Theor Appl Genet (1997) 95:573–583

H. Kuittinen · M. J. Sillanpää · O. Savolainen

Genetic basis of adaptation: flowering time in *Arabidopsis thaliana*

Received: 15 January 1997 / Accepted: 21 February 1997

Abstract We have mapped QTLs (quantitative trait loci) for an adaptive trait, flowering time, in a selfing annual, *Arabidopsis thaliana*. To obtain a mapping population we made a cross between an early-summer, annual strain, Li-5, and an individual from a late overwintering natural population, Naantali. From the backcross to Li-5 298 progeny were grown, of which 93 of the most extreme individuals were genotyped. The data were analysed with both interval mapping and composite interval mapping methods to reveal one major and six minor QTLs, with at least one QTL on each of the five chromosomes. The QTL on chromosome 4 was a major one with an effect of 17.3 days on flowering time and explaining 53.4% of the total variance. The others had effects of at most 6.5 days, and they accounted for only small portions of the variance. Epistasis was indicated between one pair of the QTLs. The result of finding one major QTL and little epistasis agrees with previous studies on flowering time in *Arabidopsis thaliana* and other species. That several QTLs were found was expected considering the large number of possible candidate loci. In the light of the suggested genetic models of gene action at the candidate loci, epistasis was to be expected. The data showed that major QTLs for adaptive traits can be detected in non-domesticated species.

Key words QTL-mapping · Adaptation · Flowering time · *Arabidopsis thaliana*

Communicated by P. M. A. Tigerstedt

H. Kuittinen (✉) · O. Savolainen
Department of Biology, University of Oulo, FIN-90570 Oulo, Finland

M. J. Sillanpää
Rolf Nevanlinna Institute, FIN-00014 University of Helsinki, Finland

Introduction

The genetic basis of adaptation is poorly understood (Orr and Coyne 1992). The view of modern synthesis has been that adaptation to novel environments is due to selection on polygenic variation. The phenotypic traits that confer adaptation have typically been discussed in a quantitative genetic context based on the infinitesimal model of Fisher (1918; Bulmer 1971; Falconer 1989; Lande 1983). However, as discussed by Orr and Coyne (1992), evidence for a polygenic nature of adaptation is surprisingly limited.

There are several reasons for a renewed interest in the genetic basis of adaptation. Theoretical developments have refined the questions that need to be asked about the nature of quantitative variation (e.g. Barton and Turelli 1989). Explanations for the maintenance of quantitative variation and predictions of the limits to selection depend on the basis of variation (Barton and Turelli 1989; Hill and Caballero 1992). We need to know about the number of loci, their mutation rates, and the distribution of allelic effects to choose the appropriate models.

Marker genes have been used to identify chromosomal regions containing QTLs (quantitative trait loci) since 1923 (Sax 1923), but the availability of a practically infinite number of DNA markers as well as new statistical methods (Lander and Botstein 1989; Zeng 1994; Jansen and Stam 1994) have made marker-based mapping efficient (e.g. Stuber et al. 1987). These studies can give detailed information on the number of loci, the effects of the alleles and epistatic interactions between them. The resolution of marker studies is better than that of traditional biometrical studies, which assume complete additivity, equal effects and independent segregation of the genes. Violations of the assumptions usually lead to an underestimate of the real number (Lande 1981).

Although many QTL mapping studies have been conducted, most of them deal with domesticated species.

These results cannot be applied directly to natural populations because artificial selection associated with domestication may fix alleles of major effect (Gottlieb 1984). Clearly more information is needed on the genetic basis of adaptive traits in natural populations.

Here we study the genetic basis of between-population differences in a quantitative trait, flowering time, in a nearly completely selfing (Abbott and Gomez 1989) annual plant *Arabidopsis thaliana*. Timing of germination and flowering in annual plants synchronizes the growth and reproduction of plants with favorable periods (Symonides 1987). Flowering time can also have an effect on the seed set of a plant, leading to possible trade-offs between timing and magnitude of reproduction (Dorn and Mitchell-Olds 1991). We chose to study a population difference in flowering time reflecting a distinction between summer annual and over-wintering life modes. The winter annual life mode in *Arabidopsis thaliana* and some other annual plants has been associated with the avoidance of drought in shallow soils during the summer (e.g. Ratcliffe 1961; Baskin and Baskin 1972). Seeds of winter annual *A. thaliana* are dormant in the summer and germinate in the autumn, the plants overwinter at the rosette stage and they flower immediately in the spring. The winter annuals have a vernalization requirement of different degrees for flowering. In the wild the requirement is met during the winter but in growth chamber experiments, with no cold treatment, the plants flower late, after growing a large rosette. Summer annuals germinate in the spring, flower and set their seeds in the same season. In growth chamber experiments they flower early, after forming a small number of rosette leaves. There is further phenotypic variation in flowering time within both major types (Napp-Zinn 1969). The latest ecotypes are often found in northern countries or high altitudes (Kranz and Kirchheim 1987). In Finland most populations are winter annual, but in Central Europe there are both winter and summer annual populations.

There are many possible candidate genes in *Arabidopsis thaliana* that could account for the different flowering-time phenotypes. The candidate genes have been identified in the last 50 years through classical crossing experiments (Härer 1951; Karlovska 1974; Napp-Zinn 1957a, 1962) and mutation and physiological studies (reviewed in Martinez-Zapater et al. 1994; Coupland 1995). In recent years linkages of molecular markers with QTLs have been searched for in crosses between different *A. thaliana* strains (Lee et al. 1993; Burn et al. 1993; Kowalski et al. 1994; Jansen et al. 1995; Mitchell-Olds 1996). Almost 20 loci influencing the timing of flowering have been characterized in mutation studies or crossing experiments, and most of them have been placed on the linkage map of *A. thaliana*. Three of the genes have been cloned (Coupland 1995). Information from the crossing experiments and QTL studies suggest the most times only 1 or a few major loci are responsible for most of the differences between *A. thaliana* strains (Härer 1951; Karlovska 1974; Napp-Zinn 1957a, 1962; Lee et al. 1993; Kowalski 1994; Clarke et al. 1995; Mitchell-Olds 1996), but sometimes many QTLs have been found (Jansen et al. 1995).

In the study presented here, we wanted to find out how many QTLs are responsible for the difference in flowering time between an early (summer annual) Central European strain and a late (winter annual) natural Finnish population, what are their relative effects on the character, and if there are epistatic interactions between the QTLs. We discuss whether the QTLs map to the same regions as the candidate loci found by artificial mutagenesis or others found in previous studies between strains.

Materials and methods

Populations and crosses

The populations were a late-flowering Finnish population (Naantali) and an early-flowering German *Arabidopsis thaliana* strain (Li-5). Li-5 was obtained from the Arabidopsis Information Service collection (acc. no. 917). The original seed sample was collected from a population growing in Limburg an der Lahn (8°5′E, 50°35′N), and the line Li-5 has been derived from it through selfing. The late-flowering Naantali population was collected in southern Finland (22°5′E, 60°40′N).

In preliminary growing experiments plants from Li-5 were homogenous with respect to flowering time. The Naantali population consists of several homozygous genotypes between which there is some genetic variation in flowering time (Kuittinen et al. 1997). One plant from Li-5 was chosen randomly, and from the Naantali population one of the latest phenotypes from the original seed collection was chosen for the crosses. Based on the results of Kuittinen et al. (1997) and because *A. thaliana* is known to be nearly completely selfing (Abbott and Gomez 1989) both of the chosen plants should be homozygous at most of their loci.

