Naturally Occurring Indel Variation in the *Brassica nigra COL1* Gene Is Associated With Variation in Flowering Time

Marita Kruskopf Österberg,*,† Oksana Shavorskaya,* Martin Lascoux† and Ulf Lagercrantz*,[1]

**Department of Plant Biology, Swedish University of Agricultural Sciences, S-750 07 Uppsala, Sweden and †Department of Conservation Biology and Genetics, Evolutionary Biology Center, Uppsala University, 752 36 Uppsala, Sweden*

Manuscript received September 24, 2001
Accepted for publication February 1, 2002

ABSTRACT

Previous QTL mapping identified a *Brassica nigra* homolog to *Arabidopsis thaliana CO* as a candidate gene affecting flowering time in *B. nigra*. Transformation of an *A. thaliana co* mutant with two different alleles of the *B. nigra CO* (Bni *COa*) homolog, one from an early-flowering *B. nigra* plant and one from a late one, did not show any differential effect of the two alleles on flowering time. The DNA sequence of the coding region of the two alleles was also identical, showing that nucleotide variation influencing flowering time must reside outside the coding region of Bni *COa*. In contrast, the nucleotide sequence of the *B. nigra COL1* (Bni *COL1*) gene located 3.5 kb upstream of Bni *COa* was highly diverged between the alleles from early and late plants. One indel polymorphism in the Bni *COL1* coding region, present in several natural populations of *B. nigra*, displayed a significant association with flowering time within a majority of these populations. These data indicate that a quantitative trait nucleotide (QTN) affecting flowering time is located within or close to the Bni *COL1* gene. The intergenic sequence between Bni *COL1* and Bni *COa* displayed a prominent peak of divergence 1 kb downstream of the Bni *COL1* coding region. This region could contain regulatory elements for the downstream Bni *COa* gene. Our data suggest that a naturally occurring QTN for flowering time affects the function or expression of either Bni *COL1* or Bni *COa*.

ADAPTIVE traits typically show quantitative phenotypic variation, and it is often assumed that this variation is shaped largely by natural selection. While there is mounting evidence that natural selection was, at least in part, involved in the evolution of adaptive traits (Kreitman and Akashi 1995), the relative importance and nature of its contribution is still a matter of debate. Because most quantitative traits are controlled by a large number of interacting loci, it is conceivable that selection intensity at each of these loci is limited, most of the variation observed within and among natural populations coming from variation among loci rather than at loci themselves (*e.g.*, Latta 1998). Polymorphism in natural populations could have also resulted mainly from selection at a few regulatory genes, variation at structural genes being predominantly neutral (Purugganan 2000). Unfortunately, there are still limited data linking the phenotypic variation of adaptively important quantitative traits to variation of underlying genes, because most studies of molecular variation so far have concentrated on a few genes.

Clearly, flowering time has a high adaptive value as reproductive success in plants closely depends on the timing of flower initiation. A correct timing ensures that flowers develop at a most favorable time for fertilization and consequently leads to a sufficient production of seeds. The process probably also includes trade-offs between timing of flowering and importance of seed production (Dorn and Mitchell-Olds 1991). Flowering time exhibits considerable variation between, but also within, natural populations and a significant part of this variation has a genetic basis (for variation between populations, see Kowalski *et al.* 1994; Clarke *et al.* 1995; Mitchell-Olds 1996; Kuittinen *et al.* 1997; Alonso-Blanco *et al.* 1998; for variation within populations, see Pors and Werner 1989; Fu and Ritland 1994).

Much progress has been made lately in understanding the genetic control of flowering time in the cruciferous model species *Arabidopsis thaliana* (Coupland 1995; Simpson *et al.* 1999; Reeves and Coupland 2000). A substantial number of genes affecting flowering time have been identified through mutagenesis studies and utilization of natural variation among ecotypes. The identification of a pathway and its underlying genes through mutagenesis studies does not in itself show that phenotypic variation in the wild is due to variation at these loci, although the nucleotide variation at genes involved in the response to environmental cues is a likely target for adaptive selection. The information on *A. thaliana* provides a large number of candidate genes for the control of flowering time in the wild and it would be of great interest to see which of the many possible

[1]*Corresponding author:* Department of Plant Biology, Swedish University of Agricultural Sciences, Box 7080, S-750 07 Uppsala, Sweden. E-mail: ulf.lagercrantz@vbiol.slu.se

alternatives has been targeted by natural selection in different species and/or populations within species. Recently, the *A. thaliana FRIGIDA* (*FRI*) gene was shown to be a major determinant of flowering time variation in natural populations (Johanson *et al.* 2000). Without vernalization, functional *FRI* alleles confer late flowering, and most early-flowering ecotypes carry loss-of-function mutations at *FRI*.

