Clade-specific positive selection on a developmental gene: *BRANCHLESS TRICHOME* and the evolution of stellate trichomes in *Physaria* (Brassicaceae)

Abigail R. Mazie^a, mazie@wisc.edu

David A. Baum^a, dbaum@wisc.edu

^aDepartment of Botany, University of Wisconsin-Madison, 132 Birge Hall, 430 Lincoln Drive, Madison, WI 53706

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

ABSTRACT

Positive selection is known to drive the evolution of genes involved in evolutionary arms races, but what role does it play in the evolution of genes involved in developmental processes? We used the single-celled epidermal trichomes of Brassicaceae as a model to uncover the molecular evolutionary processes that contributed to the transition from dendritic trichomes, as seen in most species of Brassicaceae, to the distinctive stellate trichomes of the genus *Physaria*. We explored the role of positive selection on the evolution of *BRANCHLESS TRICHOME (BLT)*, a candidate gene for changes in trichome branching pattern. Maximum likelihood models of codon evolution point to a shift in selective pressure affecting the evolution of *BLT* across the entire *Physaria* clade, and we found strong evidence that positive selection has acted on a subset of *Physaria BLT* codons. Almost all of the 10 codon sites with the highest probability of having evolved under positive selection are clustered in a predicted 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. The fact that positive selection occurred throughout the radiation of *Physaria* could reflect selection to stabilize development in response to an abrupt switch from the dendritic form to the stellate form, divergent selection for diversification of the stellate form, or both. These results point to the need for evolutionary developmental studies of *BLT* and its interacting proteins in *Physaria*.

Keywords: molecular evolution; positive selection; trichome; evo-devo; Brassicaceae; *Physaria*

1. INTRODUCTION

One of the great challenges facing biology at the start of the 21st century is understanding how selection acts upon molecular variation to cause phenotypic evolution. Methods that estimate relative rates of nonsynonymous to synonymous substitutions in protein-coding genes are commonly used to identify an evolutionary history of selection at the molecular level (recently reviewed in Anisimova and Kosiol 2009; Delpont et al. 2009). To link these changes to phenotypic evolution, molecular evolutionary approaches have been used to scan genomes for sites under positive selection as a way to test prior candidate gene hypotheses (e.g., Jarvis et al. 2014; Parker et al. 2014; Tao et al. 2015; Vargas-Pinilla et al. 2015) or to identify new candidate genes for phenotypic differences (e.g., Kim et al. 2011; Aguilera et al. 2012; Nery et al. 2013). When a gene is found to have experienced positive selection in a group that has a derived phenotype, but not in its sister group that lacks the phenotype, it is typically inferred that the gene played a role in the origin of the derived phenotype. However, a signature of positive selection could equally indicate that the gene in question did not cause the developmental change but is a downstream consequence of developmental rewiring during the trait's evolution. Here, we explore this possibility

using a potential candidate gene for the evolution of the shape of unicellular trichomes of the mustard family (Brassicaceae), *BRANCHLESS TRICHOME (BLT)*.

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). Branching trichomes 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).

A major difference between the stellate trichomes of *Physaria* and the dendritic trichomes of other members of Physarieae is the number of branches that emerge from the stalk. Dendritic trichomes have only two primary branches, while stellate trichomes have at least three primary branches, with some species having >20. A promising genetic candidate for the evolution of these differences is *BRANCHLESS TRICHOME (BLT)*. Homologues of *BLT* have been identified across most available angiosperm genomes (Kasili et al. 2011; Phytozome v. 11.0.1: <http://phytozome.jgi.doe.gov/>, last accessed 15 March 2016, Goodstein et al. 2012). 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). *BLT* plays a role in coordinating endoreplication levels with branch number (Kasili et al. 2011), which suggests a link between cell cycle regulation and cell morphogenesis. 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 is predicted to fold 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). Indeed, 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 formation of cell outgrowths (Kasili et al. 2011).

Based on the strong and specific effects of *BLT* on *A. thaliana* trichome branch initiation, it is possible that changes at this locus in *Physaria* contributed to the transition from dendritic to stellate trichomes by increasing the number of primary branch points during trichome development. This hypothesis predicts that there would be evidence of positive selection on *BLT* along the *Physaria* stem lineage – the lineage on which stellate trichomes evolved. While our results show evidence of positive selection on a subset of *BLT* codons in the *Physaria* clade, they point to positive selection primarily occurring during the radiation of *Physaria* rather than just at the time of stellate trichome origination.