The Li-5 plant was crossed with the Naantali plant, and the hybrid was backcrossed to 1 selfed progeny of the Li-5 plant. For the first growing experiment 298 backcross seeds were used together with 33 F_1 hybrids and 33 selfed seeds of the parental plants. The plants were grown in 11 completely randomized blocks. This backcross was the mapping population in the linkage and QTL analysis.

A backcross was also made to 1 selfed progeny of the late parent Naantali. This backcross population ($n = 90$) was grown together with selfed progenies of the parental plants (each $n = 60$), F_1 hybrids ($n = 60$) and F_2 generation (seeds collected from selfing F_1 plants, $n = 90$) in a second experiment in order to study the phenotypic distributions of these populations. The experimental design was completely randomized blocks.

Flowering-time experiments

The seeds were sown in pots filled with a 1 : 1 mixture of vermiculite-peat. The imbibed seeds were kept for 24 h at room temperature in the light. In the first experiment the seeds were then vernalized in darkness for 7 days at 4°C before being taken to the growth chamber. The short vernalization was used to break the possible dormancy of the seeds. On the basis of earlier experiments, a short vernalization was also expected to reduce somewhat the flowering time of the extremely late Naantali and to reduce environmental variance. The light source in the growth chamber was metal halide lamps (Osram Power star HQI-T/D) that gave photosynthetically

active radiation (PAR) at 35–42 μmol $m^{-2} s^{-1}$ during a 16-h photoperiod. The ratio of red : far-red light was approximately 2.5. Humidity in the growth chamber ranged from 50 to 65%. Flowering time and the number of rosette leaves at flowering were scored for each plant until the experiment was terminated after 244 days. Flowering time was defined as the day when the first flower opened, counted from the first day in the growth chamber.

The second experiment was carried out in the same manner except that the plants were vernalized at 2°C for 7 (all BC and F_2 plants, half of Li-5, Naantali and F_1 plants) or 70 days (half of Li-5, Naantali and F_1 plants). The long vernalization treatment was used in order to study whether it completely abolishes the difference in flowering times between Limburg, Naantali and F_1 plants. In this experiment the vernalization temperature was 2°C instead of 4°C because this was expected to give the most efficient vernalization response (Napp-Zinn 1957b). Flowering time and the number of rosette leaves were scored as before. This experiment was terminated after 134 days.

Markers

Marker data consisted of new random amplified polymorphic DNA (RAPD) markers generated for this cross and of published microsatellite, cleaved amplified polymorphic sequence (CAPS) and restriction fragment length polymorphism (RFLP) markers. RAPD markers were used because they were fast to score; the others were chosen to make the map complete and for comparisons with other *A. thaliana* maps.

Genomic DNA was isolated from the parental lines and the backcross progeny to Li-5 using leaves of either the original plants or of 20–30 of their selfed progeny. DNA for the polymerase chain reaction (PCR) was extracted according to Rogers and Bendich (1985) and that for Southern hybridisations either by the same method or following Dean et al. (1992).

To obtain RAPD sequences we used one or two random decamer (Operon Technologies) primers in amplification reactions. Reactions (10 μl) contained 0.4 pmol/μl primer(s), 50 m*M* KCl, 10 m*M* TRIS-HCl, pH 9.0, 0.1% Triton X-100, 3.5 m*M* $MgCl_2$, 100 m*M* each dNTP, 0.04 U/μl *Taq*-polymerase and 1 ng/μl template DNA. Reactions were set up in microtiter plates and overlaid with 60 μl of mineral oil. Amplification was performed in a Hybaid Omnigene thermal cycler with the following temperature profile: 92°C for 3 min; 45 cycles of 94°C for 30 s, 36°C for 15 s, 73°C for 1 min and 72°C for 8 min. The products were separated by electrophoresis on 1.2% agarose gels containing 0.5 μg/ml ethidium bromide. Because RAPDs are dominant, those fragments that were not amplified in Li-5 but were amplified in Naantali were used as markers. Their segregation in the backcross progeny should be 1 : 1. The fit to the observed segregation ratio was tested by chi-square tests. RAPD loci were named according to the recommendation of Operon Technologies. For example, $OPF04F08_{800}$ was produced with two primers, OPF04 and OPF08, and the size of the product was approximately 800 bp.

Primers for microsatellites and CAPS were obtained from Research Genetics, Alabama. The sequences and amplification methods of the microsatellites are described in Bell and Ecker (1994) and those of CAPS in Konieczny and Ausubel (1993). For CAPS, *GLI* was cleaved with *Hin*fI and *DHSI* with *Eco*RI. DNA fragments were separated by electrophoresis either on 1–4% agarose gels containing 0.5 μg/ml ethidium bromide or on 9% polyacrylamide gels (ratio of acryl-bisacrylamide 29 : 1). In the later case DNA was visualized by silver staining (Silver Stain Plus kit, Bio-Rad).

RFLPs were visualized by separating restriction enzyme-digested DNA (1.5–3 μg) by electrophoresis on 0.8% agarose gels and then transferring it to nylon membranes (Hybond N^+, Amersham) by vacuum blotting. Probe labelling, hybridization, washing and detection were performed following Karvonen et al. (1994). Probes were cosmid and lambda phages obtained from the Arabidopsis Biological Resource Center in Ohio. Cosmid DNA was extracted by the alkaline lysis method (Sambrook et al. 1989) and lambda DNA according to Chisholm (1989).

Linkage and QTL analysis

The markers (33 RAPDs, 2 CAPS, 2 RFLPs and 12 microsatellites) were scored first in the 93 most extreme backcross (to Li-5) progeny (i.e. selective genotyping). These data were used to construct a linkage map with the MAPMAKER (3.1) program (Lincoln et al. 1992a; Lander et al. 1987). A LOD score of 3.0 was used for forming linkage groups and for three-point linkage analysis. The order of the markers was verified by permuting the order of the five neighboring loci. The markers of known location (CAPS, RFLPs and microsatellites) were used to assign the linkage groups to chromosomes and also to reveal potential major deviations in the lengths of the chromosomes compared to published maps. The Kosambi (1944) mapping function was used for calculating the map distances, which allowed comparison of the map with already published maps. In QTL analysis, Haldane's (1919) mapping function was adopted.

QTL mapping was done with two methods, interval mapping (IM) (Lander and Botstein 1989) and composite interval mapping (CIM) (Zeng 1994; Jansen and Stam 1994). The assumption of normally distributed phenotypic data was not used violated by the bimodal phenotypic distribution. Log-transformation of the phenotypic data was not used because it did not improve the normality. Because block effects were not significant when tested with ANOVA, unadjusted values were used. Genotypic data for the 93 extreme and phenotypic data for all of the 298 backcross progeny were used (i.e. selective genotyping Lander and Botstein 1989). Selective genotyping should increase the efficiency of mapping when the number of genotyped individuals is a limiting factor (Lander and Bostein 1989). Data on individuals for which only phenotypic values were available could be used in the analysis because the analysis methods and associated software allow for such design. The analyses were carried out with QTL-CARTOGRAPHER (Basten et al. 1996) and MAPMAKER/QTL 1.1 (Paterson et al. 1988; Lincoln et al. 1992b). IM was first done with both software packages, but since they gave very similar likelihood curves, only QTL-CARTOGRAPHER was used to generate results for IM in further analyses.

