Copyright © 2000 by the Genetics Society of America

Variation and Selection at the *CAULIFLOWER* Floral Homeotic Gene Accompanying the Evolution of Domesticated *Brassica oleracea*

Michael D. Purugganan, Abee L. Boyles and Jane I. Suddith

Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695

Manuscript received November 11, 1999
Accepted for publication February 15, 2000

ABSTRACT

The evolution of plant morphologies during domestication events provides clues to the origin of crop species and the evolutionary genetics of structural diversification. The *CAULIFLOWER* gene, a floral regulatory locus, has been implicated in the cauliflower phenotype in both *Arabidopsis thaliana* and *Brassica oleracea*. Molecular population genetic analysis indicates that alleles carrying a nonsense mutation in exon 5 of the *B. oleracea CAULIFLOWER* (*BoCAL*) gene are segregating in both wild and domesticated *B. oleracea* subspecies. Alleles carrying this nonsense mutation are nearly fixed in *B. oleracea* ssp. *botrytis* (domestic cauliflower) and *B. oleracea* ssp. *italica* (broccoli), both of which show evolutionary modifications of inflorescence structures. Tests for selection indicate that the pattern of variation at this locus is consistent with positive selection at *BoCAL* in these two subspecies. This nonsense polymorphism, however, is also present in both *B. oleracea* ssp. *acephala* (kale) and *B. oleracea* ssp. *oleracea* (wild cabbage). These results indicate that specific alleles of *BoCAL* were selected by early farmers during the domestication of modified inflorescence structures in *B. oleracea*.

DOMESTICATED plant species provide excellent models to study and test hypotheses on the genetics and evolution of morphological diversification (Doebley 1992, 1993; Doebley *et al.* 1997; Doebley and Lukens 1998). The domestication of crop species is invariably accompanied by evolutionary changes in suites of structural traits that differentiate cultivated species from their wild relatives, or even between various crop subspecies (Schwanitz 1967; Doebley 1993). 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 (Gottlieb 1984; Doebley 1992).

One approach to understanding the evolutionary dynamics of morphological change focuses on identifying genes that underlie trait differences between domesticated and wild species and exploring the population genetics of these domestication loci. A molecular population genetic approach provides a particularly powerful framework for assessing how evolutionary forces act to shape variation at genetic loci and delineate mechanisms that may accompany evolutionary diversification during crop domestication. Indeed, studies on molecular diversity at the sequence level have begun to be utilized in inferring both the general population structure and history of crop domestication events (Eyre-Walker *et al.* 1998; Hilton and Gaut 1998; Olsen and Schaal 1999) and the selective forces acting at specific genes that underlie domestication and divergence in early agricultural crops (Hanson *et al.* 1996; Wang *et al.* 1999).

The distinct morphologies exhibited by *Brassica oleracea* subspecies represent one of the most spectacular illustrations of structural evolution in plants under domestication. *B. oleracea* is a perennial herb found largely in Europe and the Mediterranean (Tsunoda *et al.* 1980; Song *et al.* 1988; Kalloo and Bergh 1993) and is an extremely polymorphic species that includes at least six cultivated and one wild subspecies. Wild, perennial forms of *B. oleracea*, designated subspecies *oleracea* (wild cabbage), grow in coastal rocky cliffs of the Mediterranean, northern Spain, western France, and southern and southwestern Britain (Tsunoda *et al.* 1980). Selection for different characteristics during domestication, however, has resulted in extreme morphological divergence among cultivated subspecies. Of the six domesticated taxa, two subspecies, *B. oleracea* ssp. *botrytis* (cauliflower) and *B. oleracea* ssp. *italica* (broccoli), are characterized by the evolutionary modification of the inflorescence into large dense structures. The precociously large, undifferentiated inflorescence, termed the curd, is the defining characteristic of *B. oleracea* ssp. *botrytis*. The cauliflower curd consists of a dense mass of arrested inflorescence meristems, only ~10% of which will later develop into floral primordia and normal flowers (Sadik 1962).

The cauliflower phenotype characteristic of *B. oleracea* ssp. *botrytis* has been observed in mutants of the related crucifer *Arabidopsis thaliana* (Bowman *et al.* 1993; Weigel 1995; Yanofsky 1995). In Arabidopsis, the early

Corresponding author: Michael D. Purugganan, Department of Genetics, Box 7614, North Carolina State University, Raleigh, NC 27695. E-mail: michael_purugganan@ncsu.edu

acting floral meristem identity genes are a class of flower developmental regulatory loci that specify the identity of the floral meristem (as opposed to the inflorescence meristem) in developing reproductive primordia. Members of this class include the genes *APETALA1* (*AP1*; Mandel *et al.* 1992; Gustafson-Brown *et al.* 1994) and *CAULIFLOWER* (*CAL*; Kempin *et al.* 1995). Both the *APETALA1* and *CAULIFLOWER* loci have also been shown to control 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 develops, similar to the *B. oleracea* ssp. *botrytis* curd.

