 11 codon sites with >0.95 posterior probability of having experienced positive selection in either the *BLT* single-gene tree or the combined *BLT*-ITS tree (table 3). All but two of these sites (position 156 and 255) had >0.97 posterior probability with both trees. We also identified four amino acid sites that are unique to *Physaria* and shared among all (or nearly all) *Physaria* taxa studied (table 4).

A comparison of the *Physaria* amino acid state to that of the outgroup taxa shows that the chemical properties of the *Physaria*-specific and widespread residues are similar, with the following exceptions: the replacement of neutral glutamine with positively charged arginine and lysine (position 167); the replacement of negatively charged glutamic acid or aspartic acid with neutral asparagine (position 173); and the replacement of hydrophobic valine with positively charged arginine (position 174) (table 3). In addition, at position 151, negatively charged glutamic acid has been replaced with hydrophobic alanine (table 4). Interestingly, alanine is also found in one outgroup taxon known to have stellate trichomes, *Berteroa incana*. Despite chemical conservatism, *Physaria* residues under positive selection show a general tendency to have larger side-groups than their widespread counterparts: threonine for serine (position 153), isoleucine for valine (position 156), tryptophan for leucine (position 175), and phenylalanine for leucine (position 191).

Although we are unable to predict the significance of these amino acid substitutions on BLT secondary structure or interactions with other proteins, codon positions inferred to be subject to positive selection in *Physaria* might reflect changes in protein function associated with the transition to stellate trichomes. Most of the internal amino acid sequence of *A. thaliana* BLT is predicted to form a coiled-coil secondary structure (Marks et al. 2009). We were interested in determining if this domain is also present in Physarieae BLT homologs and, if so, where it is situated relative to sites under putative positive selection in *Physaria*. Coiled-coil regions were predicted in every BLT sequence we analyzed, including ten members of *Physaria*, two *Paysonia*, and two members of the DDNLS clade. The predicted Physarieae coiled-coil domains closely matched the domain in *A. thaliana*, though the C-terminal end-point was shifted upstream in a few taxa, implying shorter coiled-coil regions in *Lyrocarpa coulteri*, by 10

residues, and three out of the ten *Physaria* taxa analyzed (*P. nelsonii*, *P. mexicana*, and *P. fendleri*), by 25 residues. Given this, we were interested to see how positively selected sites were distributed across the alignment, so we graphed the posterior probability that a codon site was under positive selection given the *BLT* single-gene tree alignment (fig. 6). The results of this analysis show a cluster of sites with high posterior probabilities of being positively selected around codon position 150 to 200, which falls within the predicted coiled-coil domains.

4. DISCUSSION

Our phylogenetic analyses of *BLT* and ITS confirmed the monophyly of Physarieae and its three main clades, *Physaria*, *Paysonia*, and DDNLS, as well as the existence of a *Physaria-Paysonia* clade, as found in a previous analysis based on plastid markers (Fuentes-Soriano and Al-Shehbaz 2013) and family-wide analyses with limited sampling of Physarieae taxa (Bailey et al. 2006; Beilstein et al. 2006; Beilstein et al. 2008; Couvreur et al. 2010; Warwick et al. 2010). Our analyses further support the hypothesis that the ancestral trichome branching pattern of the tribe was dendritic, with stellate trichomes having evolved on the stem lineage of *Physaria* (Fuentes-Soriano and Al-Shehbaz 2013).

A primary difference between dendritic and stellate trichomes is the position of trichome branches, and *A. thaliana BLT* is known to play an important role in trichome branch initiation (Marks et al. 2009; Kasili et al. 2011). This led us to predict that the *Physaria BLT* gene might have been subject to a different pattern of selection during the transition to stellate trichomes and/or during the subsequent diversification of the stellate clade. Using branch-based models of codon evolution (Yang 1998), we found that the BLT protein as a whole is under purifying selection across all Brassicaceae taxa, including *Physaria*. This result is not surprising given the

presence of *BLT* in all species, which suggests the existence of conserved functions that might be disrupted by non-synonymous mutations.

Likelihood ratio tests showed that the optimal model of molecular evolution was a two-ratio model in which the *Physaria BLT* crown lineage has a different, and higher, ω value than outgroup taxa. The most likely explanation for this change in selective pressure is that, subsequent to the evolution of stellate trichomes, the selective regime acting on *BLT* differed from that seen in clades with dendritic trichomes. Models that allowed the *Physaria* stem lineage to have its own ω, although not significantly improving the overall fit to the data, estimated a higher ω for the *Physaria* stem lineage than outgroups. This suggests that at least the latter part of the *Physaria* stem lineage experienced a similar selective regime to that experienced in the crown clade, supporting the idea that the change in selective regime coincided with the origin of stellate trichomes.

We used the branch-site test of positive selection (Zhang et al. 2005) to show that the difference in ω is at least partly due to a subset of *BLT* codons evolving under positive selection in *Physaria* (but not in outgroups). Specifically, we estimated that 10% of codons experienced positive selection in *Physaria*, and these were primarily located in a predicted coiled-coil domain, a secondary structure known to mediate protein-protein interactions (Burkhard et al. 2001; Parry et al. 2008). The bias towards non-synonymous substitutions in the BLT coiled-coil domain suggests that selection favored weaker or stronger interactions between BLT and its protein-partners.

Our analyses showed that although BLT protein functionality has been maintained through purifying selection across all members of Brassicaceae, following the origin of stellate trichomes, some sites in *Physaria* came under directional selection for substitutions in formerly

conserved amino acids. Three scenarios could explain why the shift in selective pressure on *BLT* can be detected in the entire *Physaria* lineage, rather than only in the lineage on which the stellate form originally evolved. One possibility is that after the basic stellate form was established, there was selection within *Physaria* for the diversification of stellate trichome form, and this resulted in selection for amino acid substitutions in BLT. Substantial trichome diversity is seen across species of *Physaria*, including differences in the number of primary and secondary branches and the extent of branch fusion (Rollins and Shaw 1973; Rollins and Banerjee 1975; Rollins and Banerjee 1976). A second possibility is that the evolution of stellate trichomes pushed the BLT protein sequence off an adaptive optimum, and it took a long time (perhaps through to the present) for protein mutations to restore optimum protein function. This scenario would predict a return to lower levels of ω in highly nested *Physaria* lineages, but this is not apparent in these data. A third possibility is that rewiring of protein-protein interaction networks after the dendritic to stellate transition established reciprocal co-evolution with one or more interacting proteins. The best candidate for an interacting protein is another positive regulator of trichome branching, STICHEL (STI). STI and BLT are thought to form a protein complex that directs sites of trichome branch initiation (Marks et al. 2009; Kasili et al. 2011). The possibility of coevolution between BLT and STI could be tested *in vitro* by comparing the strength and specificity of binding between different *Physaria* and *Paysonia* BLT and STI proteins. It would also be interesting to look at the molecular evolution of *STI* to see if it shows a complementary pattern to *BLT*, which would further suggest that the STI-BLT protein complex played a critical role in the evolution of stellate trichomes.

Ultimately, a full understanding of the transition from dendritic to stellate trichomes in *Physaria* requires functional genetic research on *BLT*, *STI*, and perhaps other trichome

development genes. Such a research program is facilitated by the phylogenetic proximity to *A. thaliana* and by the possibility of introducing genes into *P. fendleri* by transformation (Wang et al. 2008; Chen 2011). Through such analyses of *BLT* and other trichome shape genes in Brassicaceae, we believe that Brassicaceae trichomes will emerge as a premier model system for exploring the evolution of cell shape control in multicellular eukaryotes.

ACKNOWLEDGEMENTS

The authors thank S. Fuentes-Soriano and B. Grady for providing leaf material, J.C. Pires and the USDA National Plant Germplasm System for providing seeds, and UW-Madison undergraduate student, Morgan Sell, for laboratory assistance. A. Mazie was supported through the University of Wisconsin-Madison Botany Department (Davis Summer Fellowship, Judy Croxdale Award for Women in Science, ON and EK Allen Fellowship) and University of Wisconsin-Madison Beta Chapter of Sigma Delta Epsilon-Graduate Women in Science (Ruth Dickie Grant-In-Aid).

LITERATURE CITED

Aguileta G, Lengelle J, Chiapello H, Giraud T, Viaud M, Fournier E, Rodolphe F, Marthey S, Ducasse A, Gendrault A, et al. 2012. Genes under positive selection in a model plant pathogenic fungus, *Botrytis*. Infection, Genetics and Evolution 12:987–996.

Al-Shehbaz IA, Beilstein MA, Kellogg EA. 2006. Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Plant Syst. Evol. 259:89–120.

Al-Shehbaz IA. 2012. A generic and tribal synopsis of the Brassicaceae (Cruciferae). Taxon 61:931–954.

Anisimova M, Kosiol C. 2009. Investigating protein-coding sequence evolution with probabilistic codon substitution models. Molecular Biology and Evolution 26:255–271.

Appel O, Al-Shehbaz IA. 2003. Cruciferae. In: Kubitzki K, editor. Families and genera of vascular plants. Vol. 5. Berlin (Germany): Springer Berlin Heidelberg. pp. 75–174.

Bailey CD, Koch MA, Mayer M, Mummenhoff K, O'Kane SL, Warwick SI, Windham MD, Al-Shehbaz IA. 2006. Toward a global phylogeny of the Brassicaceae. Molecular Biology and Evolution 23:2142–2160.

Baum DA, Small RL, Wendel JF. 1998. Biogeography and floral evolution of Baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. Systematic Biology 47:181–207.

Beilstein MA, Al-Shehbaz IA, Kellogg EA. 2006. Brassicaceae phylogeny and trichome evolution. American Journal of Botany 93:607–619.

Beilstein MA, Al-Shehbaz IA, Mathews S, Kellogg EA. 2008. Brassicaceae phylogeny inferred from phytochrome A and *ndhF* sequence data: tribes and trichomes revisited. American Journal of Botany 95:1307–1327.

Biegert A, Mayer C, Remmert M, Söding J, Lupas AN. 2006. The MPI Bioinformatics Toolkit for protein sequence analysis. Nucleic Acids Research 34:W335–W339.

Burkhard P, Stetefeld J, Strelkov SV. 2001. Coiled coils: a highly versatile protein folding motif. Trends Cell Biol. 11:82–88.

Chanut-Delalande H, Ferrer P, Payre F, Plaza S. 2012. Effectors of tridimensional cell morphogenesis and their evolution. Semin. Cell Dev. Biol. 23:341–349.

Chen GQ. 2011. Effective reduction of chimeric tissue in transgenics for the stable genetic transformation of *Lesquerella fendleri*. HortScience 46:86–90.

Couvreur TLP, Franzke A, Al-Shehbaz IA, Bakker FT, Koch MA, Mummenhoff K. 2010. Molecular phylogenetics, temporal diversification, and principles of evolution in the mustard family (Brassicaceae). Mol. Biol. Evol. 27(1): 55-71.

Delport W, Scheffler K, Seoighe C. 2009. Models of coding sequence evolution. Brief Bioinform 10:97–109.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32:1792–1797.

Fuentes-Soriano S, Al-Shehbaz I. 2013. Phylogenetic relationships of mustards with multiaperturate pollen (Physarieae, Brassicaceae) based on the plastid *ndhF* gene: implications for morphological diversification. Systematic Botany 38:178–191.

Grebe M. 2012. The patterning of epidermal hairs in *Arabidopsis*—updated. Current Opinion in Plant Biology 15:31–37.

Hachet O, Bendezú FO, Martin SG. 2012. Fission yeast: in shape to divide. Curr. Opin. Cell Biol. 24:858–864.

Hall JC, Iltis HH, Sytsma KJ. 2004. Molecular phylogenetics of core Brassicales, placement of orphan genera *Emblingia, Forchhammeria, Tirania,* and character evolution. Systematic Botany 29:654–669.

Hall JC, Sytsma KJ, Iltis HH. 2002. Phylogeny of Capparaceae and Brassicaceae based on chloroplast sequence data. American Journal of Botany 89:1826–1842.

Hauser M-T. 2014. Molecular basis of natural variation and environmental control of trichome patterning. Frontiers in Plant Science 5:1–7.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Ilgenfritz H, Bouyer D, Schnittger A, Mathur J, Kirik V, Schwab B, Chua N-H, Jürgens G, Hülskamp M. 2003. The Arabidopsis *STICHEL* gene is a regulator of trichome branch number and encodes a novel protein. Plant Physiology 131:643–655.

Jarvis DE, Ryu CH, Beilstein MA, Schumaker KS. 2014. Distinct roles for *SOS1* in the convergent evolution of salt tolerance in *Eutrema salsugineum* and *Schrenkiella parvula*. Molecular Biology and Evolution 31:2094–2107.

Kasili R, Huang CC, Walker JD, Simmons LA, Zhou J, Faulk C, Hülskamp M, Larkin JC. 2011. *BRANCHLESS TRICHOMES* links cell shape and cell cycle control in *Arabidopsis* trichomes. Development 138:2379–2388.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

Kiefer M, Schmickl R, German DA, Mandakova T, Lysak MA, Al-Shehbaz IA, Franzke A, Mummenhoff K, Stamatakis A, Koch MA. 2014. BrassiBase: introduction to a novel

knowledge database on Brassicaceae evolution. Plant and Cell Physiology 55:e3(1–9).

Kim KM, Kim K-W, Sung S, Kim H. 2011. A genome-wide identification of genes potentially associated with host specificity of *Brucella* species. J Microbiol. 49:768–775.

Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. Science 252:1162–1164.

Lupas A. 1996. Prediction and analysis of coiled-coil structures. In: Doolittle RF, editor. Methods in enzymology. Vol. 266. Methods in enzymology. pp. 513–525.

Marks MD, Wenger JP, Gilding E, Jilk R, Dixon RA. 2009. Transcriptome analysis of *Arabidopsis* wild-type and *gl3-sst sim* trichomes identifies four additional genes required for trichome development. Molecular Plant 2:803–822.

Marks MD. 2015. Arabidopsis trichome morphogenesis and the role of microtubules in controlling trichome branch formation. In: Fukuda H, editor. Plant cell wall patterning and cell shape. 1st ed. Hoboken, NJ, USA: John Wiley & Sons, Inc. pp. 269–285.

Nery MF, González DJ, Opazo JC. 2013. How to make a dolphin: molecular signature of positive selection in cetacean genome. PLoS ONE 8:e65491.

Parker DJ, Gardiner A, Neville MC, Ritchie MG, Goodwin SF. 2014. The evolution of novelty in conserved genes; evidence of positive selection in the *Drosophila fruitless* gene is localised to alternatively spliced exons. Heredity 112:300–306.

Parry DAD, Fraser RDB, Squire JM. 2008. Fifty years of coiled-coils and α-helical bundles: a close relationship between sequence and structure. Journal of Structural Biology 163:258–269.

Rollins R, Shaw E. 1973. The genus *Lesquerella* (Cruciferae) in North America. Cambridge, Mass.: Harvard University Press

Rollins RC, Banerjee UC. 1975. Atlas of the trichomes of *Lesquerella* (Cruciferae). Cambridge, Mass.: Publication of the Bussey Institute Harvard University.

Rollins RC, Banerjee UC. 1976. Trichomes in studies of the Cruciferae. In: Vaughan JG, Macleod AJ, Jones BMG, editors. The biology and chemistry of the Cruciferae. London: Academic Press, Inc. pp. 145–166.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Tao J, Qi Q, Kang M, Huang H. 2015. Adaptive molecular evolution of *PHYE* in *Primulina*, a Karst cave plant. PLoS ONE 10:e0127821.

Tominaga-Wada R, Ishida T, Wada T. 2011. New insights into the mechanism of development of *Arabidopsis* root hairs and trichomes. In: Jeon KW, editor. International Review of Cell

and Molecular Biology. Vol. 286. Elsevier Inc. pp. 67–106.

Vargas-Pinilla P, Paixão-Côrtes VR, Paré P, Tovo-Rodrigues L, Vieira CM de AG, Xavier A, Comas D, Pissinatti A, Sinigaglia M, Rigo MM, et al. 2015. Evolutionary pattern in the OXT-OXTR system in primates: coevolution and positive selection footprints. Proceedings of the National Academy of Sciences 112:88–93.

Wagner GJ, Wang E, Shepherd RW. 2004. New approaches for studying and exploiting an old protuberance, the plant trichome. Annals of Botany 93:3–11.

Wang W, Wang C, Huang B-L, Huang B. 2008. *Agrobacterium tumefaciens*-mediated transformation of *Lesquerella fendleri* L., a potential new oil crop with rich lesquerolic acid. Plant Cell Tiss Organ Cult 92:165–171.

Warwick SI, Mummenhoff K, Sauder CA, Koch MA, Al-Shehbaz IA. 2010. Closing the gaps: phylogenetic relationships in the Brassicaceae based on DNA sequence data of nuclear ribosomal ITS region. Plant Syst. Evol. 285:209–232.

White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, editors. PCR protocols: a guide to methods and applications. New York: Elsevier. pp. 315–322.

Whitney HM, Bennett KMV, Dorling M, Sandbach L, Prince D, Chittka L, Glover BJ. 2011. Why do so many petals have conical epidermal cells? Annals of Botany 108:609–616.

Yang Z, Wong W, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Molecular Biology and Evolution 22:1107–1118.

Yang Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Molecular Biology and Evolution 15:568–573.

Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Molecular Biology and Evolution 24:1586–1591.

Young KD. 2006. The selective value of bacterial shape. Microbiol. Mol. Biol. Rev. 70:660–703.

Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Molecular Biology and Evolution 22:2472–2479.

Table 1. Branch tests of *BLT* molecular evolution.

Tree	Model[a]	Crown ω	Stem ω	Physarieae ω	Outgroup ω	Log-L	LRT[b]	P-value[c]
BLT	A	0.24	0.24	0.24	0.24	-9108.89		
	B	0.39	0.39	0.22	0.22	-9101.38	B vs. A	<0.001
	C	0.24	0.30	0.24	0.24	-9108.80	C vs. A	NS[d]
	D	0.40	0.22	0.22	0.22	-9101.32	D vs. A	<0.001
	E	0.40	0.29	0.22	0.22	-9101.18	E vs. D	NS
	F	0.39	0.39	0.19	0.22	-9101.00	F vs. B	NS
BLT-ITS	A	0.24	0.24	0.24	0.24	-9179.87		
	B	0.37	0.37	0.22	0.22	-9173.41	B vs. A	<0.001
	C	0.24	0.27	0.24	0.24	-9179.84	C vs. A	NS
	D	0.38	0.22	0.22	0.22	-9173.19	D vs. A	<0.001
	E	0.38	0.26	0.22	0.22	-9173.12	E vs. D	NS
	F	0.37	0.37	0.19	0.23	-9172.91	F vs. B	NS

[a]For model descriptions see text and fig.5.

[b]Likelihood ratio test comparisons

[c]P-values use a X^2 distribution with 1 degree of freedom

[d]Not significant

Table 2. Full model estimates for branch-site tests of positive selection.

Tree	Site class	Proportion	$\omega_{background}$	$\omega_{foreground}$
BLT	0, purifying	0.61	0.13	0.13
	1, neutral	0.26	1.00	1.00
	2a, positive	0.09	0.13	3.22
	2b, positive	0.04	1.00	3.22
BLT-ITS	0, purifying	0.65	0.13	0.13
	1, neutral	0.27	1.00	1.00
	2a, positive	0.06	0.13	4.47
	2b, positive	0.02	1.00	4.47

Table 3. Codon sites with signatures of positive selection in *Physaria* and corresponding unambiguous
amino acids. The percentage of taxa with a particular amino acid is in parentheses. Amino acids found in only one taxon of a group are italicized.

Codon Site	***BLT* PP**	***BLT*-ITS PP**	**Outgroups,** 22 taxa	**Physarieae,** 11 taxa	***Physaria,*** 25 taxa
153	0.996	1.000	Ser (77%), Ala (18%), *Leu*	Ser (100%)	Ser (52%), Thr (44%)
156	0.830	0.983	Val (86%), *His*, *Ile*, *Lys*	Val (91%), *Ala*	Ile (56%), Val (40%)
167	1.000	0.999	Gln (77%), Glu (9%), His (9%), *Lys*	Gln (91%), *Arg*	Gln (48%), Arg (36%), Lys (12%), *Glu*
171	1.000	0.999	Leu (100%)	Leu (91%), *Gap*	Leu (52%), Gap (32%), Ser (8%), *Arg*, *Phe*
173	0.998	0.971	Glu (73%), Asp (27%)	Glu (73%), Asp (18%), *Gap*	Asp (36%), Gap (32%), Asn (28%), *Gly*
174	1.000	1.000	Val (91%), *Arg*, *Asp*	Val (100%)	Val (60%), Arg (24%), Ala (8%), *Asp*
175	1.000	1.000	Leu (100%)	Leu (100%)	Leu (76%), Trp (8%), *Lys*, *Phe*, *Ser*
180	0.999	0.998	Val (100%)	Val (91%), *Gap*	Val (88%), *Ala*, *Asp*, *Ser*
190	0.989	0.996	Phe (91%), *Gap*, *Tyr*	Phe (91%), *Gap*	Phe (68%), Leu (24%), *Ile*, *Val*
191	0.996	0.986	Leu (95%), *Gap*	Met (64%), Leu (27%), *Gap*	Leu (72%), Phe (24%), *Trp*
255	0.963	1.000	Pro (91%), *Arg*, *NI*[a]	Pro (100%)	Pro (40%), Leu (32%), Ser (20%), *Gln, Gap*

[a]no sequence information available for one taxon at this site

Table 4. Codon sites with amino acids unique and shared by the majority of *Physaria*.

Codon Site	**Outgroups** 33 taxa	***Physaria*** 25 taxa
102	Leu[a]	Val
151	Glu[b]	Ala[c]
195	Leu	Ile
278	Ile	Leu

[a]except *B. oleracea*, Met

[b]except *C. rubella*, *C. sativa*, *T. hassleriana*, Asp; *B. incana*, Ala

[c]except *P. arizonica*, Val

Fig. 1. Scanning electron micrograph images of Brassicaceae leaf trichomes. (A) *Paysonia densipila* dendritic trichome. (B) *Physaria newberryi* stellate trichome.

Fig. 2. RAxML maximum likelihood phylogram of *BLT*. Thickened branches have MLBS values of ≥75% and/or PP of ≥0.95 (for comparison, PP are multiplied by 100). MLBS values are indicated above branches, PP are below.

Fig. 3. RAxML maximum likelihood phylogram of ITS. Thickened branches have MLBS values of ≥75% and/or PP of ≥0.95 (for comparison, PP are multiplied by 100). MLBS values are indicated above branches, PP are below.

Fig. 4. RAxML maximum likelihood phylogram of concatenated *BLT*-ITS alignment. Thickened branches have MLBS values of ≥75% and/or PP of ≥0.95 (for comparison, PP are multiplied by 100). MLBS values are indicated above branches, PP are below.

Fig. 5. Summary of alternative branch models considered. Branches assigned the same number for a given model were assumed to share the same ω, whereas branches with different numbers were free to be assigned different ω values.

Fig. 6. Posterior probability that *Physaria BLT* codon is under positive selection. x-axis = codon position in *BLT* aligment. y-axis = posterior probability of positive selection (Full Model, site class 2a and 2b). As a point of comparison, the position of the predicted coiled-coil domain of *A. thaliana* BLT is indicated with a black box.

Fig. 1.

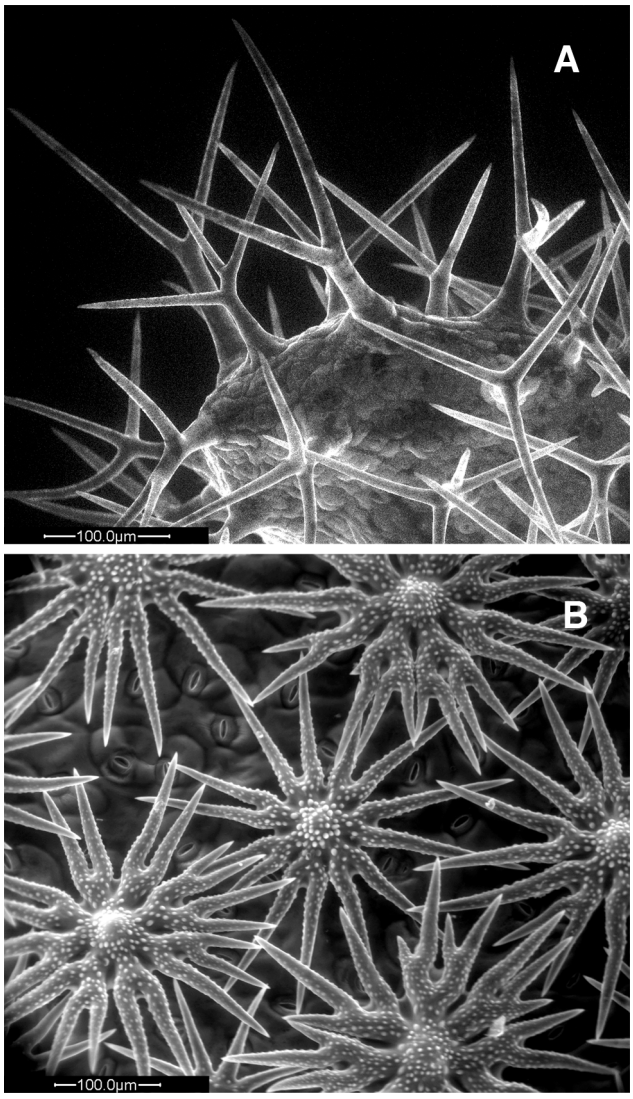


Fig. 2.


Physaria_argyraea_20816
Physaria_fendleri_596455
Physaria_mexicana
Physaria_filiformis
Physaria_pallida
Physaria_subumbellata
Physaria_arizonica
Physaria_argyraea_643169
Physaria_bellii
Physaria_acutifolia
Physaria_newberryi
Physaria_argyraea_20864
Physaria_intermedia
Physaria_pycnantha
Physaria_nelsonii
Physaria_hemiphysaria
Physaria_occidentalis_379
Physaria_goodrichii
Physaria_occidentalis_518
Physaria_humilis
Physaria_tenella
Physaria_fendleri_355037
Physaria_spatulata
Physaria_eriocarpa
Physaria_johnstonii
Paysonia_lasiocarpa
Paysonia_grandiflora
Paysonia_auriculata_345712
Paysonia_auriculata_20320
Paysonia_lescurii
Paysonia_stonensis
Paysonia_densipila_355035
Paysonia_densipila_315051
Synthlipsis_greggii
Lyrocarpa_coulteri
Nerisyrenia_johnstonii
Camelina_sativa
Capsella_rubella
Boechera_gunnisoniana
Arabidopsis_lyrata
Arabidopsis_thaliana
Malcolmia_maritima
Leavenworthia_alabamica
Berteroa_incana
Raphanus_raphanistrum
Raphanus_sativus
Brassica_oleracea
Brassica_rapa
Bivonaea_lutea
Sisymbrium_irio
Schrenkiella_parvula
Arabis_alpina
Noccaea_caerulescens
Thlaspi_arvense
Eutrema_salsugineum
Chorispora_bungeana
Aethionema_arabicum
Tarenaya_hassleriana
0.04

Fig. 3.

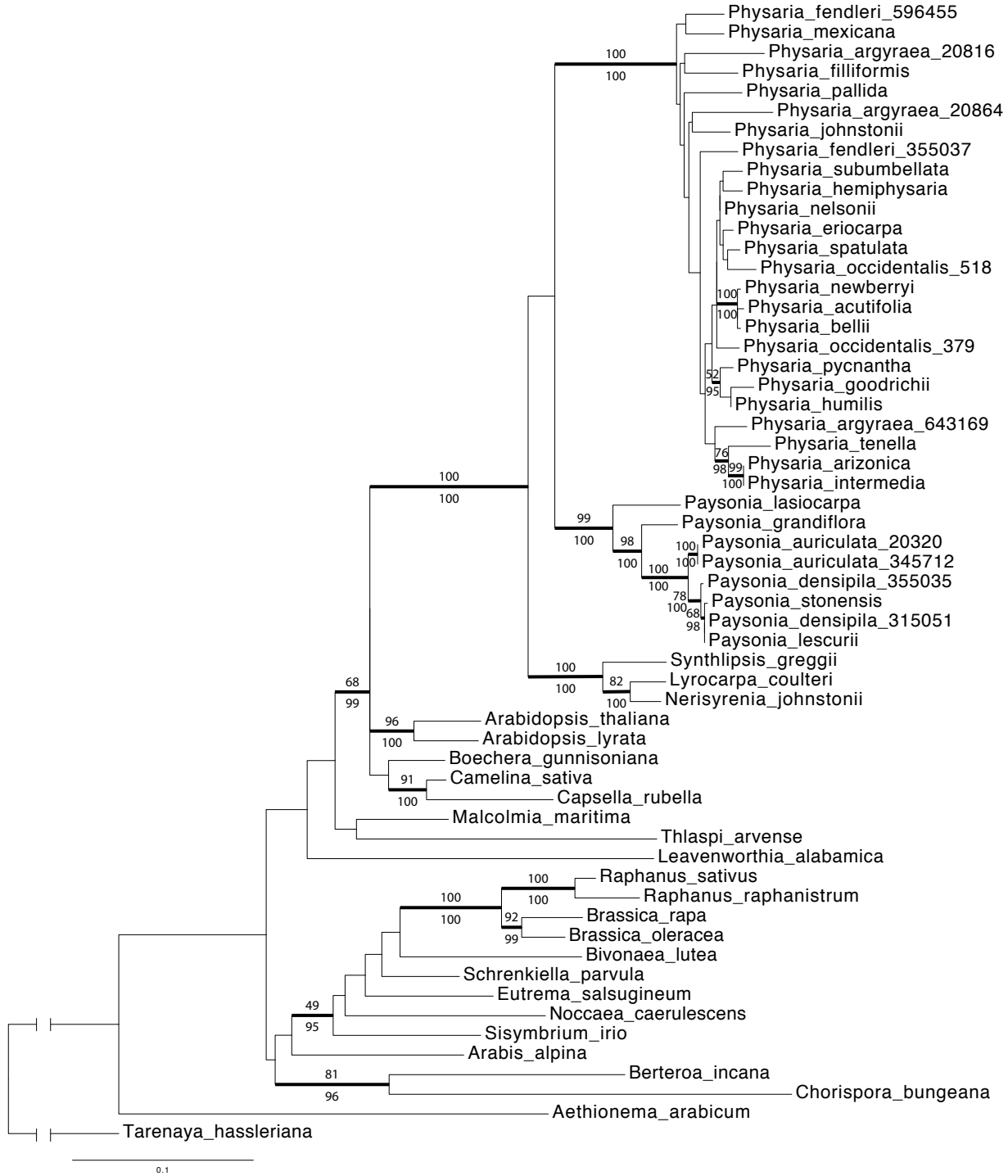

Physaria_fendleri_596455
Physaria_mexicana
Physaria_argyraea_20816
Physaria_filliformis
Physaria_pallida
Physaria_argyraea_20864
Physaria_johnstonii
Physaria_fendleri_355037
Physaria_subumbellata
Physaria_hemiphysaria
Physaria_nelsonii
Physaria_eriocarpa
Physaria_spatulata
Physaria_occidentalis_518
Physaria_newberryi
Physaria_acutifolia
Physaria_bellii
Physaria_occidentalis_379
Physaria_pycnantha
Physaria_goodrichii
Physaria_humilis
Physaria_argyraea_643169
Physaria_tenella
Physaria_arizonica
Physaria_intermedia
Paysonia_lasiocarpa
Paysonia_grandiflora
Paysonia_auriculata_20320
Paysonia_auriculata_345712
Paysonia_densipila_355035
Paysonia_stonensis
Paysonia_densipila_315051
Paysonia_lescurii
Synthlipsis_greggii
Lyrocarpa_coulteri
Nerisyrenia_johnstonii
Arabidopsis_thaliana
Arabidopsis_lyrata
Boechera_gunnisoniana
Camelina_sativa
Capsella_rubella
Malcolmia_maritima
Thlaspi_arvense
Leavenworthia_alabamica
Raphanus_sativus
Raphanus_raphanistrum
Brassica_rapa
Brassica_oleracea
Bivonaea_lutea
Schrenkiella_parvula
Eutrema_salsugineum
Noccaea_caerulescens
Sisymbrium_irio
Arabis_alpina
Berteroa_incana
Chorispora_bungeana
Aethionema_arabicum
Tarenaya_hassleriana
0.1

Fig. 4.