IM allows the calculation of likelihood scores for a putative QTL placed in any position between 2 adjacent flanking markers (interval region). CIM extends this method by also fitting the most significant markers (or all if forced) outside the interval into the model, resulting in a considerable reduction in residual variation. CIM was run with a 10-cM window in all analyses. The number of markers for background control was set to 5, which means that the 5 most significant markers, detected through stepwise regression outside the window, were fitted to the model. Background controls determined by the stepwise procedure were found to flank the individual markers lying closest to five QTLs detected in later analysis. Values of 8 and 12 background controls were also tested.

Empirical experimentwise threshold values for $P = 0.05$ (corresponding to 5% false positive rate) were constructed by permutation tests described by Churchill and Doerge (1994). The values were generated separately for the IM and the three CIM computations, with 1000 permutations for each. The permutation tests were run with QTL-CARTOGRAPHER software in a Sun UltraSparc Model 170 computer (with practically no other load during the runs). The running times were 16 h for the IM run and 51, 76 and 109 h for the CIM runs, which shows that these are very computer-intensive tasks. One-LOD support intervals recommended by Conneally et al. (1985) were calculated for the CIM run with 12 parameters. Support-interval threshold values were defined on the likelihood ratio (LR) test statistic scale by subtracting from the maximum peak value $2/\log_{10}(e)$, which corresponds to one LOD in the LR statistic scale.

Estimates for the additive effects of the QTLs given by the two methods were of the same magnitude for all QTLs found. This was not reasonable considering the bimodality of the backcross data, which suggested one or two major QTLs. To obtain better estimates,

we scored new genotype data in additional plants for the markers flanking the five main QTLs (detected in CIM using 5 background parameters), so that data from a random sample of altogether 140 backcross progeny were available. These data were used for estimating the effects of these QTLs and their interactions in an analysis of variance context. The new data were not included in QTL mapping. Estimates of the allelic effects were obtained by comparing the means of genotypes at the marker closest to each QTL. Interactions of pairs of marker loci nearest to the putative QTLs were analyzed with ANOVAs with the 5-marker genotypes as main effects and with an interaction component of one pair of markers at a time. All effects were considered to be random. F ratios for each test were constructed by dividing the difference between the regression sums of squares of the model including the interaction term to be tested and that of the model with only main effects by the mean square error for the model including both main effects and the tested interaction. The significance of each marker closest to these putative five QTLs was tested also by one-way analyses of variance using the random sample, and the R^2 of these ANOVAs were regarded as estimates of the phenotypic variances explained by the individual QTLs.

Results

Flowering time

Flowering-time distributions are shown in Fig. 1. The difference between Naantali and Li-5 was very large after 7 days vernalization at 4°C. The mean flowering time of Li-5 was 36.1 days (SD 1.3, range 34–39 days), while the first Naantali plant did not flower before 139 days. The variance in Naantali was large, and more than half of the plants had not yet flowered at 244 days when the experiment was finished. The F_1 generation was nearly intermediate in flowering at 83.7 (SD 8.1) days. In the second experiment the flowering times of Li-5 (mean 28.8 days SD 1.7), Naantali (minimum 101 days) and F_1 (mean 65.2 days, SD 10.9) were somewhat shorter, probably due to a more efficient vernalization at 2°C. The plants from the backcross to Li-5 started flowering between 30 and 82 days, and the plants from the backcross to Naantali between 56 and 134 days (only 1 plant remained in a vegetative stage until the experiment was finished). Flowering times in the F_2 population ranged from 26 to 97 days, so at least in this sample (90 plants) the latest phenotypes (identical to Naantali) were missing. The distributions of the progeny from the backcrosses to both Naantali and Li-5 were clearly bimodal, while the distribution of the F_2 population was unimodal, although skewed. The backcross distributions suggested that there is a major gene responsible for much of the flowering-time difference between the lines, but the possibility of additional genes still remains. For instance, in experiment 1, the latest part of the backcross were earlier than F_1, which would suggest some early recessive alleles in Li-5 apart from the main locus.

The long vernalization treatment (70 days at 2°C) reduced flowering times in Li-5, Naantali and the F_1 markedly compared to the 7-days vernalization at either 4°C or 2°C (Fig. 2). The variances were reduced as well. The largest change in mean flowering time was in Naantali, where the reduction was from the minimum of 101 days to the average of 34.1 (SD 2.4) days. Li-5 flowered at 25.2 (SD 1.0) days and the F_1 at 30.4 (SD 2.0) days. Even after the long vernalization treatment, a significant difference of 9 days remained between Li-5 and Naantali ($F = 323.6$, $P < 0.001$).

The linear correlation between flowering time and the number of rosette leaves was high ($r = 0.98$, $P < 0.001$) in the backcross to Li-5. Only 1 individual showed a deviant phenotype having many fewer leaves than other plants flowering at the same time. In all other cases, lateness was thus due to a delay in the onset of flowering and not to retarded growth.

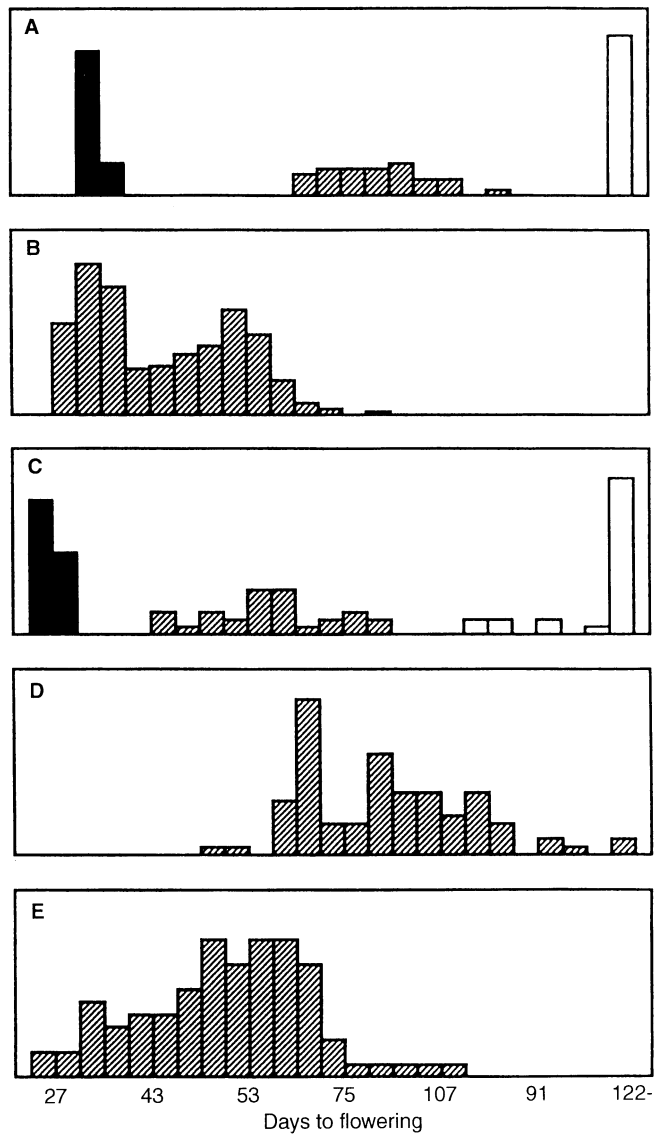


Fig. 1A–E Frequency distributions of flowering times. **A** and **B** are for the first experiment, and **C–E** are for the second experiment. **A**, **C** Li-5, F_1 and Naantali populations, indicated by *black*, *shadowed* and *white* bars respectively. **B** $F_1 \times$ Li-5, **D**, Naantali $\times F_1$, **E** F_2