2. MATERIALS AND METHODS

2.1 Sequence generation

We identified publicly available *Arabidopsis thaliana* *BLT* homologues 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* homologues. 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 25µL 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 on the single *BLT* homologue identified in this accession (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 DH5 α 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, as 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 most *BLT* sequence

coverage (i.e., the highest quality sequence that provided the most information). If all *BLT* sequences were of similar quality, we used the accession with the shortest terminal *BLT* branch length, as this represents a conservative estimate of *BLT* for a particular species. Accessions of a particular species that did not form a clade with other accessions of the same species 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-HPG BlackBox 8.1.11; Stamatakis 2014) on the CIPRES Science Gateway Version 3 (Miller et al. 2010) 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 outgroup. All analyses were continued until the average standard deviation of split frequencies was <0.02.

2.3 Analysis of BLT Molecular Evolution

We explored 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 full *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 dN/dS ratios (ω) were estimated from the data.

3. RESULTS

3.1 Sequence characteristics of *BLT* and ITS

We generated or obtained published *BLT* 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). We found no evidence of gene duplication, consistent with the fact that *BLT* is single copy in all Brassicaceae species where complete genomic sequences are available from the Phytozome database (Phytozome v. 11.0.1: <http://phytozome.jgi.doe.gov/>, last accessed 15 March 2016, Goodstein et al. 2012).

The final *BLT* alignment (after deletion of redundant 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, the Physarieae *BLT* sequences we analyzed are typically about 450 bp long (range = 375-771 bps). All indels are consistent with the addition or loss of one or more codons (i.e., indel length is always divisible by three). 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 *BLT* and ITS phylogenetic trees are in close agreement

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, Camelinae (*Arabidopsis 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 and could be an artifact associated with low taxon density in outgroups. 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. As detailed in Supplementary Information, the cases are more likely explained by misidentification than by orthology-paralogy issues. 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 briefly discuss these taxonomic issues (supplementary information 1).

3.3 Branch models suggest a change in selective pressure in the *Physaria* crown clade

Is the presence of stellate trichomes in *Physaria* associated with changes in the direction and strength of selection acting on *BLT*, a gene known 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 $\omega < 1$ 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 $\omega > 1$ 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 explore whether selective pressures on *BLT* were different in *Physaria* in comparison to outgroup lineages. The branch models allow ω to vary among preselected branches of a phylogenetic tree and can be used to identify changes in selective pressures along a single branch or across an entire clade. We were specifically interested in uncovering whether a shift in selective pressure occurred in *Physaria* and if this shift was concurrent with the shift from dendritic to stellate trichomes. We therefore considered hypotheses that explored selection in the *Physaria* crown clade, which includes all branches descended from the most recent common ancestor of living *Physaria* species (this common ancestor is depicted with a circle in fig. 5), the *Physaria* stem lineage, which includes only the branch subtending the crown clade (this branch is marked with a star in fig. 5), and the entire *Physaria* clade, including both crown clade and stem lineage. All hypotheses were tested on both the *BLT* and combined *BLT*-ITS trees, but since the results were almost identical, we report the results for the *BLT* tree only.

We first considered the null hypothesis that all branches of the tree and all codon sites in the alignment have the same ω (Hypothesis A, fig. 5, table 1). We found that, in net, *BLT* is under purifying selection, $\omega=0.24$, as expected for a gene that codes for a functionally important protein. We then considered the hypothesis that a change in selective pressure occurred in the *Physaria* stem lineage and persisted throughout the crown clade (Hypothesis B, fig. 5, table 1). Under Hypothesis B, the *Physaria* clade has an estimated ω of 0.39, as contrasted with 0.22 for outgroup taxa. The likelihood score of Hypothesis B was significantly higher than Hypothesis A, 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 the *Physaria* clade compared to outgroups and, specifically, that the proportion of non-synonymous changes is higher in *Physaria* than in other Brassicaceae.

We then considered whether the elevated ω seen in the *Physaria* clade could be due to an episodic burst of selection along the *Physaria* stem lineage, with a return to ancestral levels of ω in the crown clade (Hypothesis C, fig. 5, table 1). Although the ω of the *Physaria* stem lineage was estimated to be higher ($\omega=0.30$) than other lineages ($\omega=0.24$), a likelihood ratio test showed that Hypothesis C is not significantly better than Hypothesis A, where one ω applies to the entire tree. The complementary hypothesis (Hypothesis D, fig. 5, table 1), which estimated ω separately for the *Physaria* crown clade and all other branches on the tree, provided a significantly higher likelihood score than Hypothesis A (LRT: 1 degree of freedom, $p<0.001$). These findings suggest 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 yield a significant p-value.