We have taken a comparative mapping approach to try to identify quantitative trait nucleotides (QTNs) affecting flowering time (FT) in natural populations of Brassica species. Brassica species belong to the same family as *A. thaliana* and the induction of flowering shows great similarities with *A. thaliana* (Tommey and Evans 1991). Species in both genera are induced to flower by long days and respond in a similar way to vernalization. The first step in our approach is to map quantitative trait loci (QTL) for flowering time in Brassica species. A QTL is simply a chromosome location, and to identify candidate genes within this area, comparative mapping is then used to identify homologous chromosome segments in *A. thaliana*. These segments are searched for candidate genes on the basis of map position and knowledge about gene function in Arabidopsis. Finally, the effects on flowering time of natural variation in the Brassica homologs are specifically tested.

In the present case, previous QTL mapping in *Brassica nigra* identified two main genomic areas influencing flowering time (Lagercrantz *et al.* 1996). On the basis of comparative mapping, a *B. nigra* homolog to the Arabidopsis gene *CONSTANS* (*CO*) was suggested as a likely candidate gene for at least one of the QTL. *CO* acts in the pathway that accelerates flowering in response to long photoperiods (Putterill *et al.* 1995). The gene activates at least four early target genes with diverse biochemical function that act to promote flowering, making *CO* a key component in the regulation of flowering time in response to the environment (Samach *et al.* 2000).

As a major QTL for flowering time in *B. nigra* maps close to a *CO* homolog, we tested whether natural variation at this homolog affected flowering time. In the present study, we cloned alleles of the *B. nigra CO* homolog at the major QTL (Bni *COa*) from early- and late-flowering plants. To test any differential effect of the two alleles on flowering time, they were transformed into the *A. thaliana co* mutant. As no such effect was detected, we examined allelic variation around Bni *COa*. Notably, we analyzed variation at the *B. nigra COL1* gene, which is located 3.5 kb upstream of Bni *COa*. *COL1* displayed highly diverged alleles from early- and late-flowering plants. Studies of association between this allelic variation and flowering time were conducted in natural populations. These studies detected a significant association in several *B. nigra* populations, indicating that nucleotide variation within or close to *COL1* affects flowering time.

MATERIALS AND METHODS

Isolation of *B. nigra CO* and *COL1* alleles: A cDNA of the *A. thaliana CO* gene (Putterill *et al.* 1995) was used to screen a *B. nigra* genomic library in λEMBL3. The library was prepared from the rapid cycling line used as the early-flowering parent in our previous QTL mapping experiments (Lagercrantz *et al.* 1996). Hybridizations were performed at moderate stringency. More than 80 clones were isolated and characterized. Restriction mapping and hybridization experiments identified four different genes. One representative of each gene was subcloned and sequenced using cycle sequencing and an automatic sequencer (ABI377). The four genes were positioned on the *B. nigra* linkage map using a previously described mapping population (Lagercrantz and Lydiate 1995). A genomic library in λZap was then prepared from the late-flowering parent used to construct our mapping populations (Lagercrantz *et al.* 1996). This library was screened with the previously isolated *B. nigra* genes. Subcloned alleles were sequenced as above.

Introduction of *B. nigra CO* alleles into Arabidopsis: Alleles of Bni *COa* were cloned in pPZP211 (Hajdukiewicz *et al.* 1994). The constructs contained genomic clones comprising the coding region with a 2-kb upstream sequence and a 1-kb downstream sequence. The plasmids were transformed into Agrobacterium strain GV3101 that was used to vacuum infiltrate an *A. thaliana co* mutant (*co-2*; Nottingham stock center no. N175). Offspring from homozygous transformants were grown in controlled-environment rooms with a 16-hr photoperiod. Flowering time was measured as the number of days from sowing to the opening of the first flower, and as the number of rosette leaves after bolting. The means were calculated from 17 to 25 plants. Five independent transformants with a Bni *COa* allele from an early-flowering plant and six with an allele from a late-flowering plant were tested.

Population samples: Seed samples (population samples) originating from Ethiopia (accession no. BRA1163), Spain (accession no. BRA101), Portugal (accession no. BRA153), Italy (samples 1, 2, and 3: accession nos. BRA1045, BRA26, and BRA27), Greece (accession no. BRA185), and Germany (accession no. BRA1045) (Figure 1) were obtained from the Institute für Plantzengenetik und Kultur (Gatersleben, Germany).

Flowering time experiments: The seeds were germinated 24 hr in petri dishes and planted in pots. In the first experiment conducted with the two Italian populations (Ital 2 and Ital 3), the plants were grown under artificial light with a 16-hr photoperiod. In the second experiment plants were grown in the greenhouse with supplementary 16 hr light. Flowering time was scored as the number of days from planting to the opening of the first flower.

Marker analysis: Genomic DNA was prepared from leaf samples as described by Liscum and Oeller (1997). Two indels (*Ind1* and *Ind2*) in *COL1*, identified comparing alleles from the early- and late-flowering parents in our mapping cross, were amplified separately using PCR. The primers were for *Ind1* CO34 (5′ AGA AGA TGA AGC AGA GGC) and CO56 (5′ ACT GTA ATC GAC AAG GTC CAG) and for *Ind2* CO57 (5′ CTG GAC CTT GTC GAT TAC AGT) and CO58 (5′ GAC TCT GGC ACA AGA CTA ACC). Primers CO34 and CO57 were labeled with 6-FAM and HEX, respectively, and run on an ABI377 automatic sequencer.