Genetic analyses in *B. oleracea* suggest the involvement of the *B. oleracea CAL* gene, referred to as *BoCAL*, in the formation of the altered inflorescence in *B. oleracea* ssp. *botrytis* (Kempin *et al.* 1995). It has been demonstrated already that the *BoCAL* allele in domesticated cauliflower has a premature termination codon at position 151 (E → stop; Kempin *et al.* 1995). This nonsense mutation appears to have arisen fairly recently within *B. oleracea*. In this article, we report that haplotypes carrying this polymorphism are fixed or nearly fixed in *B. oleracea* ssp. *botrytis* and *B. oleracea* ssp. *italica*, the two subspecies that have undergone selection for altered patterns of inflorescence development. Our tests for selection suggest that the *BoCAL* gene in *B. oleracea* ssp. *botrytis* and *B. oleracea* ssp. *italica* experienced a recent adaptive sweep, consistent with the evolution of the characteristic inflorescence structures in these subspecies. These results suggest that the floral regulatory gene *BoCAL* was one of the targets of selection during the evolutionary domestication of subspecies within the vegetable crop *B. oleracea*.

MATERIALS AND METHODS

Study species: The following *B. oleracea* subspecies were used in these analyses: *B. oleracea* ssp. *oleracea* (wild cabbage), *B. oleracea* ssp. *acephala* (kale), *B. oleracea* ssp. *botrytis* (cauliflower), and *B. oleracea* ssp. *italica* (broccoli). The wild relative *B. incana* was also utilized in this study. Seeds 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 *BoCAL* alleles: Genomic DNA from young *B. oleracea* leaves was isolated using the plant DNAEASY miniprep kit (QIAGEN, Chatsworth, CA). PCR was performed with an initial 10 cycles of 15 sec at 94°, 30 sec at 48°, and 2 min at 68° followed by 25 cycles with an incremental increase of 20 sec/cycle of the extension time. The error-correcting recombinant *Pwo* polymerase (Boehringer Mannheim, Mannheim, Germany) was used to minimize nucleotide misincorporation. The error rate for this polymerase, based on multiple amplification and resequencing of known genes, is similar to other error-correcting polymerases and is <1 in 7000 bp (Purugganan and Suddith 1999). We estimate that the nonsampling variance of nucleotide diversity due to PCR misincorporation, $\mathrm{Var}_{PCR}(\pi)$, is negligible [$\mathrm{Var}_{PCR}(\pi)/\mathrm{Var}(\pi) \sim 0.14$; J. I. Suddith and M. D. Purugganan, unpublished results]. The *BoCAL*-specific primers BoCALBSF2 (for intron 2 forward; 5′-TAATCATAGGCATTATCTGG-3′) and BoCALB3R (for exon 8 reverse; 5′-TGCAGTAAATGGGTT CAAAGTC-3′) were used in PCR reactions to amplify alleles from most *B. oleracea* accessions. For two *B. oleracea* ssp. *acephala* and one *B. incana* allele, the gene was isolated in two pieces; two additional internal primers (BoCALBSR2 [5′-CACCAAGA GTGTCGGATCTA-3′] and BoCALB2F [5′-GATGCACTGTTT ACATAATGAAAAT-3′]) were constructed in addition to BoCALBSF2 and BoCALB3R to isolate these alleles. Amplified DNA was cloned into pCR2.1 using the TA cloning kit (Invitrogen, Carlsbad, CA). DNA sequencing for both genes was conducted with the ABI377 automated sequencer using a series of nested internal sense and antisense primers. All sequence polymorphisms were visually rechecked from chromatograms, with special attention to low frequency polymorphisms (Hamblin and Aquadro 1997). The DNA sequences are available from GenBank (accession nos. AF241113–AF241150).

Data analysis: Sequences used in this study were visually aligned. Phylogenetic analyses were conducted using PAUP* 4.0d54 (maximum parsimony; Swofford 1993). The heuristic search algorithm was utilized using the tree bisection-reconnection procedure, with the *B. incana* orthologue as the outgroup. Node support is assessed with 500 bootstrap replicates of the data. The polymorphism data were analyzed using the SITES (Hey and Wakeley 1997) and DNASP programs (Rozas and Rozas 1997). Levels of nucleotide diversity were estimated as mean pairwise differences (π) and number of segregating sites (θ; Nei 1987). Identification of possible recombinants utilized the four-gamete test (Hudson and Kaplan 1985). The Tajima (1989) and Fu and Li (1993) tests for distribution of nucleotide polymorphisms were conducted without specifying an outgroup. These tests are known to have low power with small sample size; we thus pooled allelic data for subspecies *B. oleracea* ssp. *botrytis* and *B. oleracea* ssp. *italica*, and for *B. oleracea* ssp. *oleracea* and *B. oleracea* ssp. *acephala*. The former group includes those subspecies that show evolutionary alterations in inflorescence morphologies.