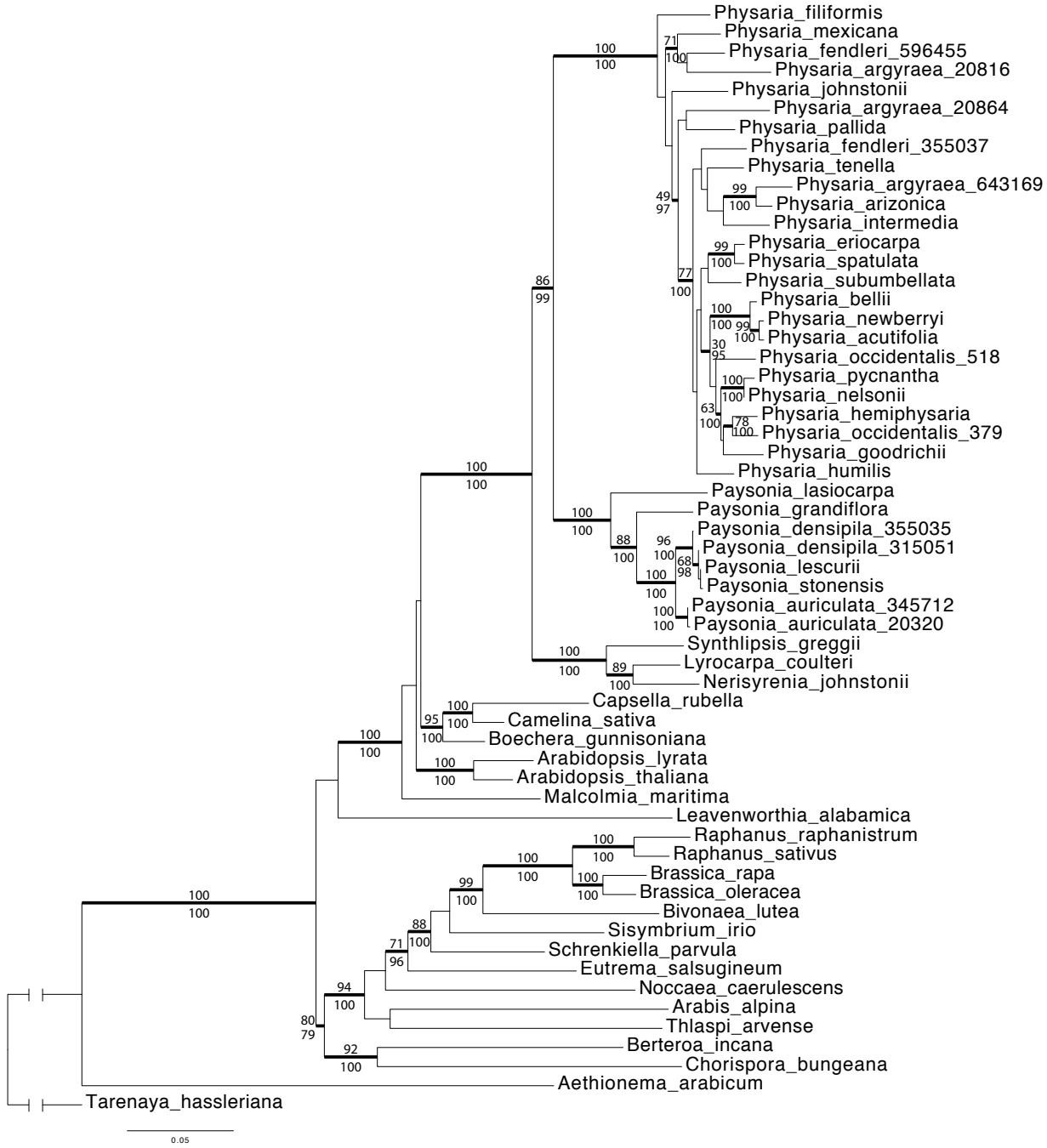
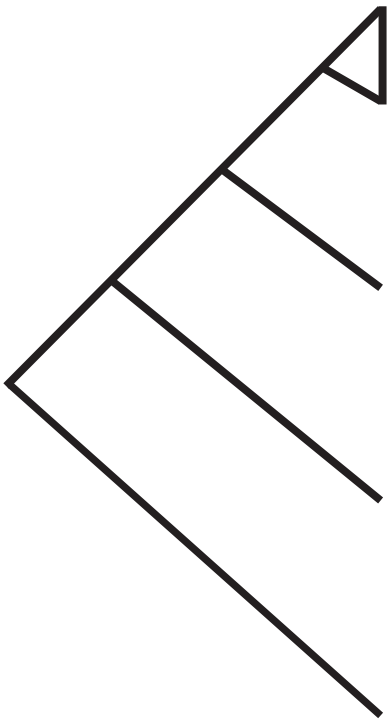
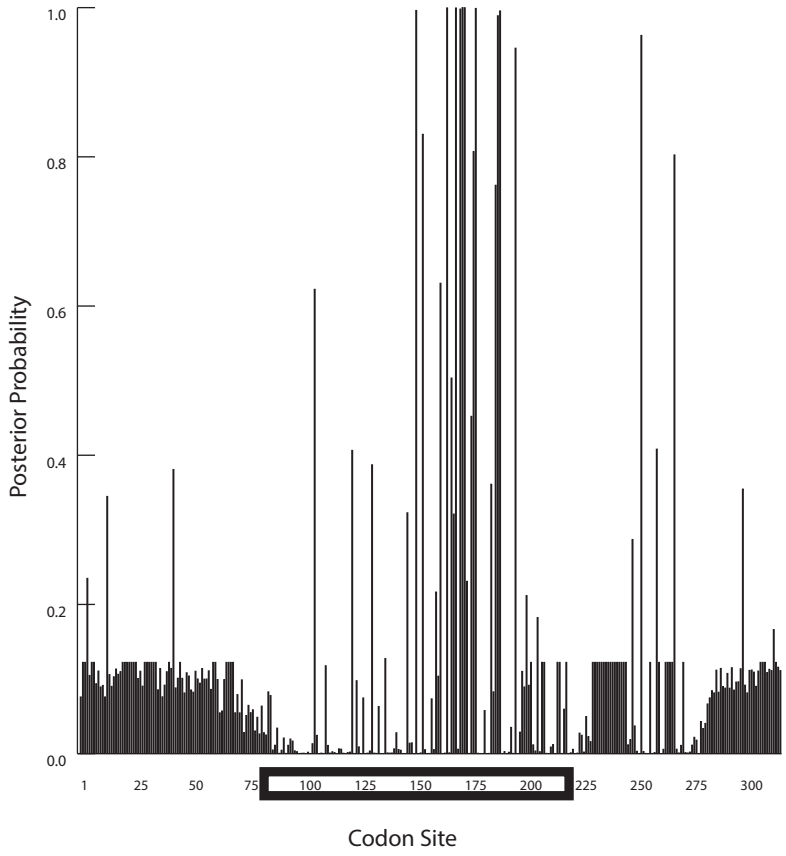

Physaria_filiformis
Physaria_mexicana
Physaria_fendleri_596455
Physaria_argyraea_20816
Physaria_johnstonii
Physaria_argyraea_20864
Physaria_pallida
Physaria_fendleri_355037
Physaria_tenella
Physaria_argyraea_643169
Physaria_arizonica
Physaria_intermedia
Physaria_eriocarpa
Physaria_spatulata
Physaria_subumbellata
Physaria_bellii
Physaria_newberryi
Physaria_acutifolia
Physaria_occidentalis_518
Physaria_pycnantha
Physaria_nelsonii
Physaria_hemiphysaria
Physaria_occidentalis_379
Physaria_goodrichii
Physaria_humilis
Paysonia_lasiocarpa
Paysonia_grandiflora
Paysonia_densipila_355035
Paysonia_densipila_315051
Paysonia_lescurii
Paysonia_stonensis
Paysonia_auriculata_345712
Paysonia_auriculata_20320
Synthlipsis_greggii
Lyrocarpa_coulteri
Nerisyrenia_johnstonii
Capsella_rubella
Camelina_sativa
Boechera_gunnisoniana
Arabidopsis_lyrata
Arabidopsis_thaliana
Malcolmia_maritima
Leavenworthia_alabamica
Raphanus_raphanistrum
Raphanus_sativus
Brassica_rapa
Brassica_oleracea
Bivonaea_lutea
Sisymbrium_irio
Schrenkiella_parvula
Eutrema_salsugineum
Noccaea_caerulescens
Arabis_alpina
Thlaspi_arvense
Berteroa_incana
Chorispora_bungeana
Aethionema_arabicum
Tarenaya_hassleriana
0.05

Fig. 5.



	A	B	C	D	E	F
Physaria crown	0	1	0	1	2	2
Physaria stem	0	1	1	0	1	2
Paysonia	0	0	0	0	0	1
DDNLS clade	0	0	0	0	0	1
Outgroups	0	0	0	0	0	0

Fig. 6.


1.0
0.8
0.6
0.4
0.2
0.0
Posterior Probability
1
25
50
75
100
125
150
175
200
225
250
275
300
Codon Site

Supplementary information 1: Possible misidentifications of taxa

Several lines of evidence indicate that *Physaria fendleri* accession 355037 is incorrectly identified. In preliminary phylogenetic analyses it was distant from all other *P. fendleri* accessions. This finding matches a recently published cluster analysis of DArTseq markers that also placed 355037 outside the *P. fendleri* group (Cruz et al. 2013). In addition, we noted that the trichome anatomy of 355037 differs from other *P. fendleri* accessions. *P. fendleri* 355037 has leaf trichomes with few (3-4) primary branches that fork one or more times, have very little to no webbing between branches, and lack an umbo, whereas all other *P. fendleri* accessions have trichomes with many (approximately 15-20) simple branches, with prominent webbing that typically extends to the middle of their branches, and an umbo. Further taxonomic work is needed to determine the correct species name for accession 355037.

The three sampled accessions labeled *Physaria argyraea* 20816, 20864, and 643169 do not form a clade, though in preliminary *BLT* and ITS analyses accession 20816 was sister to another accession ascribed to *P. argyraea* (20863). The latter two taxa have trichomes with simple branches and an umbo, while 20864 and 643169 display trichomes with forked primary branches. Based on trichome morphology, we suspect that 20816 and 20863 represent true *P. argyraea*, implying that 20864, and 643169 do not, but more work is needed. In addition, one accession labeled *P. argyraea* 643178 that was included in preliminary analyses was resolved in the *P. fendleri* clade, and we believe that this accession is actually a member of *P. fendleri*.

In addition to accessions from the USDA germplasm collection, our phylogenetic analyses indicated that further taxonomic work may be needed on accessions identified as *Physaria occidentalis*. *P. occidentalis* accessions 518 and 379 do not form a clade, although accession 518 was sister to another accession of *P. occidentalis* (533) in preliminary ITS and

BLT analyses. It remains uncertain which cluster represents *P. occidentalis*. Another taxon included in our analyses, *P. goodrichii*, has also been identified as a subspecies of *P. occidentalis* (Flora of North America Editorial Committee, eds. 2010). However, since we did not find a sister relationship between *P. goodrichii* and any other *P. occidentalis* accessions, we retained the original name (Al-Shehbaz and O'Kane 2002). Trichomes of all *P. occidentalis* and *P. goodrichii* accessions have a similar shape, with primary branches that fork one or more times.

Literature Cited

Al-Shehbaz IA, O'Kane SL Jr. 2002. *Lesquerella* is united with *Physaria* (Brassicaceae). Novon 12:319–329.

Cruz VMV, Kilian A, Dierig DA. 2013. Development of DArT marker platforms and genetic diversity assessment of the U.S. collection of the new oilseed crop lesquerella and related species. PLoS ONE 8:e64062.

Flora of North America Editorial Committee, eds. 2010. Flora of North America: North of Mexico. New York

Supplementary Table 1. Primer Sequences

Primer	Sequence (5' -> 3')
BLTF1	TGG AAR YTN TAY GAR AAT CC
BLTR3	AGC TGW RCY TTT TGR CAC TCC AGC
BLT190F	TGG GAT CTN AAT TTY ATA AAG GTN TTY ATG G
BLT740R	CAC TCC AGC TTT GAA CCC CAY TTY TCR CT
ITS5	GGA AGT AAA AGT CGT AAC AAG G
ITS4	TCC TCC GCT TAT TGA TAT GC
ITS.LEU	GTC CAC TGA ACC TTA TCA TTT AG
ITS3B	GCA TCG ATG AAG AAC GTA GC

Supplementary Table 2. Accessions used in molecular evolutionary analyses

Scientific Name	Tribe	Gene	GenBank Accession	DNA Source	Voucher/ Accession
Brassicaceae Family					
Aethionema arabicum (L.) Andrz. ex DC.	Aethionemeae	BLT	ASZG01008923		
		ITS	AY254539		
Arabidopsis lyrata (L.) O'Kane & Al-Shehbaz	Camelineae	BLT	XM_002887821		
		ITS	DQ165338		
Arabidopsis thaliana (L.) Heynh.	Camelineae	BLT	NM_105144		
		ITS	AJ232900		
Arabis alpina L.	Arabideae	BLT	CBTM010021745		
		ITS	DQ060108		

Berteroa incana (L.) DC.	Alysseae	BLT	**	1	R. Price C28
		ITS	AY254544		
Bivonaea lutea (Bivona-Bernardi) DC.	Bivonaeeae	BLT	**	1	D. Baum et al. 391, WIS
		ITS	**	1	D. Baum et al. 391, WIS
Boechera gunnisoniana (Rollins) W.A. Weber	Boechereae	BLT	GBAD01022578		
		ITS	EU274891		
Brassica oleracea L.	Brassiceae	BLT	JJMF01000464		
		ITS	AY833603		
Brassica rapa L.	Brassiceae	BLT	AENI01007419		
		ITS	KF704394		
Camelina sativa (L.) Crantz	Camelineae	BLT	JFZQ01000491		
		ITS	KC172849		
Capsella rubella Reut.	Camelineae	BLT	XM_006301680		
		ITS	AJ232912		
Chorispora bungeana Fisch. & C.A.Mey.	Chorisporeae	BLT	KA015127		

		ITS	FN821605		
Eutrema salsugineum (Pall.) Al-Shehbaz & S.I.Warwick	Eutremeae	BLT	XM_006391647		
		ITS	DQ165371		
Leavenworthia alabamica Rollins	Cardamineae	BLT	ASXC01007858		
		ITS	KC133327		
Lyrocarpa coulteri Hook. & Harv.	Physarieae	BLT	**	2	Fuentes-Soriano et al. 143, MO
		ITS	**	2	Fuentes-Soriano et al. 143, MO
Malcolmia maritima (L.) W.T.Aiton	Anastaticeae	BLT	**	3	2005-111
		ITS	AM905723		
Nerisyrenia johnstonii J.D.Bacon	Physarieae	BLT	**	2	Fuentes-Soriano et al. 98, MO
		ITS	**	2	Fuentes-Soriano et al. 98, MO
Noccaea caerulescens (J.Presl & C.Presl) F.K.Mey	Coluteocarpeae	BLT	GASZ01015331		
		ITS	DQ518397		
Paysonia auriculata (Engelm. & A.Gray) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	W6 20320

		ITS	**	4	W6 20320
Paysonia auriculata (Engelm. & A.Gray) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	PI 345712
		ITS	**	4	PI 345712
Paysonia densipila (Rollins) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	PI 315051
		ITS	**	4	PI 315051
Paysonia densipila (Rollins) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	PI 355035
		ITS	**	4	PI 355035
Paysonia grandiflora (Hook.) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	W6 20837
		ITS	**	4	W6 20837
Paysonia lasiocarpa (Hook. ex A.Gray) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	PI 643173
		ITS	**	4	PI 643173
Paysonia lescurii (A.Gray) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	PI 315053
		ITS	**	4	PI 315053
Paysonia stonensis (Rollins) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	PI 275771
		ITS	**	4	PI 275771

Physaria acutifolia Rydb.	Physarieae	BLT	**	2a	Fuentes-Soriano 275
		ITS	**	2a	Fuentes-Soriano 275
Physaria argyraea (A.Gray) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	W6 20816
		ITS	**	4	W6 20816
Physaria argyraea (A.Gray) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	W6 20864
		ITS	**	4	W6 20864
Physaria argyraea (A.Gray) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	PI 643169
		ITS	**	4	PI 643169
Physaria arizonica (S.Watson) O'Kane & Al-Shehbaz	Physarieae	BLT	**	2b	Fuentes-Soriano 277
		ITS	**	2b	Fuentes-Soriano 277
Physaria bellii G.A.Mulligan	Physarieae	BLT	**	2c	Fuentes-Soriano JC-Mb3
		ITS	**	2c	Fuentes-Soriano JC-Mb3
Physaria eriocarpa Grady & O'Kane	Physarieae	BLT	**	5	Grady 345, State: MT, Cty: Beaverhead
		ITS	**	5	Grady 345, State: MT, Cty: Beaverhead

Physaria fendleri (A.Gray) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	PI 355037
		ITS	**	4	PI 355037
Physaria fendleri (A.Gray) O'Kane & Al-Shehbaz	Physarieae	BLT	**	4	PI 596455
		ITS	**	4	PI 596455
Physaria filiformis (Rollins) O'Kane & Al-Shehbaz	Physarieae	BLT	**	2	Fuentes-Soriano et al. 240, MO
		ITS	**	2	Fuentes-Soriano et al. 240, MO
Physaria goodrichii (Rollins) O'Kane & Al-Shehbaz	Physarieae	BLT	**	5	Grady 374, State: UT, Cty: Beaver
		ITS	**	5	Grady 374, State: UT, Cty: Beaver
Physaria hemiphysaria (Maguire) O'Kane & Al-Shehbaz	Physarieae	BLT	**	5	Grady 561, State: UT, Cty: Utah
		ITS	**	5	Grady 561, State: UT, Cty: Utah
Physaria humilis (Rollins) O'Kane & Al-Shehbaz	Physarieae	BLT	**	5	Grady 466, State: MT; Cty: Ravali
		ITS	**	5	Grady 466, State: MT; Cty: Ravali
Physaria intermedia (S.Watson) O'Kane & Al-Shehbaz	Physarieae	BLT	**	2d	Fuentes-Soriano 293

		ITS	**	2d	Fuentes-Soriano 293
Physaria johnstonii (Rollins) O'Kane & Al-Shehbaz	Physarieae	BLT	**	2	Fuentes-Soriano et al. 194, MO
		ITS	**	2	Fuentes-Soriano et al. 194, MO
Physaria mexicana (Rollins) O'Kane & Al-Shehbaz	Physarieae	BLT	**	2	Fuentes-Soriano et al. 193, MO
		ITS	**	2	Fuentes-Soriano et al. 193, MO
Physaria nelsonii O'Kane & Al-Shehbaz	Physarieae	BLT	**	5	Grady 486, State: WY, Cty: Lincoln
		ITS	**	5	Grady 486, State: WY, Cty: Lincoln
Physaria newberryi A.Gray	Physarieae	BLT	**	6	71939
		ITS	**	6	71939
Physaria occidentalis (S.Watson) O'Kane & Al-Shehbaz	Physarieae	BLT	**	5	Grady 379, State: NV, Cty: White Pine
		ITS	**	5	Grady 379, State: NV, Cty: White Pine
Physaria occidentalis (S.Watson) O'Kane & Al-Shehbaz	Physarieae	BLT	**	5	Grady 518, State: CA, Cty: Modoc

		ITS	**	5	Grady 518, State: CA, Cty: Modoc
Physaria pallida (Torr. & A.Gray) O'Kane & Al-Shehbaz	Physarieae	BLT	**	2e	Fuentes-Soriano W620847
		ITS	**	2e	Fuentes-Soriano W620847
Physaria pycnantha Grady & O'Kane	Physarieae	BLT	**	5	Grady 458, State: ID, Cty: Lemhi
		ITS	**	5	Grady 458, State: ID, Cty: Lemhi
Physaria spatulata (Rydb.) Grady & O'Kane	Physarieae	BLT	**	5	Grady 473, State: WY, Cty: Crook
		ITS	**	5	Grady 473, State: WY, Cty: Crook
Physaria subumbellata (Rollins) O'Kane & Al-Shehbaz	Physarieae	BLT	**	5	Grady 597, State: CO, Cty: Rio Blanco
		ITS	**	5	Grady 597, State: CO, Cty: Rio Blanco
Physaria tenella (A.Nelson) O'Kane & Al-Shehbaz	Physarieae	BLT	**	3	P2040-SF
		ITS	**	3	P2040-SF
Raphanus raphanistrum L.	Brassiceae	BLT	EV568451		
		ITS	AY722479		
Raphanus sativus L.	Brassiceae	BLT	BAUK01069775		

		ITS	AY746463		
Schrenkiella parvula (Schrenk) D.A.German & Al-Shehbaz	Unassigned	BLT	AFAN01000006		
		ITS	AF137579		
Sisymbrium irio L.	Sisymbrieae	BLT	ASZH01004923		
		ITS	KF730127		
Synthlipsis greggii A.Gray	Physarieae	BLT	**	2	Fuentes-Soriano et al. 170, MO
		ITS	**	2	Fuentes-Soriano et al. 170, MO
Thlaspi arvense L.	Thlaspideae	BLT	GAKE01025647		
		ITS	AY662298		
Cleomaceae Family					
Tarenaya hassleriana (Chodat) Iltis		BLT	AOUI01024407		
		ITS	HM044293		

** gene sequence generated in this study

1: DNA from D. Baum

2: Dried leaf material provided by S. Fuentes-Soriano

a: Rocky Mts., USA. Plant material grown from seed bank from B&T World Seeds

b: Plant material grown from seed bank accession #2173-670, Arizona, from the National Arid Land Plant Genetic Resource Conservation (USDA-ARS)

c: Plant material grown from seed bank B&T World Seeds, *JC-1*, USA.

d: description in Tropicos (www.tropicos.org)

e: Texas, USA, Plant material grown from seed bank accession #W6-20847, from the National Arid Land Plant Genetic Resource Conservation (USDA-ARS)
3: Fresh leaf material grown from seed provided by J.C. Pires
4: Fresh leaf material grown from seed provided by USDA NPGS
5: Dried leaf material provided by B. Grady
6: Fresh leaf material grown from seed provided by B&T World Seeds

Supplementary Table 3. Physarieae accessions only included in preliminary analyses

Scientific Name	**Gene**	**GenBank Accession**	**DNA Source**	**Voucher/ Accession**
Dimorphocarpa wislizeni (Engelm.) Rollins	ITS	AF137593		
Dithyrea californica Harv.	ITS	AF137592		
Lyrocarpa coulteri Hook. & Harv.	ITS	AF137591		
	ITS	GU246182		
Nerisyrenia linearifolia (S. Watson) Greene	ITS	AF137587		
	ITS	AF055200		
Paysonia densipila (Rollins) O'Kane & Al-Shehbaz	BLT	**	1	PI 292577
	BLT	**	1	PI 309661
	BLT	**	1	PI 315052
	ITS	**	1	PI 292577
	ITS	**	1	PI 309661
	ITS	**	1	PI 315052
	ITS	AF137586		
Paysonia grandiflora (Hook.) O'Kane & Al-Shehbaz	BLT	**	1	W6 20835

	BLT	**	1	PI 293033
	BLT	**	1	PI 293034
	ITS	**	1	W6 20836
	ITS	**	1	W6 20835
	ITS	**	1	PI 293033
	ITS	**	1	PI 293034
Paysonia lasiocarpa (Hook. ex A.Gray) O'Kane & Al-Shehbaz	BLT	**	1	PI 643172
	ITS	**	1	PI 643172
Paysonia stonensis (Rollins) O'Kane & Al-Shehbaz	BLT	**	1	PI 355042
	BLT	**	2a	Fuentes-Soriano 243
	ITS	**	1	PI 355042
	ITS	**	2a	Fuentes-Soriano 243
	ITS	AF137585		
Physaria acutifolia Rydb.	ITS	AF137582		
Physaria arctica (Wormsk. ex Hornem.) O'Kane & Al-Shehbaz	ITS	JN999288		
Physaria argyraea (A.Gray) O'Kane & Al-Shehbaz	BLT	**	1	W6 20863
	BLT	**	1	PI 643178
	ITS	**	1	W6 20863
	ITS	**	1	PI 643178
Physaria didymocarpa (Hook.) A. Gray	ITS	AF137583		
Physaria fendleri (A.Gray) O'Kane & Al-Shehbaz	BLT	**	2	Fuentes-Soriano et al. 201, MO

	BLT	**	1	PI 596453
	BLT	**	1	PI 596456
	BLT	**	1	PI 596457
	BLT	**	1	PI 596458
	BLT	**	1	PI 643177
	ITS	**	2	Fuentes-Soriano et al. 201, MO
	ITS	**	1	PI 596453
	ITS	**	1	PI 643177
	ITS	AF055198		
	ITS	AF055199		
Physaria gracilis (Hook.) O'Kane & Al-Shehbaz	ITS	DQ310523		
Physaria occidentalis (S.Watson) O'Kane & Al-Shehbaz	BLT	**	3	Grady 533, State: OR, Cty: Wallowa
	ITS	**	3	Grady 533, State: OR, Cty: Wallowa
Physaria pruinosa (Greene) O'Kane & Al-Shehbaz	ITS	AF137584		
Physaria pycnantha Grady & O'Kane	BLT	**	3	Grady 460, State: ID, Cty: Lemhi
	ITS	**	3	Grady 460, State: ID, Cty: Lemhi
Physaria rosei (Rollins) O'Kane & Al-Shehbaz	ITS	**	2b	Fuentes-Soriano et al. 222
Physaria wardii (S. Watson) O'Kane & Al-Shehbaz	BLT	**	2	Fuentes-Soriano et al. 264, MO
Synthlipsis greggii A. Gray	ITS	AF137590		

** gene sequence generated in this study

1:Fresh leaf material grown from seed provided by USDA NPGS

2: Dried leaf material provided by S. Fuentes

a: Tennesse, Around Cumberland River and Dyson Ditch Wild life refuge, roadside field, 36˚ 18' 17" N, 86˚ 09' 42" W

b: description in Tropicos (www.tropicos.org)

3: Dried leaf material provided by B. Grady

Chapter III: The role of *BRANCHLESS TRICHOME* in the evolution of stellate trichomes in *Physaria* (Brassicaceae)

Abigail R. Mazie[a], mazie@wisc.edu
David A. Baum[a], dbaum@wisc.edu

[a]Department of Botany, University of Wisconsin-Madison, 132 Birge Hall, 430 Lincoln Drive, Madison, WI 53706

ABSTRACT

Brassicaceae trichomes are an excellent model to understand the genetic and developmental underpinnings of cell shape evolution. Although consisting of only a single cell, these trichomes can create an intricately branched final form. The mature trichomes of most taxa have branches arranged in a dendritic (or tree-like) pattern, but a dramatic departure from the dendritic pattern is seen in the genus *Physaria*, where trichomes have branches arranged in a stellate (or star-like) pattern. To better understand the developmental changes that took place during the evolution of the stellate form, we compared stellate trichome development in *Physaria* to dendritic trichome development in their closest relative, *Paysonia*. We found developmental differences between dendritic and stellate trichomes in the pattern of branch initiation and propose a model for how these changes led to the evolution of the stellate form. We also explored the role of *BRANCHLESS TRICHOME* (*BLT*) in contributing to the evolution of *Physaria* stellate trichomes. We used a candidate gene approach to test whether swapping the *Arabidopsis* ("dendritic") *BLT* with the *Physaria* ("stellate") *BLT* promoter or coding region or both resulted in changes to trichome branch number or morphology, and compared these results to those obtained using the homologous *A. thaliana BLT* promoter and coding regions. We found that the *Physaria BLT* promoter region was able to replace the *A. thaliana BLT* promoter, resulting in plants with normal, dendritic trichomes. On the other hand, the *P. fendleri BLT* coding region only partially rescued *A. thaliana blt* mutants, yielding dendritic trichomes with fewer branches and smaller stalks. These results suggest that *BLT* alone is not sufficient to have driven the evolution of stellate trichomes.

1. INTRODUCTION

Cell shape is an important phenotype that can have a direct effect on fitness. For instance, the elongated shape of unicellular fission yeast is necessary for proper cell division (Hachet et al. 2012), the cylindrical extensions of some aquatic bacteria increase nutrient uptake (Wagner et al. 2006), the conical shape of petal epidermal cells enhances floral attractiveness to pollinators (Whitney et al. 2011), and the length and orientation of vertebrate auditory hair-cell protrusions (stereocilia) are critical for proper hearing (Vollrath et al. 2007; Chanut-Delalande et al. 2012). Moreover, changes in the shape of individual cells can modify the gross morphology of multicellular tissues and organs. For example, variation in the length of *Aquilegia* petal spurs,

which facilitate plant-pollinator interactions, is caused by between-species variation in the anisotropic expansion of petal cells (Puzey et al. 2011). Despite the great importance of cell shape for organismal biology, relatively little is known about how cell shape evolves.

Trichomes of the mustard family (Brassicaceae) provide a useful model for studying the evolution of cell shape. Plant trichomes are hair-like extensions of the epidermal surface that can be single or multicellular, branched or unbranched, and glandular or eglandular (Werker 2000). Trichomes are thought to defend plants against biotic and abiotic stressors, including herbivory, seed predation, water loss, and UV radiation damage (Werker 2000; Wagner et al. 2004; Hanley et al. 2007). The single-celled, eglandular trichomes of Brassicaceae display extensive morphological diversity, especially in their branching pattern (Appel and Al-Shehbaz 2003; Al-Shehbaz et al. 2006; Beilstein et al. 2006; Beilstein et al. 2008). The existence of closely related clades of Brassicaceae with great differences in trichome shape opens the possibility of gaining new insights into the developmental and genetic basis of cell shape evolution.

Branching trichomes of the Brassicaceae, including those of the model species, *Arabidopsis thaliana*, most commonly have a dendritic (i.e., tree-like) architecture, with a pronounced stalk and two or more branches that angle away from the leaf surface (Beilstein et al. 2006; Beilstein et al. 2008). The development of *A. thaliana* dendritic trichomes has been carefully characterized and divided into a series of six developmental stages (Hülskamp et al. 1994; Szymanski et al. 1998; Szymanski et al. 1999). During stages one and two, trichome precursors enlarge by radial expansion, then undergo diffuse extension growth away from the leaf surface to form the main trichome stalk. During stages three and four, secondary growth foci form, typically establishing 3-4 trichome branches. The final stages of development include diffuse expansion of the entire trichome cell, sharpening of the branch tips, thickening of the cell wall, and formation of numerous cell-wall papillae.

A striking departure from the typical dendritic pattern is found in the genus *Physaria* (tribe Physarieae). All members of *Physaria* produce stellate (i.e., star-like) trichomes, with branches that radiate from a short stalk and lie parallel to the epidermis (Rollins and Shaw 1973; Rollins and Banerjee 1975; Rollins and Banerjee 1976). All other Physarieae genera have dendritic trichomes (Beilstein et al. 2006; Beilstein et al. 2008; Fuentes-Soriano and Al-Shehbaz 2013), showing that the stellate form is a synapomorphy of the *Physaria* clade. Primary

differences between stellate and dendritic trichomes appear to be the length of the trichome stalk, the orientation of primary branches in relation to the stalk, and the number of primary branches.

A promising candidate gene for the evolution of the stellate pattern is *BRANCHLESS TRICHOME* (*BLT*). In *A. thaliana*, *BLT* is a positive regulator of trichome branching, with loss-of-function mutants exhibiting completely unbranched trichomes (Marks et al. 2009; Kasili et al. 2011) and overexpression lines producing elevated branch numbers (Kasili et al. 2011). Although mutations in *BLT* dramatically change trichome shape, other aspects of plant growth and development are unaffected (Kasili et al. 2011). BLT localizes to the tips of emerging branch buds, with another positive regulator of trichome branching, STICHEL (Ilgenfritz et al. 2003; Marks et al. 2009). Together, these proteins are postulated to form a complex that directs the polarity of branch initiation sites (Kasili et al. 2011). *Physaria* BLT has been subject to a distinctive mode of molecular evolution compared to other members of Physarieae, including evidence of positive selection on as many as 10% of the gene's codons within the *Physaria* clade (Mazie and Baum, Chapter II).

In this study, we used scanning electron microscopy to compare *Physaria* stellate trichome development to *Paysonia* dendritic trichome development, the latter genus being the sister clade to *Physaria* in Physarieae (O'Kane and Al-Shehbaz 2002; Fuentes-Soriano and Al-Shehbaz 2013; Mazie and Baum, Chapter II). The aim of this work was to clarify the developmental changes that are likely to underlie the transition from a dendritic to stellate form. In addition, we tested whether changes in *Physaria* BLT protein structure or expression or both could have contributed directly to the evolution of stellate trichomes. We did this by introducing the promoter and coding regions of *P. fendleri BLT* into *A. thaliana blt* mutants and comparing the resulting trichomes to those obtained using the homologous *A. thaliana BLT* promoter and coding regions. In this way, we could test whether swapping "dendritic" *BLT* with "stellate" *BLT* resulted in changes to trichome branch number or morphology, and whether any changes discovered were a result of differences in the promoter or coding region of "stellate" *BLT*. We found that the *Physaria BLT* promoter was able to fully rescue *A. thaliana blt* mutants, resulting in dendritic trichomes with the typical *A. thaliana* branch number and morphology. Conversely, *A. thaliana blt* mutants transformed with the *Physaria BLT* coding region had trichomes with fewer branches and narrower, shorter stalks than those transformed with the *Arabidopsis BLT* coding region. Overall, the *Physaria BLT* promoter and coding regions resulted in plants with

more or less normal, dendritic trichomes. Our results therefore suggest that *BLT* alone is not sufficient to have driven the evolution of stellate trichomes. While it is still possible that *BLT* contributed to stellate trichome evolution, its effect must have been mediated through coevolution with at least one additional molecular partner.