Estimates and tests of population genetics parameters were performed using Arlequin (Schneider *et al.* 2000). Linkage disequilibrium was tested using a likelihood-ratio test, whose empirical distribution is obtained by a permutation procedure (Slatkin and Excoffier 1996). The association between flowering time and marker genotype was tested within each

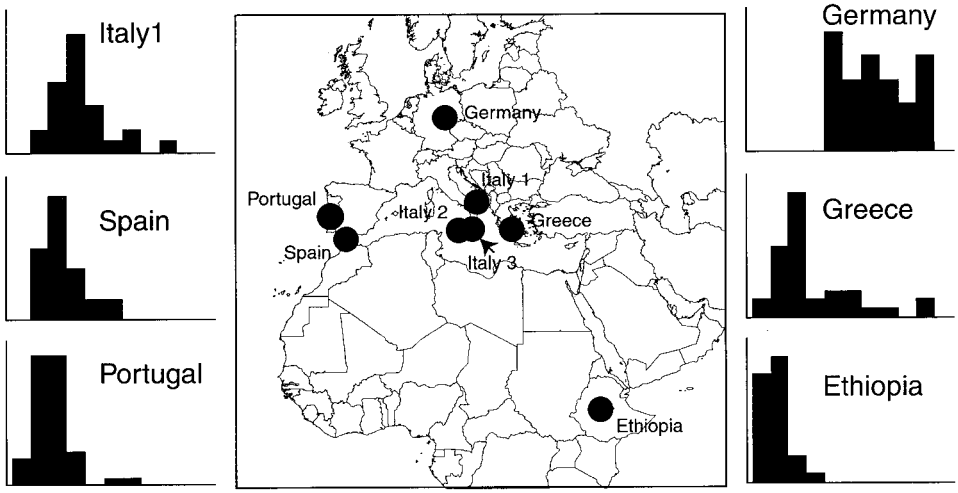

FIGURE 1.—Geographic locations of the original seed collections of *B. nigra* and distributions of flowering time from one common garden experiment.

population using analysis of variance in Statview 4.0 (Abacus Concepts, Berkeley, CA). The analyses were performed within each population as well as with all populations in a nested model with country of origin and genotype with country as factors.

RESULTS

To test the hypothesis that allelic variation in a *B. nigra* homolog to *CO* influences flowering time in natural populations, *CO* homologs were isolated from *B. nigra*. Our screens identified two *CO* homologs, Bni *COa* and Bni *COb*, corresponding to the two *CO* loci previously mapped to QTL on linkage groups 2 (LG2) and 8 (LG8), respectively (LAGERCRANTZ *et al.* 1996). In addition, one homolog each of *A. thaliana COL1* (Bni *COL1*) and *COL2* (Bni *COL2*) was identified in our screens.

We concentrated on Bni *COa* at the major QTL on LG2 and isolated alleles from the late- and early-flowering parents in our mapping cross. The DNA sequences of the coding region and 600 bp of the 5′ untranslated region of the two alleles of Bni *COa* were virtually identical in the two alleles. In total, two nucleotide substitutions were detected at positions −268 and −466 from the translation start site. To test if variation in the flanking regions or the intron might affect flowering time, we introduced the two alleles into an Arabidopsis *co* mutant (*co-2*). Both alleles were functional and contained enough regulatory sequence to restore early flowering in the *co* mutant (Table 1). However, no significant difference in flowering time was detected between alleles from early- and late-flowering *B. nigra* plants.

Thus, if diversity between the two *B. nigra CO* alleles affects flowering time variation, important nucleotide variation is likely to reside farther away from the coding region. A gene with high sequence similarity to Bni *COa* was found 3.5 kb upstream of *COa*. This gene, Bni *COL1*, is a homolog to *A. thaliana CONSTANS LIKE 1* (*COL1*; PUTTERILL *et al.* 1997), which is located 3.5 kb upstream of *A. thaliana CO*. Considering the near-identical alleles at Bni *COa*, alleles of Bni *COL1* from early- and late-flowering plants exhibited surprising sequence divergence (Figure 2). The coding region in the two alleles differed by 16 nucleotide substitutions and two in-frame indels (*Ind1* and *Ind2*) separated by 235 bp. *Ind1* is part of trinucleotide AAC repeat coding for a run of asparagine residues, and the allele from the early plant contained six additional repeat units. The 18-nucleotides-longer *Ind1* was associated with a deletion of 18 nucleotides at *Ind2* in the "early" allele. Most of the nucleotide substitutions (9) were nonsynonymous, resulting in amino acid substitutions that were concentrated around the indels and the C-terminal end of the