RESULTS AND DISCUSSION

Nucleotide variation at the *B. oleracea CAULIFLOWER* floral regulatory locus: We isolated alleles of the *BoCAL* gene from 37 worldwide accessions, representing four distinct subspecies. These include three domesticated subspecies that display differing reproductive or vegetative morphologies as a result of the selective pressures that accompanied domestication of this vegetable crop. Two subspecies, *B. oleracea* ssp. *botrytis* (cauliflower) and *B. oleracea* ssp. *italica* (broccoli), display altered inflorescence morphologies as a result of evolutionary divergence in reproductive developmental programs. The other domesticated study subspecies, *B. oleracea* ssp. *acephala*, shows the curling of leaf edges characteristic of kale, but otherwise produces a stereotypical Brassica raceme. Finally, *B. oleracea* ssp. *oleracea* accessions were included to represent the closest wild relatives of the domesticated subspecies; plants in this subspecies display no apparent changes in either vegetative or reproductive form. Most of the accessions utilized were from the Mediterranean and Northern Europe, where this

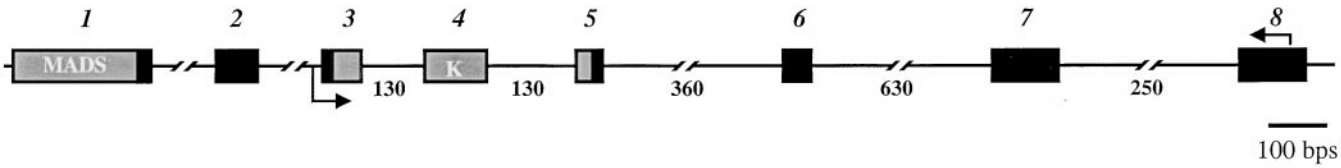

Figure 1.—Map of the *BoCAL* gene. Exons are shown as numbered boxes. Exon numbers are indicated in italics; intron sizes are also shown when known. The relative positions of the coding regions for the MADS- and K-boxes are given. Arrows depict positions of PCR primers used to isolate genomic sequences. Bar, 100 bp.

vegetable crop was believed to have originated and was cultivated historically. The *BoCAL* orthologue from the closely related congener *B. incana* was also isolated to provide an interspecific comparison of gene divergence.

Approximately 2.01 kb of the *BoCAL* gene was sequenced for each isolated allele; the sequenced region spans intron 2 to exon 8 (Figure 1) and includes the coding region for the moderately conserved K-domain of the BoCAL MADS-box transcriptional activator. The K-domain is believed to serve as a dimerization interface among MADS-box proteins. The sequenced region also encodes the C-terminal region as well as a portion of the linker I-region; the former is believed to contain the transcriptional activation domain of MADS-box proteins (Riechmann and Meyerowitz 1997).

Molecular analyses reveal a large amount of variation at the *BoCAL* locus (Figure 2 and Table 1). A total of 87 variant sites are present in these sampled alleles, of which 35 are nucleotide polymorphisms and 52 are from insertion / deletion (indel) changes of 1–12 bp in length. All of the indels are in introns. Of the 35 nucleotide polymorphisms found in *BoCAL*, 25 are located within introns while 10 are in coding regions. The coding region polymorphisms include 7 replacement and 3 silent site variants. The estimate of species-wide nucleotide diversity, π, for *BoCAL* is 0.0030, which is about half the value observed for the Arabidopsis *CAL* gene (Purugganan and Suddith 1998). The levels of nucleotide diversity at the *BoCAL* gene differ between *B. oleracea* subspecies (Table 1). Nucleotide diversity estimates within *B. oleracea* range from 0.0003 in *B. oleracea* ssp. *botrytis* to 0.0053 in the wild *B. oleracea* ssp. *oleracea*.

A total of 17 distinct *BoCAL* haplotypes are evident within *B. oleracea*. One of these haplotypes predominates in the sample and accounts for 16 of the 37 alleles (43%). Of the other haplotypes, 12 are singletons, while 3 are found twice and 1 haplotype is observed three times in the data. The genealogy of these alleles is shown in Figure 3. The phylogeny is the result of 500 bootstrap replicates under maximum parsimony, and a tree based on neighbor-joining analysis gives the same topology. There appear to be two major *BoCAL* allele classes within the sample: (i) class I alleles, which are found in one *B. oleracea* ssp. *italica* and two *B. oleracea* ssp. *oleracea* accessions; and (ii) class II alleles, which account for the majority of the observed alleles (92%). The two classes are differentiated by 28 fixed nucleotide differences (Figure 2), including three replacement changes. The two allele classes do not originate from different *BoCAL* genes. We have isolated all three *BoCAL* genes in the *B. oleracea* genome (A. L. Boyles, S. Halldorsdottir and M. D. Purugganan, unpublished results), and the alleles in this study originate from the one locus previously identified as responsible for the Brassica cauliflower phenotype (Kempin *et al.* 1995). One allele from *B. oleracea* ssp. *acephala* (accession HRI7556 from Ireland) appears to be the product of the recombination between class I and II alleles. Overall, the method of Hudson and Kaplan (1985) detects a total of two recombination events among alleles in the sampled *B. oleracea* accessions.

There is no discernible structuring of alleles along subspecific boundaries for *B. oleracea* ssp. *oleracea* and *B. oleracea* ssp. *acephala*. The gene genealogy indicates that alleles isolated from these two subspecies are interspersed in the genealogy; for example, some *B. oleracea* ssp. *acephala* alleles are more closely related to either *B. oleracea* ssp. *botrytis* or *B. oleracea* ssp. *italica* alleles than they are to those found in other kales (Figure 3). The genealogy does indicate a close relationship, however, between *B. oleracea* ssp. *botrytis* and *B. oleracea* ssp. *italica*. Except for one *B. oleracea* ssp. *italica* class I allele, the alleles in these two groups are all found in the same clade in the reconstructed genealogy.

