Theor Appl Genet (2002) 104:1173–1184
DOI 10.1007/s00122-001-0825-9

O. Loudet · S. Chaillou · C. Camilleri · D. Bouchez
F. Daniel-Vedele

Bay-0 × Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in Arabidopsis

Received: 3 July 2001 / Accepted: 10 September 2001 / Published online: 13 February 2002
© Springer-Verlag 2002

Abstract Natural genetic variation in Arabidopsis is considerable, but has not yet been used extensively as a source of variants to identify new genes of interest. From the cross between two genetically distant ecotypes, Bay-0 and Shahdara, we generated a Recombinant Inbred Line (RIL) population dedicated to Quantitative Trait Locus (QTL) mapping. A set of 38 physically anchored microsatellite markers was created to construct a robust genetic map from the 420 F6 lines. These markers, evenly distributed throughout the five chromosomes, revealed a remarkable equilibrium in the segregation of parental alleles in the genome. As a model character, we have analysed the genetic basis of variation in flowering time in two different environments. The simultaneous mapping of both large- and small-effect QTLs responsible for this variation explained 90% of the total genotypic variance. Two of the detected QTLs colocalize very precisely with *FRIGIDA* and *FLOWERING LOCUS C* genes; we provide information on the polymorphism of genes confirming this hypothesis. Another QTL maps in a region where no QTL had been found previously for this trait. This confirms the accuracy of QTL detection using the Bay-0 × Shahdara RIL population, which constitutes the largest in size available so far in Arabidopsis. As an alternative to mutant analysis, this population represents a powerful tool which is currently being used to undertake the genetic dissection of complex metabolic pathways.

Keywords Arabidopsis · RILs population · Microsatellite markers · QTL · Candidate gene

Communicated by L. Willmitzer

O. Loudet · S. Chaillou · F. Daniel-Vedele (✉)
INRA, Unité de Nutrition Azotée des Plantes,
Centre de Versailles, 78 026 Versailles, France
e-mail: vedele@versailles.inra.fr
Tel.: (33)-1-30-83-30-67, Fax: (33)-1-30-83-30-96

C. Camilleri · D. Bouchez
INRA, Station de Génétique et Amélioration des Plantes,
Centre de Versailles, 78 026 Versailles, France

Introduction

To-date, among the 25,000 coding sequences distributed throughout the *Arabidopsis thaliana* genome, only 9% have an associated physiological function. In addition, more than 30% of these putative proteins do not show any similarity with proteins from other organisms and there is no indication as to their possible role in the cell or plant (The Arabidopsis Genome Initiative 2000). The scientific challenge over the next few years will be to assign a precise function to each of these sequences.

Forward, and the more recent reverse, genetics will help in the identification of new genes/functions by physiological studies of mutants affected in the expression, under- or over-expression, of a particular gene. However, since many agronomic traits are not under the control of unique but rather multiple loci, phenotypic analyses of such mutants may not provide all the information concerning a target gene's function. In contrast, Quantitative Trait Locus (QTL) mapping is designed to dissect the genetic architecture of such characters; for plants, it was first applied to maize (Edwards et al. 1987) and tomato (Paterson et al. 1988). Moreover, it allows the integration of gene function at the level of whole-plant regulation. In Arabidopsis, this approach represents a powerful alternative to classical genetics as the small size of the Arabidopsis genome, combined to its complete sequencing, makes it possible to follow-up QTL mapping experiments by studies at the molecular level. The extensive natural variation that occurs in Arabidopsis is being increasingly exploited as a source of genetic variability for the analysis of important agronomic traits (review in Alonso-Blanco and Koornneef 2000). Indeed, genes involved for example in disease resistance, or more recently in flowering time, have been cloned (Buell and Somerville 1997; Michaels and Amasino 1999; Johanson et al. 2000), allowing investigation of the molecular basis of the allelic variation. The first step required for the successful cloning of genes associated with major QTLs is the generation of suitable mapping populations.

The first populations constructed for the analysis of marker segregation were F2/F3 families (Koornneef et al. 1983; Chang et al. 1988; Nam et al. 1989). However, due to the heterozygosity of this genetic material, the physiological analysis of a large number of individuals was generally required to reach statistically significant results, and seed stocks could be a limiting factor. Furthermore, the mapping of a unique segregating population using different markers in different laboratories allows the continuous growing-up of a central database and the accuracy of this map. This common mapping leads to a more-reliable map than that derived from the statistical integration of data obtained from similarly sized but different mapping populations. Recombinant inbred lines (RILs) provide such an "immortal" population as each individual is practically homozygous and can be propagated indefinitely as a clone. This is advantageous for both genotyping and phenotyping, which can always be carried out on the same genetic material.

