Duplication of the *Brassica oleracea APETALA1* Floral Homeotic Gene and the Evolution of Domesticated Cauliflower

A. C. Lowman and M. D. Purugganan

Development of the cauliflower phenotype in *Arabidopsis thaliana* requires mutations at both the *CAULIFLOWER* and *APETALA1* loci. *BoAP1* is the *Brassica oleracea* orthologue to the *Arabidopsis AP1* gene, and is present in two copies in *Brassica* genomes. The *BoAP1-A* gene appears to encode a full-length protein, but *BoAP1-B* alleles in *B. oleracea* contain insertions that lead to premature translation termination. The *BoAP1-B* allele found in most *B. oleracea* subspecies, including *B. oleracea* ssp. *botrytis* (domesticated cauliflower) contains a 9 bp insertion in exon 4. This insertion leads to the formation of an in-frame translation termination codon, and these alleles can encode a protein that is truncated at the K domain of this MADS-box transcriptional activator. The allele in *B. oleracea* ssp. *oleracea* (wild cabbage) lacks this insertion and instead contains a downstream 4 bp frameshift mutation resulting in the formation of a nonsense mutation. The structure of the *BoAP1-B* alleles suggests that they are impaired in their ability to perform their floral meristem identity function. These mutations, in conjunction with mutations at the *BoCAULIFLOWER* (*BoCAL*) locus, may be associated with the evolution of domesticated cauliflower.

Domesticated plant species provide excellent models to study and test hypotheses regarding the evolutionary genetics of morphological diversification (Doebley 1992, 1993; Doebley et al. 1997; Dorweiler et al. 1993). The evolution of crop species during domestication is invariably accompanied by changes in suites of structural traits that differentiate the cultivated species from its wild relatives (Doebley 1992; Schwanitz 1967). These traits include variations in inflorescence structure, plant vegetative architecture, seed shattering, and fruit and seed morphology; the evolution of these new morphologies are the result of human selection for domesticated traits or adaptation to agroecosystemic environments (Schwanitz 1967). Crop species have thus been widely regarded as providing some of the best and most dramatic examples of the degree to which plant morphologies evolve under selection pressures (Doebley 1993; Gottlieb 1984).

Among different domestic plant taxa, the distinct morphologies exhibited by *Brassica oleracea* subspecies represents one of the most spectacular illustrations of structural evolution in plants. *B. oleracea* is a perennial herb ($2n = 18$) found largely in Europe and the Mediterranean (Crisp and Tapsell 1993; Song et al. 1995; Tsunoda et al. 1980). It is an extremely polymorphic species and includes at least six cultivated and one wild subspecies. Wild, perennial forms of *B. oleracea,* designated subspecies *oleracea* (wild cabbage), normally grow in coastal rocky cliffs of the Mediterranean, northern Spain, western France, and southern and southwest Britain (Tsunoda et al. 1980). The cultivated subspecies of *B. oleracea* display extreme morphological divergence as a result of selection for different characteristics during domestication. Three subspecies display marked modification of leaf organs: subspecies *capitata* (cabbage), subspecies *acephala* (kale), and subspecies *gemmifera* (brussels sprouts). *B. oleracea* ssp. *gongylodes* (kohlrabi) has thickened stems as the edible plant portion. The subspecies *botrytis* (cauliflower) and *italica* (broccoli), on the other hand, are characterized by modification of the inflorescence into large dense structures. The precociously large, undifferentiated inflorescence, termed the curd, is the defining characteristic of the cultivated subspecies *botrytis.* The cauliflower curd consists of a dense mass of arrested inflorescence meristems, only about 10% of which will later

From the Department of Genetics, Box 7614, North Carolina State University, Raleigh, NC 27695. Address correspondence to M. D. Purugganan at the address above or e-mail: michael_purugganan@ncsu.edu. The authors would like to thank J. I. Suddith for technical assistance, and J. McFerson and D. Astley for providing seed. We are also grateful to Marty Yanofsky for helpful discussions on possible genetic models for cauliflower formation. This work was supported in part by a grant from the USDA National Research Initiative Competitive Grants program and an Alfred P. Sloan Foundation Young Investigator Award to M.D.P.

 90:514–520

develop into floral primordia and normal flowers (Sadik 1962).

Our understanding of the molecular genetics of the cauliflower phenotype has been aided by significant progress in unraveling the molecular genetic bases of floral and inflorescence development in a related crucifer, *Arabidopsis thaliana* (Bowman et al. 1993; Weigel 1995; Weigel and Meyerowitz 1994). In *Arabidopsis,* several floral homeotic genes that control flower development have been identified; mutational lesions in these genes result in the formation of aberrant organ types within flowers. Molecular studies reveal that many floral homeotic genes belong to the MADS-box regulatory gene family (Riechmann and Meyerowitz 1997; Yanofsky 1995). Members of this gene family encode proteins of about 220–260 aa in length and are characterized by the presence of a highly conserved 57-amino acid MADS box, which is also found in humans (*SRF*), *Drosophila* (*MEF2C*), and yeast (*MCM1*) transcriptional activators (Dalton and Treisman 1992; Pollock and Treisman 1991). Plant MADS-box genes also contain a moderately conserved 70-amino acid domain called the K box, which is believed to possess the capacity to form coiled-coil structures that may participate in dimerization interactions. Two other regions (the I and C regions) are also found in these proteins, but are poorly conserved in comparisons between different plant MADS-box genes (Purugganan et al. 1995).