2. MATERIALS AND METHODS

2.1 Trichome Measurements

We examined developing and mature trichomes on rosette leaves of five species of *Paysonia*: *P. auriculata* accessions 20320 and 345712, *P. densipila* accession 315051 and 315052, *P. grandiflora* 20837, *P. lasiocarpa* accession 643173, *P. stonensis* accession 355042. We also examined five species of *Physaria*: *P. argyraea* accession 20816, *P. fendleri* accession 596455, *P. "argyraea"* accession 20864, *P. "argyraea"* accession 643169, and *P. "fendleri"* accession 355037. We believe the latter three species to be misidentified based on a previous combined phylogenetic analysis of *BRANCHLESS TRICHOME* and ITS (Mazie and Baum, Chapter II). This analysis indicated a close relationship between *P. "argyraea"* 20864 and *P. pallida* and between *P. "argyraea"* 643169 and *P. arizonica,* and we will therefore refer to these taxa as *P.* aff. *pallida* and *P.* aff. *arizonica*. *P. "fendleri"* 355037 was sister to a clade that contained *P. tenella*, *P. intermedia*, and *P. arizonica*, and we will therefore refer to this taxon as *P.* aff. *tenella*. All plants were grown from seed provided by the USDA NPGS, and taxon information is listed in supplementary table 1.

We used Environmental Scanning Electron Microscopy (FEI Quanta, Hillsboro, OR) in the Newcomb Imaging Center, Botany Department, UW-Madison with a GSED detector and imaged at 4.0 °C, with 20 – 25 kV voltage, 4 - 6 Torr pressure, and a spot size of 5 - 6. Trichome measurements were taken with ImageJ64 (version 1.49u; (Rasband 2015)). We used the R statistical software package (version 3.2.2 for Mac OSX; (R Core Team 2015)) for statistical analyses and graphing, using the plyr (Wickham 2011) and ggplot2 (Wickham 2009) packages.

In describing trichome development, we refer to trichomes that have not yet initiated branches as "pre-branching trichomes", trichomes that have initiated branches, but not yet developed cell wall papillae, as "post-branching trichomes", and trichomes covered with papillae as "mature trichomes". Pre-branching trichome stalk measurements were obtained from *P. auriculata* (21 trichomes, 4 plants) and *P. fendleri* (8 trichomes, 2 plants). Mature trichome stalk

measurements were obtained from *P. auriculata* (10 trichomes, 2 plants). Post-branching trichome branch measurements were obtained from *P. auriculata* (6 branches, 3 trichomes, 3 plants), *P. fendleri* (32 branches, 6 trichomes, 4 plants), *P.* aff. *arizonica* (10 branches, 3 trichomes, 1 plant), and *P.* aff. *tenella* (14 branches, 5 trichomes, 1 plant). Mature trichome branch measurements were obtained from *P. auriculata* (71 branches, 21 trichomes, 5 plants), *P. fendleri* (63 branches, 13 trichomes, 7 plants), *P.* aff. *arizonica* (13 branches, 4 trichomes, 2 plants), *P.* aff. *tenella* (11 branches, 4 trichomes, 2 plants). Following Szymanski, Marks, and Wick (1999), we measured trichome stalk length as the distance from cell base to tip, or, if the trichome had initiated branches, from the plane of the epidermis to the lowest surface of the first branch. Stalk width was measured at the stalk midpoint. Branch length was measured from the branch tip to the branch base at the point of branch splitting. Branch width was measured at the branch base.

2.2 Isolation of *Physaria fendleri BLT*

DNA was extracted from one plant of *P. fendleri* (USDA NPGS, PI 596455; supplementary table 1) with the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). We used the GenomeWalker Universal Kit (Clontech, Mountain View, CA) to create GenomeWalker libraries with four blunt cutting restriction enzymes: *Stu*I (provided in the GenomeWalker kit) and *Psi*I, *Ssp*I, *Sca*I (New England BioLabs, Ipswich, MA). Digested DNA was purified by phenol/chloroform extraction and ethanol precipitation, and GenomeWalker Adaptors were ligated onto the ends of the fragments.

Sequence information for the majority of the *P. fendleri BLT* coding region was obtained with the degenerate primers BLT190F and BLT740R (Mazie and Baum, Chapter II). We used this sequence information to design *P. fendleri BLT* specific primers that were paired with the adaptor primers provided in the GenomeWalker kit (AP1 and AP2) to PCR "walk" up- and downstream of the known sequence. Each walk consisted of a set of primary and nested PCRs on each GenomeWalker library with Advantage 2 Polymerase (Clontech, Mountain View, CA), which is designed to amplify long DNA fragments and includes a proofreading enzyme. Multiple, sequential walks and "primer walking" on PCR fragments from initial genome walks too long to be sequenced in a single reaction allowed us to obtain sequence information upstream and downstream of the *P. fendleri BLT* coding region.

Nested PCRs that produced a major PCR product longer than 1 Kbp were extracted with the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) and cloned with the pGEM-T vector system (Promega Corp., Madison, WI) and DH5α competent cells (Invitrogen, Life Technologies, Grand Island, NY). Insert-containing colonies were identified by blue-white screening and amplified with GoTaq DNA Polymerase (Promega Corp., Madison, WI) using SP6/T7 or the nested PCR primers. PCR products were purified using enzymatic cleanup methods (ExoSAP-IT, Affymetrix, Santa Clara, CA) and sequenced with the BigDye Terminator Kit (Applied Biosystems, Life Technologies, Grand Island, NY). Agencourt CleanSEQ (Beckman Coulter, Brea, CA) was used to remove excess dye. Sequences were analyzed with an ABI 3730XL DNA Analyzer at the UW-Madison Biotechnology Center and assembled and aligned with Geneious Pro (version 5.6.6, Biomatters, Auckland, New Zealand; http://www.geneious.com; (Kearse et al. 2012)).

2.3 Allelism tests in *P. fendleri*

Two variant *P. fendleri BLT* sequences were identified during genome walking. To determine if these variant sequences represented *BLT* alleles or copies, we designed CAPS (cleaved-amplified polymorphic sequence) markers (BLT_IDF/BLT_IDR), which amplified approximately 325 bps within the coding region of both variants and overlapped a *Psi*I restriction site only found in one variant ("variant A"). As a result, the expected PCR product would be cut into 192 bp and 133 bp fragments only in variant A. We used the CAPS markers and *Psi*I (New England BioLabs, Ipswich, MA) to genotype 7 accessions of *P. fendleri* (596453, 596455, 596456, 596457, 596458, 643177, SF201; supplementary table 1) and *P.* aff. *tenella* (accession 355037). To corroborate these results, we also developed variant specific primers to amplify an approximately 1 Kbp region in the *P. fendleri* and *P.* aff. *tenella* accessions, along with six additional individuals of *P. fendleri* accession 596455. Primers BLTcdsVAR1F/BLTcdsVAR2R were designed to amplify variant A, but not variant B, while primers BLTcdsVAR2F/BLTcdsVAR1R were designed to amplify variant B, but not variant A. Primer sequences are listed in supplementary table 2.

2.4 Construction of transformation constructs

DNA was extracted from *A. thaliana* Columbia and *P. fendleri* accession 596455 with the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). Three sequence regions were amplified for each taxon: (1) the entire *BLT* coding region along with upstream and downstream noncoding regions, (2) the upstream noncoding region only, (3) the coding region and downstream noncoding regions only. Primers were designed with 5' extensions that matched either the pCAMBIA1300 vector or the opposite taxon's DNA sequence so that fragments could be ligated together in subsequent cloning reactions. Primer sequences are listed in supplementary table 2. All amplifications were performed with Advantage2 Polymerase (Clontech, Mountain View, CA), and PCR products were gel extracted with the Nucleospin Gel and PCR Cleanup Kit (Clontech, Mountain View, CA).

We made four different constructs: (1) AA, which includes the 3', 5', and coding regions of *A. thaliana BLT*; (2) PP, which includes the 3', 5', and coding regions of *P. fendleri BLT*; (3) AP, which combines the *A. thaliana* 5' region with the *P. fendleri BLT* coding and 3' regions; (4) PA, which combines the *P. fendleri* 5' region with the *A. thaliana BLT* 3' and coding regions (fig. 1). The In-fusion enzyme mix (Clontech, Mountain View, CA) was used to combine PCR fragments from *A. thaliana* and *P. fendleri* to make the AP and PA vectors and insert the fragments for all four vector types into pCAMBIA1300. Prior to the In-fusion reaction, pCAMBIA1300 was linearized with EcoRI/BamHI (New England BioLabs, Ipswich, MA) and gel extracted with the Nucleospin Gel and PCR Cleanup Kit (Clontech, Mountain View, CA). Stellar competent cells (Clontech, Mountain View, CA) were heat-shock transformed with AA, PP, AP, and PA vectors, and positive colonies were identified with kanamycin selection.

We confirmed the accuracy of clone sequences by colony PCR and sequencing of at least eight colonies of each construct type. Colony PCRs were performed as described previously using M13F/M13R and internal primers.

2.5 Plant Transformation

A. thaliana Columbia *blt-1* mutant plants (stock CS827202) were grown under standard conditions until flowering. *blt-1* plants that displayed unbranched trichomes were considered to be homozygous mutants (Marks et al. 2009).

Agrobacterium tumefaciens strain GV3101 was electro-transformed with clones AA-3, AP-2, PP-24, PP-40, PA-5, and PA-6, which had been isolated with the QIAprep Spin Miniprep

Kit (Qiagen, Valencia, CA). Positive colonies were identified by kanamycin selection. We confirmed the sequences of construct containing *A. tumefaciens* colonies with colony PCR and sequencing using M13F/M13R and internal primers.

Construct containing *A. tumefaciens* cultures were used to drip transform plants using the protocol of Correa et al. (2012). After transformation, plants were watered regularly until siliques appeared. Primary transformant (T1) seeds were collected and vapor-phase sterilized following the protocol of Clough and Bent (1998). Positive transformants were identified with hygromycin selection following the protocol of Harrison et al. (2006). Resistant seedlings were identified by their relatively long hypocotyls in comparison to non-resistant seedlings. Resistant T1 plants were grown under standard conditions until siliques appeared and T2 seed was collected.

2.6 Phenotyping

During phenotyping of transformants in the *blt-1* mutant background, we made quantitative measurements only on T1 lines (each T1 line corresponding to one plant) that showed full rescue and contained one copy of the transgene. We defined full rescue as lines in which at least 95% of trichomes were branched. Trichome branch numbers were obtained from trichomes on the adaxial surface of three rosette leaves of each T1 line with an Olympus SZX12 Stereo microscope at the Newcomb Imaging Center, Botany Department, UW-Madison. We used rosette leaf number 5 or greater and counted at least 200 trichomes per line (an average of ~400 trichomes per line).

We estimated the transgene copy number of T1 lines by counting the hygromycin resistance segregation ratio of T2 plants. We counted approximately 100 T2 seedlings from each T1 line and used chi-square tests to determine if the observed resistant:sensitive segregation ratio was significantly different from the expected 3:1 ratio. This analysis allowed us to infer the number of transgene loci found in each T1 line, although it cannot rule out the possibility that multiple gene copies are included in a locus. We discarded data from T1 lines whose T2 segregation ratio deviated significantly from 3:1 and from T1 lines that did not yield seed, and thus, whose T2 segregation ratio was unknown.

For representative T2 plants of some T1 lines, we also measured trichome stalk and branch dimensions using the same methodology as described previously for *Paysonia* and *Physaria* trichome measurements. We measured trichome stalks from plants transformed with

construct AA (19 trichomes, 8 plants, 4 lines), AP (9 trichomes, 3 plants, 1 line), and PA-6 (13 trichomes, 7 plants, 5 lines) and trichome branches on plants transformed with construct AA (81 branches, 38 trichomes, 8 plants, 4 lines), AP (19 branches, 9 trichomes, 3 plants, 1 line), and PA-6 (64 branches, 33 trichomes, 8 plants, 5 lines).

3. RESULTS

3.1 Trichome Stalk

Stalk Development: We measured trichome stalk development from pre-branching *Paysonia auriculata* and *Physaria fendleri* trichomes (fig. 2). The development of pre-branching *Paysonia* and *Physaria* stalks was nearly identical. Stalks of both genera had a mean width of 12µm (*P. auriculata* ±1.5µm standard deviation (sd), *P. fendleri* ±4.0µm sd; fig. 2A) and a mean length of 8-9µm (*P. auriculata* ±2.4µm sd, *P. fendleri* ±3.8µm sd; fig. 2B). The largest trichomes of both genera were <17µm in width and <14µm in length. Stalk length and width were positively correlated for both species (*P. auriculata*: p-value = <0.01, adjusted R^2 = 0.32; *P. fendleri*: p-value = <0.001, adjusted R^2=0.91; fig. 2C).

The Mature Stalk: In *Physaria*, the position of the trichome branches obscured the stalk at later stages, so we measured post-branching and mature trichome stalks only from *Paysonia* trichomes. During branch initiation, *Paysonia* stalk widths remained the same as at the pre-branching stage, but stalk lengths were shorter than the tallest pre-branching stalk lengths. This is expected because trichome branches are initiated subterminally. After all branches had emerged, the entire trichome underwent a diffuse growth stage, which included elongation of the stalk. Mature trichomes had an average stalk width and length of 21µm (width ±2.2µm sd, length ±3.2µm sd; fig. 3A,B). The positive correlation between stalk length and width was no longer seen in mature trichomes (fig. 3C).

3.2 Trichome Branches

Branch Patterning and Final Morphology: In *Paysonia*, the earliest post-branching trichomes had two branches of unequal size that elongated towards the two leaf margins, thus always showing a consistent orientation with respect to the leaf (fig. 4A #1&2, supplementary fig. 1). On the other hand, all species of *Physaria* had primary branches of similar size that arose from

all sides of the trichome primordium (fig. 5A, C, supplementary figs. 2, 3); therefore, branches were not oriented towards any particular leaf axis. Within *Physaria*, two different patterns of branch initiation were noted. *P. fendleri* and *P. argyraea* initiated many primary branches nearly simultaneously and symmetrically on all sides of the trichome initial, and none of these branches underwent additional branching (fig. 5A, B, supplementary fig. 2). In contrast, *P.* aff. *pallida*, *P.* aff. *arizonica*, and *P.* aff. *tenella* initiated three or four primary branches that underwent at least one additional round of branching (fig. 5C, D, supplementary fig. 3). We will refer to the former pattern as *Physaria* type 1 trichomes, and the latter as *Physaria* type 2 trichomes.

Additional branching events were always observed in *Paysonia* and *Physaria* type 2 trichomes. *Paysonia* secondary branches initially appeared as semicircular areas of cell expansion on the lateral wall of an already existing branch. A third branch was always initiated on the larger primary branch and elongated towards the leaf tip (fig. 4A, #3, supplementary fig. 4), and additional branching events often occurred on the first two branches to emerge (fig. 4A, B, supplementary fig. 5). In all cases, newly emerging branches initiated one at a time, subapically, from already established branches. On the other hand, in *Physaria* type 2 trichomes, two branch bulges formed at the apex of an extending branch, bifurcating the existing branch to create two additional branches (fig. 5C, supplementary fig. 3). These bifurcations did not always occur simultaneously between branch tips of the same tier (fig. 5C, supplementary fig. 3). After the completion of branch initiation, *Paysonia* and *Physaria* trichomes underwent a diffuse growth stage in which trichomes dramatically increased in length and width. Late in maturation, branch tips became pointed, and the cell walls came to be decorated with minute papillae (figs. 4C, 5B, 5D, supplementary figs. 6, 7, 8).

Paysonia species have mature trichomes with 3-6 external branches that typically ascend away from the leaf surface at an angle (supplementary fig. 6); however, *P. auriculata* trichome branches appear to lie more or less parallel to the leaf surface (fig. 4). *Physaria* species have mature trichomes with many more external branch tips than *Paysonia*, and trichome branches always lie parallel to the leaf surface (fig. 5, supplementary figs. 7, 8). *P. fendleri* trichomes had the most branches and the greatest degree of variation, with mature trichome branch numbers ranging from 10-20 external branches. All other *Physaria* species observed had some degree of variability in external branch number: 7-10 in *P. argyraea* and 8-10 in *P.* aff. *pallida*, *P.* aff. *tenella*, and *P.* aff. *arizonica*.

Primary Branch Measurements: We measured the length and width of primary branches from *Paysonia* (*P. auriculata*), *Physaria* type 1 (*P. fendleri*), and *Physaria* type 2 (*P.* aff. *arizonica* and *Phy.* aff. *tenella*) trichomes. To keep measurements consistent across taxa, we compared only tip measurements from primary branches that had not yet undergone any secondary branching events. *Paysonia* trichome branches had a mean width of 12 ±1.09µm sd and a mean length of 6 ±2.11µm sd. *Physaria* type 1 branches had a mean width of 5 ±1.17µm sd and a mean length of 2 ±1.24µm sd. *Physaria* type 2 trichome branches were more similar to *Paysonia* branches, with a mean width of 11 ±1.79µm sd and an average length of 5 ±1.68µm sd (figs. 6A, B). An ANOVA indicated significant differences in branch widths and lengths between taxa ($p<0.001$), with *Physaria* type 1 branches significantly smaller in width and length than *Paysonia* or *Physaria* type 2 branches (Tukey's HSD test). As in early stalk development, primary trichome branches had a generally positive relationship between branch width and length (fig. 6C), although this relationship was only significant for *Physaria* (type 1: $p<0.001$, adjusted $R^2=0.42$; type 2: $p<0.001$, adjusted $R^2=0.76$).

Mature Branch Measurements: We measured the width and length of mature, external *P. auriculata* branches from 3- and 5-branched trichomes. We compared individual branch measurements within 3- and 5-branched trichomes using the branch-free gap to categorize each branch according to its position clockwise from this gap (see supplementary fig. 9 for a visual description of how individual branches were counted). Mature, 3-branched trichomes had branches with a mean width of 19 ±2.38µm sd and length of 161 ±44.91µm sd, with no significant differences between branches (fig. 7A, B). We found a positive relationship between width and length ($p<0.01$, $R^2=0.17$), although a scatterplot of the data indicates that this result is likely to be driven by a few outliers (fig. 7C). An ANOVA of mature, 5-branched trichomes indicated significant differences in branch width across the different branch positions ($p<0.001$), with the third branch of 5-branched trichomes significantly wider than all other branches (Tukey's HSD test). The third branch of 5-branches trichomes had a mean width of 22 ±3.38µm sd, compared to 17 ±1.72µm sd for all other branches (fig. 8A). All branches had an average length of 159 ±34.75µm sd, with no significant differences between branches (fig. 8B). We found no significant relationship between branch width and length (fig. 8C).

As an example of *Physaria* type 1 trichomes, we measured the width and length of *P. fendleri* external trichome branches. In *Physaria* type 1 trichomes all external branches were primary branches (originating from the trichome stalk), though the number of external branches was very variable across trichomes, even on the same leaf. To see if there is a relationship between branch number and branch dimensions, we first examined the average width within each branch number class. Average branch width decreased as branch number increased (supplementary fig. 9), and we therefore decided to split all trichome measurements into three classes based on their average width – class A included 10 branched trichomes, class B included 11-16 branched trichomes, and class C included 18-19 branched trichomes. ANOVAs indicated significant differences between branch widths ($p<0.001$) and lengths ($p<0.001$). There were significant differences between the widths of all branch classes: class A, 14 ±2.43µm sd; class B, 11 ±1.97µm sd; class C, 8 ±1.20µm sd (Tukey's HSD tests; fig. 9A). In addition, we found that class A trichome branches were significantly longer (203 ±28.18µm sd) than those of classes B and C (158 ±20.39µm sd) (Tukey's HSD test; fig. 9B). We found no significant relationship between branch width and length in any trichome class (fig. 9C).

As an example of *Physaria* type 2 trichomes, we measured the width and length of the external branches of mature *Phy.* aff. *tenella* and *Phy.* aff. *arizonica* trichomes. All external branches were formed from bifurcations of already existing branches, some being secondary branches (formed by the bifurcation of a primary branch) and others being tertiary branches (formed by the bifurcation of a secondary branch). Two-way ANOVAs found that species identity ($p<0.001$) and branch level ($p<0.05$) had significant effects on branch length and that branch level ($p<0.05$), but not species identity, had a significant effect on branch width, but no interaction between the two variables. Secondary branches of both species had a mean width of 18 ±2.79µm sd, compared to 15 ±1.18µm sd for tertiary branches (fig. 10A), and t-tests indicated that these differences were significant ($p<0.01$). Secondary branches of both species were longer than tertiary branches (fig. 10B), although t-tests indicated that this length difference was only significant for *Phy*. aff. *arizonica* ($p<0.05$). *Phy*. aff. *tenella* secondary branches had a mean length of 220 ±54.09µm sd, while tertiary branches had a mean length of 163 ±42.96µm sd. *Phy*. aff. *arizonica* secondary branches had a mean length of 126 ±23.23µm sd, while tertiary branches had a mean length of 103 ±8.81µm sd. *Phy*. aff. *tenella* had a significant positive relationship between secondary branch width and length ($p<0.05$, adjusted $R^2=0.68$) and a

significant negative relationship between tertiary branch width and length ($p<0.05$, $R^2=0.88$), but we found no correlation between branch width and length for *Phy*. aff. *arizonica* (fig. 10C).

3.3 Sequence characteristics of two *P. fendleri BLT* alleles

We isolated the *P. fendleri* homolog of the *A. thaliana BLT* coding region (792 bps), along with approximately 2 Kbps of upstream and 2 Kbps of downstream sequence, with the aim of capturing the *BLT* promoter and any nearby *cis*-regulatory elements. Kasili et al. (2011) used a similar *A. thaliana* genomic DNA fragment to rescue homozygous *blt-3* mutant plants, consisting of the *A. thaliana BLT* coding region (822 bps), with approximately 1.4 Kbps of upstream sequence and 1.2 Kbps of downstream sequence.

In the course of isolating *P. fendleri BLT*, we identified two sequence variants. We used CAPS markers on seven *P. fendleri* accessions and one *P*. aff. *tenella* accession to determine if we had identified duplicate *BLT* genes or two alleles at a single locus. If the variant sequences were gene copies, we would expect to find three bands in every taxon. Instead, we found a mix of all three different banding patterns across taxa, suggesting that the variant sequences are most likely alleles (fig. 11). We further confirmed this finding by amplifying a 1 Kbp region in each taxon with allele specific primers. The results of this analysis largely mirrored those of our restriction digestions - taxa completely digested with *Psi*I were only amplified with primers specific to allele A and taxa not digested with *Psi*I were only amplified with primers specific to allele B. There was one exception to this pattern. *P. fendleri* accession 643177 was not digested by *Psi*I but was successfully amplified with allele A primers; however, this result is most likely due to a mutation at the cut-site, rather than the presence of two distinct loci.

The sequences of the two *P. fendleri* alleles were very similar (fig. 12). We identified only nine nonsynonymous differences in the coding region and relatively few differences between the 5' and 3' gene regions, with the exception of an eighty base-pair indel (with the deletion found in allele A) located approximately 130 bps 3' of the stop codon. In contrast, a comparison between *A. thaliana* and the *P. fendleri BLT* homologs revealed many sequence differences. The *P. fendleri* start codon appears to be shifted 15 bps downstream relative to *A. thaliana*, and we identified five indels and 37 nonsynonymous differences between *A. thaliana* BLT and the *P. fendleri* alleles.

3.4 Transgene Construct Development

We chose allele A of *P. fendleri* accession 596455 to design our four transformation constructs. Allele A appeared to be the more common allele among individuals of this accession, because allele A specific primers (but not allele B specific primers) amplified *BLT* from six out of seven individuals sampled. Transformation constructs included the full genomic constructs for *A. thaliana BLT* (AA-3: 1447 bps 5', 822 bps coding, 1228 bps 3') and *P. fendleri BLT* (PP-24 and PP-40: 2015 bps 5', 792 bps coding, 2494 bps 3'), and two chimeric constructs. The AP chimeric construct fused the 5' non-coding region of *A. thaliana* with the coding and 3' non-coding region of *P. fendleri*, whereas the PA construct fused the 5' non-coding region of *P. fendleri* with the coding and 3' non-coding region of *A. thaliana* (supplementary fig. 10).

We sequenced multiple clones to check for errors, finding that construct AA clone 3, PP clone 24, and PP clone 40 were all error-free. Unfortunately, all PA and AP clones had one or more errors located outside of the predicted gene region. For the PA construct we chose clone 5, which has one SNP (A instead of G) 806 bps upstream of the start codon and one SNP (C instead of T) 756 bps downstream of the stop codon and clone 6 which had one SNP (A instead of G) 289 bps upstream of the start codon. We chose AP clone 2, which differed from both PP constructs in having one fewer CT repeat in a series of CTs and one additional A in a series of As in a region ~2300 bps downstream of the stop codon.

3.5 Phenotyping of Transgenic Plants

To determine whether changes in *P. fendleri BLT* or its promoter result in altered trichome branching patterns in an *A. thaliana* background, we transformed *A. thaliana blt* mutant plants, which display unbranched trichomes, with AA, PP, AP, and PA constructs and counted the number of external trichome tips on rosette leaves of multiple T1 lines per construct. Unbranched trichomes have only one tip, while the number of tips in trichomes with two or more branches is equivalent to the trichome branch number. We only scored lines that showed a 3:1 segregation pattern, suggesting a single locus of insertion, and demonstrated a high frequency of branched trichomes, suggesting that the insert is intact. We excluded 1 AA, 2 PA, 1 AP, and 3 PP lines because of incomplete rescue. A chi-square test indicated no significant difference between constructs in their ability to rescue the mutant phenotype. The number of fully rescued,

single insert locus T1 plants that were scored varied across lines: 19 AA, 17 PA-5, 26 PA-6, 13 AP, 10 PP-24, and 27 PP-40.

We plotted the distribution of trichome tips across all scored lines of each construct (supplementary fig. 11). Because the two PP and PA constructs had no (or minimal) sequence differences between them and the difference between the construct means was less than 0.25 branch tips, we chose to combine the data from each pair of constructs for all further analyses. Histograms of the number of branch tips in each construct group (fig. 13) indicate that the AA and PA lines had a higher mean number of branch tips than the AP and PP lines. A two-way ANOVA suggests that whereas the different promoter regions have no effect on trichome branch number, the coding region does have a significant effect of trichome branch number ($p<0.001$), with the *P. fendleri* coding region producing fewer-branched trichomes than the *A. thaliana* region.

We used ESEM to image mature trichomes from T2 plants that contained AA, AP, and PA constructs (fig. 14). ANOVAs indicated significant effects of the identity of the coding region on stalk width ($p<0.01$) and length ($p<0.05$), with trichomes transformed with the *P. fendleri* coding region having smaller stalks than those transformed with the *A. thaliana* coding region. Trichome stalks with the *A. thaliana* coding region had a mean width of 28 ±6.31µm sd and length of 75 ±18.60µm sd; trichome stalks with the *P. fendleri* coding region had a mean width of 21 ±4.86µm sd and length of 61 ±27.54µm sd (fig. 15A, B). In addition, stalk length was positively correlated with stalk width only in constructs with the *A. thaliana coding* region ($p<0.001$, adjusted $R^2=0.29$; fig. 15C). We found no significant effect of either the promoter or coding region on the width or length of external branches. Branches of all constructs had a mean width of 21 ±4.35µm sd (fig. 16A) and length of 172 ±61.19µm sd (fig. 16B). All constructs had a significantly positive correlation between branch width and length ($p<0.001$; adjusted R^2: AA=0.60, AP=0.79, PA=0.62; fig. 16C).

4. DISCUSSION

4.1 Developmental differences between dendritic and stellate trichomes

One goal of this study was to identify developmental features that are unique to *Physaria* stellate trichomes, as contrasted with those that might be features of the broader Physarieae clade. Therefore, in addition to examining trichome development in *Physaria*, we also studied

members of the genus *Paysonia*, the closest relatives of *Physaria* with dendritic trichomes (Fuentes-Soriano and Al-Shehbaz 2013); Mazie and Baum, Chapter II). We further compared our measurements to those of *Arabidopsis* dendritic trichomes (Szymanski et al. 1999).

Paysonia and *Physaria* trichome stalk development appeared qualitatively similar to *Arabidopsis*. In all three genera, the trichome stalk is established after the nascent trichome cell expands perpendicularly to the plane of the leaf. While pre-branching stalks of all three genera were similar in width (~12μm; fig. 2, (Szymanski et al. 1999)), *Paysonia* and *Physaria* trichome branch initiation began at an earlier stage of stalk elongation than *Arabidopsis*, as evidenced by the finding that the tallest *Paysonia* and *Physaria* pre-branching trichomes were <14μm in length (fig. 2), approximately half as long as seen in *Arabidopsis* (Szymanski et al. 1999). Consistent with this, mature *Paysonia* trichomes had smaller stalks than *Arabidopsis* (*Physaria* stalks in mature trichomes could not be measured). *Paysonia* stalks were typically 20-22μm in width (fig. 3), which fell at the short end of the range of *Arabidopsis* stalks (Szymanski et al. 1999), and the tallest *Paysonia* stalks (fig. 3) were only half the length of the shortest *Arabidopsis* stalks (Szymanski et al. 1999). Our findings therefore indicate that a small (in comparison to *Arabidopsis*) stalk size was present in ancestors of both *Paysonia* and *Physaria*; therefore, this feature is not correlated with the evolution of the stellate form.

Branch growth appeared to be identical in *Arabidopsis*, *Paysonia*, and *Physaria*. Once branches were established, they underwent a stage of extension growth and tip refinement. Mature branches of all species generally reached a length of at least 100μm (figs. 7-10; (Szymanski et al. 1999)). The ability to initiate additional branches appeared to be dependent on branch width. In *Paysonia*, the central (third) branch of 5-branched trichomes was wider than the other four branches (fig. 8), and the width of all branches in 3-branched trichomes fell midway between these two sizes (fig. 7). This result could indicate that 5-branched trichomes arose from larger branch initials than 3-branched trichomes, allowing two branches to each initiate an additional branch. Because the central branch did not initiate an additional branch, it retained a larger basal width. The trend for increased branch initiation in wider branches is also seen in *Physaria*. *Physaria* type 1 trichomes had branch initials that were approximately half the size of *Physaria* type 2 or *Paysonia* initials (fig. 6), and these small primary branches never initiated additional branches. In addition, *Physaria* type 2 trichome branch widths decreased as branch tier increased, resulting in tertiary branches that were narrower than secondary branches (fig. 10).

These data suggest the possibility of a conserved mechanism of trichome development in Brassicaceae in which branch width determines whether a further branching event occurs, and branching that does occur always yields branches that are narrower then the previous branch.

In contrast to the similarities in branch growth between dendritic and stellate trichomes, we found important differences between these two trichome shape morphologies in the position of newly initiated branches. The pattern of branch initiation was nearly identical in the dendritic trichomes of *Paysonia* and *Arabidopsis*. In both taxa, two incipient branches of different size appeared in a consistent orientation with respect to the leaf, albeit a different orientation in these two genera. In *Arabidopsis*, one incipient branch is directed towards the leaf tip and the other towards the leaf base, while in *Paysonia* they point towards the two leaf margins. In *Physaria* type 1 and type 2 trichomes, in contrast, we noted almost simultaneous initiation of all primary branches (3-4 in *Physaria* type 2, up to 17 in *Physaria* type 1) in a ring around the apex of the branch stalk. In addition, we could not detect any consistent orientation in *Physaria*, where the incipient branch foci appear on all sides of the growing trichome stalk. The position of additional growth foci also differed between trichomes of the two dendritic species and *Physaria* type 2 trichomes. While additional branches arose from a position below the growing branch apex in *Paysonia* and *Arabidopsis*, secondary branches formed as near perfect bifurcations of the branch apex in *Physaria* type 2 trichomes.