TABLE 1

Flowering time in progeny of homozygous *A. thaliana co-2* plants transformed with *COa* alleles from early- and late-flowering *B. nigra* plants

Transformant lines	Average days to flowering[a]	Average no. of leaves[a,b]
Early *B. nigra* allele		
TE1	15.3 ± 0.3	4.0 ± 0.1
TE2	17.4 ± 0.7	5.2 ± 0.1
TE3	16.3 ± 0.4	4.5 ± 0.2
TE4	16.8 ± 0.2	4.6 ± 0.1
TE5	17.1 ± 0.3	5.1 ± 0.1
Late *B. nigra* allele		
TL1	15.7 ± 0.4	4.3 ± 0.1
TL2	17.1 ± 0.3	4.8 ± 0.2
TL3	17.6 ± 0.8	5.2 ± 0.1
TL4	16.3 ± 0.3	4.4 ± 0.1
TL5	17.8 ± 0.9	5.3 ± 0.2
TL6	15.6 ± 0.6	4.2 ± 0.1
Controls		
co-2	30.3 ± 0.9	17.3 ± 0.9
Landsberg erecta	18.3 ± 0.8	5.2 ± 0.1

[a] Average ± standard error (n = 17–25 plants).
[b] Number of rosette leaves counted after bolting.

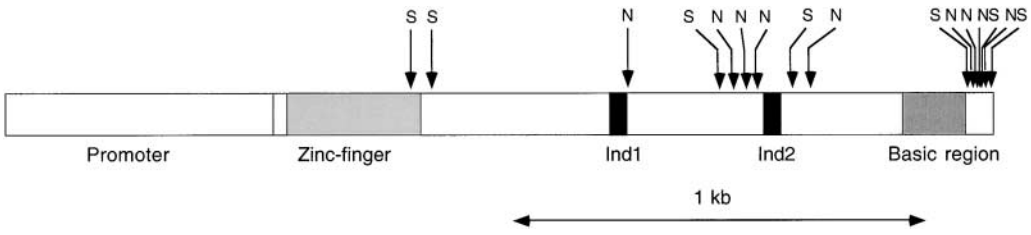

FIGURE 2.—A schematic view of the *B. nigra COL1* gene indicating nucleotide substitutions between alleles from an early- and a late-flowering *B. nigra* plant. N indicates nonsynonymous substitutions, and S denotes synonymous ones. The zinc-finger and basic regions are two highly conserved motifs in *COL1* genes (LAGERCRANTZ and AXELSSON 2000).

protein (Figure 2). The deduced amino acid variation was thus located exclusively outside the two highly conserved motifs present in *COL1* genes (Figure 2; LAGERCRANTZ and AXELSSON 2000).

Sequence analysis of *COL1* in *A. thaliana* and other Brassica species indicates that the *Ind2* polymorphism present in *B. nigra* predates the divergence of the lineage leading to *B. nigra* from the one leading to *B. rapa* and *B. oleracea*. The insertion was present in *A. thaliana* but absent in alleles sampled from *B. oleracea*, *B. rapa*, and *B. juncea* (Figure 3). A deletion of the extra 18 bp present in *A. thaliana* and some *B. nigra* alleles could have occurred independently in the lineages leading to *B. nigra*, *B. rapa*, and *B. oleracea*, but a single deletion event before the split of the *B. nigra* and *B. rapa*/*B. oleracea* lineages seems a more parsimonious explanation.

Polymorphism at *Ind1* and *Ind2*: The two indels in *COL1* were genotyped in seven *B. nigra* populations originating from Europe and Africa. At *Ind2*, only the two previously identified alleles (*L* and *S*) were segregating, while in total six alleles were found at the microsatellite locus *Ind1* (Table 2). The populations were highly differentiated, in particular at *Ind1*, which displayed an F_{ST} of 45%, while the F_{ST} estimate for *Ind2* was 20%. Within populations, no significant departure from Hardy-Weinberg equilibrium was detected (data not shown). Linkage disequilibrium was detected between *Ind1* and *Ind2* in four of the five populations where both loci were polymorphic (Table 2).

Flowering time variation: Flowering time for the plants in the seven populations was measured in two separate experiments with different environmental conditions. Thus, flowering time cannot be compared between the experiments. However, plants from six populations grown under identical long-day conditions showed significant differences in flowering time (Figure 1 and Table 2). Plants from the most southern population, Ethiopia, flowered the earliest while plants from the most northern one, Germany, flowered the latest. Plants from Spain, Portugal, Italy, and Greece displayed intermediate flowering times.

Association between flowering time and indel polymorphisms in *COL1*: Due to strong population structure, both for indel genotypes and FT, all associations between genotype and FT were tested within populations. Allelic variation at *Ind1* was observed in six populations; however, no significant association between the *Ind1* genotype FT was seen in any of these populations nor in the nested model (Table 3). At *Ind2*, variation was detected in five populations, and in four of those, FT was significantly different among genotypes (Figure 4 and Table 3). The overall effect of genotype within country was also highly significant in the nested model. In all cases, the *S* (short) allele was associated with earlier flowering. In accordance with these data, the two populations fixed for the *S* allele (Ethiopia and Portugal) flowered earlier than any of the other populations (Figure 1 and Table 2).