A nonsense polymorphism is segregating in *B. oleracea* populations: Among the 10 coding region polymorphisms observed in *B. oleracea*, 6 result in amino acid replacements in the BoCAL protein encoded by specific alleles. Three of these replacements are singletons found in class II alleles, while three others differentiate the two allele classes observed in this species. A G → T transversion in exon 5 results in a replacement polymorphism (GAG → TAG) that changes a glutamic acid to a premature termination codon in position 151 of the encoded protein (Figure 2); this previously identified nonsense mutation results in a truncated protein that includes the DNA-binding MADS-box, the I-region, and a portion of the K-domain.

This nonsense polymorphism is present at moderate frequencies in the sampled alleles. Of the 37 *B. oleracea* alleles, 23 contain this premature stop codon (62%); all nonsense alleles are found in the class II haplotypes.

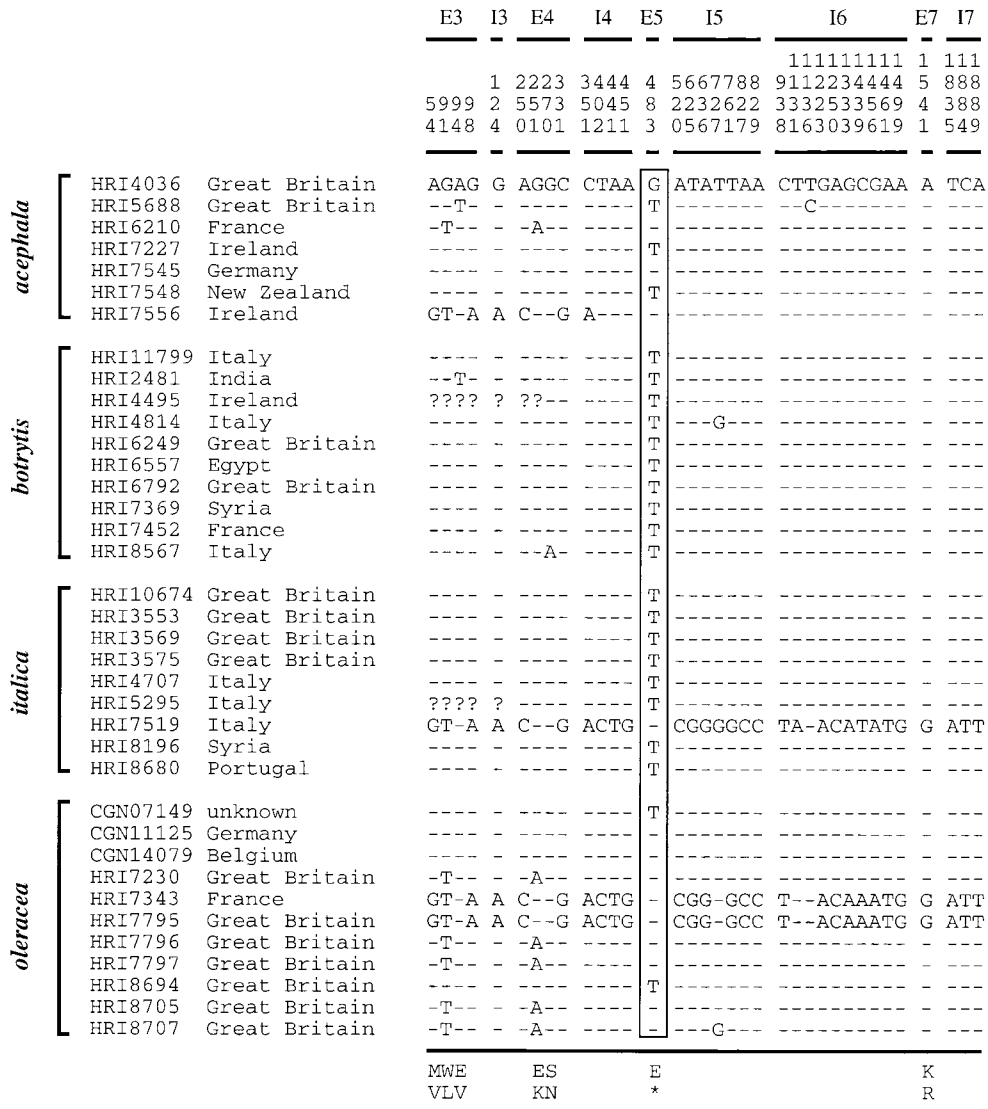

Figure 2.—Sequence of *BoCAL* alleles from different *B. oleracea* accessions. All allele sequences are compared to a reference allele from *B. oleracea* ssp. *acephala* (HRI4036). The alleles are grouped according to subspecies, indicated on the left. The position of the polymorphic sites and their locations in introns and exons are indicated at the top. The different amino acids encoded by replacement polymorphisms are shown below. The nonsense polymorphism leading to truncated BoCAL proteins is boxed. The question marks indicate missing data.

The nonsense mutation that gave rise to this polymorphism appears to be of fairly recent origin; all alleles that contain this substitution differ from each other by fewer than two nucleotide substitutions.