Based on our developmental analyses, we hypothesize that a change in the pattern of branch initiation led to the evolution of the stellate form from a dendritic ancestral form and propose a model for how such a change could lead to the stellate shape (fig. 17). Both dendritic and stellate trichomes initially have one growth focus positioned away from the leaf surface, which results in the trichome initial undergoing polarized expansion perpendicular to the plane of the leaf epidermis, establishing the trichome stalk (fig. 17A, B). In dendritic trichomes, a secondary growth focus is established at a position beneath the growing apex. The original growth focus continues to grow, albeit at an increase angle, as the secondary focus begins its own polarized expansion in the opposite direction, creating two primary branches. Additional growth foci reiterate this pattern, and can be initiated subterminally on any elongating branch of sufficient width (fig. 17A). In the stellate trichomes of *Physaria*, however, the pattern is quite different. Growth of the primary focus ceases and instead all growth is redirected into secondary growth foci, which are initiated more or less simultaneously around the apex of the pre-branching stalk

(fig. 17B). Additional bifurcations occur in *Physaria* type 2 trichomes, apparently by the formation of dichotomous growth foci at secondary growth apices of sufficient diameter (fig. 17C). The small initial size of *Physaria* type 1 branches prevents the establishment of additional branches, and these primary branches simply expand to form the multi-branched, stellate shape of the mature trichome (fig. 17D).

4.2 The role of *BRANCHLESS TRICHOME* in the control of trichome morphology

A secondary goal of our research was to better understand the genetic changes that underlie the evolution of *Physaria* stellate trichomes. Stellate trichomes appear to have arisen as a result of developmental changes to branch patterning, in conjunction with an increase in branch number. In the dendritic trichomes of *Arabidopsis*, *BRANCHLESS TRICHOME* (*BLT*) is postulated to play a role in positioning branch initiation sites and determining the ultimate number of branches a trichome initiates (Marks et al. 2009; Kasili et al. 2011). We explored the contribution of *BLT* to the altered branch morphology and increased branch number of stellate trichomes by means of a candidate gene approach where we introduced the *P. fendleri BLT* coding and promoter region into an *A. thaliana blt* mutant background and compared the results with those obtained by rescuing the *blt* mutant with the *A. thaliana BLT* gene.

We found that the *P. fendleri BLT* promoter region could replace the *A. thaliana* homolog, with no detectable differences in trichome morphology or branch number. In contrast, the *P. fendleri* BLT protein only partially rescued the mutant phenotype of *A. thaliana blt* mutant plants. *A. thaliana blt* plants transformed with the *P. fendleri BLT* coding region had fewer branches than those transformed with the *A. thaliana BLT* coding region, and trichomes of T2 plants that contained the *P. fendleri BLT* coding region had significantly shorter and narrower trichome stalks than those with the *A. thaliana* coding region. Although *P. fendleri BLT* led to a reduced trichome branch number, it did not affect branch morphology. The branch dimensions of all T2 lines measured were the same, regardless of the whether they expressed the *P. fendleri* or *A. thaliana BLT* coding region.

Two hypotheses could explain the partial rescue of *A. thaliana blt* lines transformed with the *P. fendleri BLT* coding region. The first is that *P. fendleri BLT* is hypomorphic in the *A. thaliana* background. Although *P. fendleri BLT* has some level of functionality, its ability to interact with other *A. thaliana* trichome branch-promoting proteins is reduced, which results in trichomes with

fewer branches than seen in plants rescued with *A. thaliana BLT*. The reduced activity of *P. fendleri BLT* may also reduce the rate of stalk growth, resulting in stalks that are shorter and narrower at the time of branch initiation. In this scenario, the early retardation of *P. fendleri BLT* stalk growth results in a shifted growth trajectory that leads to mature trichomes with narrower and shorter stalks than those of plants with *A. thaliana BLT*. A second hypothesis is that the small stalk phenotype and reduced branch number of plants transformed with *P. fendleri BLT* is a hypermorphic response. In this scenario, increased activity of *P. fendleri BLT* leads to branch initiation that occurs earlier in stalk development than with *A. thaliana BLT*. This narrower, shorter stalk initiates fewer branches, and, just as in the hypomorphic scenario, shifts the growth trajectory of *P. fendleri BLT* stalks, resulting in mature stalks that are narrower and shorter than those of plants transformed with *A. thaliana BLT*. This second hypothesis is intriguing as it implies a possible function for *BLT* in regulating stalk size by controlling the timing of branch initiation. In addition, the transgenic phenotype of trichomes with small stalks but normal sized branches is shared with the donor species, *Physaria*, although not specific to the stellate form, as it is also a feature of *Paysonia* trichomes.

Physaria BLT was able to promote some trichome branching in the *Arabidopsis* background, but it did not increase branch number or modify branch initiation patterns. Overall, our results can conclusively rule out changes in *BLT* as having driven *Physaria* stellate trichome evolution. These findings are consistent with molecular evolutionary analyses that found that *BLT* is under positive selection in the *Physaria* crown lineage, rather than on the stem lineage where stellate trichomes evolved (Mazie and Baum, Chapter II). The results of molecular evolutionary and transgenic analyses suggest that changes in *Physaria BLT* occurred after the evolution of the stellate form, possibly in response to other genetic changes that drove stellate trichome evolution. In *Arabidopsis*, BLT is postulated to form a protein complex with STICHEL to promote branch initiation (Kasili et al. 2011). Perhaps changes to the STI protein or the BLT-STI protein complex were necessary for the evolution of the stellate form. Future studies that explore the possible contribution of these and other branch promoting genes to the evolution of trichome shape have the potential to provide novel insights into not only the evolutionary mechanisms that underlie the evolution of Brassicaceae trichomes, but also general principles of cell shape evolution.

ACKNOWLEDGEMENTS

The authors thank S. Fuentes-Soriano for providing leaf material, the USDA National Plant Germplasm System for providing seeds, M. Otegui for providing pCAMBIA1300 and *A. tumefaciens* strain GV3101, and UW-Madison undergraduate students Emily Kief, Amanda Schussman, Morgan Sell, and Young me Yoon for laboratory assistance. A. Mazie was supported through the University of Wisconsin-Madison Botany Department (Davis Summer Fellowship, Judy Croxdale Award for Women in Science, ON and EK Allen Fellowship) and University of Wisconsin-Madison Beta Chapter of Sigma Delta Epsilon-Graduate Women in Science (Ruth Dickie Grant-In-Aid).

LITERATURE CITED

Al-Shehbaz IA, Beilstein MA, Kellogg EA. 2006. Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Plant Syst. Evol. 259:89–120.

Appel O, Al-Shehbaz IA. 2003. Cruciferae. In: Kubitzki K, editor. Families and genera of vascular plants. Vol. 5. Berlin (Germany): Springer Berlin Heidelberg. pp. 75–174.

Beilstein MA, Al-Shehbaz IA, Kellogg EA. 2006. Brassicaceae phylogeny and trichome evolution. American Journal of Botany 93:607–619.

Beilstein MA, Al-Shehbaz IA, Mathews S, Kellogg EA. 2008. Brassicaceae phylogeny inferred from phytochrome A and *ndhF* sequence data: tribes and trichomes revisited. American Journal of Botany 95:1307–1327.

Chanut-Delalande H, Ferrer P, Payre F, Plaza S. 2012. Effectors of tridimensional cell morphogenesis and their evolution. Semin. Cell Dev. Biol. 23:341–349.

Clough, Steven J. and Andrew F. Bent. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. The Plant Journal 16 (6): 735-743.

Correa, Raul, John Stanga, Bret Larget, Aaron Roznowski, Guoping Shu, Brian Dilkes, and David A. Baum. 2012. An assessment of transgenomics as a tool for identifying genes involved in the evolutionary differentiation of closely related plant species. New Phytologist 193: 494-503.

Fuentes-Soriano S, Al-Shehbaz I. 2013. Phylogenetic relationships of mustards with multiaperturate pollen (Physarieae, Brassicaceae) based on the plastid *ndhF* gene: implications for morphological diversification. Systematic Botany 38:178–191.

Hachet O, Bendezú FO, Martin SG. 2012. Fission yeast: in shape to divide. Curr. Opin. Cell Biol. 24:858–864.

Hanley ME, Lamont BB, Fairbanks MM, Rafferty CM. 2007. Plant structural traits and their role in anti-herbivore defence. Perspectives in Plant Ecology, Evolution and Systematics 8:157–178.

Harrison, Samuel J, Ellie K Mott, Kate Parsley, Sue Aspinall, John C Gray, and Amanda Cottage. 2006. A rapid and robust method of identifying transformed *Arabidopsis thaliana* seedlings following floral dip transformation. Plant Methods 2:19.

Hülskamp M, Miséra S, Jürgens G. 1994. Genetic dissection of trichome cell development in Arabidopsis. Cell 76:555–566.

Ilgenfritz H, Bouyer D, Schnittger A, Mathur J, Kirik V, Schwab B, Chua N-H, Jürgens G, Hülskamp M. 2003. The Arabidopsis *STICHEL* gene is a regulator of trichome branch number and encodes a novel protein. Plant Physiology 131:643–655.

Kasili R, Huang CC, Walker JD, Simmons LA, Zhou J, Faulk C, Hülskamp M, Larkin JC. 2011.

BRANCHLESS TRICHOMES links cell shape and cell cycle control in *Arabidopsis* trichomes. Development 138:2379–2388.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

Marks MD, Wenger JP, Gilding E, Jilk R, Dixon RA. 2009. Transcriptome analysis of *Arabidopsis* wild-type and *gl3-sst sim* trichomes identifies four additional genes required for trichome development. Molecular Plant 2:803–822.

O'Kane SL Jr, Al-Shehbaz IA. 2002. Paysonia, a new genus segregated from Lesquerella (Brassicaceae). Novon.

Puzey JR, Gerbode SJ, Hodges SA, Kramer EM, Mahadevan L. 2011. Evolution of spur-length diversity in Aquilegia petals is achieved solely through cell-shape anisotropy. Proceedings of the Royal Society B 279:1640-1645.

R Core Team. R: A language and environment for statistical computing. Available from: https://www.R-project.org/

Rasband W. ImageJ. Available from: http://imagej.nih.gov/ij/

Rollins R, Shaw E. 1973. The genus *Lesquerella* (Cruciferae) in North America. Cambridge, Mass.: Harvard University Press

Rollins RC, Banerjee UC. 1975. Atlas of the trichomes of *Lesquerella* (Cruciferae). Cambridge: Mass.

Rollins RC, Banerjee UC. 1976. Trichomes in studies of the Cruciferae. In: Vaughan JG, Macleod AJ, Jones BMG, editors. The biology and chemistry of the Cruciferae. London: the Biology and Chemistry of the Cruciferae. …. pp. 145–166.

Szymanski DB, Jilk RA, Pollock SM, Marks MD. 1998. Control of GL2 expression in Arabidopsis leaves and trichomes. Development 125:1161–1171.

Szymanski DB, Marks MD, Wick SM. 1999. Organized F-Actin Is Essential for Normal Trichome Morphogenesis in Arabidopsis. The Plant Cell 11:2331–2347.

Vollrath MA, Kwan KY, Corey DP. 2007. The Micromachinery of Mechanotransduction in Hair Cells. Annual review of neuroscience 30:339–365.

Wagner GJ, Wang E, Shepherd RW. 2004. New approaches for studying and exploiting an old protuberance, the plant trichome. Annals of Botany 93:3–11.

Wagner JK, Setayeshgar S, Sharon LA, Reilly JP, Brun YV. 2006. A nutrient uptake role for bacterial cell envelope extensions. Proc. Natl. Acad. Sci. U.S.A. 103:11772–11777.

Werker E. 2000. Trichome diversity and development. Advances in botanical research 31:1–35.

Whitney HM, Bennett KMV, Dorling M, Sandbach L, Prince D, Chittka L, Glover BJ. 2011. Why do so many petals have conical epidermal cells? Annals of Botany 108:609–616.

Wickham H. 2009. ggplot2: elegant graphics for data analysis.

Wickham H. 2011. The Split-Apply-Combine Strategy for Data Analysis. J. Stat. Soft. 40.

FIGURE LEGENDS

Fig. 1. Diagrammatic representation of transformation constructs. Dark grey boxes indicate genomic region from *A. thaliana*. Lights grey boxes indicate genomic region from *P. fendleri*.

Fig. 2. *Paysonia* and *Physaria* pre-branching stalk measurements. Boxplots of stalk width (A) and length (B). Diamonds indicate mean values. (C) Scatterplot of stalk length as a function of width. Trend lines indicate statistically significant linear regressions.

Fig. 3. *Paysonia* mature stalk measurements. Boxplots of stalk width (A) and length (B). Diamonds indicate mean values. (C) Scatterplot of stalk length as a function of width.

Fig. 4. *Paysonia auriculata* trichome branch development. (A) Post-branching trichome with five branches. The first, second, and third emerging branches are labeled. We believe that the first branch corresponds to the continued growth of the primary axis. (B) Post-branching trichomes at several stages of branch initiation. (C) Mature trichomes.

Fig. 5. *Physaria* type 1 (represented by *P. fendleri*) and *Physaria* type 2 (represented by *P.* aff. *arizonica*) trichome branch development. (A) Several stages of *Physaria* type 1 pre-branching and post-branching trichomes. (B) Mature *Physaria* type 1 trichome. (C) Several stages of *Physaria* type 2 pre-branching and post-branching trichomes. Black arrow points to early bifurcating branches. White arrow points to pre-branching trichome with symmetrical branch bifurcations. White arrowhead points to pre-branching trichome with asymmetrical branch bifurcations. (D) Mature *Physaria* type 2 trichome.

Fig. 6. *Paysonia*, *Physaria* type 1, and *Physaria* type 2 primary trichome branch measurements. Boxplots of branch width (A) and length (B). Diamonds indicate mean values. (C) Scatterplot of trichome branch length as a function of branch width. Trend lines indicate statistically significant linear regressions.

Fig. 7. *Paysonia* three-branched, mature trichome measurements, separated by branch position. Boxplots of branch width (A) and length (B). Diamonds indicate mean values. (C) Scatterplot of trichome branch length as a function of branch width.

Fig. 8. *Paysonia* five-branched, mature trichome measurements, separated by branch position. Boxplots of branch width (A) and length (B). Diamonds indicate mean values. (C) Scatterplot of trichome branch length as a function of branch width.

Fig. 9. *Physaria* type 1 mature trichome measurements, separated by trichome branch number – class A included 10 branched trichomes, class B included 11-16 branched trichomes, and class C included 18-19 branched trichomes. Boxplots of branch width (A) and length (B). Diamonds indicate mean values. (C) Scatterplot of trichome branch length as a function of branch width.

Fig. 10. *Physaria* type 2 mature trichome measurements, separated by species and branch level. Boxplots of branch width (A) and length (B). Diamonds indicate mean values. (C) Scatterplot of trichome branch length as a function of branch width. Trend lines indicate linear regressions (statistically significant for *P.* aff. *tenella*).

Fig. 11. CAPs analysis of *Physaria BLT* variant sequences. The banding pattern across taxa indicates that the variant sequences are alleles. Molecular weight marker band size is indicated to the left in bps. Lanes with a single band have allele B, lanes with two bands have allele A, lanes with three bands are heterozygous. (1) *P.* aff. *tenella*, (2) *P. fendleri* acc. 596453, (3) *P. fendleri* acc. 596455, (4) *P. fendleri* acc. 596456, (5) *P. fendleri* acc. 596457, (6) *P. fendleri* acc. 596458, (7) *P. fendleri* acc. 643177, (8) *P. fendleri* acc. SF201.

Fig. 12. Coding region alignment of *A. thaliana BLT* to *P. fendleri BLT* alleles A and B.

Fig. 13. Histogram of trichome tip counts for the four construct classes (AA, AP, PA, PP). Red, dashed lines indicate mean values.

Fig. 14. Comparison of (A) *A. thaliana blt* mutant trichomes with 3-branched trichomes from T2 plants transformed with construct (B) AA, (C) AP, and (D) PA.

Fig. 15. Comparison of trichome stalk measurements for T2 lines expressing either the *A. thaliana* or *P. fendleri* coding region. Boxplots of stalk width (A) and length (B). Diamonds indicate mean values. (C) Scatterplot of trichome stalk length as a function of width. Trend line indicates statistically significant linear regression for *A. thaliana* coding region.

Fig. 16. Comparison of T2 trichome branch measurements from T1 lines transformed with construct AA, AP, or PA. Boxplots of branch width (A) and length (B). Diamonds indicate mean values. (C) Scatterplot of branch length as a function of width. Trend lines are statistically significant linear regressions.

Fig. 17. Model of developmental differences in trichome branch initiation. Side view of *Paysonia* dendritic development (A) and *Physaria* stellate development (B). Arrowhead indicates original growth focus, which continues to elongate and initiate new branches in dendritic trichomes, but ceases growth in stellate trichomes as soon as branches are initiated. Top view of *Physaria* type 2 (C) and *Physaria* type 1 (D) trichome development. *Physaria* type 2 trichome branches bifurcate, creating multiple secondary branch points (although not indicated, additional branch bifurcations are possible). *Physaria* type 1 trichome branches are smaller, more numerous, and do not initiate additional branches.

FIGURES

Fig. 1.

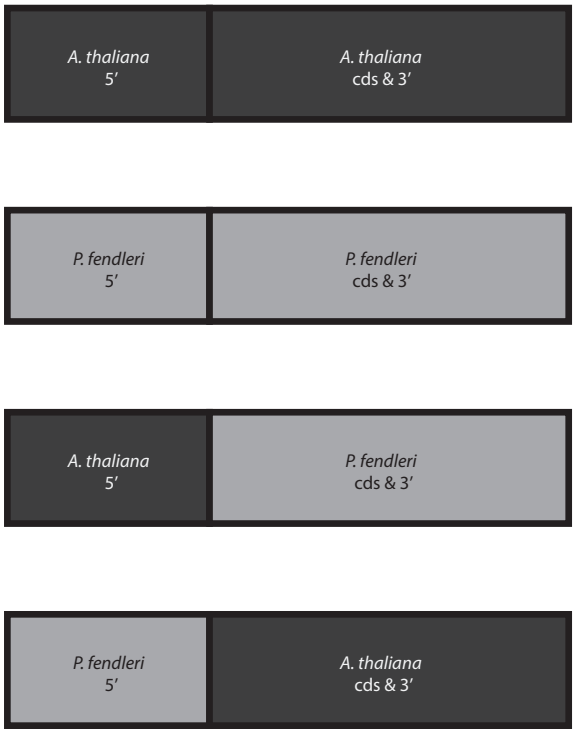


Fig. 2.

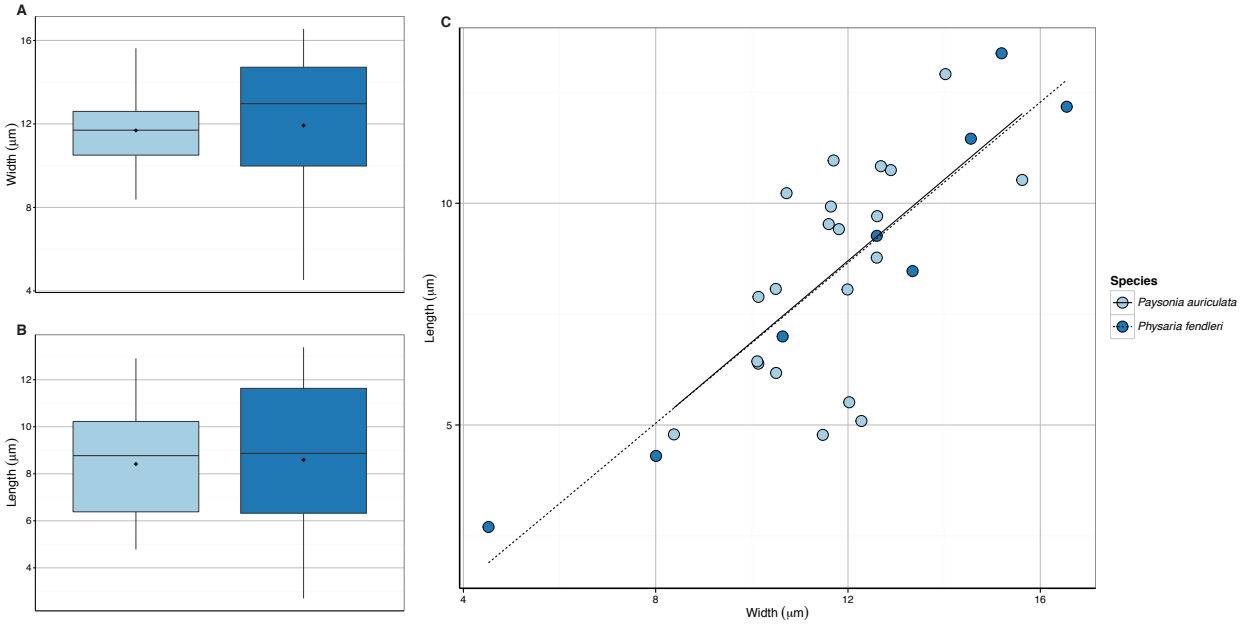


Fig.3.

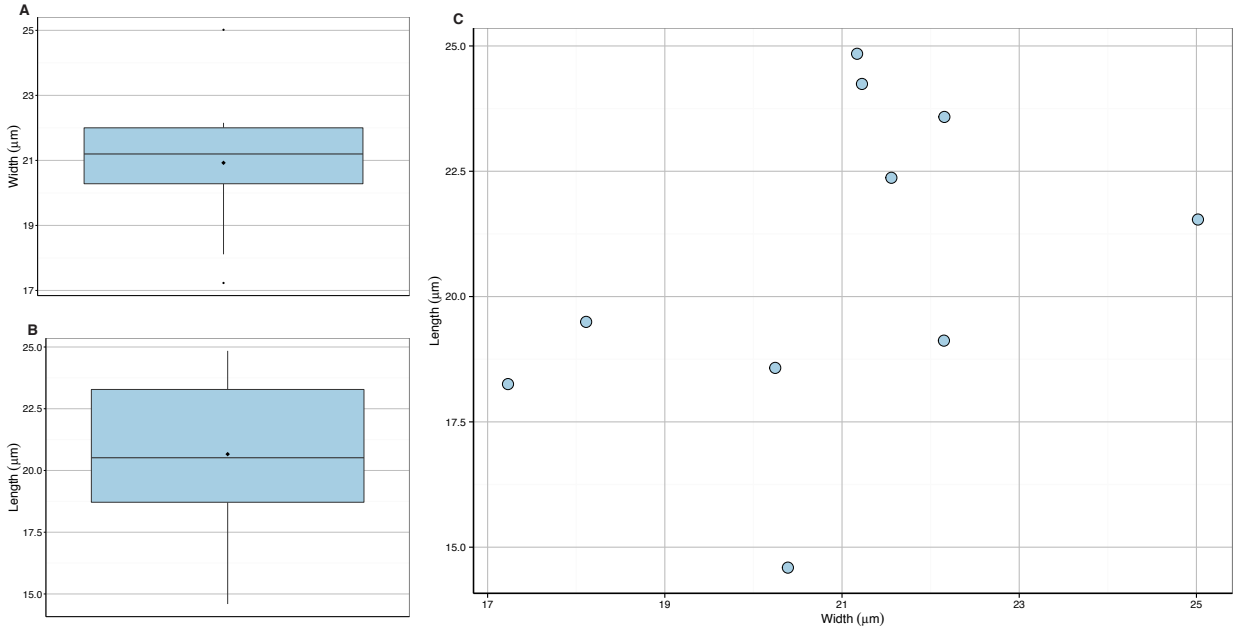

A
Width (µm)
B
Length (µm)
C
Length (µm)
Width (µm)

Fig.4.

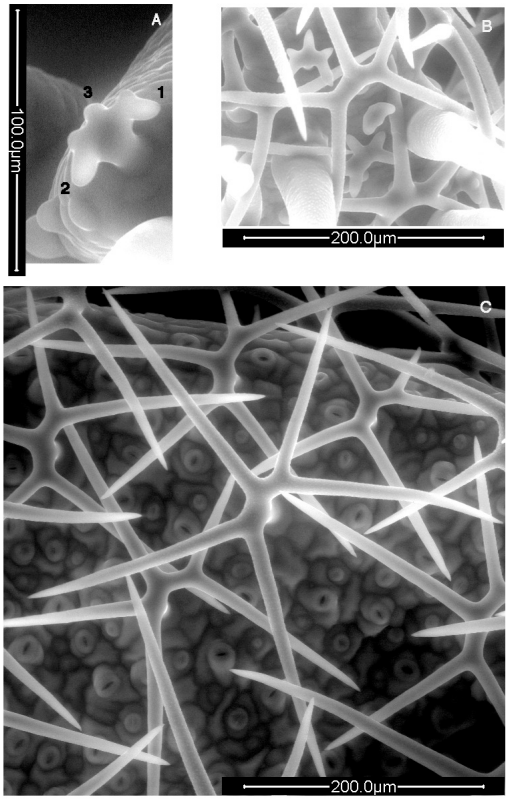

A
1
2
3
100.0µm
B
200.0µm
C
200.0µm

Fig. 5.

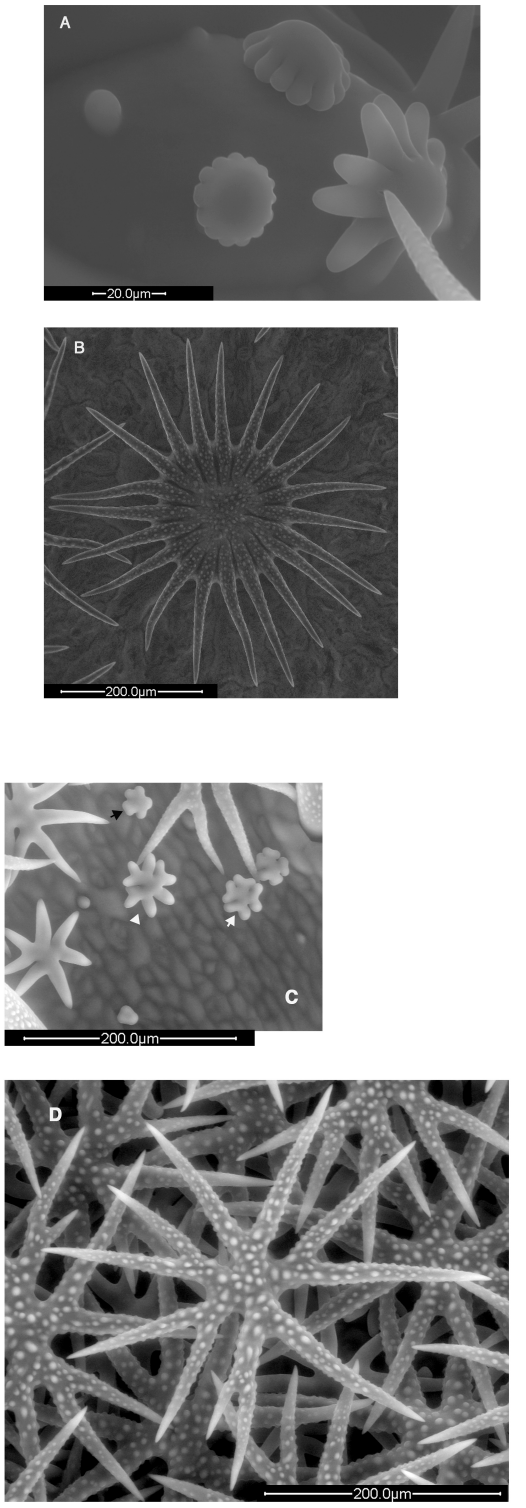


Fig. 6.

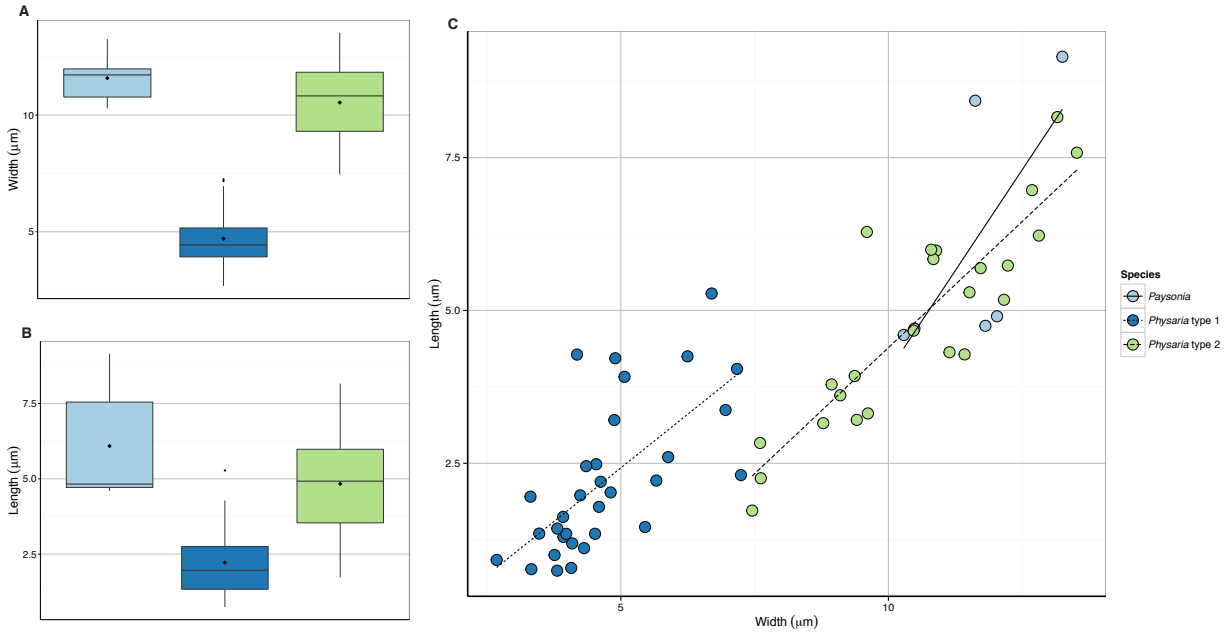

A
Width (µm)
10
5
B
Length (µm)
7.5
5.0
2.5
C
Length (µm)
7.5
5.0
2.5
5
10
Width (µm)
Species
Paysonia
Physaria type 1
Physaria type 2

Fig. 7.

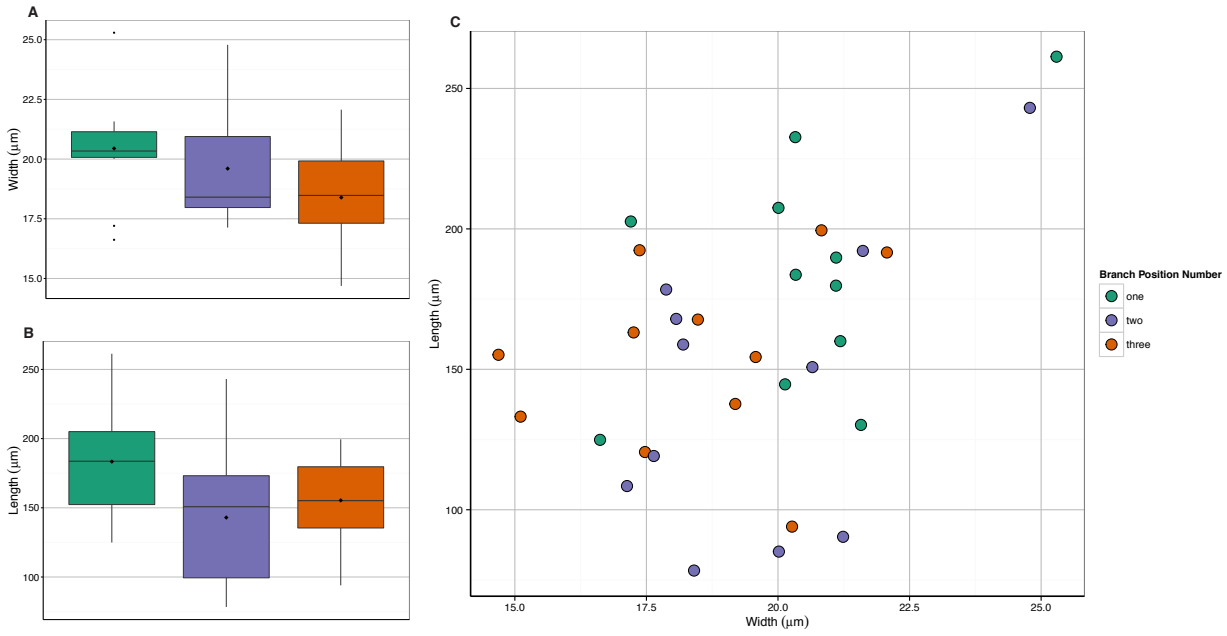

A
Width (µm)
25.0
22.5
20.0
17.5
15.0
B
Length (µm)
250
200
150
100
C
Length (µm)
250
200
150
100
15.0
17.5
20.0
22.5
25.0
Width (µm)
Branch Position Number
one
two
three

Fig. 8.