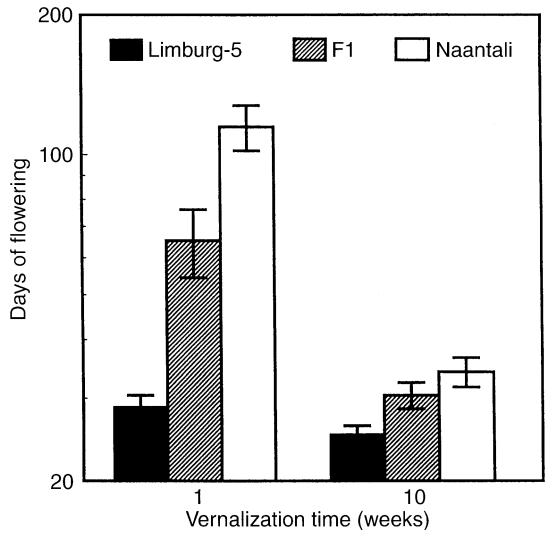


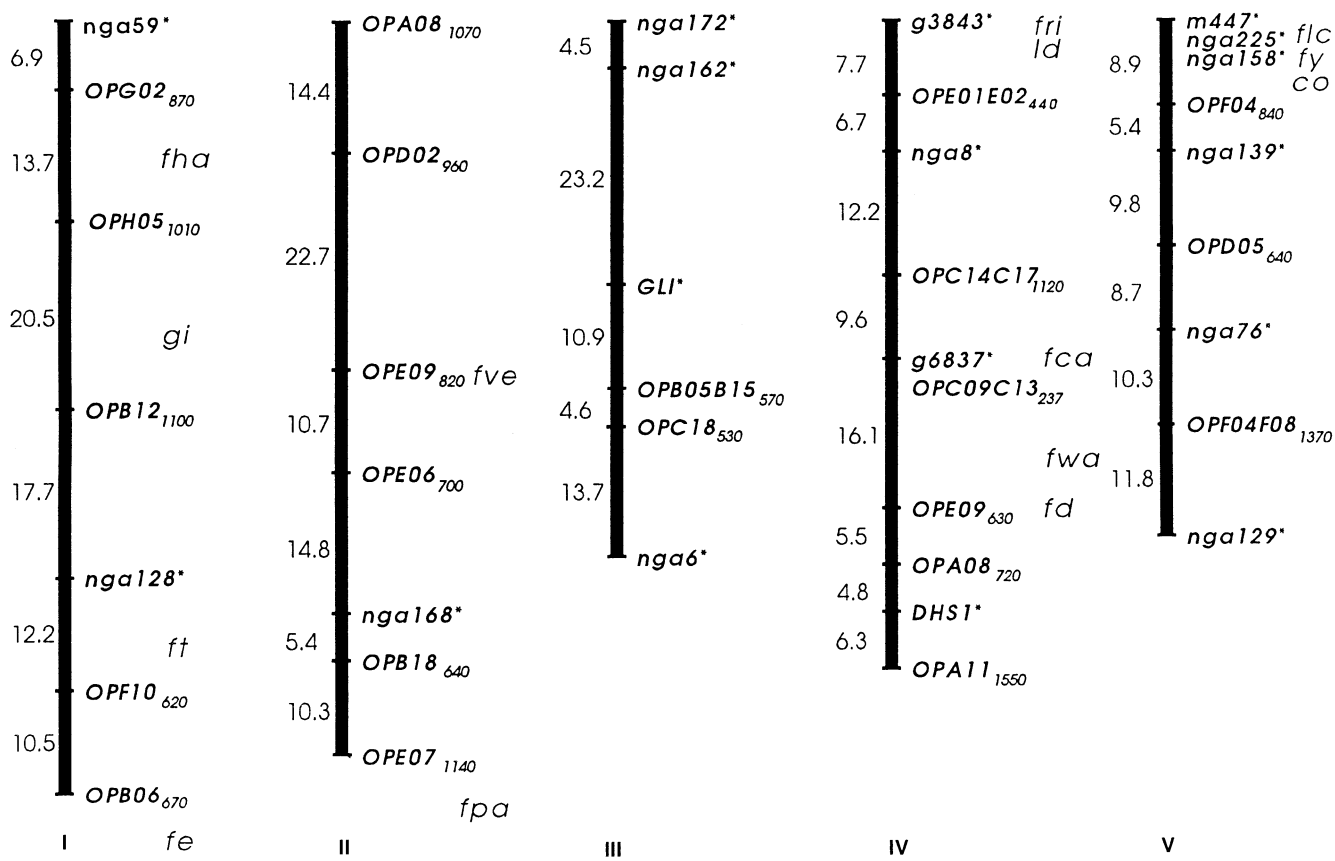
Fig. 2 Average flowering times with their standard deviations in Li-5, Naantali and F_1 after 1 week and 10 weeks vernalization at 2°C. The flowering time in Naantali after 1 week's vernalization is the average of the plants that flowered before the experiment was terminated at 134 days