One class of floral homeotic genes are the early acting floral meristem identity genes, which specify the identity of the floral meristem (as opposed to the inflorescence meristem). Members of this class include the genes *APETALA1* (*AP1*) (Gustafson-Brown et al. 1994; Mandel et al. 1992) and *CAULIFLOWER* (*CAL*) (Kempin et al. 1995]. The *APETALA1* locus has primary functions in specifying floral meristem identity, but is also necessary for the correct specification of sepal and petal organs in the flower. Mutations in *AP1* result in the loss of sepals and petals, with leaflike organs developing in place of sepals. At the axils of these leaflike organs, new *ap1* flowers form which recapitulate the whole *ap1* floral ontogeny and result in the development of a nested series of flowers. Another gene, *CAULIFLOWER,* has also been shown to participate in the specification of floral meristem identity. *Arabidopsis* individuals that are mutant for both *AP1* and *CAL* are arrested in development at the inflorescence meristem stage (Kempin et al. 1995). In these plants, a dense mass of inflorescence meristems develop similar to the *B. oleracea* ssp. *botrytis* curd.

Genetic analysis in *Arabidopsis* has suggested the involvement of the *B. oleracea CAL* and *AP1* orthologues, which are referred to as *BoCAL* (Kempin et al. 1995) and *BoAP1* (Anthony et al. 1993, 1996; Carr and Irish 1997), in the formation of the altered inflorescence in *B. oleracea* ssp. *botrytis.* It has already been demonstrated that the *BoCAL* allele in domesticated cauliflower has a premature termination codon at position 151 (E → stop) (Kempin et al. 1995). Previous studies, however, have failed to determine whether a mutant *BoAP1* allele also exists in ssp. *botrytis* (Anthony et al. 1993, 1996). In *Arabidopsis,* mutations at both loci are necessary for the cauliflower phenotype. In this article we report that *BoAP1* is present in two copies in *Brassica* genomes. Only one copy encodes a full-length *BoAP1* protein; the other *BoAP1* copy encodes a prematurely terminated protein. The structure of this mutant duplicate copy structure suggests that it has a reduced ability to specify floral meristem identity and may be associated with the development of the cauliflower curd.

Materials and Methods

Study Species

The following *B. oleracea* subspecies were used in these analyses: ssp. *oleracea* (wild cabbage), ssp. *capitata* (cabbage), ssp. *acephala* (kale), ssp. *gongylodes* (kohlrabi), ssp. *botrytis* (cauliflower), and ssp. *italica* (broccoli). The wild relative *B. insularis* was also utilized in this study. We were unable to obtain material for *B. oleracea* ssp. *gemmifera* (brussels sprouts) with a clear provenance, and excluded this subspecies from the analyses. Seed from these species/subspecies were obtained from the HRI Genetic Resources Unit at Wellesbourne, UK, the Center for Genetics Resources in the Netherlands, and the USDA-ARS Plant Genetic Resources Unit at Geneva, NY.

Isolation and Sequencing of *B. Oleracea AP1* (*BoAP1*) Alleles

Genomic DNA from young *B. oleracea* plant leaves was isolated using a modification of a procedure widely utilized in obtaining DNA from *A. thaliana* leaves (Dellaporta et al. 1983). The isolated genomic DNA (1 μg) was used as a template in PCR amplification reactions to isolate the *BoAP1* gene. PCR primers were initially designed, based on published sequences (Carr and Irish 1997), to recognize sequences at exon 3 and exon 8 of *BoAP1.* PCR amplification was undertaken using the error-correcting rTth polymerase XL formulation (Perkin-Elmer, Branchburg, NJ). The introduced nucleotide errors using this polymerase formulation are less than 1 in 8000 bp (Purugganan MD and Suddith JI, unpublished observations). The reaction conditions were those recommended by the manufacturer, and included 40 cycles at 94°C for 1 min, 42°C–52°C for 1 min (optimized for specific reactions), and 70°C for 3 min. Amplification products were cloned into plasmid vector using the TA cloning kit (InVitrogen, San Diego, CA).

Preliminary analysis indicated the presence of two copies of *BoAP1* within *Brassica* genomes. Based on unambiguous nucleotide differences between the two copies, gene-specific primers were constructed to isolate the segments between exon 3 and intron 7 of each copy. PCR amplification reactions using genomic DNA were undertaken as above, and amplified products cloned as above. Nested primers in both directions were used to sequence cloned PCR products. Sequencing was undertaken using the ABI 377 automated sequencer at the North Carolina State University (NCSU) DNA sequencing facility. Ambiguous and low-frequency polymorphisms within *B. oleracea* were visually corrected against the chromatograms and/or checked with resequencing. The sequences are available in GenBank (accession numbers AF126722–AF126734).

Reverse Transcription Polymerase Chain Reaction (RT-PCR) Expression Studies

Expression of different copies of *BoAP1-A* and *BoAP1-B* was assayed by RT-PCR techniques. Total RNA was isolated from ssp. *botrytis* curds and ssp. *italica* floral sprouts using an RNA plant extraction kit (Qiagen, Santa Clarita, CA). Poly(A+) RNA was purified from total RNA with the Oligotex suspension system (Qiagen, Santa Clarita, CA), and oligo(dT)-primed first strand cDNA was synthesized using AMV reverse transcriptase (Promega, Madison, WI). Products of the cDNA syntheses were used as a template in RT-PCR reactions using a universal exon 1 forward and gene-specific exon 4 (*BoAP1-B*) or exon 8 (*BoAP1-A* and *BoAP1-B*) PCR primers.

Alignments and Phylogenetic Analyses

Sequences were edited and aligned visually using the ASSEMBLYALIGN program

(Kodak/IBI, Rochester, NY). Sequences of other Brassicaceae *AP1* genes (from *A. thaliana, Arabis lyrata,* and *Sinapis alba*) were obtained from GenBank and aligned to the *BoAP1* coding region using previous *AP1* alignments as a guide (Purugganan et al. 1995). The gene genealogy of *BoAP1* alleles was estimated using both maximum parsimony (MP) (Swofford 1993) and neighbor-joining (NJ) (Saitou and Nei 1987) techniques. For the MP technique, the heuristic search algorithm with the tree bisection-reconnection procedure of the PAUP 3.0 program was used, with random addition of genes, and with the MULPARS and collapse options in effect. Node confidence was assessed with 500 bootstrap replicates of the data. For the NJ analysis, genetic distances were estimated using the Kimura 2P model. Confidence estimates were assessed with 500 bootstrap replicates of the data. The NJ analysis was undertaken using the MEGA program package (Kumar et al. 1994). Brassicaceae *CAL* genes were used as the outgroup genes in the analyses.