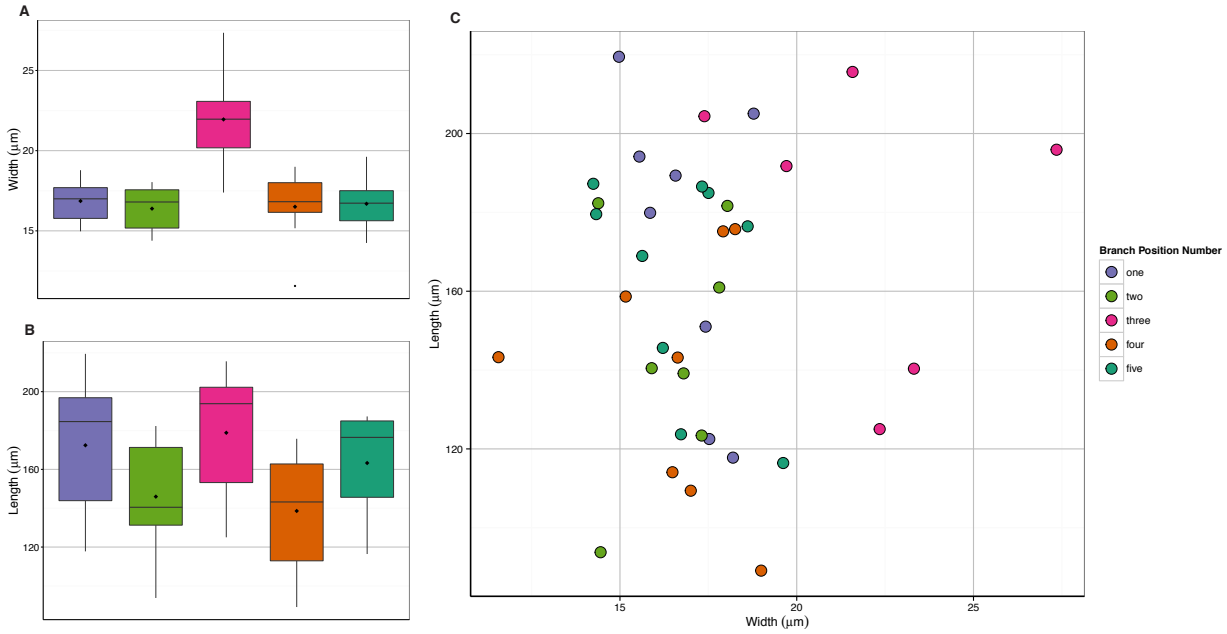

A
Width (µm)
25
20
15
B
Length (µm)
200
160
120
C
Length (µm)
200
160
120
15
20
25
Width (µm)
Branch Position Number
one
two
three
four
five

Fig. 9.

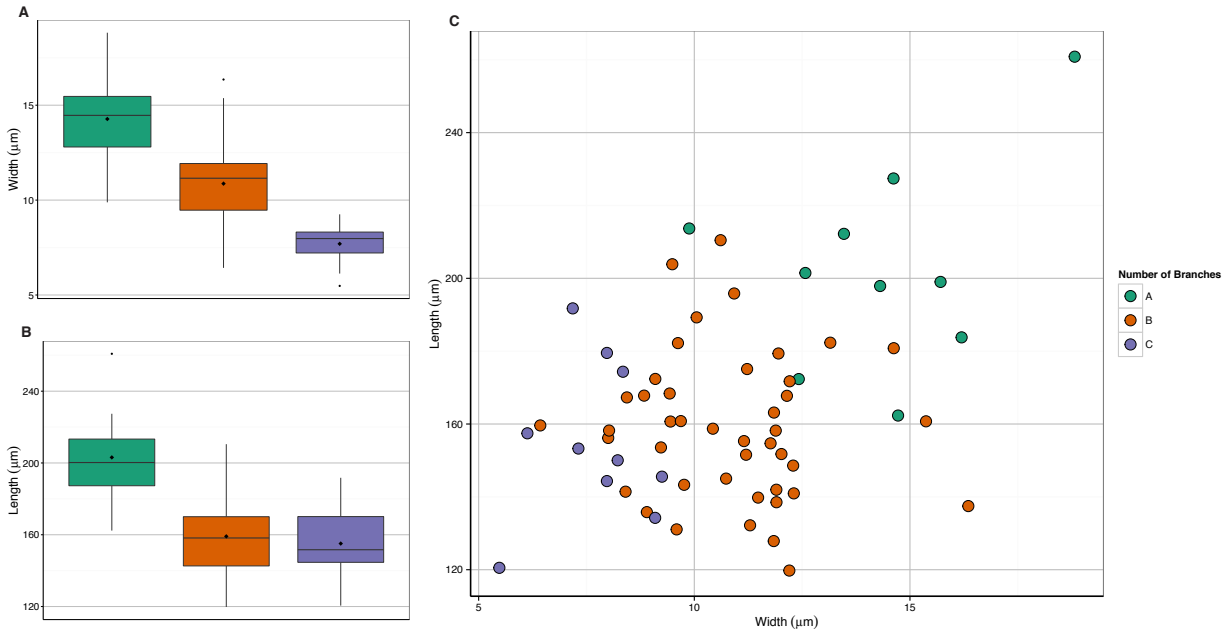

A
Width (µm)
15
10
5
B
Length (µm)
240
200
160
120
C
Length (µm)
240
200
160
120
5
10
15
Width (µm)
Number of Branches
A
B
C

Fig. 10.

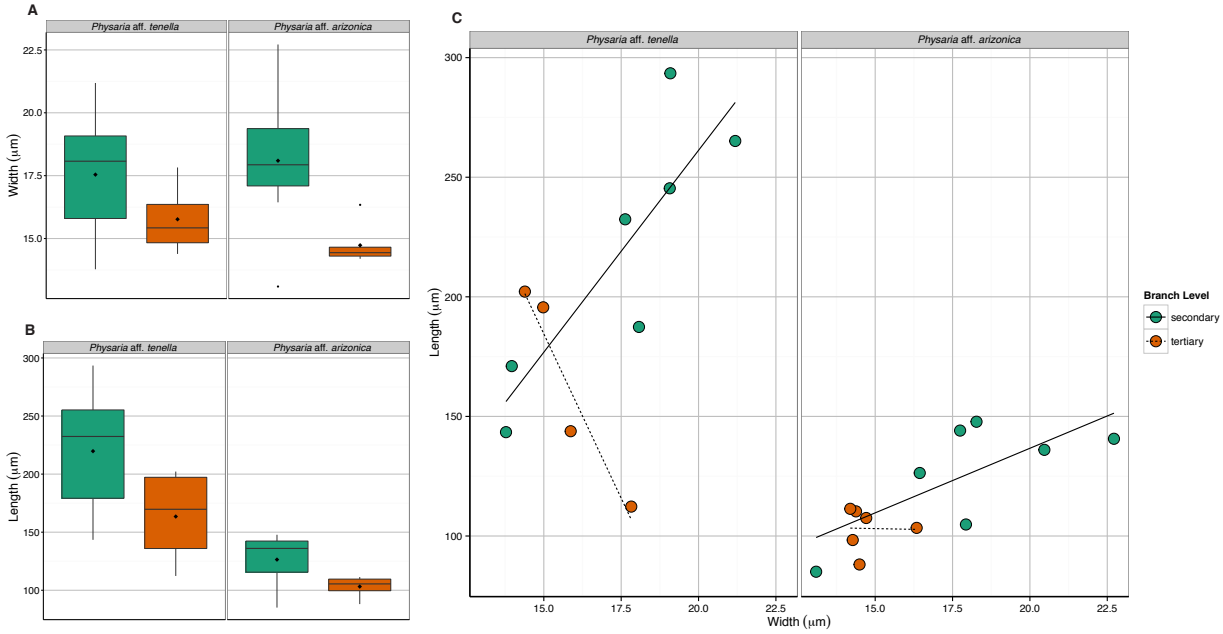

A
Physaria aff. tenella
Physaria aff. arizonica
Width (μm)
22.5
20.0
17.5
15.0
B
Physaria aff. tenella
Physaria aff. arizonica
Length (μm)
300
250
200
150
100
C
Physaria aff. tenella
Physaria aff. arizonica
Length (μm)
Width (μm)
15.0
17.5
20.0
22.5
Branch Level
secondary
tertiary

Fig. 11.

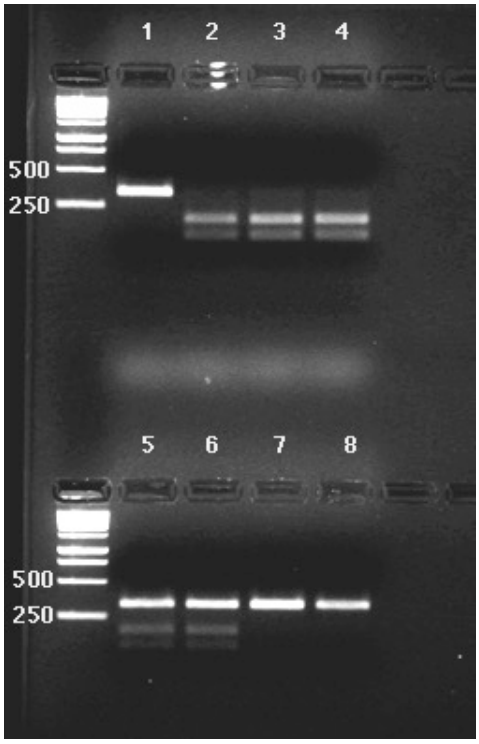

1
2
3
4
500
250
5
6
7
8
500
250

Fig. 12.

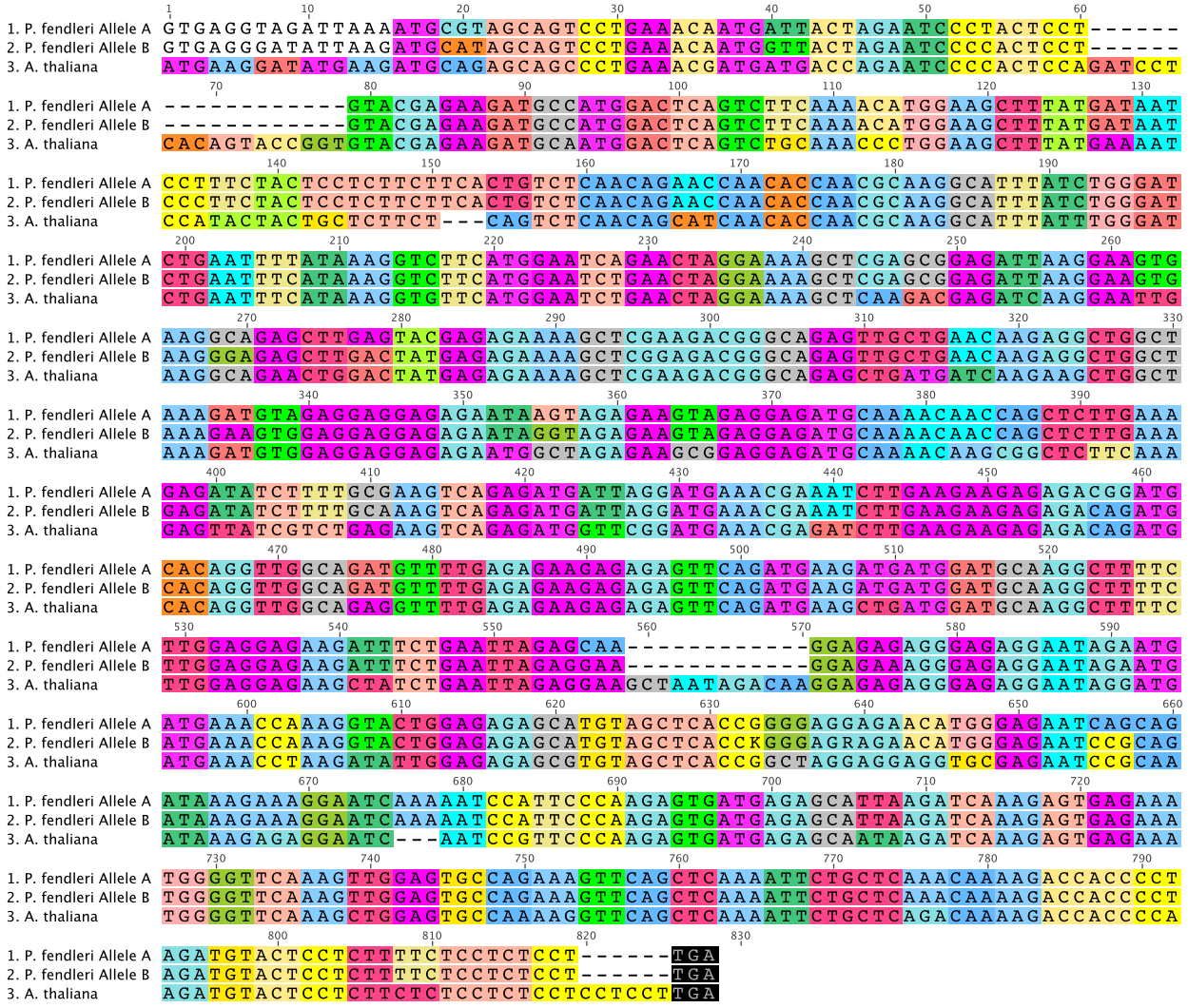


Fig. 13.

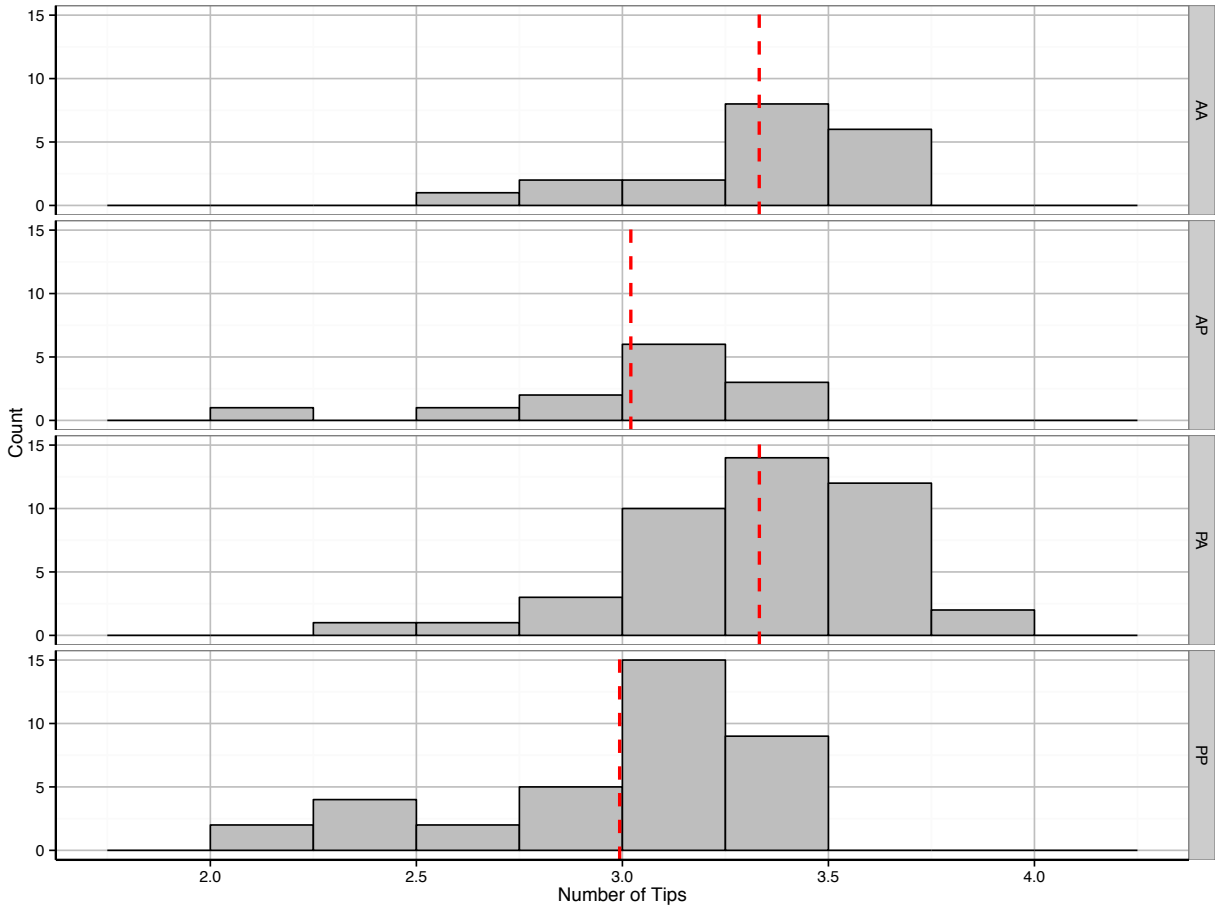


Fig. 14.

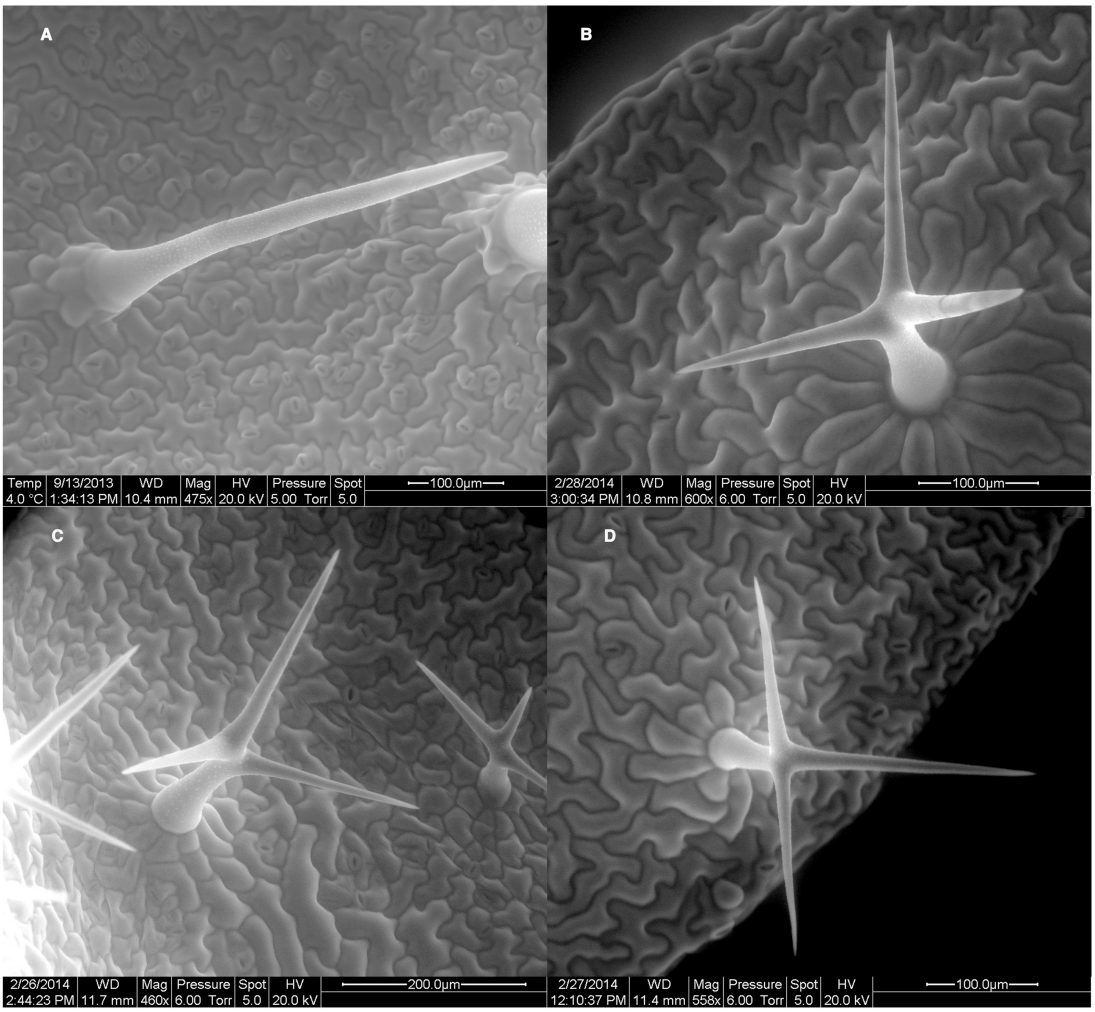

A
Temp 4.0 °C
9/13/2013 1:34:13 PM
WD 10.4 mm
Mag 475x
HV 20.0 kV
Pressure 5.00 Torr
Spot 5.0
100.0µm
B
2/28/2014 3:00:34 PM
WD 10.8 mm
Mag 600x
Pressure 6.00 Torr
Spot 5.0
HV 20.0 kV
100.0µm
C
2/26/2014 2:44:23 PM
WD 11.7 mm
Mag 460x
Pressure 6.00 Torr
Spot 5.0
HV 20.0 kV
200.0µm
D
2/27/2014 12:10:37 PM
WD 11.4 mm
Mag 558x
Pressure 6.00 Torr
Spot 5.0
HV 20.0 kV
100.0µm

Fig. 15.

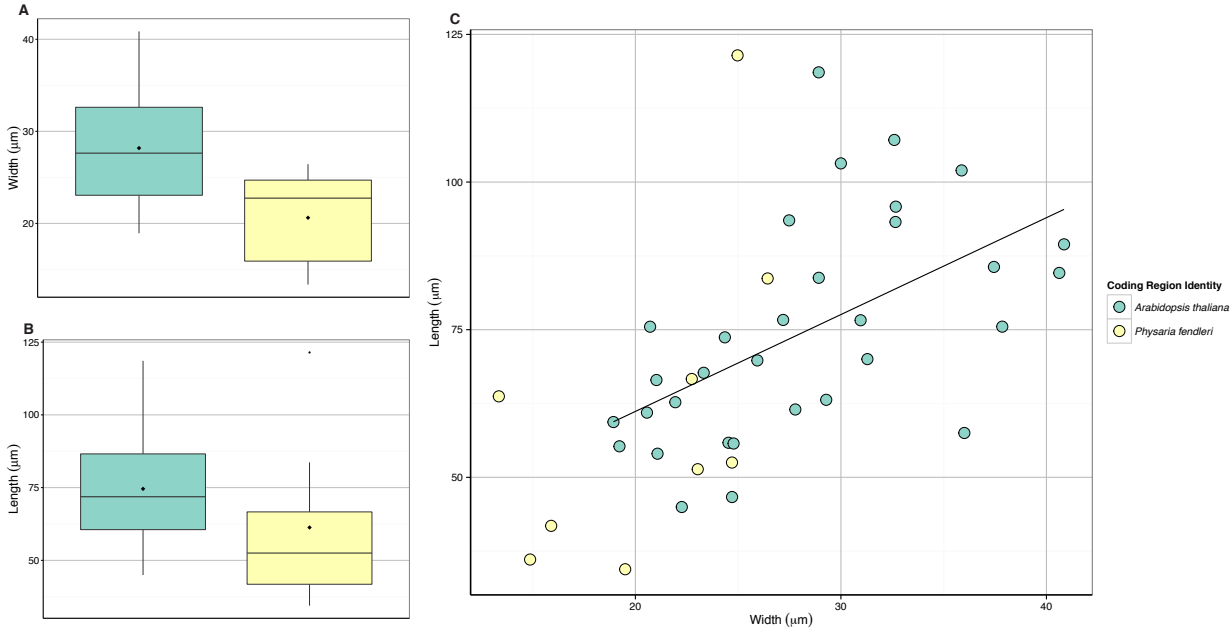

A
Width (µm)
B
Length (µm)
C
Length (µm)
Width (µm)
Coding Region Identity
Arabidopsis thaliana
Physaria fendleri

Fig. 16.

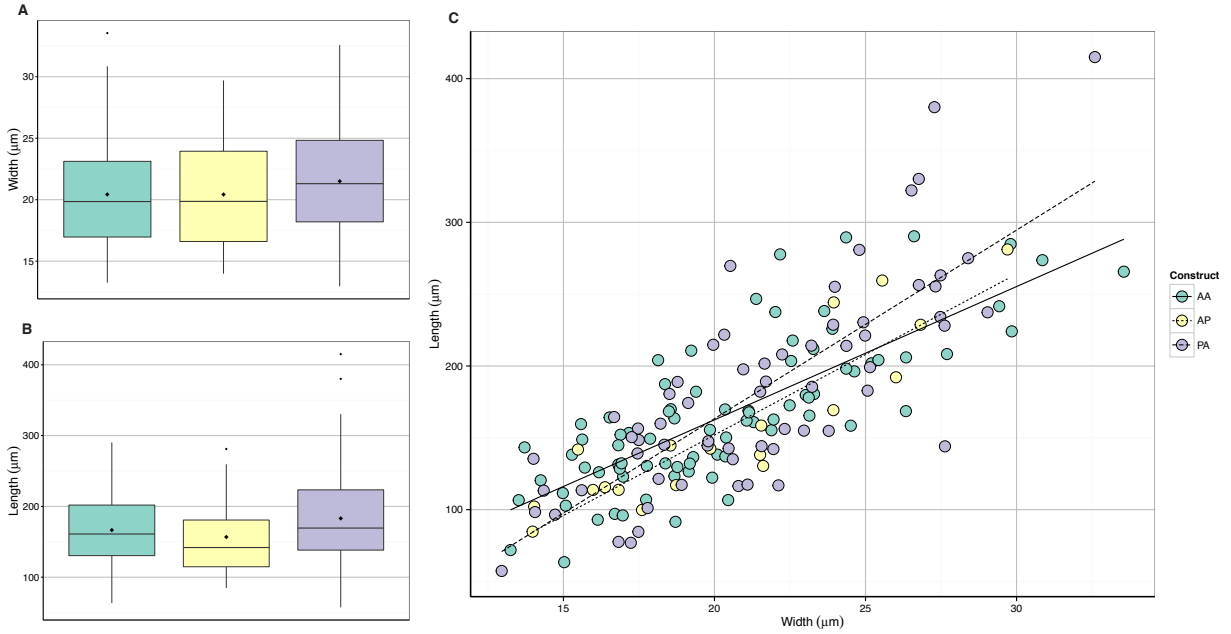
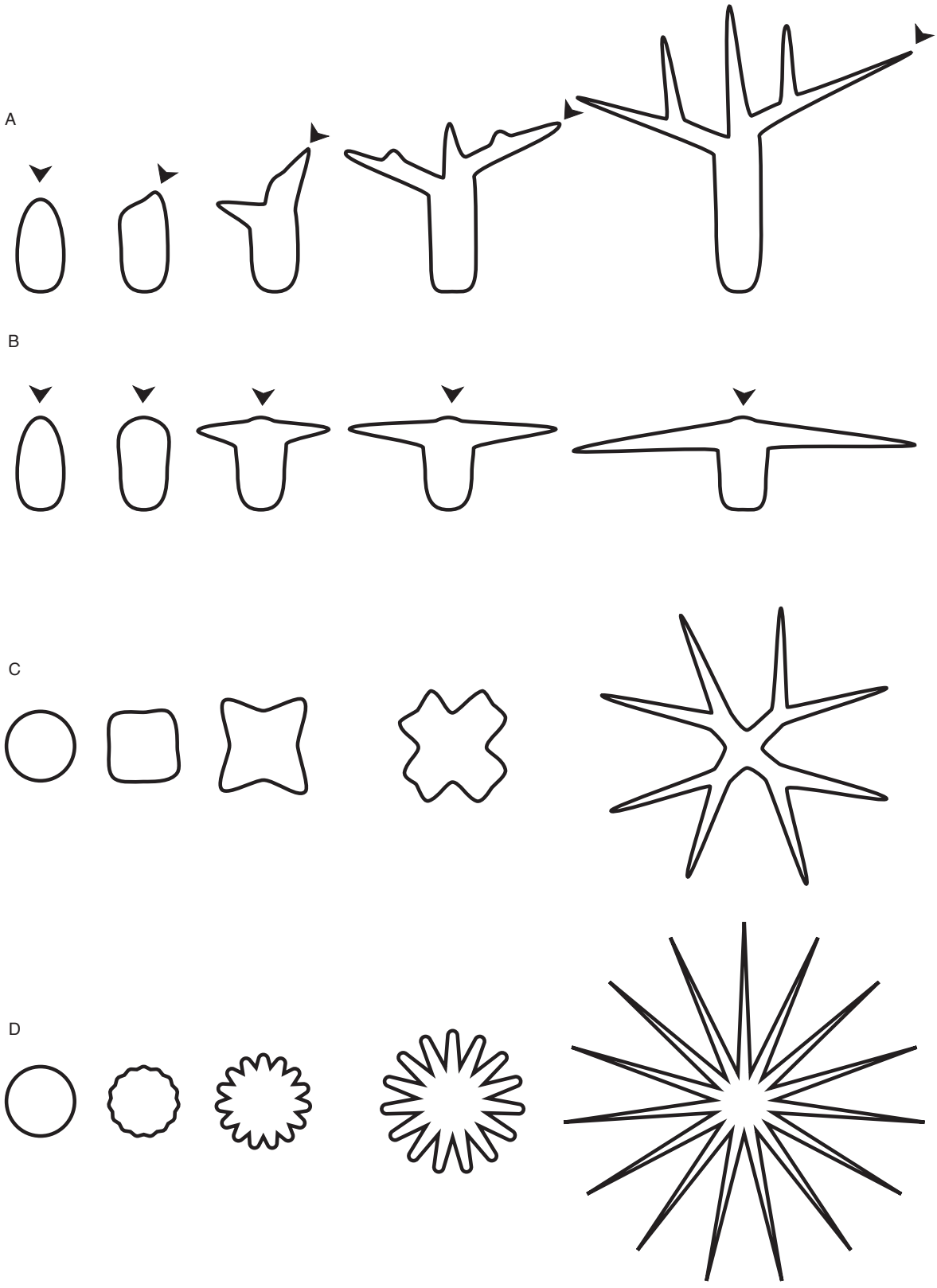

A
Width (µm)
B
Length (µm)
C
Length (µm)
Width (µm)
Construct
AA
AP
PA

Fig. 17.



SUPPLEMENTARY TABLES

Supplementary Table 1. Accessions included in analyses

Scientific Name	DNA Source, accession number in parentheses	Voucher/Accession
Arabidopsis thaliana (L.) Heynh. ecotype Columbia	1	
Arabidopsis thaliana (L.) Heynh. ecotype Columbia *blt-1*	2	
Paysonia auriculata (Engelm. & A.Gray) O'Kane & Al-Shehbaz	3	W6 20320, PI 345712
Paysonia densipila (Rollins) O'Kane & Al-Shehbaz	3	PI 315051, PI 315052
Paysonia grandiflora (Hook.) O'Kane & Al-Shehbaz	3	W6 20837
Paysonia lasiocarpa (Hook. ex A.Gray) O'Kane & Al-Shehbaz	3	PI 643173
Paysonia stonensis (Rollins) O'Kane & Al-Shehbaz	3	PI 355042
Physaria argyraea (A.Gray) O'Kane & Al-Shehbaz	3	W6 20816
Physaria fendleri (A.Gray) O'Kane & Al-Shehbaz	4	Fuentes-Soriano et al. 201, MO
	3	PI 596453, PI 596455, PI 596456, PI 596457, PI 596458, PI 643177
Physaria aff. *arizonica*	3	PI 643169
Physaria aff. *pallida*	3	W6 20864
Physaria aff. *tenella*	3	PI 355037

1: Seed provided by Otegui Lab, Botany Department, UW-Madison
2: ABRC stock #CS827202
3: Seed provided by USDA NPGS
4: Dried leaf material provided by S. Fuentes-Soriano

Supplementary Table 2. Primer Sequences

Primer	Sequence (5' -> 3')
CAPS markers	
BLT_IDF	CCTGTACGAGAAGATGCCATGGACTCAG
BLT_IDR	TCTCTTTCAAGAGCTGGTTGTTTTGCAT
***P. fendleri BLT* allele A Variant specific**	
BLTcdsVAR1F	CTCCTCTCTATTCTGAGGAAGAATGAATAAAAC
BLTcdsVAR2R	CAGCTAGAGAATAAGATGAGAAGATTATTGAGAAAG
***P. fendleri BLT* allele B Variant specific**	
BLTcdsVAR2F	GCAAGAAAAAGTGAGGGATATTAAG
BLTcdsVAR1R	GGAGGATGAACACCGAAAGTAGGAATGGA
AA Construct	
kasiliBLTF	CCATGATTACGAATTCACCTGTGACAGACTTGCTCCTAC
kasiliBLTR	CGACTCTAGAGGATCCAGCCACATAAGGACCCGAC
AP Construct	
kasiliBLTF	CCATGATTACGAATTCACCTGTGACAGACTTGCTCCTAC
Ath5'R	CTTTTTGGTTGGCCTTCTATATTTGAG
Pfend3'RwExt	CGACTCTAGAGGATCCCTGGAGAATCGGATGGTAAGGCGT
PfendcdsFAthExt	AGGCCAACCAAAAAGGTGAGGTAGATTAAAATGCGTAGCAG
PA Construct	
Pfend5'Fno2wExt	CCATGATTACGAATTAGCACTTCTTCAACAACAGCTTCACTCA
Pfend5'R	TTTTTCTTGCTTTAGTTGTAAAAAGAGGGTAAAG
AthcdsFPfeExt	CTAAAGCAAGAAAAAATGAAGGATATGAAGATGCAGAGCAG
kasiliBLTR	CGACTCTAGAGGATCCAGCCACATAAGGACCCGAC
PP Construct	
Pfend5'Fno2wExt	CCATGATTACGAATTAGCACTTCTTCAACAACAGCTTCACTCA
Pfend3'RwExt	CGACTCTAGAGGATCCCTGGAGAATCGGATGGTAAGGCGT

SUPPLEMENTARY FIGURE LEGENDS

Sup. Fig. 1. ESEM images of developing *Paysonia* trichomes with two branch tips. (A-D) *Pay. auriculata*, (E-J) *Pay. densipila*, (K) *Pay. grandiflora*, (L) *Pay. lasiocarpa*, (M-N) *Pay. stonensis*. (O-S) *A. thaliana* 2-branched trichomes included for comparison.