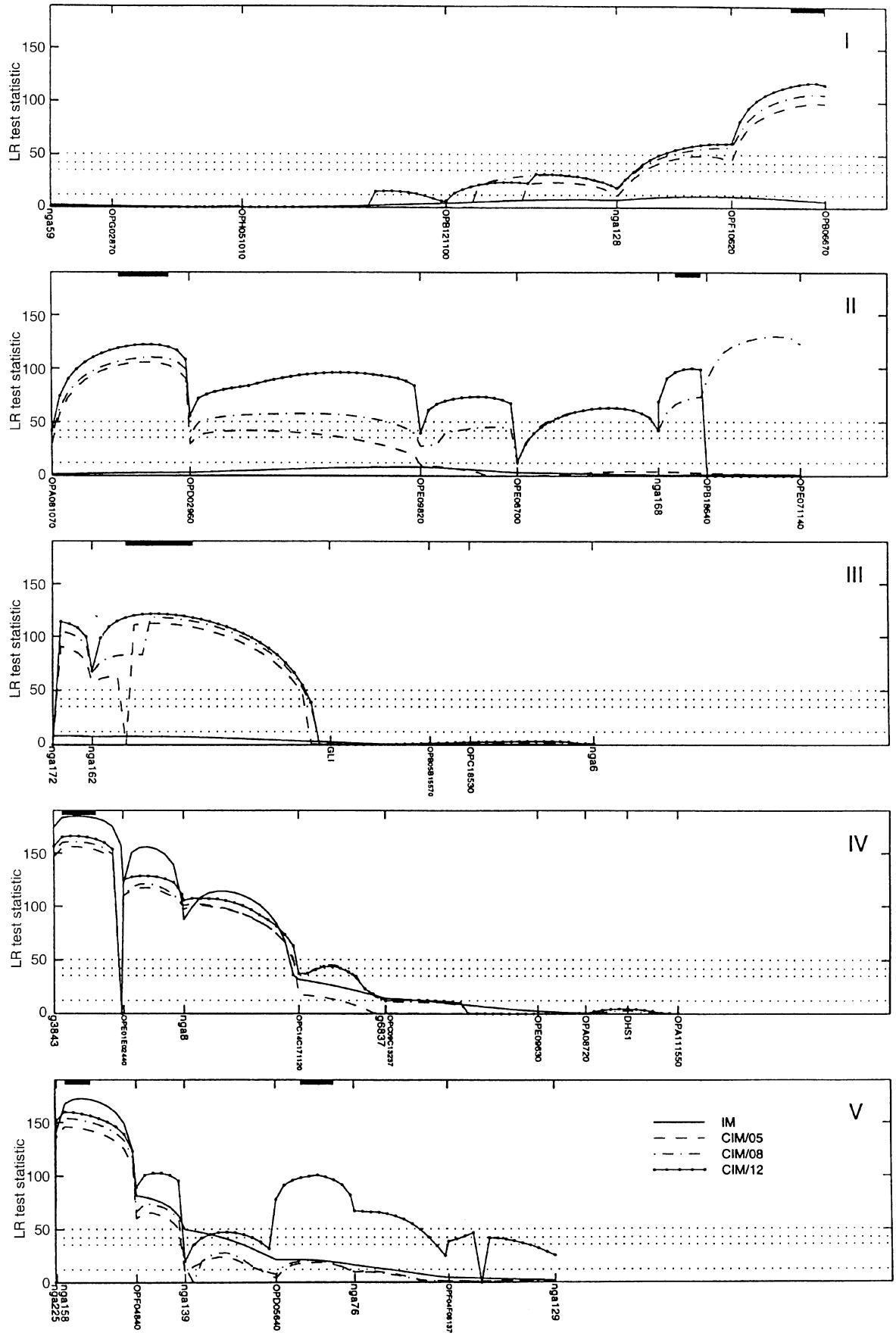
Fig. 3 The linkage map based on the segregation of markers in the backcross to Li-5. Previously mapped markers are indicated with an *asterisk*. Flowering time loci found by mutagenesis have been placed in their approximate positions to the *right* of the markers

Linkage map

The linkage map constructed with previously mapped markers (microsatellites CAPS and RFLPs) and new random markers (RAPDs) is shown in Fig. 3. Some RAPD loci that were close to others have been excluded. Two of the RAPD loci could not be assigned in any of the five linkage groups. At 5 of the 47 loci (11%) assigned to linkage groups the segregation of the alleles differed significantly from the Mendelian 1:1 ratio. This is somewhat more than expected by chance alone (5%). These loci are located in three unlinked areas. In two of these areas a Li-5 allele and in one area a Naantali allele was in excess. However, all of the loci were used for the linkage map to avoid large gaps.

The lengths of the five chromosomes were 82, 78, 57, 69, and 55 cM for chromosomes 1–5, respectively. Altogether, the markers spanned 341 cM. In the map used for QTL analysis the average spacing between the markers was 11 cM. The previously mapped markers and the new RAPD markers covered the chromosomes rather well. Only chromosome 5 lacked markers for about a 30 cM length at one end, *nga129* being the most distal marker. Compared to the F_2 map of Hauge et al. (1993), which spans 520 cM, all the chromosomes in this study showed less recombination. The shortness of the chromosomes was only partly due to an incomplete coverge of the markers because the distances between the common markers in our map were




I
II
III
IV
V
LR test statistic
0
50
100
150
nga59
OPG02870
OPH05 1010
OPB12 1100
nga128
OPF10620
OPB06670
OPA08 1070
OPD02960
OPE09820
OPE06700
nga168
OPB18640
OPE07 1140
nga172
nga162
GLI
OPB05 15570
OPC18530
nga6
g3843
OPE01 E02440
ngaB
OPC14 C17120
OPD05 G13237
g0837
OPE09630
OPA08720
OHS1
OPA11 1550
nga158
nga225
OPF04 B40
nga139
OPD05 B40
nga76
OPF04 F04 1137
nga129
IM
CIM/05
CIM/08
CIM/12

consistently smaller compared to the map of Hauge et al. (1993). A decrease in the lengths of chromosomes is also expected when selective genotyping is used if there are QTLs in coupling phase within the chromosomes. An indication of coupling QTLs was, indeed, found in the QTL analysis in chromosome 5. This is supported by the fact that the map distances increased in chromosome 5 after the map distances at the top of chromosome 5 were recalculated in a small random set of backcross progeny. The distance was also recalculated in the same manner for the top of chromosome 4 without any change. Most of the difference in the recombination frequency may be due to the different types of crosses (backcross vs. F_2) (Vizir and Korol 1990) and the genetic backgrounds used (Säll et al. 1993).

QTLs for flowering time

Two putative QTLs were found in interval mapping, and from five to seven in composite interval mapping, depending on the number of background parameters used (Fig. 4). The experimentwise threshold obtained from 1000 permutations under $P = 0.05$ for interval mapping was T = 12.12, which can be rescaled to a LOD score of 2.63. Similarly, experimentwise thresholds under $P = 0.05$ for CIM having 5, 8 and 12 background controls were T5 = 35.52, T8 = 42.39, and T12 = 50.50, respectively. The thresholds increased linearly with the increase in the number of parameters, as did running times.

The location of the QTLs can be seen in Fig. 4. The main QTLs in chromosomes 4 and 5 were revealed by both interval mapping and composite interval mapping, and the likelihood curves surrounding the QTLs were similar. IM did not find the QTLs on chromosomes 1, 2 and 3 and the minor QTL on chromosome 5, all of which were revealed only with CIM. Increasing the number of background parameters in CIM revealed additional QTLs. The second QTL in chromosome 2 was found with 8 and 12 parameters, and the second QTL in chromosome 5 with 12 parameters. In some positions the likelihood curve decreased steeply at marker positions. In interval-type QTL analyses, the true QTL genotypes are unknown except in marker positions, where at least some proportion of the genotypes are observed. Therefore, the likelihood ratio curve can decrease at marker positions, which show more evidence against a QTL than the nearby positions between markers. In other words, a decrease in the curve at marker positions may reflect the proportion of observed genotypes at the marker loci. However, this decrease can be seen only in those parts of the marker map where the likelihood ratio had high values. Misscoring of the markers could also be responsible (Lincoln et al. 1993b).

In the ANOVA of the random sample of backcross progeny the interaction component was significant in 1 marker combination, indicating epistasis between the QTLs (Table 1). The Naantali allele at marker *OPD02* near the first QTL on chromosome 2 had an effect only when there was also the Naantali allele at marker *OPE01E02* near the QTL on chromosome 4. We did not correct for multiple testing. The interaction estimates represent lower bounds, and we would not be able to exclude epistasis with confidence anyway.

The difference of the means of the marker genotypes at the loci closest to the five QTLs detected by CIM with 5 background parameters are shown in Table 2. Naantali alleles caused lateness in all cases. The delay in flowering caused by the Naantali allele in the chromosome 4 QTL was 17.3 days and that of the chromosome 5 QTL, 6.5 days. The other QTLs had effects of 1.7–3 days. R^2 of the ANOVA using the model with the markers flanking the five main QTLs as main effects together with the one significant interaction term was 61.5%. Based on the single classification ANOVAs, the major QTL on chromosome 4 explained most (53.5%) of the variance, with the other QTLs and the interaction term explaining much smaller portions of it (Table 2). The two additional QTLs detected by CIM with 12 background parameters have probably even smaller effects than the others and would not contribute much to the explained variance if they were taken into the model.