Genetic diversity in the intergenic region between *COL1* and *COa*: The alleles of COL1 from early- and late-flowering plants were highly diverged, in particular toward the C-terminal part, while the corresponding alleles of *COa* 3.5 kb downstream were virtually identical. We thus wanted to check if the allelic diversity also extended into the noncoding region between the two genes. A comparison of the 3.5-kb intergenic region

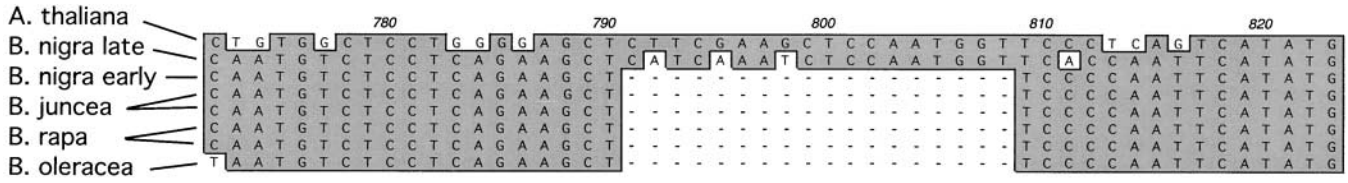

FIGURE 3.—Alignment of a part of the *COL1* gene around the *B. nigra Ind2* indel from *A. thaliana*, *B. juncea*, *B. rapa*, and *B. oleracea*.

TABLE 2

Allele frequencies and linkage disequilibrium at two indels (*Ind1* and *Ind2*) in the *B. nigra COL1* gene and average flowering time in natural populations

		Ind1						*Ind2*			
Population	n_a	*134*	*137*	*140*	*143*	*147*	*153*	*S*	LD[a]	n_{FT}	FT
Germany	134	0.99	0	0.01	0	0	0	0.80	0.04	61	21.7 ± 0.6
Greece	82	0.82	0	0.07	0	0	0.10	0.65	0.01	40	12.8 ± 1.3
Italy 1	0									26	12.6 ± 1.0
Spain	56	0.32	0.09	0	0.21	0	0.38	0.86	0.01	28	9.6 ± 0.6
Portugal	58	0.68	0	0.27	0	0	0.05	1		51	8.2 ± 0.4
Ethiopia	32	0	0	0	0	1	0	1		30	5.2 ± 0.4
Italy 2	40	0.88	0.05	0.02	0	0	0.05	0.28	0.05		NC
Italy 3	46	0.93	0.04	0	0	0	0.02	0.13	0.30		NC

n_a, number of alleles scored at *Ind1* and *Ind2*; n_{FT}, number of plants scored for flowering time; LD, linkage disequilibrium; FT, flowering time; NC, not comparable (flowering time for these populations was estimated in a separate experiment and not comparable to the rest).

[a] *P* value from linkage disequilibrium test.

identified 42 indels. Most of those were smaller than 10 bp, but three indels were larger than 30 bp (98, 147, and 400 bp; Figure 5). The two alleles also differed by a large number of nucleotide substitutions, with a peak of divergence ~1 kb downstream of the *COL1* coding region (Figure 5).

DISCUSSION

Our previous QTL mapping identified a *CO* homolog as a candidate for a gene affecting naturally occurring flowering time variation in *B. nigra*. In the present study we could not detect any effect of allelic variation in Bni *COa* on flowering time. Introduction of allelic variants from early- and late-flowering *B. nigra* plants into a *A. thaliana co* mutant resulted in plants flowering earlier than the wild type, although transformants with either of the two alleles displayed similar flowering times.

In contrast to the near sequence identity found between Bni *COa* alleles from early- and late-flowering plants, we detected a surprising sequence divergence between the corresponding Bni *COL1* alleles located only 3.5 kb upstream of Bni *COa*. Furthermore, the genotype at *Ind2* located within the Bni *COL1* coding region showed a strong association with flowering time in several populations. In these populations the *S* allele was consistently associated with early flowering. These data indicate that flowering time is affected by some nucleotide variation (QTN) close to *Ind2*, possibly within the *COL1* gene. From the present data it is not possible to say how close the QTN really is. When a mutation occurs, it has a strong positive disequilibrium with the carrier haplotype and a negative disequilibrium with other haplotypes in the population. Recombination in each generation will then decrease linkage disequilibrium. For old mutations in large equilibrium

TABLE 3

Association between flowering time and indels in population samples of *B. nigra*

	Ind1			*Ind2*		
Population	d.f.	*F*-ratio	*P*	d.f.	*F*-ratio	*P*
Germany	NP	NP	NP	2, 47	0.88	0.42
Greece	4, 35	1.1	0.37	2, 37	14.9	<0.0001
Portugal	4, 23	0.28	0.89	NP	NP	NP
Spain	7, 20	0.93	0.51	1, 26	5.4	0.028
Italy 2	3, 15	0.54	0.66	1, 20	5.0	0.037
Italy 3	2, 20	1.1	0.35	1, 21	17.9	0.0004
Overall[a]	20, 177	0.88	0.61	7, 192	12.0	<0.0001

Analysis of variance was used to test the association between flowering time and indels (*Ind1* and *Ind2*) in the Bni *COL1* gene. The tests were performed separately for each population and in a nested model with country of origin and genotype within country as factors. NP, not polymorphic among plants assessed for flowering time.