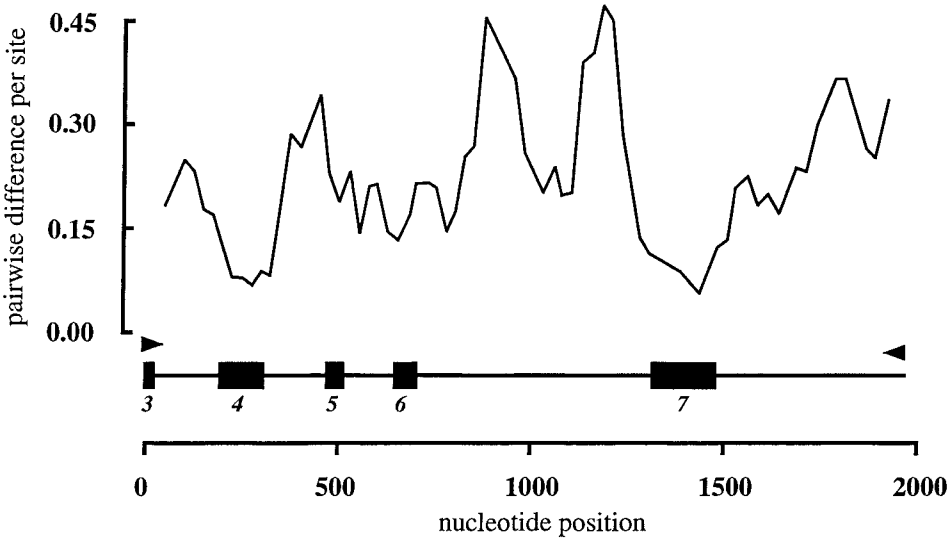


Figure 1. Sliding window analysis of nucleotide differences between *BoAP1-A* and *BoAP1-B.* The diagram indicates the levels of mean pairwise difference per site in a 100 bp window translated in steps of 25 bp. A schematic of the *BoAP1* gene showing introns and exons is included, with exons numbered and the position of gene-specific PCR primers indicated by arrows.

Results and Discussion

The *BoAPETALA1* Gene Is Present in Duplicate Copies

The *APETALA1* gene in *A. thaliana* is part of the regulatory pathway that controls the transition from inflorescence to floral meristems, as well as the specification of sepals and petals in developing flowers (Yanofsky 1995). Orthologues to the *AP1* locus have been isolated in other Brassicaceae species, including *S. alba* (Menzel et al. 1995) and *B. oleracea* (Anthony et al. 1993, 1996; Carr and Irish 1997). *BoAP1* is the *B. oleracea* orthologue of *AP1* and is expressed in developing *B. oleracea* inflorescence structures, including the cauliflower curd found in *B. oleracea* ssp. *botrytis* (Anthony et al. 1996; Carr and Irish 1997).

Our results indicate, however, that *BoAP1* is present in two copies within the genomes of *B. oleracea* and its wild relative *B. insularis.* These two copies are referred to as *BoAP1-A* and *BoAP1-B.* The sequence of *BoAP1-A* is almost identical to the *BoAP1* gene sequence previously isolated by other workers. Using gene-specific PCR primers, we have isolated a segment of both *BoAP1-A* and *BoAP1-B* in both wild and cultivated subspecies of *B. oleracea*; although PCR analysis indicates that *B. insularis* contains both copies, sequence data was successfully produced only from *BoAP1-B.* The gene sequences obtained span exon 3 to the 3′ end of intron 7 and were approximately 1.8 kb for each copy of these duplicate loci. The isolated sequence includes the coding region that specifies the K-box and C-terminal domain of the BoAP1 transcriptional activator.

Comparison of the sequence from the two copies found in *B. oleracea* ssp. *acephala* (kale) reveals that 52 insertion/deletions ranging in size from 1 to 23 bp differentiate *BoAP1-A* and *BoAP1-B.* The genetic distance (K) between the duplicates is 0.26 substitutions/site using the Kimura-2P model. Sliding window analysis (see Figure 1) indicates that most of the divergence between duplicates is associated with intron sequences. In the exon sequences, 19 nucleotide differences are found between *BoAP1-A* and *BoAP1-B.* Fourteen of these are synonymous changes while five are nonsynonymous differences that result in amino acid replacements between the duplicate copies.

The presence of two copies of *BoAP1* in *B. oleracea* genomes is not unexpected. Recent genetic mapping and cytogenetic studies have provided evidence that most *Brassica* species are paleopolyploids and have experienced an increase in chromosome number during their evolution within the Brassicaceae (Lagercrantz 1998; Slocum et al. 1990; Song et al. 1995; Truco et al. 1996). RFLP analysis indicates that most *Brassica* genomes, including *B. oleracea* ($n = 9$), may be the result of polyploidization events from an ancestral species of smaller genome size ($n = 6$) (Truco et al. 1996). As a result, many of the genes within *B. oleracea* are present in duplicate copies; the *BoAP1* gene isolated by previous workers appears to be only one of two copies of this floral regulatory gene.