Fig. 4 Interval and composite interval mapping. The chromosome number is in the *upper right corner* of each *panel*; marker names are on the *x-axis*. The value of the LR test statistic was calculated at every 1-cM position on each chromosome. Explanation of the *different line types* are given in the panel of chromosome 5. The *four horizontal dotted lines* are experimental threshold values for IM and CIM with 5, 8 and 12 background parameters, from lowest to highest, respectively. One-LOD support intervals for CIM run with 12 parameters are drawn as *black rectangles* on the *upper part* of the *panels*

Table 1 Interactions of the genotypes at marker loci flanking the QTLs. Interaction was analyzed with ANOVA using a model with the five markers flanking the QTLs as main effects and an interaction component of two markers at a time, based on a random sample of backcross individuals. *P* values of the interaction component are presented

	$OPB06_{670}$	$OPD02_{960}$	*nag162*	$OPE01E02_{440}$
$OPD02_{960}$	0.2739			
nga162	0.0749	0.4845		
$OPE01E02_{440}$	0.7247	0.0488	0.8969	
nga158	0.6197	0.9576	0.8437	0.8249

Table 2 One-way ANOVA for genotypes at marker loci flanking QTLs revealed in CIM analysis, based on a random sample of backcross individuals

	R^2(%)	Pr < F	Class means[a]		Delay caused by the Naantali allele
			LL	LN	
OPB06$_{670}$ (Chrom. 1)	1.6	0.118	45.9	48.9	3.0
OPD02$_{960}$ (Chrom. 2)	0.5	0.362	46.1	47.8	1.7
nga162 (Chrom. 3)	1.0	0.209	45.5	47.9	2.4
OPE01E02$_{440}$ (Chrom. 4)	53.4	0.001	39.1	56.4	17.3
nga158 (Chrom. 5)	7.3	0.001	43.6	50.1	6.5

[a] Genotype classes are coded as LL for a homozyote for the Limburg-5 allele and LN for a heterozygote for the Limburg-5 and Naantali alleles

Discussion

A small number of genes for population differences was found

We studied quantitative variation that is found in nature using a cross between individuals from an early-flowering strain and a late-flowering natural population. While most of this difference could be eliminated with a long vernalization treatment, a significant difference still remained. We found up to seven QTLs, of which the five main ones were responsible for 62% of the phenotypic variance. One of the QTLs had a major effect on flowering time, while the others had much smaller effects. Evidence for interaction between two of these QTLs was found. There is a possibility that we have missed some minor QTLs because mapping in the backcross population presumes that the late genes are at least partly dominant. Also, without fine-scale mapping it is not possible to distinguish between tightly linked QTLs, which would here appear as one QTL.

Other studies between summer and winter annual *Arabidopsis thaliana* strains have revealed from one to a few loci. Based on crosses between Li-5 and late St (Stockholm, Sweden) Napp-Zinn (1957a, 1962) concluded that two relatively strong flowering-time genes and at least two minor genes were responsible for the flowering-time difference. The two major genes (*FRI* and *KRYO*) acted epistatically and the other genes additively (Napp-Zinn 1987). *FRI* was later located at the upper end of chromosome 4 (Clark and Dean 1994). Clarke et al. (1995) found 5 genes in a cross between early-summer annual Landsberg *erecta* (L*er*) and an inbred line (homozygous for the late *FRI* allele) derived from the cross Li-5 × St.

FRI is probably allelic to the gene (*FLA*) responsible for most of the strong vernalization demand of the late strain Sf-2 (Spain) and Leiden (Netherlands) analyzed in crosses against the early-flowering strain Columbia (Lee et al. 1993). It may also be allelic to the flowering-time gene found in two late-vernalization-responsive strains (Pitztal and Innsbruck, Austria) (Burn et al. 1993). *FRI* has been postulated as the most common gene responsible for the vernalization demand of Central European and Scandinavian late ecotypes (Napp-Zinn 1987; Clarke and Dean 1994). On the basis of its location and mode of action, it is possible that *FRI* is the QTL in chromosome 4 in our cross.

The distribution of allelic effects seems to be more even in crosses between two summer annual strains than between summer and winter annuals, although major genes have also been found. Kowalski et al. (1994) studied flowering-time variation in a cross between two summer annuals, WS (Wassilewskaja, Belorussia) and HM (Hannover/Munden, Germany), and found two unlinked QTLs accounting for approximately similar amounts of the variation (12% and 16%). Two different QTLs were found in the cross between early strains Col and L*er* accounting for 16% and 22% of the variation (Mitchell-Olds 1996). Eleven loci were revealed in a study where QTLs were mapped in many environments in a cross between Col and L*er*, but the relative effects of the loci were not analyzed (Jansen et al. 1995).

Thus, the differences between the winter and summer annuals seem to be controlled by one major gene and several minor ones. The major gene may be the same (FRI) in all of the studied cases. Studies between summer annuals have typically detected some genes with equal effects, but presumably many more minor genes may be contributing as in the study of Jansen et al. (1995).

QTL mapping methods

Composite interval mapping showed its strength in our analysis. It proved to have better precision in detecting QTLs over simple interval mapping due to its ability to

simultaneously take into account effects of some other QTLs lying outside the interval region. Aside from the overall greater power of QTL detection of CIM (when the number of background controls was kept small), the largest difference between the results of the two methods was seen in chromosomes 2 and 5. IM gave an insignificant LOD score peak in the middle of chromosome 2, while CIM gave two significant peaks at the ends of the chromosomes when a high number of background parameters was used. The second QTL on chromosome 5 was found only by CIM with 12 background parameters. The shortening of the chromosome that was observed when the map distances were calculated based on extreme individuals indicates that these two QTLs may have effects of similar sign.

The fact that the IM and CIM methods did not give reasonable estimates of phenotypic effects should not be due solely to selective genotyping of extreme individuals. Expectation Maximization (EM)-type algorithms (Dempster 1977) used in IM and CIM should be able to handle this kind of data by replacing missing values with their expectations in the iterative scheme. The bimodality under extreme genotyping seems to be responsible for the problem. Analysis of variance of the random sample gave more reasonable results, even if normality assumptions were violated.

Estimates of phenotypic effects obtained from the analyses of variance were lower bounds because recombination will have taken place between the markers and the QTLs. Epistatic effects are usually weaker than main effects. They are downward-biased, as above, because of recombination between markers and QTLs, and are sensitive to the amount of data used for estimation. In addition to the one significant interaction detected, the possibility of other epistatic effects cannot be totally excluded.