[a] Calculated from nested ANOVA (genotype within country).

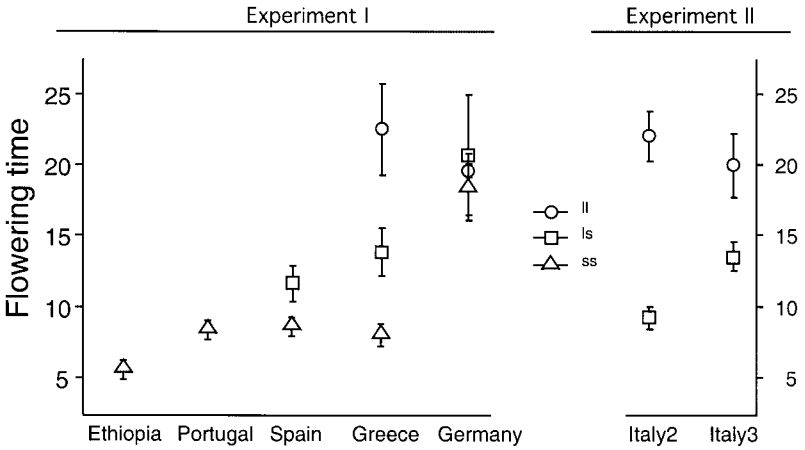

FIGURE 4.—Average flowering times for *LL*, *LS*, and *SS* genotypes at *Ind2* in different *B. nigra* populations. Vertical lines indicate standard errors. Experiments I and II were conducted under different environmental conditions, and the flowering times are not comparable between experiments.

populations, strong linkage disequilibrium is expected only over distances as short as a few kilobases or less (KRUGLYAK 1999; LANGLEY *et al.* 2000; NORDBORG 2000; THORNSBERRY *et al.* 2001).

However, we do not know the age of mutations at *Ind2* and the QTN, or the number of meioses that have occurred, and cannot therefore estimate the distance between the two loci. There is some indication that the *Ind2* polymorphism is old because the insertion was present in Arabidopsis but absent in samples of more closely related Brassica species. These observations suggest that the *Ind2* polymorphism in *B. nigra* was present before the split of the lineage leading to *B. nigra* and the one leading to *B. oleracea* and *B. rapa*. This split most likely occurred several million years ago (R. PRICE, personal communication). Furthermore, the *Ind1* polymorphism located only 250 bp from *Ind2* did not show any significant association with the QTN for flowering time. Either recombination has been frequent enough to break the association between the QTN and *Ind1*, or the lack of association is due to the accumulation of new mutations at *Ind1*. Even if the mutation rate at the simple sequence repeat *Ind1* locus could be relatively high, the data indicate that a considerable number of generations have passed since the occurrence of the QTN mutation. In conclusion, our data suggest that recombination in multiple generations of meiosis should have broken disequilibria over larger distances, indicating that the QTN is probably close to *Ind2*.

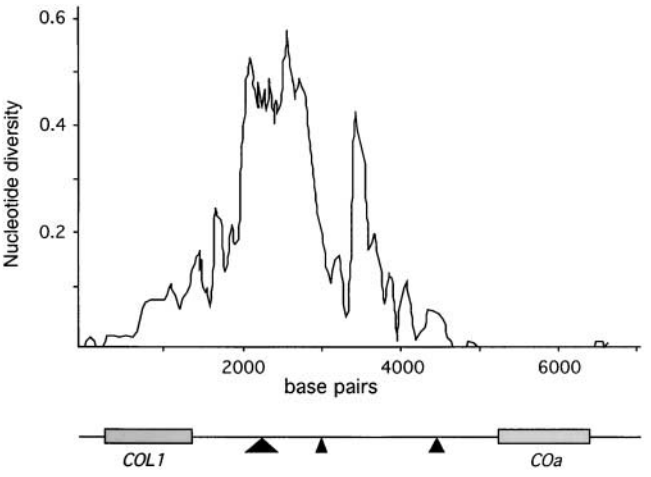

FIGURE 5.—Distribution of silent/synonymous substitutions along the *COL1–COa* region between two alleles from early- and late-flowering *B. nigra* plants. A sliding window of 100 sites was used. The organization of the *B. nigra COL1–COa* region is given below the plot. Boxes indicate genes, lines depict intergenic regions, and solid triangles denote large indels.