Sup. Fig. 2. ESEM images of developing *Physaria* type 1 trichomes. (A-B) *Phy. argyraea*, (C-F) *Phy. fendleri*.

Sup. Fig. 3. ESEM images of developing *Physaria* type 2 trichomes. (A-H) *Phy*. aff. *tenella*, (I-L) *Phy*. aff. *arizonica*, (M-O) *Phy*. aff. *pallida*.

Sup. Fig. 4. ESEM images of developing *Paysonia* trichomes with three branch tips. (A-C) *Pay. auriculata*, (D) *Pay. densipila*, (E-G) *Pay. lasiocarpa*. (H-I) *A. thaliana* 3-branched trichomes included for comparison.

Sup. Fig. 5. ESEM images of developing *Paysonia* trichomes with four or more branch tips. (A-C) *Pay. auriculata*, (D-H) *Pay. densipila*, (I-M) *Pay. grandiflora*, (N-P) *Pay. lasiocarpa*.

Sup. Fig. 6. ESEM images of mature *Paysonia* trichomes. (A) *Pay. auriculata*, (B) *Pay. densipila*, (C) *Pay. grandiflora*, (D) *Pay. lasiocarpa*, (E) *Pay. stonensis*. (F) *A. thaliana* mature trichomes included for comparison.

Sup. Fig. 7. ESEM images of mature *Physaria* type 1 trichomes. (A,C) *Phy. fendleri*, (B) *Phy. argyraea*.

Sup. Fig. 8. ESEM images of mature *Physaria* type 2 trichomes. (A-B) *Phy*. aff. *tenella*, (C) *Phy*. aff. *arizonica*, (D) *Phy*. aff. *pallida*.

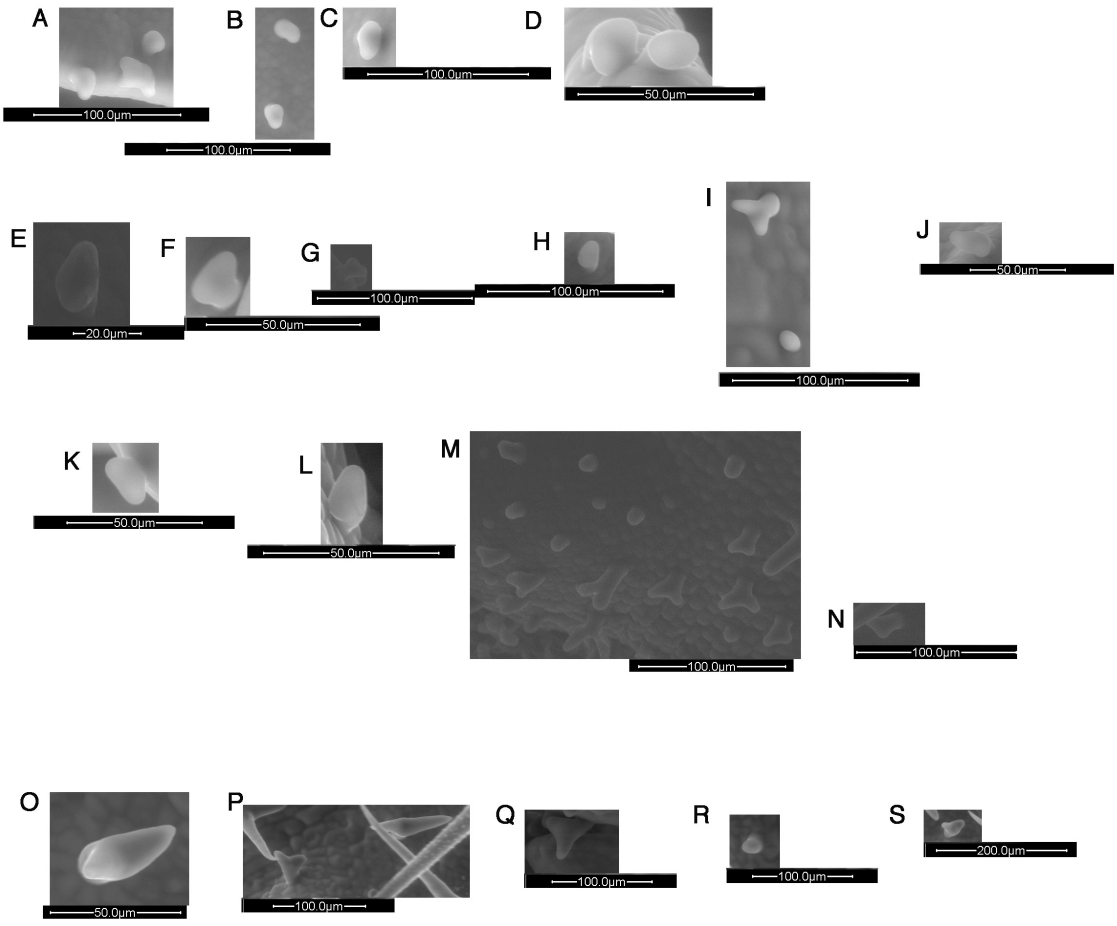
Sup. Fig. 9. ESEM image of mature *Paysonia auriculata* trichomes with three and five branches. Branches were numbered according to their position clockwise from the branch-free gap.

Sup. Fig. 10. *Physaria* type 1 (*Phy. fendleri*) mature external branch widths.

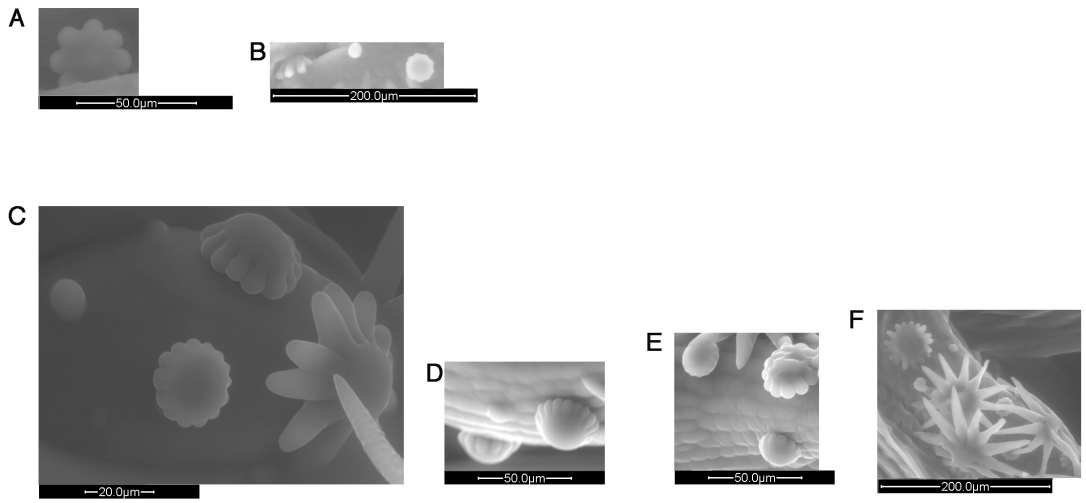
Sup. Fig. 11. Fasta file of transformation constructs.

Sup. Fig. 12. Histogram of trichome tip counts for all six constructs (AA, AP, PA-5, PA-6, PP-24, PP-40). Red, dashed line indicates mean.

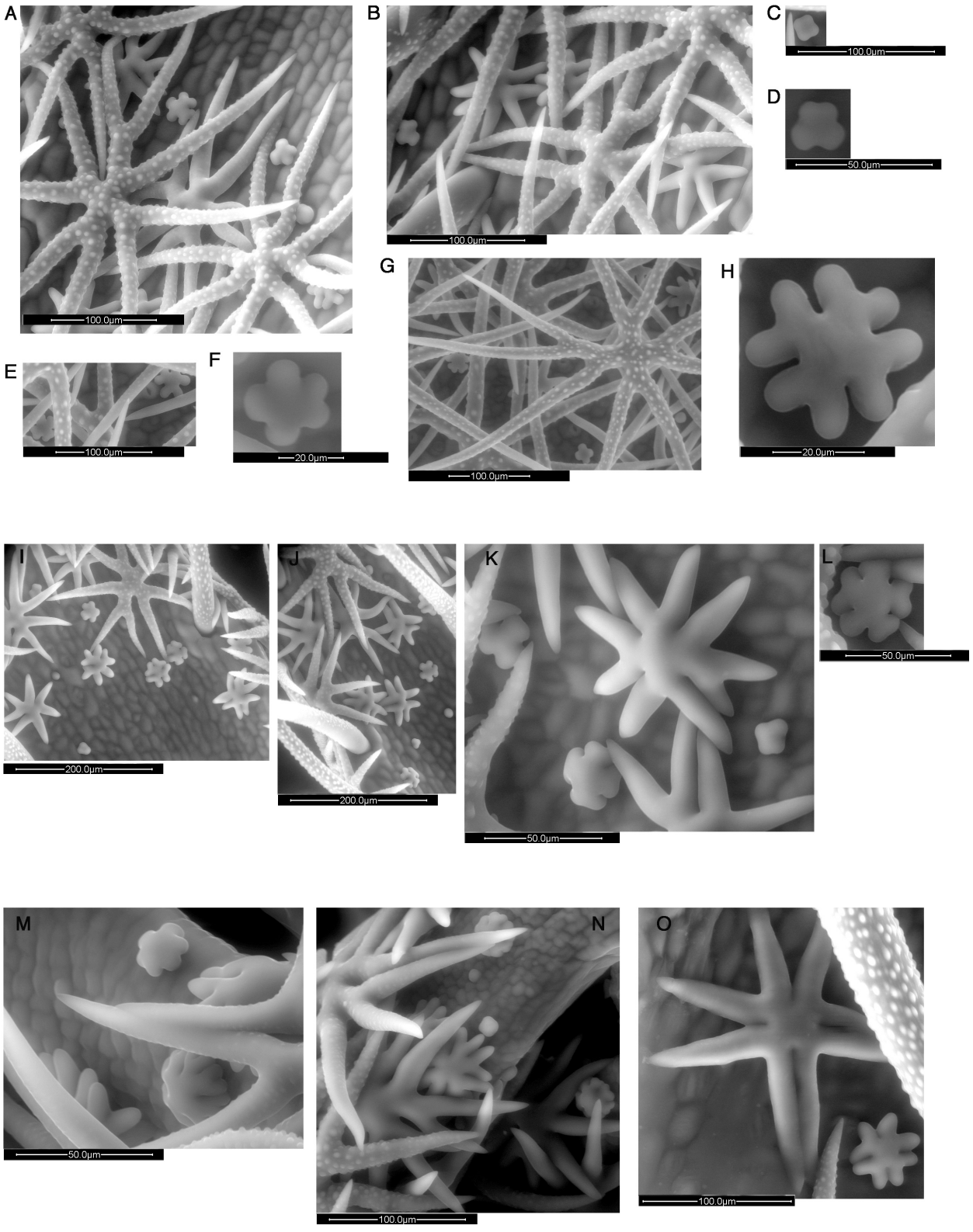
SUPPLEMENTARY FIGURES
Sup.Fig.1.



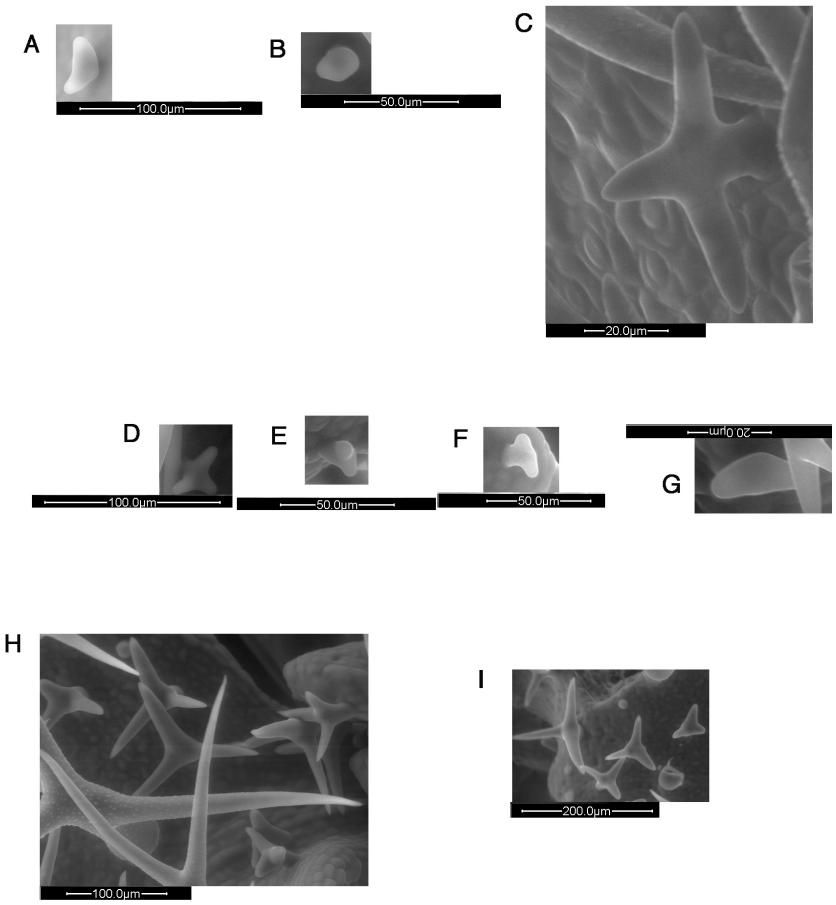
Sup. Fig. 2.



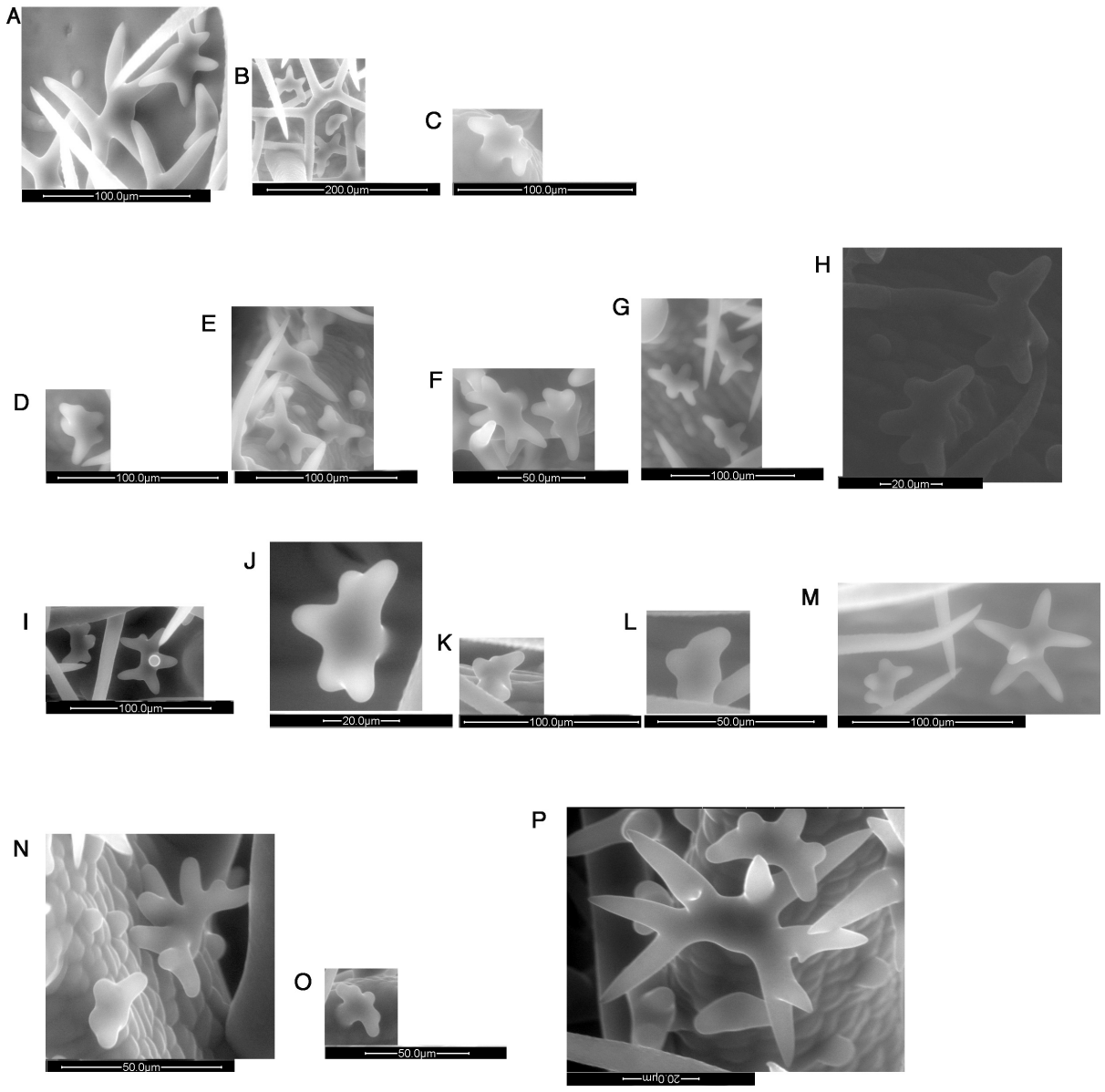
Sup.Fig.3


A
100.0µm
B
100.0µm
C
100.0µm
D
50.0µm
E
100.0µm
F
20.0µm
G
100.0µm
H
20.0µm
I
200.0µm
J
200.0µm
K
50.0µm
L
50.0µm
M
50.0µm
N
100.0µm
O
100.0µm

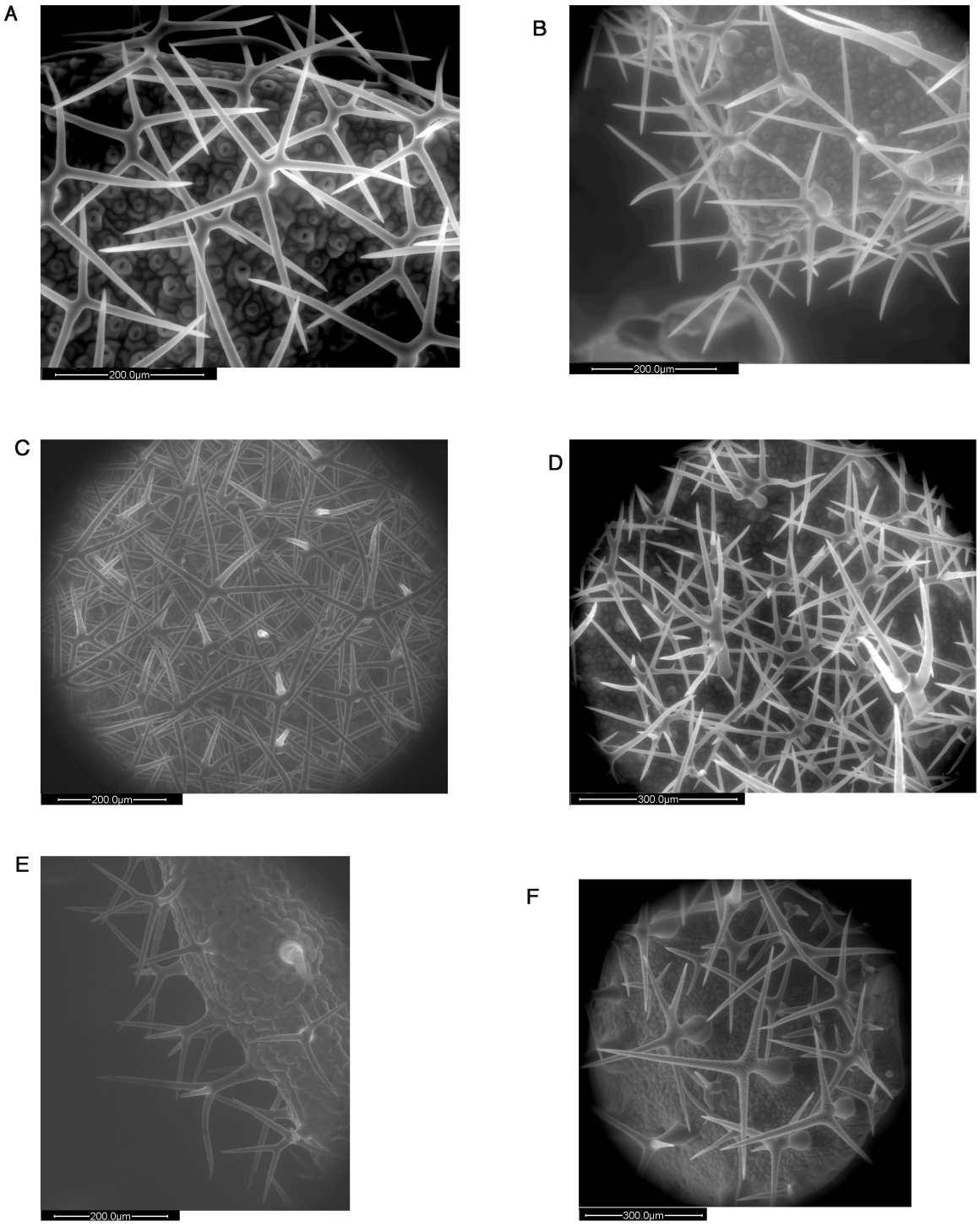
Sup.Fig.4.


A
100.0µm
B
50.0µm
C
20.0µm
D
100.0µm
E
50.0µm
F
50.0µm
20.0µm
G
H
100.0µm
I
200.0µm

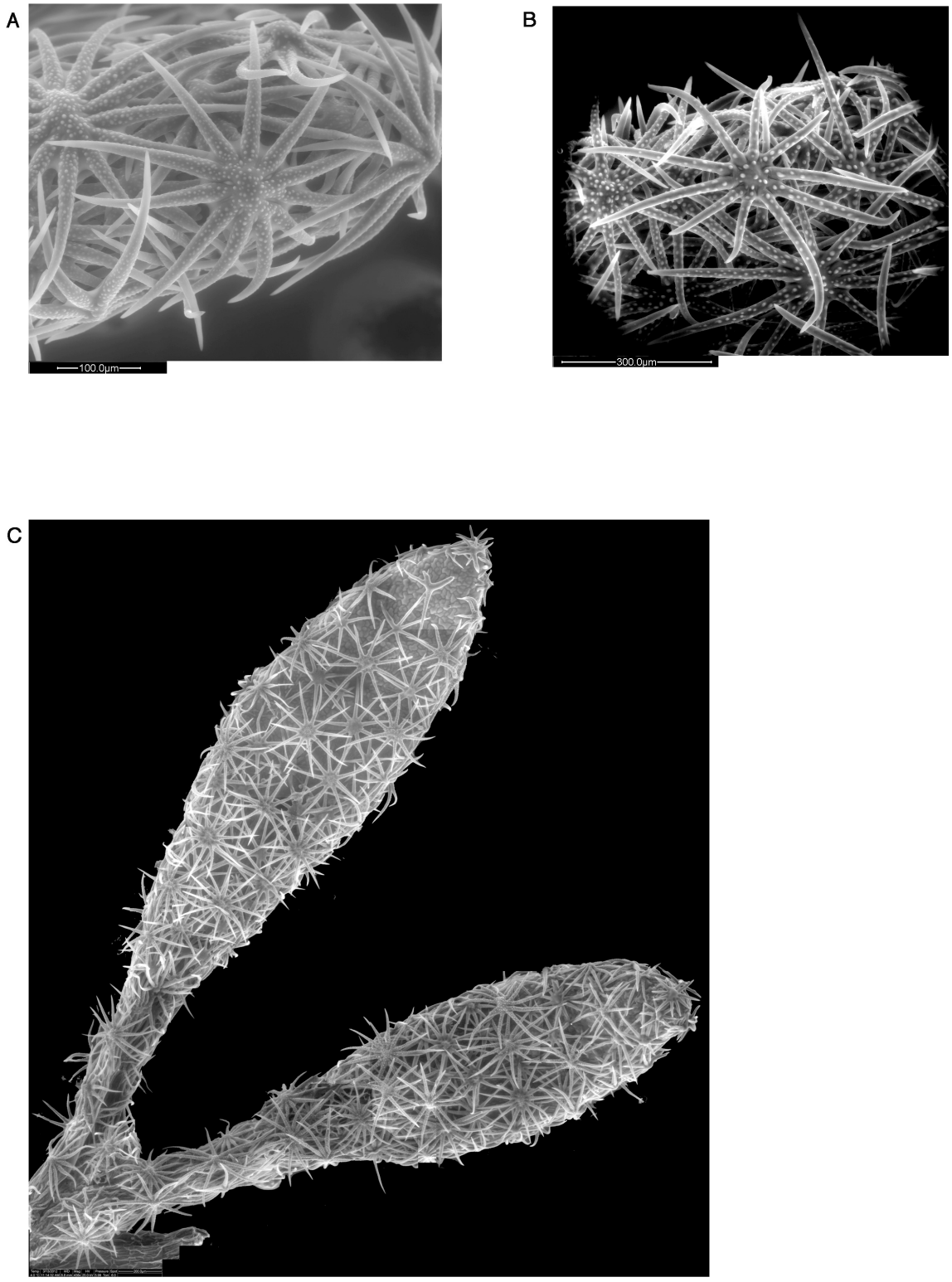
Sup.Fig.5.


A
100.0µm
B
200.0µm
C
100.0µm
D
100.0µm
E
100.0µm
F
50.0µm
G
100.0µm
H
20.0µm
I
100.0µm
J
20.0µm
K
100.0µm
L
50.0µm
M
100.0µm
N
50.0µm
O
50.0µm
P
20.0µm

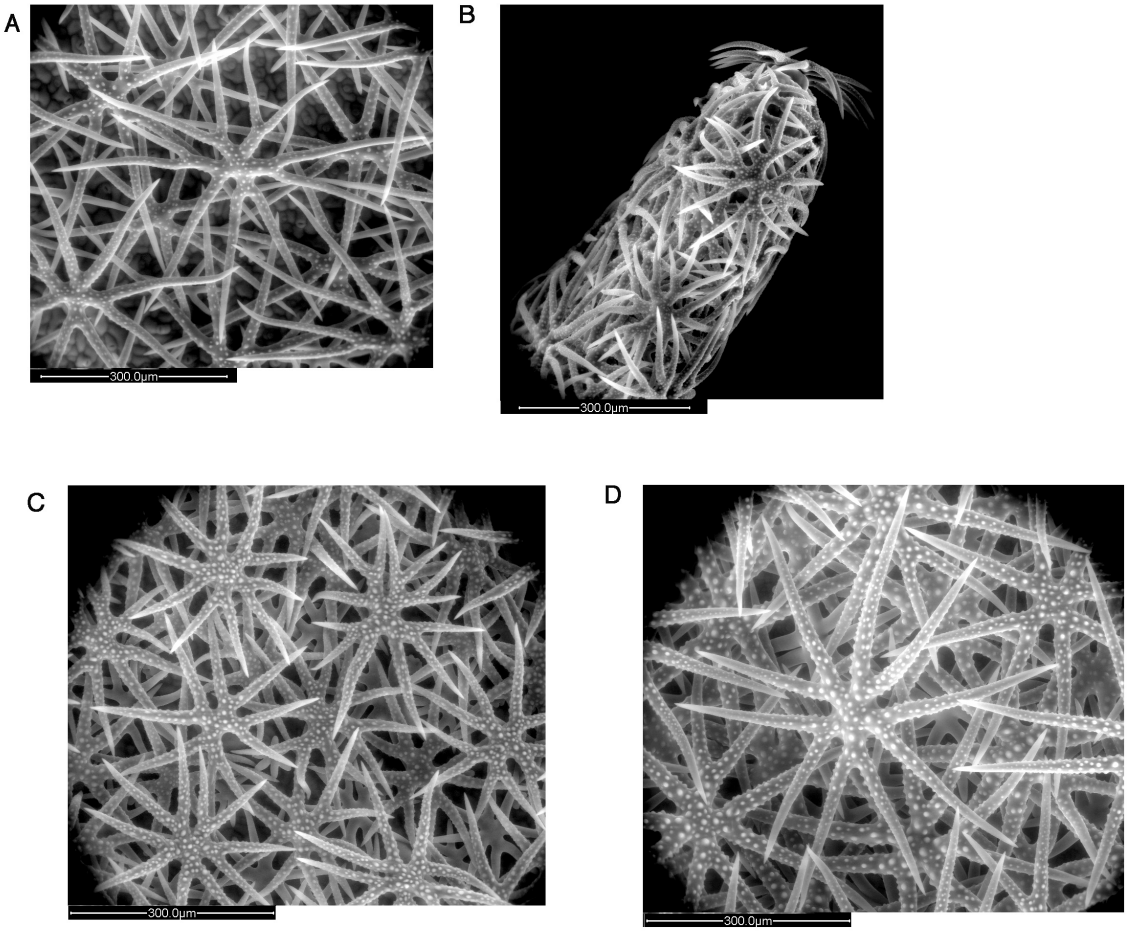
Sup.Fig.6.



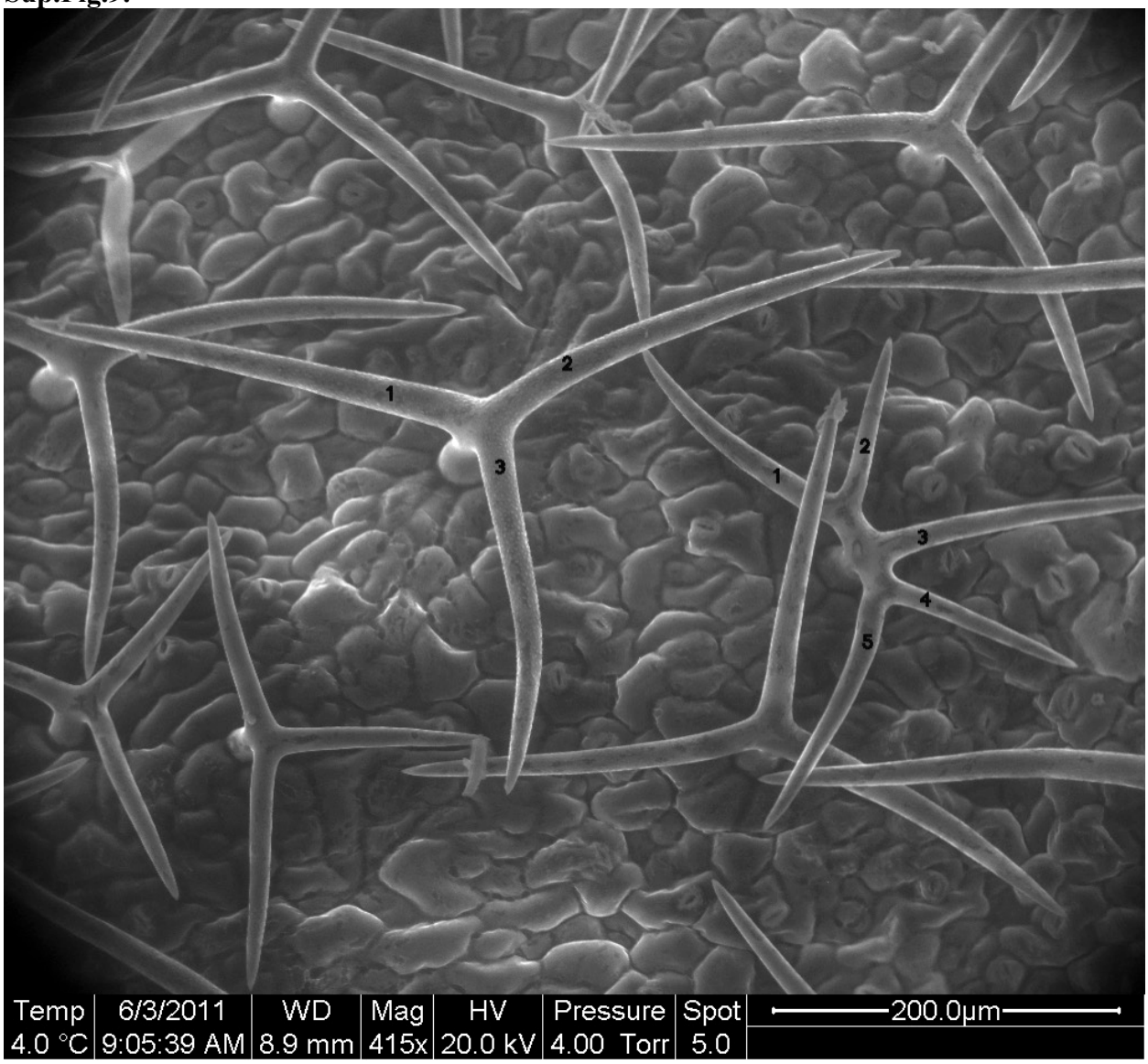
Sup.Fig.7.