Evolutionary Relationships Between *AP1* Genes Within the Brassicaceae

Phylogenetic analysis provides a framework for dissecting the evolutionary relationships between the *AP1* genes in the Brassicaceae. The coding regions (exons 3–7) of the *BoAP1-A* and *BoAP1-B* genes in *B. oleracea* and *B. insularis* were aligned with sequences of the *A. thaliana* and *S. alba AP1* loci, and a gene tree estimated using maximum parsimony analysis. Four equally parsimonious trees were estimated (length = 174 steps), which differed only in the placement of the *BoAP1-A* and *BoAP1-B* alleles at the tips of the tree. The consistency index of the data to the tree is 0.943. A bootstrap analysis with 500 replicates provided support for the various nodes within the phylogeny (see Figure 2A), and the estimated tree was largely congruent with a phylogeny based on neighbor-joining reconstruction methods (see Figure 2B).

The *BoAP1-A* and *BoAP1-B* alleles in *B. oleracea* and *B. insularis* are phylogenetically distinct and each form separate monophyletic gene clades with 100% bootstrap support in both maximum parsimony and neighbor-joining trees (see Figure 2). The phylogeny indicates that the divergence between *BoAP1-A* and *BoAP1-B* occurred within the Brassicaceae; the *A.*

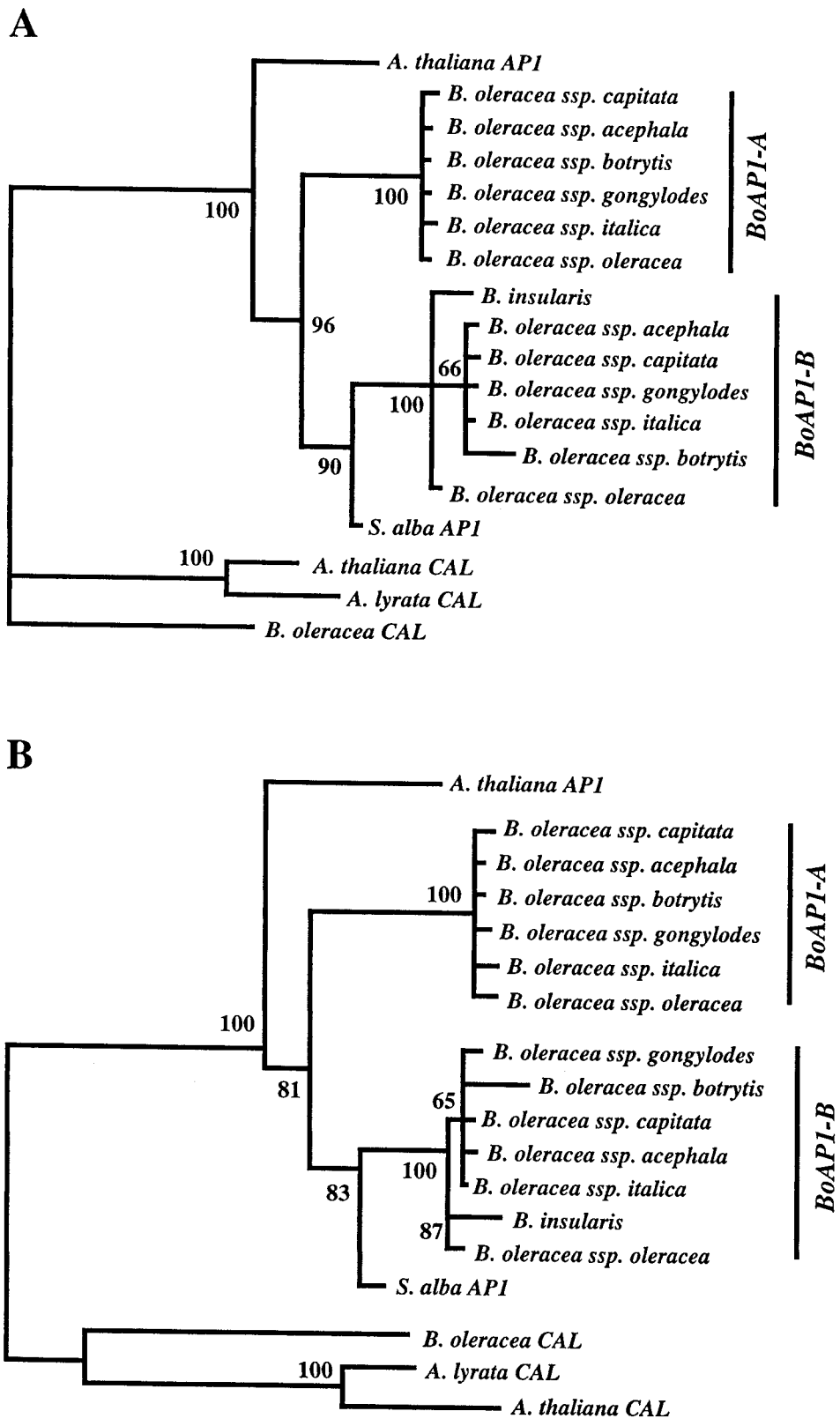


Figure 2. Phylogenies of coding regions for *AP1* loci in the Brassicaceae. The phylogenetic trees were obtained using both **(A)** maximum parsimony and **(B)** neighbor-joining methods. The genes in these analyses included *AP1* loci from *B. oleracea, A. thaliana,* and *S. alba.* The paralogous *CAULIFLOWER* genes were used as the outgroup sequences. Numbers next to the nodes show bootstrap support from 500 replicated pseudodatasets. All nodes that are either near-zero length and/or with less than 60% support are collapsed in the trees.

thaliana AP1 locus is found to be phylogenetically basal to the *BoAP1* gene copies. Thus the hypothetical polyploidization events that may have led to the duplication of *BoAP1-A* and *BoAP1-B* appears to have transpired after the split of the *Arabidopsis* line from the *Brassica/Sinapis* lineage. Only one *AP1* gene has been reported from *S. alba,* and it groups strongly (90% bootstrap support in the maximum parsimony analysis) with the *BoAP1-B* duplicate copy.