Candidate genes

Many artificially generated mutations influence flowering time. The induced late mutations in the early background strain (L*er*) have been assigned to 12 different loci (Koornneef et al. 1991; Martinez-Zapater et al. 1994). A few early-flowering mutants have also been characterized (e.g. Zagotta et al. 1992). In addition to these, several mutants have a pleiotropic effect on flowering time (reviewed by Coupland 1995). In many cases flowering-time loci found in population studies are not located near the loci induced by mutation (Lee et al. 1993; Kowalski et al. 1994; Jansen et al. 1995; Clarke et al. 1995; Mitchell-Olds 1996). There were several candidate genes in the vicinity of the QTLs found in this study. *FRI* and *LD* are located at the top of chromosome 4 near the major QTL. Both of these are vernalization-sensitive late genes, the known late *FRI* alleles and early *LD* alleles being dominant. The partial dominance in the F_1 population suggests that the major locus is *FRI* rather than *LD*. On the upper part of chromosome 5 there are three possible candidate genes near the distal QTL, *FLC*, *FY* and *CO*. The L*er* allele at *FLC* has been shown to suppress many of the late genotypes, like late *FRI* and *LD* alleles (Koornneef et al. 1994; Lee et al. 1994). Li-5 is known to have an *FLC* allele that has an effect similar to that of the L*er* allele (Koornneef et al. 1994), suppressing the lateness of *FRI*. Thus, *FLC* probably does not segregate in our cross. Without more detailed mapping and allelism tests a definite determination is, however, not possible. The same is true for the other QTLs found. The QTL in chromosome 1 is near two flowering-time loci, *FE* and *FT*, and the second QTL on chromosome 2 is near *FPA*. Of all the putative QTLs, the second QTL on chromosme 5 is the only one not situated near QTLs found in other mapping studies between *A. thaliana* lines.

Our data provided evidence for epistasis between only one pair of the QTLs, although our physiological studies of the mutants suggest many possibilities for epistatic interactions between the flowering-time genes other than between *FLC* and *FRI* or *LD* (reviewed in Coupland 1995; Martinez-Zapater et al. 1994). In the former crossing experiments there is one reported case of epistasis (Napp-Zinn 1987) and a few examples of the F_1 being later than the later parent (Karlovska 1974; Burn et al. 1993), which may indicate interaction between the loci. Mapping studies have reported epistatic interactions in one case (Clarke et al. 1995). The lack of epistasis may in many cases be an artefact because statistical power may not allow weak epistasis to be detected in the standard mapping populations and between loci of minor effect.

Flowering-time determination in *A. thaliana* is similar to that in domesticated crop species

In crop species flowering time seems to be controlled principally in the same manner as in *A. thaliana*. Major genes are found, but polygenic effects also exist sometimes. Among the species studied, the closest relatives to *A. thaliana*, *Brassica napus* and *B. rapa*, have one and two (additive) major vernalization genes (Ferreira et al. 1995; Teutonico and Osborn 1995). In addition, there are two minor genes responsible for flowering-time difference in *B. napus*. In *B. oleracea* the annual versus biennial growth habit has been reported to be both monogenic and polygenic (reviewed in Teutonico and Osborn 1995). A comparison between *B. oleracea* and broccoli detected two major loci controlling the vernalization demand with strong epistatic interaction between them (Kennard et al. 1994).

The general picture from classical studies in grasses and legumes is that several loci control the vernalization requirement or flowering time, sometimes with

epistatic interactions (Murfet 1977). The winter versus spring habit in wheat varieties is controlled by one to four major loci, and within each class, variation is controlled by several minor loci (Murfet 1977). Flowering time in *Pisum* is controlled by four major loci with some interactions (Murfet 1977).

Recently, Paterson et al. (1995) proposed that the domestication of sorghum, rice and maize is mainly due to major genes that were selected convergently in the three species. This is in accordance with Gottlieb's (1984) idea that domestication tends to fix major genes, but the results from *A. thaliana* show that it is possible to also find major genes in non-domesticated species. Additional data should be collected before more general conclusions are drawn. In our study a major QTL was found between populations, but there is evidence that this kind of variation can also be found within populations. Within some English *A. thaliana* populations there was little variation, but others consisted of families differing widely in flowering times, probably summer and winter annuals (Jones 1971). In many cases, though, there is not much variation in flowering time between individuals, and the populations may even consist of one genotype (e.g. Kuittinen et al. 1997).

Environmental factors are known to have a large effect on flowering time in *A. thaliana* (e.g. Koornneef et al. 1991). Genotype-by-environment interactions can give rise to different sets of QTLs detected in different environments (Jansen et al. 1995; Clarke et al. 1995). The actual performance and significance of the genetic variants should be studied by carrying out QTL and physiological studies in the field. Correlations of many traits connected with life history, such as flowering time, vernalization response, seed dormancy and winter hardiness, could also be examined. It is possible that they are partly controlled by pleiotropic genes. In barley QTLs associated with winter hardiness (Hayes et al. 1993) and in *B. rapa* a QTL associated with freezing tolerance (Teutonico and Osborn 1995) were located near a vernalization-requirement QTL.

Acknowledgements We thank Anne Karjalainen and Elena Baena-Gonzalez for technical assistance, and Maarten Koornneef, Hans Koelewijn and Thomas Mitchell-Olds for comments on the manuscript. This work was supported by the Foundation for Biotechnical and Industrial Fermentation Research, the Research Council for Environment and Natural Resources, and the University of Oulu.

References

Abbott RJ, Gomes MF (1989) Population genetic structure and outcrossing rate of *Arabidopsis thaliana* (L.) Heynh. Heredity 62:411–418

Barton NH, Turelli M (1989) Evolutionary quantitative genetics: how little do we know? Annu Rev Genet 23:337–370

Baskin JM, Baskin CC (1972) Ecological life cycle and physiological ecology of seed germination of *Arabidopsis thaliana*. Can J Bot 50:353–360

Basten CJ, Weir BS, Zeng Z-B (1996) QTL CARTOGRAPHER, reference manual and tutorial for QTL mapping. North Carolina State University, anonymous ftp: esssjp.stat.ncsu.edu/pub/qtlcart

Bell CJ, Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19:137–144

Bulmer MG (1971) The effect of selection on genetic variability. Am Nat 105:201–211

Burn JE, Smyth DR, Peacock WJ, Dennis ES (1993) Genes conferring late flowering in *Arabidopsis thaliaa*. Genetica 90:147–155

Chisholm D (1989) A convenient moderate-scale procedure for obtaining DNA from bacteriophage lambda. Biotechnology 2:21–23

Churchill GA, Doerge RW (1994) Empirical threshold values for quantative trait mapping. Genetics 138:963–971

Clarke JH, Dean C (1994) Mapping *FRI*, a locus controlling flowering time and vernalization response in *Arabidopsis thaliana*. Mol Gen Genet 242:81–89

Clarke JH, Mithen R, Brown JKM, Dean C (1995) QTL analysis of flowering time in *Arabidopsis thaliana*. Mol Gen Genet 248:278–286

Conneally PM, Edwarks JH, Kidd KK, Lalouel J-M, Morton NE, Ott J, White R (1985) Report of the committee on methods of linkage analysis and reporting. Cytogenet Cell Genet 40:356–359

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

Dean C, Sjodin C, Page T, Jones, J, Lister C (1992) Behaviour of the maize transposable element *Ac* in *Arabidopsis thaliana*. Plant J 2:69–81

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc Ser B 39:1–38

Dorn LA, Mitchell-Olds T (1991) Genetics of *Brassica compestris*. 1. Genetic constraints on evolution of life-history characters. Evolution 445:371–379

Falconer DS (1989) Introduction to quantitative genetics, 3rd. edn. Longman Scientific & Technical, London

Ferreira ME, Satagopan J, Yandell BS, Williams PH, Osborn TC (1995) Mapping loci controlling vernalization requirement and flowering time in *Brassica napus*. Theor Appl Genet 90:727–732

Fisher RA (1918) The correlation between relatives under the supposition of Mendelian inheritance. Trans R Soc Edinburgh 52:399–433

Gottlieb LD (1984) Genetics and morphological evolution in plants. Am Nat 123:681–709