The nonsense polymorphism in *BoCAL* predominates, and indeed is close to fixation, in subspecies that have evolved altered inflorescence structures under domestication. All *B. oleracea* ssp. *botrytis* alleles contain this premature termination codon, and it is also found in 8 of the 9 alleles sampled from *B. oleracea* ssp. *italica*

TABLE 1

Nucleotide diversity of *BoCAL* in *B. oleracea*

Group	Common name	Modification	n^a	S^b	π	Θ
All	Cole crop	NA	37	35	0.0030	0.0043
oleracea	Wild cabbage	None	11	31	0.0053	0.0054
acephala	Kale	Leaf edges	7	11	0.0018	0.0023
botrytis	Cauliflower	Inflorescence	10	3	0.0003	0.0006
italica	Broccoli	Inflorescence	9	31	0.0035	0.0059
oleracea / *acephala*	NA	None / leaf	18	33	0.0040	0.0049
botrytis / *italica*	NA	Inflorescence	19	33	0.0018	0.0049

NA, not applicable.
[a] Number of sampled accessions.
[b] Number of segregating sites.

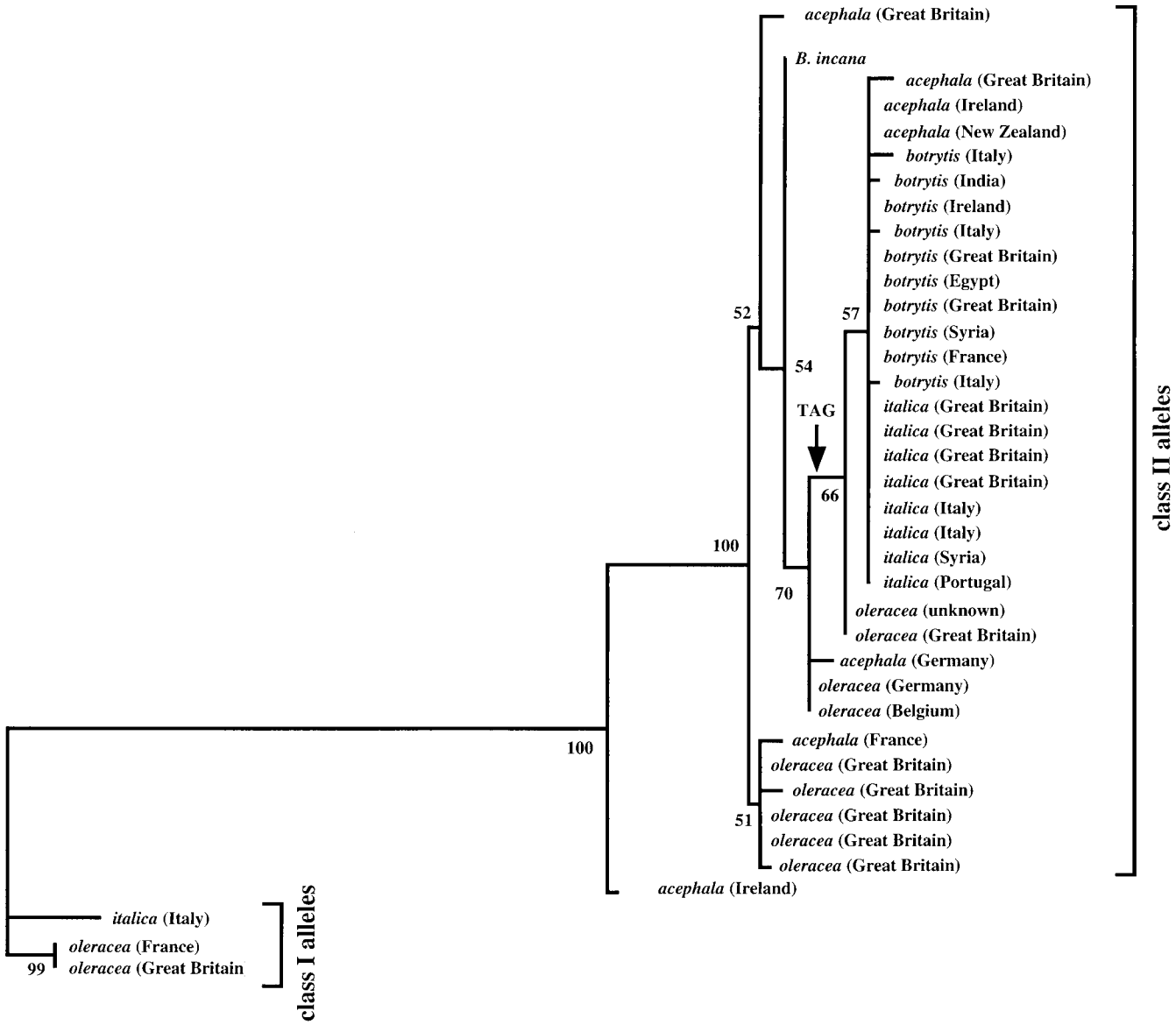

Figure 3.—Gene genealogy of *BoCAL* alleles. All nodes with <50% bootstrap support are collapsed; the other bootstrap values are indicated next to relevant nodes. Class I and II alleles for *BoCAL* are indicated. The arrow shows the probable branch associated with the origin of the nonsense mutation found in many *BoCAL* alleles. The total tree length is 101 steps, and the consistency index is 0.8713.