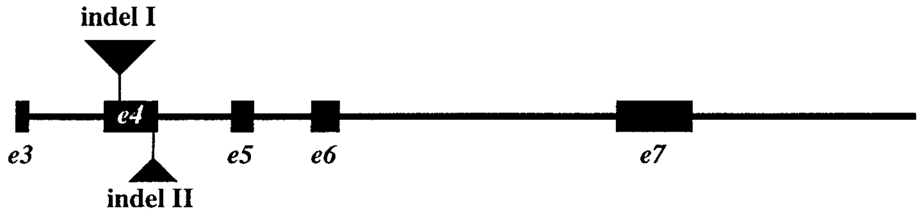
Insertion/Deletion Mutations in Exon 4 of *BoAP1-B* Alleles in *B. oleracea*

All alleles isolated from both wild and cultivated *B. oleracea* subspecies contain insertion/deletion (indel) mutations in exon 4 of *BoAP1-B* (see Figure 3). Different alleles, however, possess one of two possible insertion mutations. A 9 bp insertion is found in the *BoAP1-B* alleles of all domesticated *B. oleracea* subspecies. This insertion is located in exon 4 of the gene, just 35 bp downstream of the splice acceptor site; all alleles that carry this insertion are designated *BoAP1-B(I).* This insertion is missing, however, in the one *BoAP1-B* allele sampled in the wild subspecies *oleracea* (see Figure 3B). The *oleracea* allele instead contains a unique 4 bp frameshift insertion within the same exon. This unique indel is downstream of the insertion in *BoAP1-B(I),* and is located 9 bp before the intron 4 splice donor site. We refer to the allele found in subspecies *oleracea* as *BoAP1-B(II).* A restriction enzyme site survey of *BoAP1-B* alleles in different *B. oleracea* subspecies, using a unique *Cfo*I site generated by the 4 bp insertion in the *BoAP1-B(II)* allele, indicates that this insertion is found only in ssp. *oleracea* accessions and not in domesticated subspecies. The allele isolated from the *BoAP1-B* orthologue found in wild *B. insularis* lacks both insertions.

Both insertions/deletions create stop codons that result in premature termination of translation for the encoded proteins (see Figure 3). The *BoAP1-B(I)* allele can encode a protein that terminates within the K box, and includes the first 32 amino acids of this 68 aa domain (Riechmann and Meyerowitz 1997). The *BoAP1-B(II)* allele also encodes a truncated protein, but this putative protein would be 19 amino acids larger than that encoded by *BoAP1-B(I).*

BoAP1-A and *BoAP1-B* also display differing patterns of nucleotide substitution in their coding regions. The coding region from exons 3–7 were aligned for *BoAP1-A*

A



B

```
                                149                                         I           204
B. oleracea italica        ACACTATCTTGGGGAAGACTTGCAAGCAATGAGC |TAA| ATAGGCTCTAAGGAGCTCC
B. oleracea acephala       ACACTATCTTGGGGAAGACTTGCAAGCAATGAGC |TAA| ATAGGCTCTAAGGAGCTCC
B. oleracea botrytis       ACACTATCTTGGGGAAGACTTGCAAGCAATGAGC |TAA| ATAGGCTCTAAGGAGCTCC
B. oleracea gongylodes     ACACTATCTTGGGGAAGACTTGCATGCAATGAGC |TAA| ATAGGCTCTAAGGAGCTCC
B. oleracea capitata       ACACTATCTTGGGGAGGACTTGCAAGCAATGAGC |TAA| ATAGGCTCTAAGGAGCTCC
B. oleracea oleracea       ACACTATCTTGGGGAAGACTTGCAAGCAATGAGC T..  .......CTAAGGAGCTCC
B. insularis               ACACTATCTTGGGGAAGACTTGCAAGCAGTGAGC T..  .......CTAAGGAGCTCC

                                205                                         II          261
B. oleracea italica        AGAATCTGGAGCAGCAGCTTGACACAGCTCTTAAGCACATCCGC....TC TAG AAAG
B. oleracea acephala       AGAATCTGGAGCAGCAGCTTGACACAGCTCTTAAGCACATCCGC....TC TAG AAAG
B. oleracea botrytis       AGAATCTGGAGCAGCAGCTTGACACAGCTCTTAAGCACATCCGC....TC TAG AAAG
B. oleracea gongylodes     AGAATCTGGAGCAGCAGCTTGACACAGCTCTTAAGCACATCCGC....TC TAG AAAG
B. oleracea capitata       AGAATCTGGAGCAGCAGCTTGACACAGCTCTTAAGCACATCCGC....TC TAG AAAG
B. oleracea oleracea       AGAATCTGGAGCAGCAGCTTGACACAGCTCTTAAGCACATCCGCGCGCTC |TAG| AAAG
B. insularis               AGAATCTGGAGCAGCAGCTTGACACAGCTCTTAAGCACATCCGC....TC TAG AAAG
```

Figure 3. Insertion/deletion mutations in *BoAP1-B*. **(A)** Schematic of *BoAP1-B* gene showing positions of 9 bp (indel I) and 4 bp insertions (indel II). **(B)** Exon 4 sequence of *BoAP1-B* from different *B. oleracea* subspecies. The indels are overlined and the in-frame premature translation termination codons produced by the insertions are shown in boxes.

and *BoAP1-B* genes in all *B. oleracea* subspecies, ignoring the exon 4 insertions in the latter gene. This region encompasses the sequence that encodes most of the K-box and C-terminal domain. For *BoAP1-A,* one replacement and two silent nucleotide site differences are observed between all *B. oleracea* subspecies alleles. In contrast, the same region contains seven replacement and three nucleotide site differences in the *BoAP1-B* gene. There appears to be an excess of replacement nucleotide site changes within the *BoAP1-B* gene, although this excess is not statistically significant ($p = .315$, one-tailed Fisher's exact test).