Is the QTN for flowering time likely to reside within *COL1*? Alleles of *COL1* from early- and late-flowering plants showed considerable sequence divergence with a total of nine amino acid substitutions in addition to the two indels. Most of the substitutions were radical and occurred in regions of the protein characterized by a high evolutionary rate (Figure 2; LAGERCRANTZ and AXELSSON 2000). However, studies of *COL1* in Arabidopsis have not suggested a role for *COL1* in the control of flowering time. LEDGER *et al.* (2001) did not detect any effect on flowering time when over- or under-expressing *COL1* in *A. thaliana*, but overexpression of *COL1* did shorten the period of two circadian rhythms. Although these data do not support the location of a QTN for flowering time in Bni *COL1*, further studies are needed to test this hypothesis.

The Bni *COa* gene is still an attractive candidate for the QTN: *CO* has been shown to be essential for the induction of flowering in Arabidopsis (PUTTERILL *et al.* 1995). One of the pathways for induction of flowering promotes flowering in response to day length and includes the *CO* gene. The *co* mutant flowers considerably later than wild type under long days, but not under short days, indicating that the *CO* gene is required to promote flowering under long-day conditions. The *CO* mRNA is also more abundant under long days than under short days, indicating that transcriptional regulation of *CO*, at least in part, is likely to regulate flowering time (SUÁREZ-LÓPEZ *et al.* 2001). Flowering time in long days is also correlated with amount of *CO* expression

(Putterill *et al.* 1995). One possibility is that the *B. nigra* QTN for flowering time resided in the intergenic region between *COL1* and *CO* and that it affects flowering time through regulation of *COa* expression levels. Our transformation experiments with *CO* alleles from early- and late-flowering *B. nigra* plants into Arabidopsis *co* did not show any differences in flowering time. These constructs included 2 kb of DNA upstream of the *CO* coding region, which is likely to contain the most important promoter elements in Arabidopsis. However, we cannot exclude that *cis*-acting elements are located farther away. The intergenic region between *COL1* and *CO* is highly diverged between alleles from early- and late-flowering *B. nigra* plants (Figure 5). Polymorphic enhancer elements for *COa* could potentially reside close to *COL1*. An obvious peak of divergence between early and late alleles was observed 2.5 kb upstream of the *COa* coding region. The very high level of sequence diversity at this position could be due to the action of balancing selection, because theoretical models predict a peak of polymorphism near a site that is under balancing selection (Hudson and Kaplan 1988).

Data from Drosophila suggest that variation affecting quantitative traits can be complex at the nucleotide level, with multiple interacting sites. A minimum of three polymorphic sites in one 2.3-kb region of the *ADH* gene contributes to naturally occurring variation in the concentration of ADH protein (Stam and Laurie 1996). To locate the QTN/QTNs for flowering time in the *COL1/COa* region we need to extend the association studies with more markers to establish the extent of linkage disequilibrium in the region and more individuals to increase the statistical power of the tests. We also need to conduct additional transformation experiments, where specific parts of the *COL1–COa* region from early and late alleles are introduced into a common genetic background (*e.g.*, an Arabidopsis *col1, co* double mutant).

We thank Agneta Ottosson for help with Arabidopsis transformation experiments. This work was supported by the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning, and by the Swedish Research Council. M.K.Ö. was supported by a fellowship from the Lawski foundation.

LITERATURE CITED

Alonso-Blanco, C., S. E. El-Assal, G. Coupland and M. Koornneef, 1998 Analysis of natural allelic variation at flowering time loci in the Landsberg *erecta* and Cape Verde Islands ecotypes of *Arabidopsis thaliana*. Genetics **149:** 749–764.

Clarke, J. H., R. Mithen, J. K. M. Brown and C. Dean, 1995 QTL analysis of flowering time in *Arabidopsis thaliana*. Mol. Gen. Genet. **248:** 278–286.

Coupland, G., 1995 Genetic and environmental control of flowering time in *Arabidopsis*. Trends Genet. **11:** 393–397.

Dorn, L. A., and T. Mitchell-Olds, 1991 Genetics of *Brassica campestris*. 1. Genetic constraints on evolution of life-history characters. Evolution **445:** 371–379.

Fu, Y., and K. Ritland, 1994 Marker-based inference about the genetic basis of flowering time in *Mimulus guttatus*. Hereditas **121:** 267–272.

Hajdukiewicz, P., Z. Svab and P. Maliga, 1994 The small versatile pPZP family of Agrobacterium binary vectors for plant transformation. Plant Mol. Biol. **25:** 989–994.

Hudson, R. R., and N. Kaplan, 1988 The coalescent process in models with selection and recombination. Genetics **120:** 831–840.

Johanson, U., J. West, C. Lister, S. Michaels, R. Amasino *et al.*, 2000 Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. Science **290:** 344–347.

Kowalski, S., T. Lan, K. Feldman and A. Paterson, 1994 QTL-mapping of naturally occurring variation in flowering time of *Arabidopsis thaliana*. Mol. Gen. Genet. **245:** 548–555.