(95%). This nonsense polymorphism, however, is not confined to taxa that display altered inflorescence morphologies. This mutation is also found in 3 of the 7 *B. oleracea* ssp. *acephala* (43%) and 2 of the 11 *B. oleracea* ssp. *oleracea* alleles (18%; Figure 2). The widespread distribution of this polymorphism in *B. oleracea* subspecies suggests either that (i) it arose prior to the origin of cauliflower and broccoli or that (ii) it originated in *B. oleracea* ssp. *botrytis* and/or *B. oleracea* ssp. *italica*, but spread to other groups via hybridization. The frequency of this allele also suggests that there is no strong negative selection for this mutation in these subspecies.

Selective sweep at the *BoCAL* gene in *B. oleracea* ssp. *botrytis* and *B. oleracea* ssp. *italica* is associated with fixation of nonsense haplotypes: The extent and patterning of nucleotide variation along the *BoCAL* locus suggests that this regulatory gene is evolving in a nonneutral fashion. Specifically, it appears that alleles containing the nonsense polymorphism in exon 5 have been one of the targets of selection in subspecies displaying altered inflorescence morphologies as a result of domestication.

Allelic variation is expected to be reduced in a gene under selection (Aquadro 1997). Indeed, levels of molecular variation for this floral meristem identity gene are markedly reduced in those subspecies that show evolutionary alterations in inflorescence development. The value of π for the combined data from *B. oleracea* ssp. *botrytis* and *B. oleracea* ssp. *italica* alleles is less than

half of that estimated for *B. oleracea* ssp. *acephala* and *B. oleracea* ssp. *oleracea* ($\pi = 0.0018$ *vs.* 0.0040). *B. oleracea* ssp. *botrytis BoCAL* alleles are nearly identical to one another, with only four polymorphic nucleotide sites within the cultivated group. In *B. oleracea* ssp. *italica*, all of the variation is contributed by the presence in the sample of one class I allele; all the other alleles in this subspecies are identical to one another and all contain the nonsense mutation. The reduction in polymorphism at *BoCAL* within these two subspecies does not appear to be due to a population bottleneck during domestication. Both randomly amplified polymorphic DNA and isozyme studies indicate a significant level of polymorphism within these two subspecies at other molecular markers (Hu and Quiros 1991; Simonsen and Heneen 1995).

The action of historical adaptive sweeps in genes can also be detected by a number of tests for selection. Two tests, proposed by Tajima (1989) and Fu and Li (1993), compare the nucleotide diversity with the distribution of segregating sites expected under the neutral model of molecular evolution (Simonsen *et al.* 1995). Both tests reveal that the *BoCAL* gene is not evolving according to the predictions of the neutral model, and the pattern of nucleotide variation in this regulatory locus is consistent with a hypothesis of positive selection within some *B. oleracea* subspecies. Specifically, there is evidence of selection in subspecies that display altered inflorescence morphologies. In the combined alleles of *BoCAL* from *B. oleracea* ssp. *botrytis* and *B. oleracea* ssp. *italica*, the skewness in the frequency distribution of polymorphisms is significant in both tests. The Tajima test statistic D is -2.4418 ($P < 0.001$) for the combined alleles in these two subspecies; the negative value of the D statistic indicates that sampled alleles have an excess of low-frequency nucleotide polymorphisms over that expected in a neutrally evolving population. The excess of rare polymorphisms in these subspecies is due primarily to the presence of the single, divergent class I allele in *B. oleracea* ssp. *italica*. The Fu and Li test statistic D^* is also significantly negative for this gene ($D^* = -3.56997$, $P < 0.02$). In contrast, results of both the Tajima ($D = -0.7070$, $P > 0.10$) and Fu and Li tests ($D^* = 1.1645$, $P > 0.10$) with the combined alleles in both *B. oleracea* ssp. *acephala* and *B. oleracea* ssp. *oleracea* reveal that these genes are evolving according to the predictions of the equilibrium-neutral theory.

Molecular population genetics of regulatory genes associated with evolution under crop domestication: A comprehensive understanding of the process by which plant morphologies evolve under domestication requires us to (i) isolate genes that were the targets of selection by early agriculturalists and (ii) dissect the evolution of these domestication loci. Understanding the molecular genetics of a developmental system allows us to identify candidate genes and gene-gene interactions that may be the focus of selection during the process of morphological diversification. Subsequent studies on the molecular population genetics of morphological loci can then provide us with crucial information on the origin, history, and evolutionary forces that underlie the transformation in plant morphologies that accompany crop domestication events.

We have focused our attention on the genes that underlie the transformation in inflorescence morphologies observed in some subspecies within *B. oleracea*. Specifically, it has been suggested that the presence of a nonsense mutation at position 151 of the *BoCAL* floral regulatory locus is responsible in part for the evolution of the cauliflower curd in *B. oleracea* ssp. *botrytis* (Kempin *et al.* 1995). Alleles containing this nonsense mutation are expected to produce proteins of 150 amino acids in length (as compared to the 254-amino-acid full-length protein), which are truncated in the middle of the K-domain of the encoded MADS-box transcriptional activator.