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 for Physarieae as a whole. We therefore explored a model that estimated a separate ω for the *Physaria* clade ($\omega=0.39$), other sampled members of Physarieae (*Paysonia*, *Lyrocarpa*, *Nerisyrenia*, and *Synthlipsis*, $\omega=0.19$), and the remaining outgroup lineages ($\omega=0.22$) (Hypothesis E, fig. 5, table 1). A likelihood-ratio test did not support the addition of a separate ω for the latter two clades, thus suggesting that the elevated ω seen in *Physaria* is specific to this clade, not to the broader Physarieae.

3.4 Evidence for positive selection on *BLT* in the *Physaria* crown clade

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 in which a subset of codon sites in a designated lineage evolve under positive selection to a null model where codon sites only evolve under purifying selection or neutrally (Zhang et al. 2005). When conducting the branch-site test, the lineage suspected of experiencing positive selection is labeled as the foreground, while 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). We designated the *Physaria* crown clade, but not its stem branch, as the foreground lineage, as this was the best-fitting branch hypothesis (Hypothesis D, fig. 5, table 1). Our results strongly support a role for positive selection driving the evolution of a subset of *BLT* codons in the *Physaria* crown clade (likelihood ratio = 16.38, $p<0.0001$). The preferred model estimates that 13% of codon sites have been under positive selection in the *Physaria* crown clade, but neutrally evolving or under purifying selection in other Brassicaceae (table 2).

In the previously described model, the *Physaria* crown clade was designated as the foreground, whereas the rest of the tree was designated as the background. To identify sites under positive selection in other Brassicaceae taxa that are evolving under purifying or neutral selection in *Physaria*, we flipped these designations (i.e., identifying *Physaria*, including its stem branch, as the background lineage and the rest of the taxa as foreground). However, we failed to find evidence of positive selection outside of *Physaria*: ω was estimated to be ≤ 1 at all sites in the foreground lineage even though its value was unconstrained.

3.5 Positively selected sites are clustered in the coiled-coil domain of BLT

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 has been subject to purifying selection, positive selection, or evolving neutrally in the *Physaria* crown clade. We identified 10 codon sites with >0.95 posterior probability of having experienced positive selection (table 3). Several sites identified as candidates for positive selection in *Physaria* BLT encode amino acids that are chemically dissimilar to the presumed ancestral state, as seen in the outgroups: 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). Likewise, *Physaria* residues with signatures of positive selection show a general tendency to have larger side-groups than outgroups: threonine for serine (position 153), tryptophan for leucine (position 175), and phenylalanine for leucine (position 191).

In addition to the codon sites that show a signature of positive selection in *Physaria*, we identified four codon sites that consistently distinguish *Physaria* from outgroups (table 4). Although none of these additional sites appear to be subject to positive selection, three of the four sites are unique to *Physaria* and shared among all *Physaria* taxa studied (positions 102, 195, 278). It is intriguing to note that at position 151 the negatively charged glutamic acid seen in most Brassicaceae has been replaced in most *Physaria* taxa with the hydrophobic alanine (table 4), which is also found in one outgroup taxon, *Berteroa incana*, that has independently evolved stellate trichomes.

Although direct experimental tests would be needed to clarify the significance of these amino acid substitutions on BLT function, a plot showing the posterior probability that a codon site is under positive selection in the *Physaria* crown clade (fig. 6) shows 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 domain that is thought to mediate protein-protein interactions (Marks et al. 2009).

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 in the stem lineage of *Physaria* (Fuentes-Soriano and Al-Shehbaz 2013).

A primary difference between dendritic and stellate trichomes is the number of trichome branches that emerge from the central stalk, 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, the subsequent diversification of the stellate clade, or both. 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 one in which the *Physaria* *BLT* crown clade 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. Applying a model 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 was under a similar selective regime to that experienced by the crown clade, supporting the possibility 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 the *Physaria* crown clade. Specifically, we estimated that 13% of codons experienced positive selection in *Physaria*, and showed that these sites are 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 altered 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. Furthermore, this positive selection was not transient, but persisted throughout the radiation of the *Physaria* crown clade. Most prior examples of genes experiencing positive selection encode proteins involved in antagonistic arms races, such as immune response or gamete recognition, or genes that have acquired novel functions after gene duplication (reviewed in Yang 2014). Less well known are cases of single-copy developmental genes experiencing persistent positive selection following the evolution of a novel phenotype like stellate trichomes.

We can suggest three scenarios to explain why the shift in selective pressure on *BLT* is detected across the *Physaria* crown clade, rather than only in the stem lineage where the stellate form originally evolved. One possibility is that after the basic stellate form was established in *Physaria*, there was selection 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 (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 *Physaria* 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 and for linking such phenotypic evolution to selection acting at the level of gene sequences.