BoAP1-A, but not *BoAP1-B,* Is Expressed in *B. oleracea* ssp. *botrytis* and ssp. *italica* Inflorescences

The mutations at *BoAP1-B* suggest that this locus, if expressed, could encode a truncated protein. Alternatively, the premature translation termination signals introduced in alleles of this gene, as well as the increase in replacement site substitutions, suggests that *BoAP1-B* may be a pseudogene that is not expressed in the developing inflorescence of *B. oleracea.* RT-PCR analysis using gene-specific primers was undertaken to determine whether *BoAP1-A* and *BoAP1-B* genes are expressed in the curd of *B. oleracea* ssp. *botrytis* and the floral sprouts of *B. oleracea* ssp. *italica* (broccoli).

The *BoAP1-A* gene is transcribed in both subspecies *botrytis* and *italica* (see Figure 4), confirming earlier expression studies in these two subspecies (Anthony et al. 1996; Carr and Irish 1997). RT-PCR analyses of floral cDNAs results in the amplification of an expected 650 bp band from both ssp. *italica* and *botrytis,* which spans most of the gene's coding region (exons 1–7). Sequencing of *BoAP1-A* cDNAs from both subspecies also demonstrates that no other premature translation termination codons are present in the coding region of this gene. No expression of the *BoAP1-B* gene, however, was detected in both ssp. *italica* and *botrytis* during inflorescence development using two sets of gene-specific PCR primers. The results suggest that *BoAP1-B* is transcriptionally silenced in subspecies *botrytis* and *italica,* and appears to be a pseudogene in these domesticated *B. oleracea* groups.

BoAP1-B Alleles and the Evolution of Domesticated Cauliflower

Based on genetic studies in *A. thaliana,* it has been surmised that the cauliflower phenotype observed in *B. oleracea* ssp. *botrytis* requires mutations in both the *BoCAL* and *BoAP1* loci (Anthony et al. 1996; Kempin et al. 1995). Previous studies have already identified a unique nonsense mutation in exon 5 of the *BoCAL* allele in domesticated cauliflower (Kempin et al. 1995), but it remains unclear whether mutations in *BoAP1* are also present to condition the inflorescence phenotype in subspecies *botrytis* (Anthony et al. 1993, 1996).

We have shown that *BoAP1* is present in duplicate copies, *BoAP1-A* and *BoAP1-B,* within *B. oleracea* and its close relative *B. insularis. BoAP1-B* alleles in *B. oleracea* all contain insertion/deletion mutations that result in premature translation termination (see Figure 3). These insertions, which lead to nonsense mutations, are not present in the closely related wild species *B. insularis* or *S. alba* (Menzel et al. 1995). Moreover, no expression from this gene has been detected, suggesting that it persists as a pseudogene in domesticated taxa. In contrast, expression of the *BoAP1-B* orthologue in *S. alba* has been detected in the apical meristem of this plant (Menzel et al. 1995). It is unclear whether these two genes differ in tissue-specific expres-

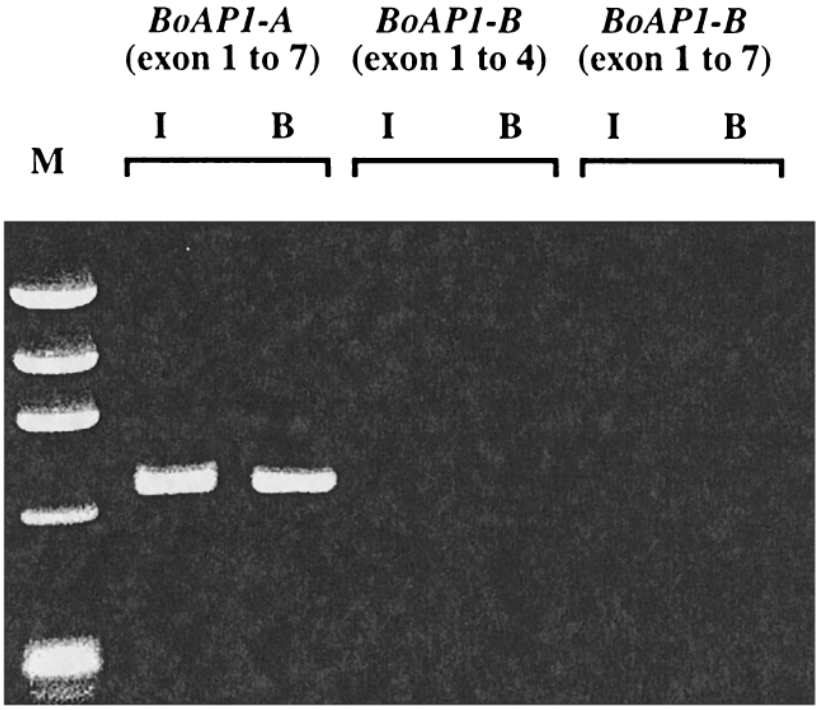


Figure 4. Expression analysis of *BoAP1-A* and *BoAP1-B* in developing inflorescences. RT-PCR analysis indicates that *BoAP1-A* genes are expressed in both *B. oleracea* ssp. *italica* (I) and ssp. *botrytis* (B). Two different primer pairs (*BoAP1-B* exons 1–4 and exons 1–7) failed to detect expression of *BoAP1-B* genes in these two subspecies. M is the marker lane (ΦX174 *Hae*III fragments).

sion patterns during floral development; the high similarity between the coding regions of the two genes makes it difficult to discriminate the expression of the two loci using in situ RNA expression assays.