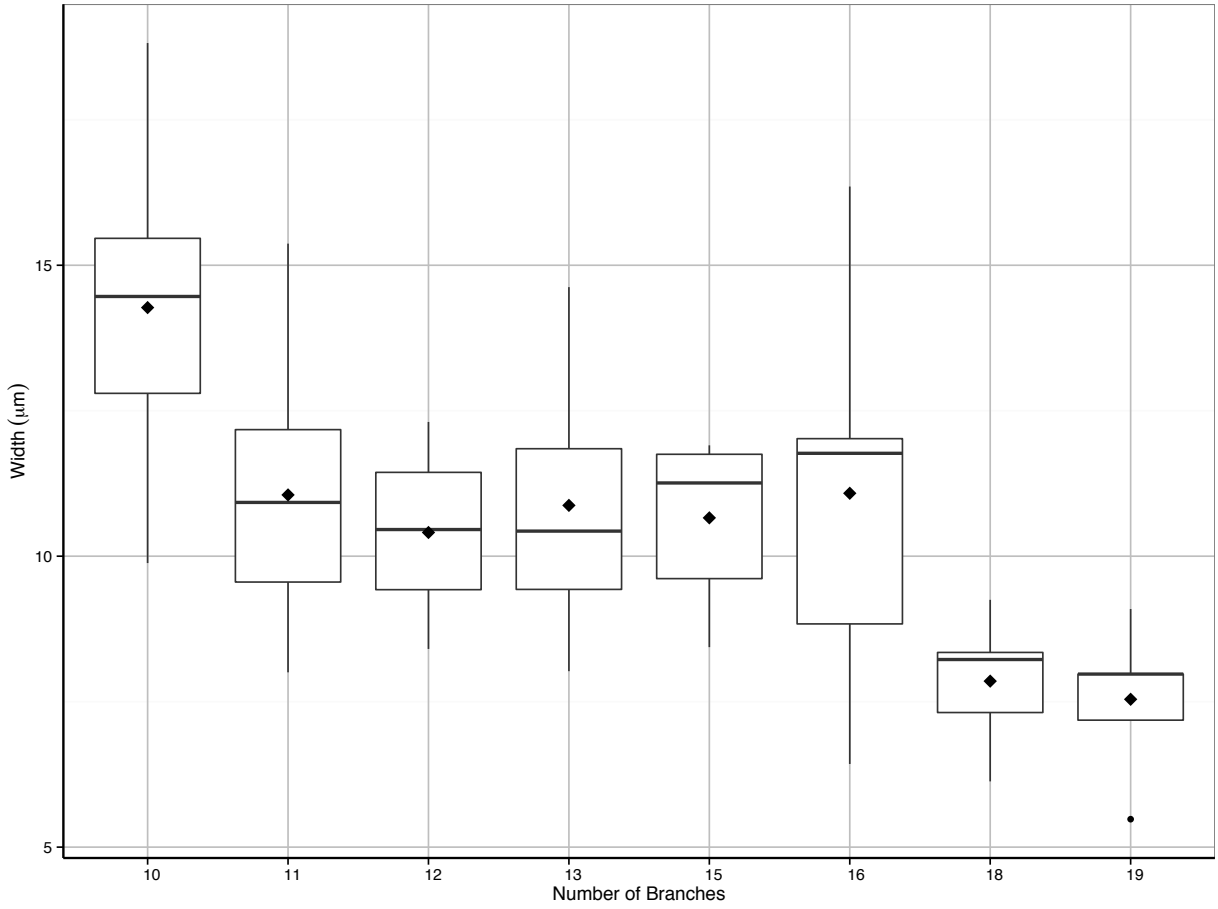
Sup.Fig.8.



Sup.Fig.9.


1
2
3
1
2
3
4
5
Temp
4.0 °C
6/3/2011
9:05:39 AM
WD
8.9 mm
Mag
415x
HV
20.0 kV
Pressure
4.00 Torr
Spot
5.0
200.0µm

Sup.Fig.10.



Sup. Fig. 11.

>AA_3
TGTGACAGACTTGCTCCTACTGTTGATTGAAGGTATCTGGAGAACAGTGGGTGTGTAGGAATGTGTGT
GAATATGTGCAAGATTCCAACCCAAGATTTCTTCACCAATGAGTTTGGCCTCCCACTCACCATGAACCC
AAGTATGTAACATATCTCCTTTTGACCATCCTTATACATTGGTGTTTGTTTTCATTAGCCATCTCTCCAT
TTTGGTTTGTAGATTATGAAGACATGAGCTGCGAGATGATATACGGGCAAGCACCTCCGGCCTTTGAG
GAGGATGTAGCCACCAAGCAACCTTGTCTAGCAGATATATGTAAGTCTTACAATGCTCAGATTGCATA
CATGCATGCAAGATGTATCTAACATGATGGTGATGATGTTGCCTGCAGGTTCTATGTCGAATCCAAGCT
CCCCAATCTGCCCTAAACTAGAGGCATGAGTGGAGTGGATCAGCAGCCACTTCTTTGACGGGATGTGC
AATTTCCAAGCGATATACATACTTTTTTTCTTTTCCAAACACATACATTACATATAAAAGACAGTTGTG
AAGACGAGTTAGTGTTCTTTGTTACTATTCTGGTTGGGTTGTGTAAAGTATACGATAATACACTTGAGC
TTACTCATCTTCTTTATGTTTTAAGTTTGAACTCGCCATGTTGGTTATAACTTGGATCTCAAGTGTTACT
TGCAGTTTGTATTGTAATATAAAACCAAAAACTGAGATTAGCTTAGCTTTAAAGTAACCCTCTCTGCTT
ATATGGATATCAGATACTTACTGTAGTTTTGCTTTAACGACTTACTTTCATGGAATGTTTAGTTTGCGGT
TTGGTGCTTTGATACACACTGAAAGACTTAACTTTCATTTGACTTGCACACACAAAAAAAGAAAAAAG
AATTCTGTTCCTTGTTTTGTTTCCTGCAACTCCAACTGACTAGGGAGTGAACAGAACAGAGTAAATGTT
TAGATAAAGAAAATCTGTTCATAGTTTAATTTATAGAGAGACATATAGCTAGGAGAAGTGGAACAGAT
GAGGCATTAATGTTTGAATGGGCCAATTAATTGACACCCCCAAGTGTACGAACACACAAAATGAGAG
GTGGTTCATTGAAAGCACAGAGAAGTTGTCGGCTTTTCAAAGATACACTGCCCAACAAACAATCGTCA
TGAAACTACAACAAAGGTCAGTTTCATATATCTTCTAATCGCATTTGTTAGTTTACCTAAGAGTTAGAC
GCCATCTGTTTGTTTCTGCTACTTTTCCTTACTCTTTAACGACGAGGTCCTGAGGAAGAATGATTGAAG

CTGGTAATCATGTGATGACTCACTCTTTCAGCTTTACAACACACTTCCATCTTCTAAAGTGATATTTAAT
AACCCTTACTTTTTTTCTTTCTTTTTCCCTTTTTTTTCCCTTTATCTCAAATATAGAAGGCCAACCAAAAA
GATGAAGGATATGAAGATGCAGAGCAGCCCTGAAACGATGATGACCAGAATCCCCACTCCAGATCCT
CACAGTACCGGTGTACGAGAAGATGCAATGGACTCAGTCTGCAAACCCTGGAAGCTTTATGAAAATCC
ATACTACTGCTCTTCTCAGTCTCAACAGCATCAACACCAACGCAAGGCATTTATTTGGGATCTGAATTT
CATAAAGGTGTTCATGGAATCTGAACTAGGAAAAGCTCAAGACGAGATCAAGGAATTGAAGGCAGAA
CTGGACTATGAGAGAAAAGCTCGAAGACGGGCAGAGCTGATGATCAAGAAGCTGGCTAAAGATGTGG
AGGAGGAGAGAATGGCTAGAGAAGCGGAGGAGATGCAAAACAAGCGGCTCTTCAAAGAGTTATCGTC
TGAGAAGTCAGAGATGGTTCGGATGAAACGAGATCTTGAAGAAGAGAGACAGATGCACAGGTTGGCA
GAGGTTTTGAGAGAAGAGAGAGTTCAGATGAAGCTGATGGATGCAAGGCTTTTCTTGGAGGAGAAGC
TATCTGAATTAGAGGAAGCTAATAGACAAGGAGAGAGGGAGAGGAATAGGATGATGAAACCTAAGAT
ATTGGAGAGAGCGTGTAGCTCACCGGCTAGGAGGAGGTGCGAGAATCCGCAAATAAAGAGAGGAATC
AATCCGTTCCCAAGAGTGATGAGAGCAATAAGATCAAAGAGTGAGAAATGGGGTTCAAAGCTGGAGT
GCCAAAAGGTTCAGCTCAAAATTCTGCTCAGACAAAAGACCACCCCAAGATGTACTCCTCTTCTCTCCT
CTCCTCCTCCTTGATACTGTATGCTTGTCTTTTGTTCGCCTCGCTCAATCTCAATAATTCACATCTAAGA
CTGAAGATGTACTAATAAGATTCTCTTTGTCTGTCCGTTAGATAATATTTTATTCACATGGATGTGTTAT
ACATAGCATATATATTGACGTCTATTAAAGATGGATTGGAAGAAAATGATTATCCCATTATCCAAATTT
GTTATGTTGAGTCGTCTCGGACAAGATCAGACAAATATACATAGTTTTCTGCACCTAAAGCTTTCTTCT
ATCGCTGGTATAGTTTTGAAGGATCTTATAAGGATCTATAAAGAATTAACAAAAGTTTTTGAAAGATT
GAATGAATGTTTTGAAACGATTTACAAGAGATTTTAATGATCTATAAGATTTTAAAATGATTTACAAAT
GGTTTTCAACGGATTTATAAGAGGTTTAAAAGATTTATAAAAAGTTTTACAAGAGTTTAAAAAAAGAT
TTTCAAGAATTTTTAAAAAATAAGGGGTTATTTAAAGGATTTACAAAGATTTTATAAGAGGTTTAAAA
AGATTTACAAGAGGTTTTAATAGAATTGTATAAGAGTTTATAAATGATTTACAAGAATTTTAACAAGA
TTTTTTTTTAAGTAATTGCAGAAAATTTTAGTGTAAAATTTTCACGTTTGCGGATAATTTTTGACGAAG
AAATTTCACCTTCACATTATTTGGTCTCACTAAAAAAAGGTGATATAGTCATTTTGTATATAAAAATGT
GTATTGTGAAAATAGACTAGAAGAGAGGAAGCTCATTTTATTGTCCAACTATATGCAATTAACACAAT
TATTCGCCAACTATCGTGTGCACAAATTTAATTTTTAACTTGTCGTTGACTTAACAATTTTACTAGTTAA
TAGAAAAACGTTAAATGACTCAATTAAAACCTAATAAATTTAGAAAATCTCCAAATTTCCTCCCAAAT
TGGTGCAAGATAACATACTTCATTACTTATAAGTTATAACATTAACCACACATTTAAATATAGGGAATT
AAAATCTACCAGAAATCAAATTCTAGGTCTGTACCACCGACAAGAAATTCTCGGACGAACTACTTGTG
ATGCTCGATGTTTCACCCTTGGCTACGGCTCTCACACTCTCTATTAGCTCATTTGTTCTTCCATCTTCAG
ACAGAATCTGTAACAGCTCTTCAGTGTTTATGATCAGATTCACTTTCCATATTGCACCTTTGCTGCTCGT
AC

>PA_6

AATTGGTAACTCCACAAGCAGTCCACACTTTCTGTGTTTGTTCTTCCATCTTCTCAACAACCATATTCAC
TTTCTTTTTTTGTTCTTCCATCATCTCAACAACCTTCTTCTCCACCTCAAGACTTCTAGCTTCACTTTTAG
CCACATTGTTTGCAAGAATCTCAAACAATGGTGACATTTTAAGCACTTCTTCAACAACAGCTTCACTCA
ACCACTTGAATAGGTGATTGTCATCTGCTAACTACATTATCACATGAACTGACTAAGCATAACCAGGG
ACTAGCCAAAAACACTATCAAGCTATAAACACACCGAAAATCGCAAACCTTATTAGCTGCTGCATATT
CACACCGAAAATATTTTCGGCCTGGATTTGCTTTTGATCTTGAGATCAATAACTCCACCTTCCTATCAC
AAAACGAGCATTTCGTAGGGATGCCATCACAAGAGGAACCTCTTCGGACAAATGAAGATCCACTGCTC
ATCTCAGATTCGAAACTCATGTTCAATCTGCAAAAAAAAATTAGAGATTCATATTAAACACAAATCAT
CTCAGAAATGCATGTAGAACAACTCGAAATCAACATAGAAAGATAACCCAGAAGTGAATTGAACTCA
ATCCTATTACACTAAAAATTCCCAATTCCAGACCTCAATTTCGCGACTTAGGGTTCTAAGTTCACAAAT
GCACAAAAAATTTCGATTTTAATGACCCTAACGGTAACAAACGGGTCGGATTTTTAAAAGACAAACGG
GTCGGAATTTTAAAAGAGAAACGGGTCGGAATTTTCGGATGAAATTGGTATGGTTCAGGGCAATAGGT
TCGATTTTCGGTTTTCCAGTCGAAATCATGTGATTCGAACATTGTGTGTTAATTGGCAGATGATTTTCA
ATCTGTGTGTAATAATGAACACCATTATTTTCGGTGTGTAAATTTTCTGATATGCTAATTCCTGTGTGTA
AAATGACCGTCAACCCGTTTTTAAAAGGTAATACTTGGTAATTTTCAGATCACACACAAATCGAGGAT
ATAAACTGGCCACTTATCTTATGCTTTCGAATTAAATATATAGTTCCAAGAGATCATATGTTATATCTC
TTAACTCACGAGAGTGCTTGTTAAGAATGTTGCAATATATTTAACATCTGCTCTAATGTTTGCTTATGG
AGGGAGATCAATAAAAGCGTGTTCCGGCATATTAGACTCTCGGATCCCTGGGAATAGTAATTGCTGTC
GAGATATTGAAGTAGTTAATATTTGTGATTGAATTTGAGATAGATTTTCAATTAGTCTCTGAAAAGTTG
TTAGGTCTTTAGATTTCAATTCGTTGATATATTCTCTAACCCTTAGCACTTTTAACTATTGGTATTTCAT
ATAGTTTTTATTTGCTCTTTACAATTGCATTCTCCAATCTCTTATTATGTATTGTCTTGATTTCAGTTTGA
AAATTCAATAGTACCTTGTTCTCTGTTGGTCGATCCCATTATATACCACATGTCCCTAACTCTGCAGTTA

ATGGATCGGGTTACTCTCAAAAATACTTCTCTTGGATGAAGTCTAGTGGTGTTTTCTCAAGTATTAATG
AATTTGTTTCGTGCAACTCTAGCTAAGGTAGAATTTTTGATAATAAAGGAGTGGCCCATAATAGACAT
AAATAAAAATTAAATTGAGAAAATCCGTTGTTAGTTTATAGAGACAGACATGTATATATGGTTAGGAG
AAGTGAAACAAATGAGGCACTCACTAATGTTTTGAGTGGTCCACTAAATTAGTTGGCACACCTAAGTC
GAAGAGATAGACAATGTCCAACAAACAATCGTCTTCAAAATGTTAGTTTCATATTATTAATTGCATTTT
GTTAGTTACATAAAAGAGTAGCACGGCCTTCGTTGCTGCTCCTCTCTATTCTGAGGAAGAATGAATAA
AACTAATACAACACATGTGATGATGATTCACTCTTGAAGAAGAAGCCATATATATCTTTACCCTCTTTT
TACAACTAAAGCAAGAAAAAATGAAGGATATGAAGATGCAGAGCAGCCCTGAAACGATGATGACCAG
AATCCCCACTCCAGATCCTCACAGTACCGGTGTACGAGAAGATGCAATGGACTCAGTCTGCAAACCCT
GGAAGCTTTATGAAAATCCATACTACTGCTCTTCTCAGTCTCAACAGCATCAACACCAACGCAAGGCA
TTTATTTGGGATCTGAATTTCATAAAGGTGTTCATGGAATCTGAACTAGGAAAAGCTCAAGACGAGAT
CAAGGAATTGAAGGCAGAACTGGACTATGAGAGAAAAGCTCGAAGACGGGCAGAGCTGATGATCAA
GAAGCTGGCTAAAGATGTGGAGGAGGAGAGAATGGCTAGAGAAGCGGAGGAGATGCAAAACAAGCG
GCTCTTCAAAGAGTTATCGTCTGAGAAGTCAGAGATGGTTCGGATGAAACGAGATCTTGAAGAAGAG
AGACAGATGCACAGGTTGGCAGAGGTTTTGAGAGAAGAGAGAGTTCAGATGAAGCTGATGGATGCAA
GGCTTTTCTTGGAGGAGAAGCTATCTGAATTAGAGGAAGCTAATAGACAAGGAGAGAGGGAGAGGAA
TAGGATGATGAAACCTAAGATATTGGAGAGAGCGTGTAGCTCACCGGCTAGGAGGAGGTGCGAGAAT
CCGCAAATAAAGAGAGGAATCAATCCGTTCCCAAGAGTGATGAGAGCAATAAGATCAAAGAGTGAGA
AATGGGGTTCAAAGCTGGAGTGCCAAAAGGTTCAGCTCAAAATTCTGCTCAGACAAAAGACCACCCC
AAGATGTACTCCTCTTCTCTCCTCTCCTCCTCCTTGATACTGTATGCTTGTCTTTTGTTCGCCTCGCTCAA
TCTCAATAATTCACATCTAAGACTGAAGATGTACTAATAAGATTCTCTTTGTCTGTCCGTTAGATAATA
TTTTATTCACATGGATGTGTTATACATAGCATATATATTGACGTCTATTAAAGATGGATTGGAAGAAAA
TGATTATCCCATTATCCAAATTTGTTATGTTGAGTCGTCTCGGACAAGATCAGACAAATATACATAGTT
TTCTGCACCTAAAGCTTTCTTCTATCGCTGGTATAGTTTTGAAGGATCTTATAAGGATCTATAAAGAAT
TAACAAAAGTTTTTGAAAGATTGAATGAATGTTTTGAAACGATTTACAAGAGATTTTAATGATCTATA
AGATTTTAAAATGATTTACAAATGGTTTTCAACGGATTTATAAGAGGTTTAAAAGATTTATAAAAAGTT
TTACAAGAGTTTAAAAAAAGATTTTCAAGAATTTTTAAAAAATAAGGGGTTATTTAAAGGATTTACAA
AGATTTTATAAGAGGTTTAAAAAGATTTACAAGAGGTTTTAATAGAATTGTATAAGAGTTTATAAATG
ATTTACAAGAATTTTAACAAGATTTTTTTTTAAGTAATTGCAGAAAATTTTAGTGTAAAATTTTCACGT
TTGCGGATAATTTTTGACGAAGAAATTTCACCTTCACATTATTTGGTCTCACTAAAAAAAGGTGATATA
GTCATTTTGTATATAAAAATGTGTATTGTGAAAATAGACTAGAAGAGAGGAAGCTCATTTTATTGTCC
AACTATATGCAATTAACACAATTATTCGCCAACTATCGTGTGCACAAATTTAATTTTTAACTTGTCGTT
GACTTAACAATTTTACTAGTTAATAGAAAAACGTTAAATGACTCAATTAAAACCTAATAAATTTAGAA
AATCTCCAAATTTCCTCCCAAATTGGTGCAAGATAACATACTTCATTACTTATAAGTTATAACATTAAC
CACACATTTAAATATAGGGAATTAAAATCTACCAGAAATCAAATTCTAGGTCTGTACCACCGACAAGA
AATTCTCGGACGAACTACTTGTGATGCTCGATGTTTCACCCTTGGCTACGGCTCTCACACTCTCTATTA
GCTCATTTGTTCTTCCATCTTCAGACAGAATCTGTAACAGCTCTTCAGTGTTTATGATCAGATTCACTTT
CCATATTGCACCTTTGCTGCTCGTAC
>PA_5
AATTGGTAACTCCACAAGCAGTCCACACTTTCTGTGTTTGTTCTTCCATCTTCTCAACAACCATATTCAC
TTTCTTTTTTTGTTCTTCCATCATCTCAACAACCTTCTTCTCCACCTCAAGACTTCTAGCTTCACTTTTAG
CCACATTGTTTGCAAGAATCTCAAACAATGGTGACATTTTAAGCACTTCTTCAACAACAGCTTCACTCA
ACCACTTGAATAGGTGATTGTCATCTGCTAACTACATTATCACATGAACTGACTAAGCATAACCAGGG
ACTAGCCAAAAACACTATCAAGCTATAAACACACCGAAAATCGCAAACCTTATTAGCTGCTGCATATT
CACACCGAAAATATTTTCGGCCTGGATTTGCTTTTGATCTTGAGATCAATAACTCCACCTTCCTATCAC
AAAACGAGCATTTCGTAGGGATGCCATCACAAGAGGAACCTCTTCGGACAAATGAAGATCCACTGCTC
ATCTCAGATTCGAAACTCATGTTCAATCTGCAAAAAAAAATTAGAGATTCATATTAAACACAAATCAT
CTCAGAAATGCATGTAGAACAACTCGAAATCAACATAGAAAGATAACCCAGAAGTGAATTGAACTCA
ATCCTATTACACTAAAAATTCCCAATTCCAGACCTCAATTTCGCGACTTAGGGTTCTAAGTTCACAAAT
GCACAAAAAATTTCGATTTTAATGACCCTAACGGTAACAAACGGGTCGGATTTTTAAAAGACAAACGG
GTCGGAATTTTAAAAGAGAAACGGGTCGGAATTTTCGGATGAAATTGGTATGGTTCAGGGCAATAGGT
TCGATTTTCGGTTTTCCAGTCGAAATCATGTGATTCGAACATTGTGTGTTAATTGGCAGATGATTTTCA
ATCTGTGTGTAATAATGAACACCATTATTTTCGGTGTGTAAATTTTCTGATATGCTAATTCCTGTGTGTA
AAATGACCGTCAACCCGTTTTTAAAAGGTAATACTTGGTAATTTTCAGATCACACACAAATCGAGGAT
ATAAACTGGCCACTTATCTTATGCTTTCGAATTAAATATATAGTTCCAAGAGATCATATGTTATATCTC
TTAACTCACGAGAGTGCTTGTTAAGAATGTTGCAATATATTTAACATCTGCTCTAATGTTTGCTTATGG

AGGGAGATCAATAAAAGCGTGTTCCGGCATATTAGACTCTCAGATCCCTGGGAATAGTAATTGCTGTC
GAGATATTGAAGTAGTTAATATTTGTGATTGAATTTGAGATAGATTTTCAATTAGTCTCTGAAAAGTTG
TTAGGTCTTTAGATTTCAATTCGTTGATATATTCTCTAACCCTTAGCACTTTTAACTATTGGTATTTCAT
ATAGTTTTTATTTGCTCTTTACAATTGCATTCTCCAATCTCTTATTATGTATTGTCTTGATTTCAGTTTGA
AAATTCAATAGTACCTTGTTCTCTGTTGGTCGATCCCATTATATACCACATGTCCCTAACTCTGCAGTTA
ATGGATCGGGTTACTCTCAAAAATACTTCTCTTGGATGAAGTCTAGTGGTGTTTTCTCAAGTATTAATG
AATTTGTTTCGTGCAACTCTAGCTAAGGTAGAATTTTTGATAATAAAGGAGTGGCCCATAATAGACAT
AAATAAAAATTAAATTGAGAAAATCCGTTGTTAGTTTATAGAGACAGACATGTATATATGGTTAGGAG
AAGTGGAACAAATGAGGCACTCACTAATGTTTTGAGTGGTCCACTAAATTAGTTGGCACACCTAAGTC
GAAGAGATAGACAATGTCCAACAAACAATCGTCTTCAAAATGTTAGTTTCATATTATTAATTGCATTTT
GTTAGTTACATAAAAGAGTAGCACGGCCTTCGTTGCTGCTCCTCTCTATTCTGAGGAAGAATGAATAA
AACTAATACAACACATGTGATGATGATTCACTCTTGAAGAAGAAGCCATATATATCTTTACCCTCTTTT
TACAACTAAAGCAAGAAAAAATGAAGGATATGAAGATGCAGAGCAGCCCTGAAACGATGATGACCAG
AATCCCCACTCCAGATCCTCACAGTACCGGTGTACGAGAAGATGCAATGGACTCAGTCTGCAAACCCT
GGAAGCTTTATGAAAATCCATACTACTGCTCTTCTCAGTCTCAACAGCATCAACACCAACGCAAGGCA
TTTATTTGGGATCTGAATTTCATAAAGGTGTTCATGGAATCTGAACTAGGAAAAGCTCAAGACGAGAT
CAAGGAATTGAAGGCAGAACTGGACTATGAGAGAAAAGCTCGAAGACGGGCAGAGCTGATGATCAA
GAAGCTGGCTAAAGATGTGGAGGAGGAGAGAATGGCTAGAGAAGCGGAGGAGATGCAAAACAAGCG
GCTCTTCAAAGAGTTATCGTCTGAGAAGTCAGAGATGGTTCGGATGAAACGAGATCTTGAAGAAGAG
AGACAGATGCACAGGTTGGCAGAGGTTTTGAGAGAAGAGAGAGTTCAGATGAAGCTGATGGATGCAA
GGCTTTTCTTGGAGGAGAAGCTATCTGAATTAGAGGAAGCTAATAGACAAGGAGAGAGGGAGAGGAA
TAGGATGATGAAACCTAAGATATTGGAGAGAGCGTGTAGCTCACCGGCTAGGAGGAGGTGCGAGAAT
CCGCAAATAAAGAGAGGAATCAATCCGTTCCCAAGAGTGATGAGAGCAATAAGATCAAAGAGTGAGA
AATGGGGTTCAAAGCTGGAGTGCCAAAAGGTTCAGCTCAAAATTCTGCTCAGACAAAAGACCACCCC
AAGATGTACTCCTCTTCTCTCCTCTCCTCCTCCTTGATACTGTATGCTTGTCTTTTGTTCGCCTCGCTCAA
TCTCAATAATTCACATCTAAGACTGAAGATGTACTAATAAGATTCTCTTTGTCTGTCCGTTAGATAATA
TTTTATTCACATGGATGTGTTATACATAGCATATATATTGACGTCTATTAAAGATGGATTGGAAGAAAA
TGATTATCCCATTATCCAAATTTGTTATGTTGAGTCGTCTCGGACAAGATCAGACAAATATACATAGTT
TTCTGCACCTAAAGCTTTCTTCTATCGCTGGTATAGTTTTGAAGGATCTTATAAGGATCTATAAAGAAT
TAACAAAAGTTTTTGAAAGATTGAATGAATGTTTTGAAACGATTTACAAGAGATTTTAATGATCTATA
AGATTTTAAAATGATTTACAAATGGTTTTCAACGGATTTATAAGAGGTTTAAAAGATTTATAAAAAGTT
TTACAAGAGTTTAAAAAAAGATTTTCAAGAATTTTTAAAAAATAAGGGGTTATTTAAAGGATTTACAA
AGATTTTATAAGAGGTTTAAAAAGATTTACAAGAGGTTTTAATAGAATTGTATAAGAGTTTATAAATG
ATTTACAAGAATTTTAACAAGATTTTTTTTTAAGTAATTGCAGAAAATTTTAGTGTAAAATTTTCACGT
TTGCGGATAATTTTTGACGAAGAAATTTCACCTTCACATTATTTGGTCTCACTAAAAAAAGGTGATATA
GTCATTTTGTATATAAAAATGTGTATTGTGAAAACAGACTAGAAGAGAGGAAGCTCATTTTATTGTCC
AACTATATGCAATTAACACAATTATTCGCCAACTATCGTGTGCACAAATTTAATTTTTAACTTGTCGTT
GACTTAACAATTTTACTAGTTAATAGAAAAACGTTAAATGACTCAATTAAAACCTAATAAATTTAGAA
AATCTCCAAATTTCCTCCCAAATTGGTGCAAGATAACATACTTCATTACTTATAAGTTATAACATTAAC
CACACATTTAAATATAGGGAATTAAAATCTACCAGAAATCAAATTCTAGGTCTGTACCACCGACAAGA
AATTCTCGGACGAACTACTTGTGATGCTCGATGTTTCACCCTTGGCTACGGCTCTCACACTCTCTATTA
GCTCATTTGTTCTTCCATCTTCAGACAGAATCTGTAACAGCTCTTCAGTGTTTATGATCAGATTCACTTT
CCATATTGCACCTTTGCTGCTCGTAC

>AP_2

TGTGACAGACTTGCTCCTACTGTTGATTGAAGGTATCTGGAGAACAGTGGGTGTGTAGGAATGTGTGT
GAATATGTGCAAGATTCCAACCCAAGATTTCTTCACCAATGAGTTTGGCCTCCCACTCACCATGAACCC
AAGTATGTAACATATCTCCTTTTGACCATCCTTATACATTGGTGTTTGTTTTCATTAGCCATCTCTCCAT
TTTGGTTTGTAGATTATGAAGACATGAGCTGCGAGATGATATACGGGCAAGCACCTCCGGCCTTTGAG
GAGGATGTAGCCACCAAGCAACCTTGTCTAGCAGATATATGTAAGTCTTACAATGCTCAGATTGCATA
CATGCATGCAAGATGTATCTAACATGATGGTGATGATGTTGCCTGCAGGTTCTATGTCGAATCCAAGCT
CCCCAATCTGCCCTAAACTAGAGGCATGAGTGGAGTGGATCAGCAGCCACTTCTTTGACGGGATGTGC
AATTTCCAAGCGATATACATACTTTTTTTCTTTTCCAAACACATACATTACATATAAAAGACAGTTGTG
AAGACGAGTTAGTGTTCTTTGTTACTATTCTGGTTGGGTTGTGTAAAGTATACGATAATACACTTGAGC
TTACTCATCTTCTTTATGTTTTAAGTTTGAACTCGCCATGTTGGTTATAACTTGGATCTCAAGTGTTACT
TGCAGTTTGTATTGTAATATAAAACCAAAAACTGAGATTAGCTTAGCTTTAAAGTAACCCTCTCTGCTT
ATATGGATATCAGATACTTACTGTAGTTTTGCTTTAACGACTTACTTTCATGGAATGTTTAGTTTGCGGT