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), Graduate School (flexible research funds), and the 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.
- Burkhard P, Stetefeld J, Strelkov SV. 2001. Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol.* 11:82–88.
- Chen GQ. 2011. Effective reduction of chimeric tissue in transgenics for the stable genetic transformation of *Lesquerella fendleri*. *HortScience* 46:86–90.
- Couvreux 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.
- Delpont 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.
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U,

- Putnam N, Rokhsar DS. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40 (Database issue): D1178-D1186.
- 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, Thierer T, Ashton B, Mentjies P, Drummond A. 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.
- 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.
- Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*, 14 Nov. 2010, New Orleans, LA pp 1 - 8.
- 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.
- Vargas-Pinilla P, Paixão-Côrtés VR, Paré P, Tovo-Rodrigues L, Vieira CM, Xavier A, Comas D, Pissinatti A, Sinigaglia M, Rigo MM, Vieira GF, Lucion AB, Salzano FM, Bortolini MC. 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.
- 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.

Yang Z. 2014. Molecular evolution: a statistical approach. Oxford: Oxford University Press

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.

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

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

Fig. 5. Simplified phylogenetic tree and summary of selection tests for the *BLT* single-gene tree. The star represents the *Physaria* stem lineage, and the circle represents the most recent common ancestor of *Physaria*. Adjacent to the phylogenetic tree, we provide a summary of alternative branch hypotheses considered. Branches assigned the same number for a given hypothesis 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* alignment. 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.

Figure 1

Figure 2

Figure 3

0.05

Figure 5

	A	B	C	D	E
<i>Physaria</i> crown	0	1	0	1	2
<i>Physaria</i> stem	0	1	1	0	2
<i>Paysonia</i>	0	0	0	0	1
DDNLS clade	0	0	0	0	1
Outgroups	0	0	0	0	0

Figure 6

Table 1

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

Model	Crown ω	Stem ω	Physarieae ω	Outgroup ω	Log-L	LRT ^a	P-value ^b
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 ^c
D	0.40	0.22	0.22	0.22	-9101.32	D vs. A	<0.001
E	0.39	0.39	0.19	0.22	-9101.00	F vs. B	NS

^aLikelihood ratio test comparisons

^bP-values use a X^2 distribution with 1 degree of freedom

^cNot significant

Table 2. Full model estimates for the branch-site test of positive selection in the *Physaria* crown clade.

Site class	Proportion	$\omega_{\text{background}}$	$\omega_{\text{foreground}}$
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

Table 3

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.

Site	PP ^a	Outgroups, 22 taxa	Physarieae, 11 taxa	<i>Physaria</i> , 25 taxa
153	0.996	Ser (77%), Ala (18%), <i>Leu</i>	Ser (100%)	Ser (52%), Thr (44%)
167	1.000	Gln (77%), Glu (9%), His (9%), <i>Lys</i>	Gln (91%), <i>Arg</i>	Gln (48%), Arg (36%), Lys (12%), <i>Glu</i>
171	1.000	Leu (100%)	Leu (91%)	Leu (52%), Ser (8%), <i>Arg</i> , <i>Phe</i>
173	0.998	Glu (73%), Asp (27%)	Glu (73%), Asp (18%)	Asp (36%), Asn (28%), <i>Gly</i>
174	1.000	Val (91%), <i>Arg</i> , <i>Asp</i>	Val (100%)	Val (60%), Arg (24%), Ala (8%), <i>Asp</i>
175	1.000	Leu (100%)	Leu (100%)	Leu (76%), Trp (8%), <i>Lys</i> , <i>Phe</i> , <i>Ser</i>
180	0.999	Val (100%)	Val (91%)	Val (88%), <i>Ala</i> , <i>Asp</i> , <i>Ser</i>
190	0.989	Phe (91%), <i>Tyr</i>	Phe (91%)	Phe (68%), Leu (24%), <i>Ile</i> , <i>Val</i>
191	0.996	Leu (95%)	Met (64%), Leu (27%)	Leu (72%), Phe (24%), <i>Trp</i>
255	0.963	Pro (91%), <i>Arg</i> , <i>NI</i> ^b	Pro (100%)	Pro (40%), Leu (32%), Ser (20%), <i>Gln</i>

^aposterior probability

^bno sequence information available for one taxon at this site

Table 4

Table 4. Distinctive *Physaria* codon sites.

Site	Outgroups, 33 taxa	<i>Physaria</i> , 25 taxa
102	Leu ^a	Val
151	Glu ^b	Ala ^c
195	Leu	Ile
278	Ile	Leu

^aexcept *B. oleracea* , Met

^bexcept *C. rubella* , *C. sativa* ,*T. hassleriana* , Asp; *B. incana* , Ala

^cexcept *P. arizonica* , Val

Graphical Abstract

POSITIVE SELECTION