A survey of nucleotide variation at *BoCAL* in four subspecies of *B. oleracea* indicates that this nonsense polymorphism appears to have originated once in this species and that alleles containing this mutation are close to fixation in groups that display alterations in inflorescence morphology. In *B. oleracea* ssp. *botrytis*, all alleles isolated contain this nonsense polymorphism, while only one *B. oleracea* ssp. *italica* allele did not have this premature stop codon. Based on tests of selection, the near fixation of nonsense haplotypes in *B. oleracea* ssp. *botrytis* and *B. oleracea* ssp. *italica* is consistent with a selective sweep that, based on genetic studies in *A. thaliana*, is likely associated with the evolution of the altered inflorescence morphologies in these cultivated groups.

Although previous studies in *A. thaliana* strongly implicate mutations at the *BoCAL* gene in the cauliflower phenotype, it is possible that mutations at this locus may play a role in the evolution of broccoli as well. Indeed, at least one naturally occurring allele in the *A. thaliana CAL* gene appears to produce a high density of floral meristems reminiscent of that seen in domestic *B. oleracea* ssp. *italica* (Purugganan and Suddith 1998). The presence of a highly divergent haplotype in *B. oleracea* ssp. *italica* that does not contain the nonsense polymorphism does suggest, however, that this mutation is not necessary for the formation of the broccoli phenotype. This could suggest that other mutations at this or related genes may also be associated with the evolution of the distinct inflorescence phenotype of *B. oleracea* ssp. *italica* (Kempin *et al.* 1995; Lowman and Purugganan 1999). Moreover, there is a wide range of variation to the density, size, and degree of floral differentiation between different cultivars of *B. oleracea* ssp. *italica*, and this may reflect the greater variation observed at the *BoCAL* locus within this subspecies.

It is also clear from our analyses that the nonsense mutation at the *BoCAL* locus is not sufficient to condi-

tion the cauliflower phenotype in *B. oleracea*. This nonsence polymorphism is also found at moderate frequencies in both *B. oleracea* ssp. *oleracea* and *B. oleracea* ssp. *acephala*, two groups that produce normal inflorescences. Genetic studies in *A. thaliana* indicate that mutations at both the *CAL* and *AP1* floral meristem identity genes are necessary to produce the cauliflower phenotype in this model plant (Bowman *et al.* 1993; Kempin *et al.* 1995). The *AP1* orthologues in *B. oleracea* have been identified and exist in at least two copies (*BoAP1-A* and *-B*), and no mutation in either copy is clearly associated with either *B. oleracea* ssp. *botrytis* or *B. oleracea* ssp. *italica* (Carr and Irish 1997; Lowman and Purugganan 1999). Although inheritance in *B. oleracea* is disomic, comparative gene mapping studies indicate that this species is a polyploid with two to three copies of each genetic locus (Bohuon *et al.* 1998). It may be that mutations at another, as yet unidentified *BoAP1* gene, may act in concert with the nonsense mutation at *BoCAL* to condition the altered inflorescence development observed in *B. oleracea* ssp. *botrytis*. It is also possible that polymorphisms at other *BoCAL* gene copies may be involved in the cauliflower phenotype. Indeed, we have identified two other *BoCAL* copies that appear to have arisen as a result of the ancient polyploidization event that led to the present-day *B. oleracea* genome (A. C. Lowman, S. Hallsdordottir and M. D. Purugganan, unpublished results). Moreover, the presence of this nonsense polymorphism at moderate frequencies in *B. oleracea* ssp. *oleracea* and *B. oleracea* ssp. *acephala* suggests that negative selection is not acting strongly at this locus and that these *BoCAL* copies may be genetically redundant to one another.

The molecular population genetics of developmental loci that control morphological traits are poorly understood, particularly within crop species groups that exhibit marked morphological divergence as a result of domestication. There has been recent work on the population genetics of loci such as *teosinte-branched1* (*tb1*; Wang *et al.* 1999) and *C1* (Hanson *et al.* 1996), both of which control traits that differ between Zea subspecies. It has been suggested that variation at promoter regions in these and other developmental control genes may be responsible for the evolution of plant developmental patterns (Doebley and Lukens 1998), and there is indeed evidence that selection at the *tba1* promoter may have been associated with maize evolution (Wang *et al.* 1999). Our study suggests, however, that evolutionary changes in regulatory proteins themselves may also permit the diversification of plant morphologies. Although the altered infloresences of subspecies within *B. oleracea* may be extreme, they nonetheless illustrate the possible role that regulatory proteins may play during crop domestication. The study of the molecular population genetics of these loci provides insights into the extent to which variation at these key regulatory loci accompany morphological divergence in crop plant species.

The authors thank S. Halldorsdottir and A. W. Womack for technical assistance and J. McFerson and D. Astley for providing seed. This work was supported in part by a grant from the United States Department of Agriculture National Research Initiative Competitive Grants program and an Alfred P. Sloan Foundation Young Investigator Award to M.D.P.

LITERATURE CITED

Aquadro, C. F., 1997 Insights into the evolutionary process from patterns of DNA sequence variability. Curr. Opin. Genet. Dev. **7:** 835–840.