In *B. oleracea* ssp. *botrytis,* curd formation is known to be controlled by several loci, including at least one major locus and several modifier genes (Crisp and Tapsell 1993). Our results suggest that domesticated cauliflower may be associated with variation in at least two loci: a nonsense mutation in *BoCAL,* and a 9 bp insertion at *BoAP1-B.* The widespread distribution of the insertion mutation in *BoAP1-B* among different subspecies indicates that this polymorphism probably arose first (see Figure 5). Ancestral-state reconstruction suggests that the 9 bp insertion event occurred after the divergence of *B. oleracea* ssp. *oleracea* from most of the domesticated *B. oleracea* subspecies. The 4 bp frameshift insertion in the *BoAP1-B(II)* allele appears to have arisen in the *B. oleracea* ssp. *oleracea* lineage after the establishment of most of the domestic groups. The inability to detect mRNA from *BoAP1-B* may be due to transcript instability as a result of the coding region insertions, or the presence of distinct mutations in the promoter region that have abolished transcription of this gene. It appears that the conversion of *BoAP1-B* into a pseudogene was a relatively recent event. Unlike in *B. oleracea,* the *BoAP1-B* copy in its wild relative *B. insularis* does not contain any frameshift mutations in the coding region.

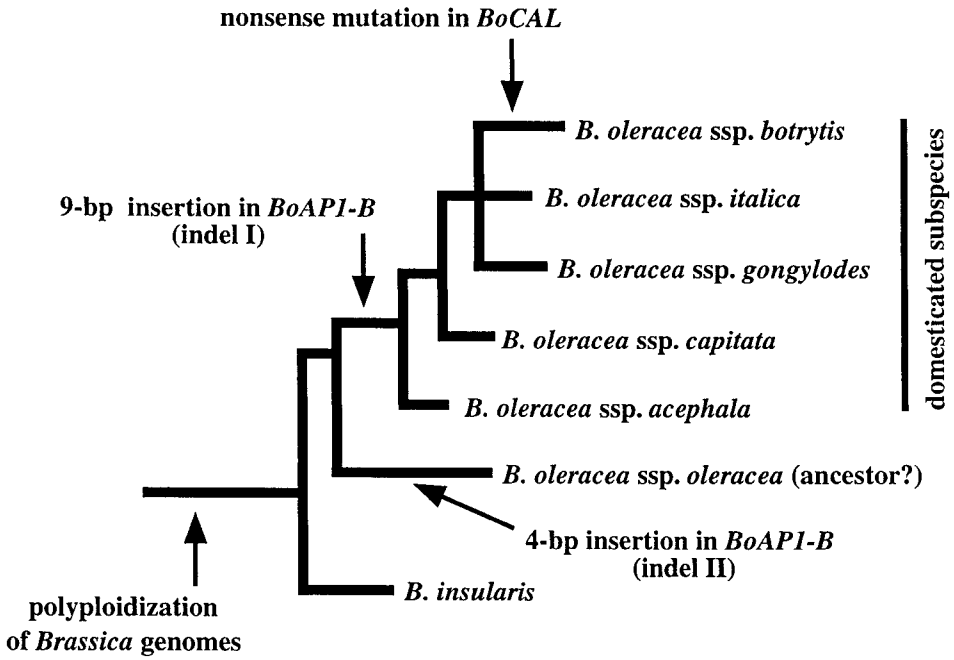


Figure 5. Evolutionary hypothesis of genetic changes associated with origin of domesticated cauliflower (*B. oleracea* ssp. *botrytis*). The placement of events in the phylogeny is based on character state reconstruction of gene sequences. The presumed phylogeny of *B. oleracea* subspecies assumes that ssp. *oleracea* is ancestral to all domesticated subspecies and that ssp. *acephala* and *capitata* were the first domesticates. The scenario proposed is not dependent on the actual relationships between domesticated taxa.

The pattern of variation between these two genes indicates two possible genetic models for the evolutionary origin of domesticated cauliflower. A divergence model suggests that the *BoAP1* and *BoCAL* loci may have acquired distinct functional specializations, with both *BoAP1-A* and *BoAP1-B* losing their abilities to specify floral meristem identity. Under this model, mutations at *BoCAL* alone, such as those reported in ssp. *botrytis* (Kempin et al. 1995), are sufficient for the formation of cauliflower heads. Alternatively, a redundant gene model assumes that the activities of at least two floral meristem identity genes, *BoAP1-B* and *BoCAL,* are necessary for flower formation in this species. Under this model, all subspecies of *B. oleracea* posses mutant alleles for *BoAP1-B,* but redundant floral meristem identity functions in *BoCAL* compensate for the partial loss of activity. In subspecies *botrytis,* however, additional mutations at *BoCAL* result in greater loss of meristem specification activity and presumably leads to the cauliflower phenotype. In this model, *BoAP1-A* may still retain meristem identity activity, but activity of this gene alone is insufficient to program normal floral primordia initiation.

It is only under the redundant gene model that the mutations at *BoAP1-B* may be considered to be directly associated with the evolution of domesticated cauliflower. The duplication of the *AP1* ortho-

logue in *B. oleracea,* however, may explain at least one difference in the genetic architecture of the cauliflower trait in *A. thaliana* and *B. oleracea* ssp. *botrytis.* Both the *Arabidopsis* cauliflower mutants (Bowman et al. 1993) and *B. oleracea* ssp. *botrytis* (Sadik 1962) will eventually form some flowers after prolonged growth. In *Arabidopsis,* the flowers are apetalous and display homeotic transformations in sepal structures. These organ identity defects presumably arise from the lack of a functional *AP1* gene in *Arabidopsis* cauliflower plants, which are *ap1 cal* double mutants. In *B. oleracea* ssp. *botrytis,* however, the flowers that eventually form have normal sepals and petals (Sadik 1962), which suggests that floral organ identity functions are intact in this subspecies. Our models indicate that subspecies *botrytis* should possess a functional *BoAP1-A* gene. Since the *BoAP1-A* gene in *B. oleracea* ssp. *botrytis* is still functional, the flowers that eventually form should be normal. Intriguingly, this particular genetic architecture of cauliflower formation in domesticated cauliflower may have been selected by primitive farmers, who would have needed fully fertile flowers to produce seed for further efficient propagation of this subspecies.