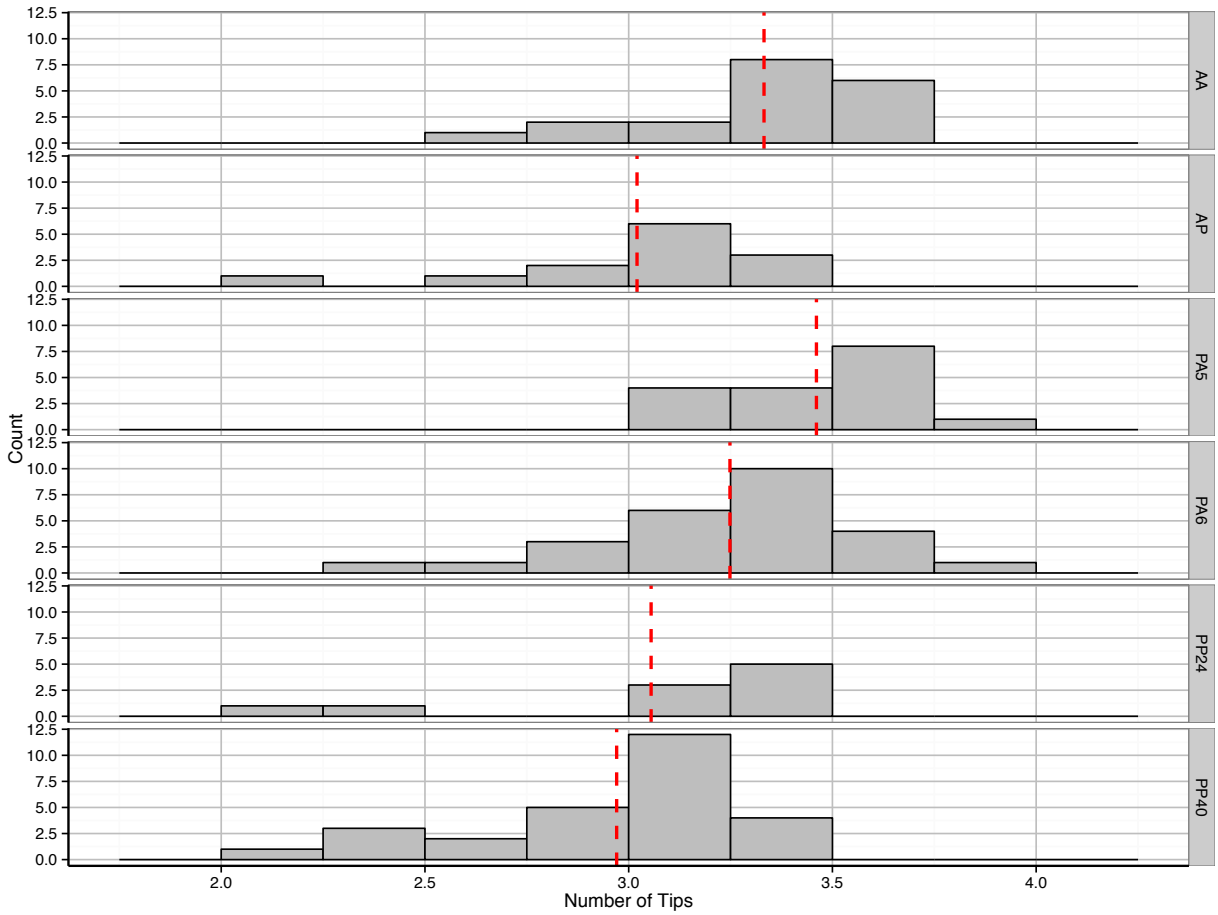
TTGGTGCTTTGATACACACTGAAAGACTTAACTTTCATTTGACTTGCACACACAAAAAAAGAAAAAAG
AATTCTGTTCCTTGTTTTGTTTCCTGCAACTCCAACTGACTAGGGAGTGAACAGAACAGAGTAAATGTT
TAGATAAAGAAAATCTGTTCATAGTTTAATTTATAGAGAGACATATAGCTAGGAGAAGTGGAACAGAT
GAGGCATTAATGTTTGAATGGGCCAATTAATTGACACCCCCAAGTGTACGAACACACAAAATGAGAG
GTGGTTCATTGAAAGCACAGAGAAGTTGTCGGCTTTTCAAAGATACACTGCCCAACAAACAATCGTCA
TGAAACTACAACAAAGGTCAGTTTCATATATCTTCTAATCGCATTTGTTAGTTTACCTAAGAGTTAGAC
GCCATCTGTTTGTTTCTGCTACTTTTCCTTACTCTTTAACGACGAGGTCCTGAGGAAGAATGATTGAAG
CTGGTAATCATGTGATGACTCACTCTTTCAGCTTTACAACACACTTCCATCTTCTAAAGTGATATTTAAT
AACCCTTACTTTTTTTCTTTCTTTTTCCCTTTTTTTTCCCTTTATCTCAAATATAGAAGGCCAACCAAAAA
GGTGAGGTAGATTAAAATGCGTAGCAGTCCTGAAACAATGATTACTAGAATCCCTACTCCTGTACGAG
AAGATGCCATGGACTCAGTCTTCAAAACATGGAAGCTTTATGATAATCCTTTCTACTCCTCTTCTTCAC
TGTCTCAACAGAACCAACACCAACGCAAGGCATTTATCTGGGATCTGAATTTTATAAAGGTCTTCATG
GAATCAGAACTAGGAAAAGCTCGAGCGGAGATTAAGGAAGTGAAGGCAGAGCTTGAGTACGAGAGA
AAAGCTCGAAGACGGGCAGAGTTGCTGAACAAGAGGCTGGCTAAAGATGTAGAGGAGGAGAGAATA
AGTAGAGAAGTAGAGGAGATGCAAAACAACCAGCTCTTGAAAGAGATATCTTTTGCGAAGTCAGAGA
TGATTAGGATGAAACGAAATCTTGAAGAAGAGAGACGGATGCACAGGTTGGCAGATGTTTTGAGAGA
AGAGAGAGTTCAGATGAAGATGATGGATGCAAGGCTTTTCTTGGAGGAGAAGATTTCTGAATTAGAGC
AAGGAGAGAGGGAGAGGAATAGAATGATGAAACCAAAGGTACTGGAGAGAGCATGTAGCTCACCGG
GGAGGAGAACATGGGAGAATCAGCAGATAAAGAAAGGAATCAAAAATCCATTCCCAAGAGTGATGA
GAGCATTAAGATCAAAGAGTGAGAAATGGGGTTCAAAGTTGGAGTGCCAGAAAGTTCAGCTCAAAAT
TCTGCTCAAACAAAAGACCACCCCTAGATGTACTCCTCTTTTCTCCTCTCCTTGATACTCTTTAAAAGCT
TCTTTTTTCGTTTTCAATCTCGCTCTTTCTCAATAATCTTCTCATCTTATTCTCTAGCTGTACTAATAAGA
TTTTTCTTTGTTTGTCTGTTAGAATAACATTTTCTTGAATCGAATAAAAATCAGAAAAAAGGTGTAACT
TTAGTATTTTGAAATGAATATGTTTCAATTTCACAAACACAAATGTTCAATTTTTTTTTTTTTTAACAAA
ATACAAATTGTTGATATTAAAACCGTAAAGAATAGAGTATTGCTCATTTTTTAAGAATGAAATTGATAT
ATTTTTCCGTAGATGATGGTCATCTAGATGTCATGACCAGTGTATATATTACATGAGCTGGGATTGTAC
GTACATTTAATAGATTAATTAGTTAAAAGAGTTTGATGACCTCATTATCCTAATTTGTTATGTTGTTAA
GTTTGTTCTGAGTTCAAACACATACAAAAGAGTTTTCTTCCAAGTAGAGCTATCTTTTTCTTTCTTTTTG
GCTGATATATAACTCTAAAAAAACACAACTATAAAAGGAAGATAACAAAATCTTTCTCCTCATTCACA
ATCATAGAAATAAATCTGATTTTTAAGAGGGGCATATCCAGTGGCGGATCCAGGACTAATTTCTAATG
GGGGCATACATTAAAATATATAACAACTAGATTTGATAAACCCGTGCGATGCACGGGACTTAATCATT
AATAAAATAATACAAATATATTTATTGATTATTTTGAAATTTTGCGTAATAACAAATCTTTTGAACAAC
ATTGCTACCTTCGAACTAACAGAAAATCAGATTCTTGAAGAAAGCGCAAAGCAGACATATAATATAGA
TAAGATTCTCATCAATTAATAATAATTCTATCCCCATGAGAGTATTGTTTGAATGAAAAATCAAAAGA
AAGATGAATGAAAGGTTGAATGTCAATAATCAAAAGATGAATGCAACCGAAAAGATGAATGCAAACT
AATCTATTTATACTTGAATTAGGTATAGAATTTTAATTGCAAATATGGAAAGACGTATATAAGAAATA
TGGCTACTTTGCAAATATGGAAAGACGTATATAAGAAAGATGGCTTTAAATATTCGAATTTTAGTCTA
GAATTTTAGGAAAATTTCGAATTTTAGGAAAATATATAGGCATGATATGCAAACATGGAAAGACGTAT
ATAAGAAAGATGGCTGCTTTGCAAATATGGAAAGACGTATATAAGAAAGATGGCTTTTAATATTCTAA
TTTTAGGTCTAGAATTTTAGGAAAATAATTTTAGGAAAATTTCGAATTTTAGGTTTTTTTAATTAAAAA
TTTCATTAAAAGATTTAAATATTTAGTTTCCAATTTTTTAATTAAAAATCAATTCTTAATAATCAAAATT
AGGAAACTACAAATTAACGGCTGAGATCTCAATCACTTGACACGTGGCGCAATTTTTATCAAGTTTCA
AATCCAAATTTATATAATAAGATAAGAAAGATGGCTTTTAATATTCTAATTTTAGGTCTAAAATTTTAG
GAAAATAATTTTAGGAAAATTTCAAATTTTAGGTTTTTTTAATTAAAAATTTCATTAAAAGATATAAAT
ATTTAGTTTCCAATTTTTTAATAAAAAATCAATTCTTAATAATCTAAATTAGGAAACTACAAATTAACG
GCAGAGATCTCAATCACTTGACACGTGGCGCAATCCTGAATTTTATCAAGTTTCAAACCCAACTTTATA
TAATAAGATAGTTGGATTTAGTTACAGATTTTAATTTTAATGGGGTCATTTTAGCCTTTTCTATGGGAG
CATTCAAATTTAACCTTAATAACTAATATTGATTTCTAAACAACCACTGGGTCATATGCTCACCAAAAT
CTATGTACGTCCACCACTGGGCATATTAGGTAATAGTTCGAACAATAAGACCATCCACATTGGTGATA
CCCTGAAATTGGGATTGTAAAACAAAGAAAAATAATATTTTAATAAAATTAATCATTAAATTAAATTA
AAATATTTGTTAAAAAGAATCAACTAATAATAGTTTGATATCCGGAAGAATAAGTCCTTAACCTAATC
GTTAAAGTCTCTAAATAAGAGGCGTATCAACCTTCTCTCACTTTCTTTTTTATATTTTATTATTAAAAAT
TATTAACAATCCCTCTAAAATACTACCATTACCATTGGAGCATTGGAGGTGCTCTAAGGTAAGTTAGA
AATTGATCCGAAGATTTGCATTGAACGGCCAAAGATGGACCTATAATAATCCAACCATCTTCCCAAAT
CCTAATCTTTAGGATCTTCTCTCCTTCGTGCCAAAAAAATAAAAAAGAAAAATAAATAAATAAATAAA
TAGGATCTTCTCTCCTCTTTTATTGACAACCTTTTTCAATTAATCTCATTATTCTCTGACTCCTCATGGG

GCCCGTAACCCTATTCAAGACTCAATCCTCGTCCAATTGTTGTGCTGTGACGCCTTACCATCCGATTCT
CCAGG
>P_24
AATTGGTAACTCCACAAGCAGTCCACACTTTCTGTGTTTGTTCTTCCATCTTCTCAACAACCATATTCAC
TTTCTTTTTTTGTTCTTCCATCATCTCAACAACCTTCTTCTCCACCTCAAGACTTCTAGCTTCACTTTTAG
CCACATTGTTTGCAAGAATCTCAAACAATGGTGACATTTTAAGCACTTCTTCAACAACAGCTTCACTCA
ACCACTTGAATAGGTGATTGTCATCTGCTAACTACATTATCACATGAACTGACTAAGCATAACCAGGG
ACTAGCCAAAAACACTATCAAGCTATAAACACACCGAAAATCGCAAACCTTATTAGCTGCTGCATATT
CACACCGAAAATATTTTCGGCCTGGATTTGCTTTTGATCTTGAGATCAATAACTCCACCTTCCTATCAC
AAAACGAGCATTTCGTAGGGATGCCATCACAAGAGGAACCTCTTCGGACAAATGAAGATCCACTGCTC
ATCTCAGATTCGAAACTCATGTTCAATCTGCAAAAAAAAATTAGAGATTCATATTAAACACAAATCAT
CTCAGAAATGCATGTAGAACAACTCGAAATCAACATAGAAAGATAACCCAGAAGTGAATTGAACTCA
ATCCTATTACACTAAAAATTCCCAATTCCAGACCTCAATTTCGCGACTTAGGGTTCTAAGTTCACAAAT
GCACAAAAAATTTCGATTTTAATGACCCTAACGGTAACAAACGGGTCGGATTTTTAAAAGACAAACGG
GTCGGAATTTTAAAAGAGAAACGGGTCGGAATTTTCGGATGAAATTGGTATGGTTCAGGGCAATAGGT
TCGATTTTCGGTTTTCCAGTCGAAATCATGTGATTCGAACATTGTGTGTTAATTGGCAGATGATTTTCA
ATCTGTGTGTAATAATGAACACCATTATTTTCGGTGTGTAAATTTTCTGATATGCTAATTCCTGTGTGTA
AAATGACCGTCAACCCGTTTTTAAAAGGTAATACTTGGTAATTTTCAGATCACACACAAATCGAGGAT
ATAAACTGGCCACTTATCTTATGCTTTCGAATTAAATATATAGTTCCAAGAGATCATATGTTATATCTC
TTAACTCACGAGAGTGCTTGTTAAGAATGTTGCAATATATTTAACATCTGCTCTAATGTTTGCTTATGG
AGGGAGATCAATAAAAGCGTGTTCCGGCATATTAGACTCTCGGATCCCTGGGAATAGTAATTGCTGTC
GAGATATTGAAGTAGTTAATATTTGTGATTGAATTTGAGATAGATTTTCAATTAGTCTCTGAAAAGTTG
TTAGGTCTTTAGATTTCAATTCGTTGATATATTCTCTAACCCTTAGCACTTTTAACTATTGGTATTTCAT
ATAGTTTTTATTTGCTCTTTACAATTGCATTCTCCAATCTCTTATTATGTATTGTCTTGATTTCAGTTTGA
AAATTCAATAGTACCTTGTTCTCTGTTGGTCGATCCCATTATATACCACATGTCCCTAACTCTGCAGTTA
ATGGATCGGGTTACTCTCAAAAATACTTCTCTTGGATGAAGTCTAGTGGTGTTTTCTCAAGTATTAATG
AATTTGTTTCGTGCAACTCTAGCTAAGGTAGAATTTTTGATAATAAAGGAGTGGCCCATAATAGACAT
AAATAAAAATTAAATTGAGAAAATCCGTTGTTAGTTTATAGAGACAGACATGTATATATGGTTAGGAG
AAGTGGAACAAATGAGGCACTCACTAATGTTTTGAGTGGTCCACTAAATTAGTTGGCACACCTAAGTC
GAAGAGATAGACAATGTCCAACAAACAATCGTCTTCAAAATGTTAGTTTCATATTATTAATTGCATTTT
GTTAGTTACATAAAAGAGTAGCACGGCCTTCGTTGCTGCTCCTCTCTATTCTGAGGAAGAATGAATAA
AACTAATACAACACATGTGATGATGATTCACTCTTGAAGAAGAAGCCATATATATCTTTACCCTCTTTT
TACAACTAAAGCAAGAAAAAGTGAGGTAGATTAAAATGCGTAGCAGTCCTGAAACAATGATTACTAG
AATCCCTACTCCTGTACGAGAAGATGCCATGGACTCAGTCTTCAAAACATGGAAGCTTTATGATAATC
CTTTCTACTCCTCTTCTTCACTGTCTCAACAGAACCAACACCAACGCAAGGCATTTATCTGGGATCTGA
ATTTTATAAAGGTCTTCATGGAATCAGAACTAGGAAAAGCTCGAGCGGAGATTAAGGAAGTGAAGGC
AGAGCTTGAGTACGAGAGAAAAGCTCGAAGACGGGCAGAGTTGCTGAACAAGAGGCTGGCTAAAGAT
GTAGAGGAGGAGAGAATAAGTAGAGAAGTAGAGGAGATGCAAAACAACCAGCTCTTGAAAGAGATA
TCTTTTGCGAAGTCAGAGATGATTAGGATGAAACGAAATCTTGAAGAAGAGAGACGGATGCACAGGT
TGGCAGATGTTTTGAGAGAAGAGAGAGTTCAGATGAAGATGATGGATGCAAGGCTTTTCTTGGAGGA
GAAGATTTCTGAATTAGAGCAAGGAGAGAGGGAGAGGAATAGAATGATGAAACCAAAGGTACTGGA
GAGAGCATGTAGCTCACCGGGGAGGAGAACATGGGAGAATCAGCAGATAAAGAAAGGAATCAAAAA
TCCATTCCCAAGAGTGATGAGAGCATTAAGATCAAAGAGTGAGAAATGGGGTTCAAAGTTGGAGTGC
CAGAAAGTTCAGCTCAAAATTCTGCTCAAACAAAAGACCACCCCTAGATGTACTCCTCTTTTCTCCTCT
CCTTGATACTCTTTAAAAGCTTCTTTTTTCGTTTTCAATCTCGCTCTTTCTCAATAATCTTCTCATCTTAT
TCTCTAGCTGTACTAATAAGATTTTTCTTTGTTTGTCTGTTAGAATAACATTTTCTTGAATCGAATAAAA
ATCAGAAAAAAGGTGTAACTTTAGTATTTTGAAATGAATATGTTTCAATTTCACAAACACAAATGTTC
AATTTTTTTTTTTTTTTAACAAAATACAAATTGTTGATATTAAAACCGTAAAGAATAGAGTATTGCTCAT
TTTTTAAGAATGAAATTGATATATTTTTCCGTAGATGATGGTCATCTAGATGTCATGACCAGTGTATAT
ATTACATGAGCTGGGATTGTACGTACATTTAATAGATTAATTAGTTAAAAGAGTTTGATGACCTCATTA
TCCTAATTTGTTATGTTGTTAAGTTTGTTCTGAGTTCAAACACATACAAAAGAGTTTTCTTCCAAGTAG
AGCTATCTTTTTCTTTCTTTTTGGCTGATATATAACTCTAAAAAAACACAACTATAAAAGGAAGATAAC
AAAATCTTTCTCCTCATTCACAATCATAGAAATAAATCTGATTTTTAAGAGGGGCATATCCAGTGGCGG
ATCCAGGACTAATTTCTAATGGGGGCATACATTAAAATATATAACAACTAGATTTGATAAACCCGTGC
GATGCACGGGACTTAATCATTAATAAAATAATACAAATATATTTATTGATTATTTTGAAATTTTGCGTA
ATAACAAATCTTTTGAACAACATTGCTACCTTCGAACTAACAGAAAATCAGATTCTTGAAGAAAGCGC

AAAGCAGACATATAATATAGATAAGATTCTCATCAATTAATAATAATTCTATCCCCATGAGAGTATTG
TTTGAATGAAAAATCAAAAGAAAGATGAATGAAAGGTTGAATGTCAATAATCAAAAGATGAATGCAA
CCGAAAAGATGAATGCAAACTAATCTATTTATACTTGAATTAGGTATAGAATTTTAATTGCAAATATG
GAAAGACGTATATAAGAAATATGGCTACTTTGCAAATATGGAAAGACGTATATAAGAAAGATGGCTTT
AAATATTCGAATTTTAGTCTAGAATTTTAGGAAAATTTCGAATTTTAGGAAAATATATAGGCATGATAT
GCAAACATGGAAAGACGTATATAAGAAAGATGGCTGCTTTGCAAATATGGAAAGACGTATATAAGAA
AGATGGCTTTTAATATTCTAATTTTAGGTCTAGAATTTTAGGAAAATAATTTTAGGAAAATTTCGAATT
TTAGGTTTTTTTAATTAAAAATTTCATTAAAAGATTTAAATATTTAGTTTCCAATTTTTTAATTAAAAAT
CAATTCTTAATAATCAAAATTAGGAAACTACAAATTAACGGCTGAGATCTCAATCACTTGACACGTGG
CGCAATTTTTATCAAGTTTCAAATCCAAATTTATATAATAAGATAAGAAAGATGGCTTTTAATATTCTA
ATTTTAGGTCTAAAATTTTAGGAAAATAATTTTAGGAAAATTTCAAATTTTAGGTTTTTTTAATTAAAA
ATTTCATTAAAAGATATAAATATTTAGTTTCCAATTTTTTAATAAAAAATCAATTCTTAATAATCTAAA
TTAGGAAACTACAAATTAACGGCAGAGATCTCAATCACTTGACACGTGGCGCAATCCTGAATTTTATC
AAGTTTCAAACCCAACTTTATATAATAAGATAGTTGGATTTAGTTACAGATTTTAATTTTAATGGGGTC
ATTTTAGCCTTTTCTATGGGAGCATTCAAATTTAACCTTAATAACTAATATTGATTTCTAAACAACCAC
TGGGTCATATGCTCACCAAAATCTATGTACGTCCACCACTGGGCATATTAGGTAATAGTTCGAACAAT
AAGACCATCCACATTGGTGATACCCTGAAATTGGGATTGTAAAACAAAGAAAAATAATATTTTAATAA
AATTAATCATTAAATTAAATTAAAATATTTGTTAAAAAGAATCAACTAATAATAGTTTGATATCCGGA
AGAATAAGTCCTTAACCTAATCGTTAAAGTCTCTAAATAAGAGGCGTATCAACCTTCTCTCACTTTCTT
TTTTATATTTTATTATTAAAAATTATTAACAATCCCTCTAAAATACTACCATTACCATTGGAGCATTGG
AGGTGCTCTAAGGTAAGTTAGAAATTGATCCGAAGATTTGCATTGAACGGCCAAAGATGGACCTATAA
TAATCCAACCATCTTCCCAAATCCTAATCTTTAGGATCTTCTCTCTCCTTCGTGCCAAAAAATAAAAAA
GAAAAATAAATAAATAAATAAATAGGATCTTCTCTCCTCTTTTATTGACAACCTTTTTCAATTAATCTC
ATTATTCTCTGACTCCTCATGGGGCCCGTAACCCTATTCAAGACTCAATCCTCGTCCAATTGTTGTGCT
GTGACGCCTTACCATCCGATTCTCCAGG
>P_40
AATTGGTAACTCCACAAGCAGTCCACACTTTCTGTGTTTGTTCTTCCATCTTCTCAACAACCATATTCAC
TTTCTTTTTTTGTTCTTCCATCATCTCAACAACCTTCTTCTCCACCTCAAGACTTCTAGCTTCACTTTTAG
CCACATTGTTTGCAAGAATCTCAAACAATGGTGACATTTTAAGCACTTCTTCAACAACAGCTTCACTCA
ACCACTTGAATAGGTGATTGTCATCTGCTAACTACATTATCACATGAACTGACTAAGCATAACCAGGG
ACTAGCCAAAAACACTATCAAGCTATAAACACACCGAAAATCGCAAACCTTATTAGCTGCTGCATATT
CACACCGAAAATATTTTCGGCCTGGATTTGCTTTTGATCTTGAGATCAATAACTCCACCTTCCTATCAC
AAAACGAGCATTTCGTAGGGATGCCATCACAAGAGGAACCTCTTCGGACAAATGAAGATCCACTGCTC
ATCTCAGATTCGAAACTCATGTTCAATCTGCAAAAAAAAATTAGAGATTCATATTAAACACAAATCAT
CTCAGAAATGCATGTAGAACAACTCGAAATCAACATAGAAAGATAACCCAGAAGTGAATTGAACTCA
ATCCTATTACACTAAAAATTCCCAATTCCAGACCTCAATTTCGCGACTTAGGGTTCTAAGTTCACAAAT
GCACAAAAAATTTCGATTTTAATGACCCTAACGGTAACAAACGGGTCGGATTTTTAAAAGACAAACGG
GTCGGAATTTTAAAAGAGAAACGGGTCGGAATTTTCGGATGAAATTGGTATGGTTCAGGGCAATAGGT
TCGATTTTCGGTTTTCCAGTCGAAATCATGTGATTCGAACATTGTGTGTTAATTGGCAGATGATTTTCA
ATCTGTGTGTAATAATGAACACCATTATTTTCGGTGTGTAAATTTTCTGATATGCTAATTCCTGTGTGTA
AAATGACCGTCAACCCGTTTTTAAAAGGTAATACTTGGTAATTTTCAGATCACACACAAATCGAGGAT
ATAAACTGGCCACTTATCTTATGCTTTCGAATTAAATATATAGTTCCAAGAGATCATATGTTATATCTC
TTAACTCACGAGAGTGCTTGTTAAGAATGTTGCAATATATTTAACATCTGCTCTAATGTTTGCTTATGG
AGGGAGATCAATAAAAGCGTGTTCCGGCATATTAGACTCTCGGATCCCTGGGAATAGTAATTGCTGTC
GAGATATTGAAGTAGTTAATATTTGTGATTGAATTTGAGATAGATTTTCAATTAGTCTCTGAAAAGTTG
TTAGGTCTTTAGATTTCAATTCGTTGATATATTCTCTAACCCTTAGCACTTTTAACTATTGGTATTTCAT
ATAGTTTTTATTTGCTCTTTACAATTGCATTCTCCAATCTCTTATTATGTATTGTCTTGATTTCAGTTTGA
AAATTCAATAGTACCTTGTTCTCTGTTGGTCGATCCCATTATATACCACATGTCCCTAACTCTGCAGTTA
ATGGATCGGGTTACTCTCAAAAATACTTCTCTTGGATGAAGTCTAGTGGTGTTTTCTCAAGTATTAATG
AATTTGTTTCGTGCAACTCTAGCTAAGGTAGAATTTTTGATAATAAAGGAGTGGCCCATAATAGACAT
AAATAAAAATTAAATTGAGAAAATCCGTTGTTAGTTTATAGAGACAGACATGTATATATGGTTAGGAG
AAGTGGAACAAATGAGGCACTCACTAATGTTTTGAGTGGTCCACTAAATTAGTTGGCACACCTAAGTC
GAAGAGATAGACAATGTCCAACAAACAATCGTCTTCAAAATGTTAGTTTCATATTATTAATTGCATTTT
GTTAGTTACATAAAAGAGTAGCACGGCCTTCGTTGCTGCTCCTCTCTATTCTGAGGAAGAATGAATAA
AACTAATACAACACATGTGATGATGATTCACTCTTGAAGAAGAAGCCATATATATCTTTACCCTCTTTT
TACAACTAAAGCAAGAAAAAGTGAGGTAGATTAAAATGCGTAGCAGTCCTGAAACAATGATTACTAG

AATCCCTACTCCTGTACGAGAAGATGCCATGGACTCAGTCTTCAAAACATGGAAGCTTTATGATAATC
CTTTCTACTCCTCTTCTTCACTGTCTCAACAGAACCAACACCAACGCAAGGCATTTATCTGGGATCTGA
ATTTTATAAAGGTCTTCATGGAATCAGAACTAGGAAAAGCTCGAGCGGAGATTAAGGAAGTGAAGGC
AGAGCTTGAGTACGAGAGAAAAGCTCGAAGACGGGCAGAGTTGCTGAACAAGAGGCTGGCTAAAGAT
GTAGAGGAGGAGAGAATAAGTAGAGAAGTAGAGGAGATGCAAAACAACCAGCTCTTGAAAGAGATA
TCTTTTGCGAAGTCAGAGATGATTAGGATGAAACGAAATCTTGAAGAAGAGAGACGGATGCACAGGT
TGGCAGATGTTTTGAGAGAAGAGAGAGTTCAGATGAAGATGATGGATGCAAGGCTTTTCTTGGAGGA
GAAGATTTCTGAATTAGAGCAAGGAGAGAGGGAGAGGAATAGAATGATGAAACCAAAGGTACTGGA
GAGAGCATGTAGCTCACCGGGGAGGAGAACATGGGAGAATCAGCAGATAAAGAAAGGAATCAAAAA
TCCATTCCCAAGAGTGATGAGAGCATTAAGATCAAAGAGTGAGAAATGGGGTTCAAAGTTGGAGTGC
CAGAAAGTTCAGCTCAAAATTCTGCTCAAACAAAAGACCACCCCTAGATGTACTCCTCTTTTCTCCTCT
CCTTGATACTCTTTAAAAGCTTCTTTTTTCGTTTTCAATCTCGCTCTTTCTCAATAATCTTCTCATCTTAT
TCTCTAGCTGTACTAATAAGATTTTTCTTTGTTTGTCTGTTAGAATAACATTTTCTTGAATCGAATAAAA
ATCAGAAAAAAGGTGTAACTTTAGTATTTTGAAATGAATATGTTTCAATTTCACAAACACAAATGTTC
AATTTTTTTTTTTTTTAACAAAATACAAATTGTTGATATTAAAACCGTAAAGAATAGAGTATTGCTCAT
TTTTTAAGAATGAAATTGATATATTTTTCCGTAGATGATGGTCATCTAGATGTCATGACCAGTGTATAT
ATTACATGAGCTGGGATTGTACGTACATTTAATAGATTAATTAGTTAAAAGAGTTTGATGACCTCATTA
TCCTAATTTGTTATGTTGTTAAGTTTGTTCTGAGTTCAAACACATACAAAAGAGTTTTCTTCCAAGTAG
AGCTATCTTTTTCTTTCTTTTTGGCTGATATATAACTCTAAAAAAACACAACTATAAAAGGAAGATAAC
AAAATCTTTCTCCTCATTCACAATCATAGAAATAAATCTGATTTTTAAGAGGGGCATATCCAGTGGCGG
ATCCAGGACTAATTTCTAATGGGGGCATACATTAAAATATATAACAACTAGATTTGATAAACCCGTGC
GATGCACGGGACTTAATCATTAATAAAATAATACAAATATATTTATTGATTATTTTGAAATTTTGCGTA
ATAACAAATCTTTTGAACAACATTGCTACCTTCGAACTAACAGAAAATCAGATTCTTGAAGAAAGCGC
AAAGCAGACATATAATATAGATAAGATTCTCATCAATTAATAATAATTCTATCCCCATGAGAGTATTG
TTTGAATGAAAAATCAAAAGAAAGATGAATGAAAGGTTGAATGTCAATAATCAAAAGATGAATGCAA
CCGAAAAGATGAATGCAAACTAATCTATTTATACTTGAATTAGGTATAGAATTTTAATTGCAAATATG
GAAAGACGTATATAAGAAATATGGCTACTTTGCAAATATGGAAAGACGTATATAAGAAAGATGGCTTT
AAATATTCGAATTTTAGTCTAGAATTTTAGGAAAATTTCGAATTTTAGGAAAATATATAGGCATGATAT
GCAAACATGGAAAGACGTATATAAGAAAGATGGCTGCTTTGCAAATATGGAAAGACGTATATAAGAA
AGATGGCTTTTAATATTCTAATTTTAGGTCTAGAATTTTAGGAAAATAATTTTAGGAAAATTTCGAATT
TTAGGTTTTTTTAATTAAAAATTTCATTAAAAGATTTAAATATTTAGTTTCCAATTTTTTAATTAAAAAT
CAATTCTTAATAATCAAAATTAGGAAACTACAAATTAACGGCTGAGATCTCAATCACTTGACACGTGG
CGCAATTTTTATCAAGTTTCAAATCCAAATTTATATAATAAGATAAGAAAGATGGCTTTTAATATTCTA
ATTTTAGGTCTAAAATTTTAGGAAAATAATTTTAGGAAAATTTCAAATTTTAGGTTTTTTTAATTAAAA
ATTTCATTAAAAGATATAAATATTTAGTTTCCAATTTTTTAATAAAAAATCAATTCTTAATAATCTAAA
TTAGGAAACTACAAATTAACGGCAGAGATCTCAATCACTTGACACGTGGCGCAATCCTGAATTTTATC
AAGTTTCAAACCCAACTTTATATAATAAGATAGTTGGATTTAGTTACAGATTTTAATTTTAATGGGGTC
ATTTTAGCCTTTTCTATGGGAGCATTCAAATTTAACCTTAATAACTAATATTGATTTCTAAACAACCAC
TGGGTCATATGCTCACCAAAATCTATGTACGTCCACCACTGGGCATATTAGGTAATAGTTCGAACAAT
AAGACCATCCACATTGGTGATACCCTGAAATTGGGATTGTAAAACAAAGAAAAATAATATTTTAATAA
AATTAATCATTAAATTAAATTAAAATATTTGTTAAAAAGAATCAACTAATAATAGTTTGATATCCGGA
AGAATAAGTCCTTAACCTAATCGTTAAAGTCTCTAAATAAGAGGCGTATCAACCTTCTCTCACTTTCTT
TTTTATATTTTATTATTAAAAATTATTAACAATCCCTCTAAAATACTACCATTACCATTGGAGCATTGG
AGGTGCTCTAAGGTAAGTTAGAAATTGATCCGAAGATTTGCATTGAACGGCCAAAGATGGACCTATAA
TAATCCAACCATCTTCCCAAATCCTAATCTTTAGGATCTTCTCTCTCCTTCGTGCCAAAAAATAAAAAA
GAAAAATAAATAAATAAATAAATAGGATCTTCTCTCCTCTTTTATTGACAACCTTTTTCAATTAATCTC
ATTATTCTCTGACTCCTCATGGGGCCCGTAACCCTATTCAAGACTCAATCCTCGTCCAATTGTTGTGCT
GTGACGCCTTACCATCCGATTCTCCAGG

Sup.Fig.12.


Count
Number of Tips
AA
AP
PA5
PA6
PP24
PP40
0.0
2.5
5.0
7.5
10.0
12.5
2.0
2.5
3.0
3.5
4.0