The two Arabidopsis RIL populations that have been used most frequently have been derived mainly from laboratory accessions, namely Ler/Col (Lister and Dean 1993) and Ler/Cvi (Alonso-Blanco et al. 1998c). Genetic variation has been found in these populations for traits such as freezing, drought or ozone tolerance, flowering time, plant or seed size, seed dormancy and pathogen resistance (review in Alonso-Blanco and Koornneef 2000). The generation of other RILs from exotic accessions will be interesting in order to reveal variation for other agronomic traits and to analyse the persistence and stability of previously identified QTLs in different genetic backgrounds. Creating such populations requires significant economic and time investments. For this reason, care must be taken over the choice of the parental accessions, taking into account observed phenotypic variations as well as genetic distance and phylogenetic studies. Recent studies show, in particular, that Central-Asian accessions represent an original material, genetically distant from globally unstructured European accessions (Innan et al. 1997; Breyne et al. 1999; Erschadi et al. 2000; Sharbel et al. 2000). For this reason, two ecotypes, Bay-0 and Shahdara, were chosen for the construction of a new RIL population: the phenotypic variation resulting from such a cross is expected to reflect the adaptation to their specific habitat and the genetic distance between them.

The mapping of segregating populations is generally achieved using RFLPs (restriction fragment length Polymorphisms) or PCR-based markers. As opposed to the former, which are co-dominant, reliable but laborious (Nam et al. 1989), the latter can be used routinely in any laboratory. RAPD markers (random amplified polymorphic DNAs) are dominant and difficult to reproduce (Reiter et al. 1992), whereas AFLP markers (amplified fragment length polymorphism) are theoretically unlimited, consistently efficient and provide a multi-locus polymorphism (Vos et al. 1995). These markers were used to generate linkage maps of Ler/Col and Ler/Cvi Arabidopsis populations (Alonso-Blanco et al. 1998c). From this study, it appears that, with some exceptions, it is difficult to consider AFLP markers as co-dominant; and, unless the allele has been sequenced, AFLP markers are not physically anchored. Furthermore, despite a regular distribution of AFLPs among the five chromosomes, their distribution within each linkage group is not always uniform, and this clustering can prevent the complete ordering of all the markers. The completion of the sequencing of the Arabidopsis genome now allows the selection of highly polymorphic codominant markers, such as microsatellites, evenly distributed throughout the genome (Casacuberta et al. 2000). This type of marker seems to be particularly suitable for the construction of a high-confidence framework map, that could also be used to generate a high-density linkage map following, for example, the strategy of selective mapping (Vision et al. 2000).

Here, we describe the generation and genetic mapping of a large family of Arabidopsis RILs dedicated to QTL analysis. As an illustration of the potential of the population, the mapping of seven QTLs responsible for most of the observed variability in flowering time is presented.

Materials and methods

Generation of the RIL population

The ecotypes Bay-0 and Shahdara were crossed using Bay-0 as the female parent. Bay-0 originates from a fallow-land habitat near Bayreuth in Germany (Reference N954 and CS954 respectively in the Nottingham Arabidopsis Stock Centre and Arabidopsis Biological Resource Center catalogs), whereas Shahdara grows in the Pamiro-Alay mountains in Tadjikistan (Reference N929 and CS929 respectively in the NASC and ABRC catalogs). Note: Shahdara is also spelled Shakdara in the NASC catalog, and sometimes Shokhdara. Seeds obtained from the stock centre were initially sown and a unique plant for each ecotype selfed to ensure full homozygosity of the parental plants involved in the crossing. The F1 seeds were sown and the plants obtained were checked with two markers showing polymorphism between Bay-0 and Shahdara (NGA128 and NGA172; see Table 1). One heterozygous F1 plant was chosen and self-fertilized. F2 seeds were sown individually and each plant was self-fertilized. Then, three additional cycles of single seed descent (SSD) were performed using a design aimed at minimizing any bias in the selection of the plant to be selfed: 15 seeds of each line were sown in a pot and a single plant was randomly selected from its position on the pot and allowed to self-fertilize. A last cycle of SSD was performed to produce F7 seeds for each F6 recombinant inbred line: this generation of selfing was performed between November 1999 and February 2000, in a greenhouse with a controlled temperature and an additional light supply providing long-day growing conditions. All the RILs were thus multiplied in the same conditions, ensuring a homogeneous material for phenotypic analyses.

Genetic mapping

All the markers used to construct our genetic map are PCR-based markers which reveal a length-polymorphism at microsatellite loci (SSR, simple sequence repeats). Table 1 presents the information concerning the 38 microsatellite markers used to construct a genetic map of the Bay-0 × Shahdara RIL population. Most of these markers (MSATx.x) have been newly defined from the numerous sequenced microsatellite loci, according to their position on the chromosomes, using the strategy described below. Primers sequences were placed on the whole-chromosome sequences (called

Table 1 The 38 microsatellite markers used to construct the Bay-0 × Shahdara genetic map