Kreitman, M., and H. Akashi, 1995 Molecular evidence for natural selection. Annu. Rev. Ecol. Syst. **26:** 403–422.

Kruglyak, L., 1999 Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat. Genet. **22:** 139–144.

Kuittinen, H., M. J. Sillanpää and O. Savolainen, 1997 Genetic basis of adaptation: flowering time in Arabidopsis thaliana. Theor. Appl. Genet. **95:** 573–583.

Lagercrantz, U., and T. Axelsson, 2000 Rapid evolution of the family of *CONSTANS LIKE* genes in plants. Mol. Biol. Evol. **17:** 1499–1507.

Lagercrantz, U., and D. Lydiate, 1995 RFLP mapping in *B. nigra* indicate differing recombination rates in male and female meiosis. Genome **38:** 255–264.

Lagercrantz, U., J. Putterill, G. Coupland and D. Lydiate, 1996 Comparative mapping in Arabidopsis and Brassica: congruence of genes controlling flowering time. Plant J. **9:** 13–20.

Langley, C. H., B. P. Lazzaro, W. Phillips, E. Heikkinen and J. M. Braverman, 2000 Linkage disequilibria and the site frequency spectra in *su(s)* and *su(w^a)* regions of the *Drosophila melanogaster X* chromosome. Genetics **156:** 1837–1852.

Latta, R. G., 1998 Differentiation of allelic frequences at quantitative trait loci affecting locally adaptive traits. Am. Nat. **151:** 283–292.

Ledger, S., C. Stryer, F. Ahton, S. A. Kay and J. Putterill, 2001 Analysis of the function of two circadian-regulated *CONSTANS-LIKE* genes. Plant J. **26:** 15–22.

Liscum, M., and P. Oeller, 1997 AFLP: not only for fingerprinting, but for positional cloning (http://carnegiedpb.stanford.edu/methods/aflp.html).

Mitchell-Olds, T., 1996 Genetic constraints on life-history evolution: quantitative-trait loci influencing growth and flowering in *Arabidopsis thaliana*. Evolution **50:** 140–145.

Nordborg, M., 2000 Linkage disequilibrium, gene trees, and selfing: an ancestral recombination graph with partial selfing. Genetics **154:** 923–929.

Pors, B., and P. Werner, 1989 Individual flowering time in a goldenrod (*Solidago canadensis*): field experiments show genotype more important than environment. Am. J. Bot. **76:** 1681–1688.

Purugganan, M. D., 2000 The molecular population genetics of regulatory genes. Mol. Ecol. **9:** 1451–1461.

Putterill, J., F. Robson, K. Lee and G. Coupland, 1995 The *CONSTANS* gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell **80:** 847–857.

Putterill, J., S. Ledger, K. Lee, F. Robson, G. Murphy *et al.*, 1997 The flowering time gene *CONSTANS* and homologue *CONSTANS LIKE 1* (accession no. Y10555 and Y10556) exist as a tandem repeat on chromosome 5 of Arabidopsis. Plant Physiol. **114:** 396.

Reeves, P. H., and G. Coupland, 2000 Response of plant development to environment: control of flowering by daylength and temperature. Curr. Opin. Plant Biol. **3:** 37–42.

Samach, A., H. Onouschi, S. E. Gold, G. S. Ditta, Z. Schwarz-Sommer *et al.*, 2000 Distinct roles of *CONSTANS* target genes in reproductive development of Arabidopsis. Science **288:** 1613–1616.

Schneider, S., D. Roessli and L. Excoffier, 2000 Arlequin: a software for population genetics data analysis, version 2.000. Genetics and Biometry Lab, Department of Anthropology, University of Geneva, Geneva.

Simpson, G. G., A. R. Gendall and C. Dean, 1999 When to switch to flowering. Annu. Rev. Cell Dev. Biol. **15:** 519–550.

Slatkin, M., and L. Excoffier, 1996 Testing for linkage disequilibrium in genotypic data using the EM algorithm. Heredity **76:** 377–383.

Stam, L. F., and C. C. Laurie, 1996 Molecular dissection of a major gene effect on a quantitative trait: the level of alcohol dehydrogenase expression in *Drosophila melanogaster*. Genetics **144:** 1559–1564.

Suárez-Lopés, P., K. Wheatley, F. Robson, H. Onouchi, F. Valverde *et al.*, 2001 *CONSTANS* mediates between the circadian clock and the control of flowering in *Arabidopsis*. Nature **410:** 1116–1120.

Thornsberry, J. M., M. M. Goodman, J. Doebley, S. Kresovich, D. Nielsen *et al.*, 2001 Dwarf8 polymorphisms associate with variation in flowering time. Nat. Genet. **28:** 286–289.

Tommey, A., and E. Evans, 1991 Temperature and daylength control of flower initiation in winter oilseed rape. Ann. Appl. Biol. **118:** 201–208.

Communicating editor: O. Savolainen