Härer L (1951) Die Vererbung des Blühalters früher und später sommereinjähriger Rassen von *Arabidopsis thaliana* (L.) Heynh. Beitr Biol Pflanz 28:1–34

Haldane JBS (1919) The combination of linkage values and the calculation of distances between the loci or linked factors. J Genet 8:299–309

Hauge BM, Hanley SM, Cartinhour S, Cherry JM, Goodman HM, Koornneef M, Stam P, Chang C, Kempin S, Medrano L, Meyerowitz EM (1993) An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. Plant J 3:745–754

Hayes PM, Blake T, Chen THH, Tragoonrung S, Chen F, Pan A, Liu B (1993) Quantitative trait loci on barley (*Hordeum vulgare* L.) chromosome 7 associated with components of winter hardiness. Genome 36:66–71

Hill WG, Caballero A (1992) Artificial selection experiments. Annu Rev Ecol Syst 23:287–310

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136:1447–1455

Jansen RC, Van Ooijen JW, Stam P, Lister C, Dean C (1995) Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. Theor Appl Genet 91:33–37

Jones ME (1971) The population genetics of *Arabidopsis thaliana* II. Population structure. Heredity 27:51–58

Karlovska V (1974) Genotypic control of the speed of development in *Arabidopsis thaliana* (L.) Heynh. Lines obtained from natural populations. Biol Plant 16:107–117

Karvonen P, Szmidt AE, Savolainen O (1994) Length variation in the internal transcribed spacers of ribosomal DNA in *Picea abies* and related species. Theor Appl Genet 89 : 969–974

Kennard WC, Slocum MK, Figdore SS, Osborn (1994) Genetic analysis of morphological variation in *Brassica oleracea* using molecular markers. Theor Appl Genet 87 : 721–723

Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4 : 403–410

Koornneef M, Hanhart CJ, van der Veen JH (1991) A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. Mol Gen Genet 229 : 57–66

Koornneef M, Blankestijn-de Vries H, Hanhart C, Soppe W, Peeters T (1994) The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg *erecta* wild-type. Plant J 6 : 911–919

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12 : 172–175

Kowalsi SP, Lan T-H, Feldmann KA, Paterson AH (1994) QTL-mapping of naturally occurring variation in flowering time of *Arabidopsis thaliana*. Mol Gen Genet 245 : 548–555

Kranz AR, Kirchheim B (1987) Genetic resources in *Arabidopsis*. Arabidopsis Inf Serv 24

Kuittinen H, Mattila A, Savolainen O (1997) Genetic variation at marker loci and in quantitative traits in Scandinavian natural populations of *Arabidopsis thaliana*. Heredity (in press)

Lande R (1981) The minimum number of genes contributing to quantitative variation between and within populations. Genetics 99 : 541–553

Lande R (1983) The response to selection on major and minor mutations affecting a metrical trait. Heredity 50 : 47–65

Lander E, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121 : 185–199

Lander E, Green P, Abrahamson J, Barlow A, Daley M, Lincoln S, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1 : 174–181

Lee I, Amasino RM (1995) Effect of vernalization, photoperiod, and light quality on the flowering phenotype of *Arabidopsis* plants containing the *FRIGIDA* gene. Plant Physiol 108 : 157–162

Lee I, Bleecker A, Amasino R (1993) Analysis of naturally occurring late flowering in *Arabidopsis thaliana*. Mol Gen Genet 237 : 171–176

Lee I, Michaels SD, Masshardt AS, Amasino RM (1994) The late-flowering phenotype of *FRIGIDA* and mutations in *LUMINIDEPENDENDS* is suppressed in the Landsberg *erecta* strain of *Arabidopsis*. Plant J 6 : 903–909

Lincoln S, Daly M, Lander E (1992a) Constructing genetic linkage maps with MAPMAKER/EXP Version 3.0. Whitehead Institute for Biomedical Research Technical Report, 3rd edn, Cambridge, Mass.

Lincoln S, Daly M, Lander E (1992b) Mapping genes controlling quantitative traits using MAPMAKER/QTL version 1.1 Whitehead Institute for Biomedical Research Technical Report, 2nd edn, Cambridge, Mass.

Martinez-Zapater JM, Coupland G, Dean C, Koornneef M (1994) The transition to Flowering in *Arabidopsis*. In: Somerville CR, Meyerowitz EM (eds) *Arabidopsis*. Cold Spring Harbor Laboratory, Press, New York, pp 403–434

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

Murfet IC (1977) Environmental interaction and the genetics of flowering. Annu Rev Plant Physiol 28 : 253–278

Napp-Zinn K (1957a) Untersuchungen zur Genetik des Kältebedürfnisses bei *Arabidopsis thaliana*. Zeitschrift für indukt. Abstammungs Vererbundslähre 88 : 253–285

Napp-Zinn K (1957b) Untersuchungen über das Vernalisationsverhalten einer winterannuellen Rasse von *Arabidopsis thaliana*. Planta 50 : 177–210

Napp-Zinn K (1962) Über die Genetischen Grundlagen des Vernalisationsbedürfnisses bei *Arabidopsis thaliana*. I. Die Zahl der beteiligten Faktoren. Z Vererbungs 93 : 154–163

Napp-Zinn K (1969) *Arabidopsis thaliana* (L.) Heynh. In: Evans LT (ed) The induction of flowering: some case histories. MacMillan, Melbourne, pp 291–304

Napp-Zinn K (1987) Vernalization – environmental and genetic regulation. In: Atherton JG (ed) Manipulation of flowering. Butterworths, London pp 123–132

Orr HA, Coyne JA (1992) The genetics of adaptation: a reassessment. Am Nat 140 : 725–742

Paterson A, Lander E, Lincoln S, Hewitt J, Peterson S, Tanksley S (1988) Resolution of quantitative traits into Mendelian factors using a complete RFLP linkage map. Nature 335 : 721–726

Paterson AH, Lin Y-R, Li Z, Schertz KF, Doebley JF, Pinson SRM, Liu S-C, Stansel JW, Irvine JE (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science 269 : 1714–1718

Ratcliffe D (1961) Adaptation to habitat in a group of annual plants. J Ecol 49 : 187–203

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5 : 69–76

Säll T, Nilsson N-O, Bengtsson B-O (1993) When everyone's map is different. Curr Biol 3 : 631–633

Sambrook J, Fritch EF, Maniatis T (1989) Molecular cloning. A laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, New York

Sax K (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8 : 552–560

Stuber CW, Edwards MD, Wendell JF (1987) Molecular marker-facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. Crop Sci 27 : 639–648

Symonides E (1987) Population dynamics of annual plants. In: Davy AJ, Hutchings MJ, Watkinson AR (eds) 28th Symp Br Ecol Soc. Blackwell, London pp 221–248

Teutonico RA, Osborn TC (1995) Mapping loci controlling vernalization requirement in *Brassica rapa*. Theor Appl Genet 91 : 1279–1283

Vizir IY, Korol AB (1990) Sex difference in recombination frequency in *Arabidopsis*. Heredity 65 : 379–383

Zagotta MT, Shannon S, Jacobs C, Meeks-Wagner DR (1992) Early-flowering mutants of *Arabidopsis thaliana*. Aust J Plant Physiol 19 : 411–418

Zeng Z-B (1994) Precision mapping of quantitative trait loci. Genetics 136 : 1457–1468