Recent studies in the genetic architecture of morphological evolution in crop plants suggest that mutations at a few loci may be responsible for much of the observed structural divergence between cultivated and wild taxa (Doebley 1992, 1993; Gottlieb 1984). The evolution of the *BoAP1* loci in *B. oleracea* illustrates one of the evolutionary genetic pathways that characterize domestication in this plant species. These results suggest that it may be possible to dissect the genetic basis of evolutionary transformations in crop plant morphology through the use of both molecular evolutionary and molecular genetic analyses (Doebley 1992, 1993), and provides a window to understanding the origins of structural diversification between species.

References

Anthony RG, James PE, and Jordan BR, 1993. Cloning and sequence analysis of a FLO/LFY homologue isolated from cauliflower. Plant Mol Biol 22:1163–1166.

Anthony RG, James PE, and Jordan BR, 1996. Cauliflower curd development—the expression of meristem identity genes. J Exp Bot 47:181–188.

Bowman JL, Alvarez J, Weigel D, Meyerowitz EM, and Smyth D, 1993. Control of flower development in *Arabidopsis thaliana* by apetala1 and interacting genes. Development 119:721–743.

Carr SM and Irish VF, 1997. Floral homeotic gene expression defines developmental arrest stages in *B. oleracea* L vars. *botrytis* and *italica.* Planta 201:179–188.

Crisp P and Tapsell CR, 1993. Cauliflower. In: Genetic improvement of vegetable crops (Kalloo G and Bergh BO, eds). Oxford: Pergamon Press.

Dalton S and Treisman R, 1992. Characterization of SAP-1, a protein recruited by serum response factor to the C-fos serum response element. Cell 68:597–612.

Dellaporta SL, Wood J, and Hicks JB, 1983. Isolation of plant DNA. Plant Mol Biol Rep 1:19–21.

Doebley J, 1992. Mapping the genes that made maize. Trends Genet 8:302–307.

Doebley J, 1993. Genetics, development and plant evolution. Curr Opin Genet Dev 3:865–872.

Doebley J, Stec A, and Hubbard L, 1997. The evolution of apical dominance in maize. Nature 386:485–488.

Dorweiler J, Stec A, Kermicle J, and Doebley J, 1993. Teosinte-glume architecture 1—a genetic locus controlling a key step in maize evolution. Science 262:233–235.

Gottlieb LD, 1984. Genetic and morphological evolution in plants. Am Nat 123:681–709.

Gustafson-Brown C, Savidge B, and Yanofsky MF, 1994. Regulation of the *Arabidopsis* floral homeotic gene *APETALA1.* Cell 76:131–143.

Kempin S, Savidge S, and Yanofsky MF, 1995. Molecular basis of the cauliflower phenotype in *Arabidopsis.* Science 267:522–525.

Kumar S, Tamura K, and Nei M, 1994. Molecular evolutionary genetics analysis package 1.1. State College, PA: Pennsylvania State University.

Lagercrantz C, 1998. Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150:1217–1228.

Mandel MA, Gustafson-Brown C, Savidge B, and Yanofsky MF, 1992. Molecular characterization of the *Arabidopsis* floral homeotic gene *APETALA1.* Nature 360:273–277.

Menzel G, Apel K, and Melzer S, 1995. Identification of two MADS box genes that are expressed in the apical meristem of the long-day plant *Sinapis alba* in transition to flowering. Plant Physiol 108:853–854.

Pollock R and Treisman R, 1991. Human SRF-related proteins—DNA-binding properties and potential regulatory targets. Genes Dev 5:2327–2341.

Purugganan MD, Rounsley SD, Schmidt RJ, and Yanofsky MF, 1995. Molecular evolution of flower development—diversification of the plant MADS-box regulatory gene family. Genetics 140:345–356.

Riechmann JL and Meyerowitz EM, 1997. MADS domain proteins in plant development. Biol Chem 378:1079–1101.

Sadik S, 1962. Morphology of the curd of cauliflower. Am J Bot 49:290–297.

Saitou N and Nei M, 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425.

Schwanitz F, 1967. Die evolution der kulturpflanzen. Munich: Bayerischer Landwirtschaftverlag.

Slocum MK, Figdore SS, Kennard WC, Suzuki JY, and Osborn TC, 1990. Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea.* Theor Appl Genet 80:57–64.

Song KM, Lu P, Tang KL, and Osborn TC, 1995. Rapid genomic change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc Natl Acad Sci USA 92:7719–7723.

Swofford D, 1993. Phylogenetic analysis using parsimony. Champaign, IL: Illinois Natural History Survey.

Truco MJ, Hu J, Sadowski J, and Quiros CF, 1996. Inter- and intragenomic homology of the *Brassica* genomes: implications for their origin and evolution. Theor Appl Genet 93:1225–1233.

Tsunoda S, Hirata K, and Gomez-Campo C, 1980. *Brassica* crops and wild allies. Tokyo: Japan Scientific Societies Press.

Weigel D, 1995. The genetics of flower development—from floral induction to ovule morphogenesis. Annu Rev Genet 29:19–39.

Weigel D and Meyerowitz EM, 1994. The ABCs of floral homeotic genes. Cell 78:203–209.

Yanofsky MF, 1995. Floral meristems to floral organs—genes controlling early events in *Arabidopsis* flower development. Annu Rev Plant Physiol Plant Mol Biol 46: 167–188.

Received September 1, 1998
Accepted May 31, 1999

Corresponding Editor: James L. Hamrick