Bohuon, E. J., L. D. Ramsay, J. Craft, A. Arthur, D. F. Marshall *et al.*, 1998 The association of flowering time QTL with duplicated regions and candidate loci in *Brassica oleracea*. Genetics **150:** 393–401.

Bowman, J. L., J. Alvarez, D. Weigel, E. M. Meyerowitz and D. Smyth, 1993 Control of flower development in *A. thaliana* by *AP1* and interacting genes. Development **119:** 721–743.

Carr, S. M., and V. F. Irish, 1997 Floral homeotic gene expression defines developmental arrest stages in *Brassica oleracea L.* vars. botrytis and italica. Planta **201:** 179–188.

Doebley, J., 1992 Mapping the genes that made maize. Trends Genet. **8:** 203–207.

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

Doebley, J., and L. Lukens, 1998 Transcriptional regulation and the evolution of plant form. Plant Cell **10:** 1075–1082.

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

Eyre-Walker, A., R. L. Gaut, H. Hilton, D. L. Feldman and B. S. Gaut, 1998 Investigation of the bottleneck leading to the domestication of maize. Proc. Natl. Acad. Sci. USA **95:** 4441–4446.

Fu, Y.-X., and W.-H. Li, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 693–709.

Gottlieb, L. D., 1984 Genetics and morphological evolution in plants. Am. Nat. **123:** 681–709.

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

Hamblin, M. T., and C. F. Aquadro, 1997 Contrasting patterns of nucleotide variation at the *Gld* locus in different populations of *Drosophila melanogaster*. Genetics **145:** 1053–1062.

Hanson, M. A., B. S. Gaut, A. Stec, S. I. Fuerstenberg, M. I. Goodman *et al.*, 1996 Evolution of anthocyanin biosynthesis in maize kernels: the role of regulatory and enzymatic loci. Genetics **143:** 1398–1407.

Hey, J., and J. Wakeley, 1997 A coalescent estimator of the population recombination rate. Genetics **145:** 833–846.

Hilton, H., and B. S. Gaut, 1998 Speciation and domestication in maize and its wild relatives: evidence from the *Glb-1* gene. Genetics **150:** 863–872.

Hu, J., and C. Quiros, 1991 Identification of broccoli and cauliflower cultivars with RAPD markers. Plant Cell Rep. **10:** 505–511.

Hudson, R. R., and N. Kaplan, 1985 Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics **111:** 147–164.

Kalloo, G., and B. O. Bergh, 1993 *Genetic Improvement of Vegetable Crops*. Pergamon Press, Oxford.

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

Lowman, A. C., and M. D. Purugganan, 1999 Duplication of the *Brassica oleracea APETALA1* floral homeotic gene and the evolution of domesticated cauliflower. J. Hered. **90:** 514–520.

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

Nei, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Olsen, K. M., and B. Schaal, 1999 Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. Proc. Natl. Acad. Sci. USA **96:** 5586–5591.

Purugganan, M. D., and J. I. Suddith, 1998 Molecular population genetics of the *Arabidopsis CAULIFLOWER* regulatory gene: non-

neutral evolution and naturally-occurring variation in floral homeotic function. Proc. Natl. Acad. Sci. USA **95:** 8130–8134.

Purugganan, M. D., and J. I. Suddith, 1999 Molecular population genetics of floral homeotic loci: departures from the equilibrium-neutral model at the *AP3* and *PI* genes of *Arabidopsis thaliana*. Genetics **151:** 839–848.

Riechmann, J. L., and E. M. Meyerowitz, 1997 MADS-domain proteins in plant development. Biol. Chem. **378:** 1079–1101.

Rozas, J., and R. Rozas, 1997 DNASP2.0: a novel software package for extensive molecular population genetic analysis. Comput. Appl. Biosci. **13:** 307–311.

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

Schwanitz, F., 1967 *Die Evolution der Kulturpflanzen*. Bayerischer Landwirtschaftverlag, Munich.

Simonsen, K. L., G. Churchill and C. Aquadro, 1995 Properties of statistical tests of neutrality for DNA polymorphism data. Genetics **141:** 413–429.

Simonsen, V., and W. K. Heneen, 1995 Genetic variation within and among different cultivars and landraces of *B. campestris* and *B. oleracea* based on isozymes. Theor. Appl. Genet. **91:** 346–352.

Song, K., T. C. Osborn and P. H. Williams, 1988 Brassica taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs). Theor. Appl. Genet. **79:** 497–506.

Swofford, D. L., 1993 *Phylogenetic Analysis Using Parsimony (PAUP) Version 4.0d64.* University of Illinois, Champaign, IL.

Tajima, F., 1989 Statistical methods for testing the neutral mutation hypothesis by DNA polymorphisms. Genetics **123:** 585–595.

Tsunoda, S., K. Hirata and C. Gomez-Campo, 1980 *Brassica Crops and Wild Allies.* Japan Scientific Societies Press, Tokyo.

Wang, R. L., A. Stec, J. Hey, L. Lukens and J. Doebley, 1999 The limits of selection in maize. Nature **398:** 236–239.

Weigel, D., 1995 The genetics of flower development. Annu. Rev. Genet. **29:** 19–39.

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

Communicating editor: M. K. Uyenoyama