Marker[a]	Chrom.	BAC/P1[b]	Pattern	5′ –3′ Forward Primer	5′ –3′ Reverse Primer	Col Length (bp)[c]	Position (cM)[d]
T1G11	1	T1G11	A	GAAGACAAAGCTCTGCAGTAATG	AATTGCATAAGGCACTTGAAAG	207	0.0
F21M12	1	F21M12	GAAA	GGCTTTCTCGAAATCTGTCC	TTACTTTTTGCCTCTTGTCATTG	201	5.7
MSAT1.10	1	F9H16	AT	ATGGTGAGATACTGAGATTAT	CGAGAAGGTCTAAAGGTA	235	15.7
NGA248	1	F3H9	AG	TACCGAACCAAAACACAAAGG	TCTGTATCTCGGTGAATTCTCC	142	26.9
T27K12	1	F7F22	AT	GGACAACGTCTCAAACGGTT	GGAGGCTATACGAATCTTGACA	146	43.6
NGA128	1	F7A10	AG	GGTCTGTTGATGTCGTAAGTCG	ATCTTGAAACCTTTAGGGAGGG	180	56.3
F5I14	1	F5I14	A	CTGCCTGAAATTGTCGAAAC	GGCATCACAGTTCTGATTCC	196	63.7
MSAT1.13	1	F24J5	AT	CAACCACCAGGCTC	GTCAAACCAGTTCAATCA	221	70.4
MSAT1.5	1	T14N5	AG	GCATCGCTCTTAAACAACCAT	CGTTGCAAAACCGTATCAGAA	157	84.1
MSAT2.5	2	F2I9	AG	TGAGAGGGACAGATAGGAA	ATCAAAAGGGATACTGACAA	227	0.0
MSAT2.38	2	F18P14	AT	TGTAACGCTAATTTAATTGG	CGCTCTTTCGCTCTG	180	13.0
MSAT2.36	2	T2G17	AG	CCAAGAACTCAAAACCGTT	GATCTGCCTCTTGATCAGC	158	26.8
MSAT2.41	2	T19L18	AT	ACAAACCATTGTTGGTCGTG	GACTGTTTCATCGGATCCAT	144	34.5
MSAT2.7	2	F7F1	AG	CTCAAATCAAGAACGCTGAC	CCCGATATAGACAACGACAA	251	42.7
MSAT2.10	2	T1O24	AG	ACAAACATGTTCTGGGTTA	ATTCTTCATTATCTGCTGCT	299	55.9
MSAT2.22	2	F17A22	AT	CGATCCAATCGGTCTCTCT	TGGTAACATCCCGAACTTC	248	62.5
NGA172	3	T21P5	AG	AGCTGCTTCCTTATAGCGTCC	CATCCGAATGCCATTGTTC	166	0.0
ATHCHIB2	3	T2E22	AT	GGATCCAAGTGCTCATATATAC	CTTTCGTTTCTAAATATGAGAAGC	110	6.8
MSAT3.19	3	K7M2	AT	TAATTCGATCCAATTGACAT	TGGCTTGGCACAAAC	171	23.9
MSAT3.32	3	MX021	AT	GCACTTGCAGCTTAACTT	CGTGACTGTCAAACCG	173	41.1
MSAT3.21	3	T6H20	AT	TTACCCCGAGCTTGA	TGAATCATGGTGCTTCTA	179	49.4
MSAT3.18	3	F15B8	AT	TCATACCTACATATTGCCCT	TACCTCAAAAGAGCAAACA	267	65.7
MSAT4.39	4	F6N15	AT	GTTATCACATTAAAATCACC	CCAATTGTAATATATGAACA	161	0.0
MSAT4.8	4	T18A10	AG	CGGGTAAAGACAGAGCAT	GTTGGGTTTAGTTGGTAACA	202	2.0
NGA8	4	T32A17	AG	GAGGGCAAATCTTTATTTCGG	TGGCTTTCGTTTATAAACATCC	157	16.4
MSAT4.35	4	F25G13	AT	CCCATGTCTCCGATGA	GGCGTTTAATTTGCATTCT	217	24.8
MSAT4.15	4	FCA6	AG	TTTCTTGTCTTTCCCCTGAA	GACGAAGAAGGAGACGAAAA	174	34.1
MSAT4.18	4	T12H17	AT	TGTAAATATCGGCTTCTAAG	CTGAAACAAATCGCATTA	159	47.4
MSAT4.9	4	F4D11	AG	AAGTAATTAAGACGCTGAGA	GAAATCAACGGCTGAG	235	55.9
MSAT4.37	4	F23K16	AT	CGTTTCATCAAGTTCCGA	TAGGAGGTTATCATGCGTG	139	69.9
NGA225	5	MUG13	AG	GAAATCCAAATCCCAGAGAGG	TCTCCCCACTAGTTTTGTGTCC	120	0.0
NGA249	5	MAH20	AG	TACCGTCAATTTCATCGCC	GGATCCCTAACTGTAAAATCCC	125	3.1
MSAT5.14	5	MQJ16	AT	AACAACCCTATCTTCTTCTG	TGTGACCCCTTACTCAATA	221	18.0
NGA139	5	K18P6	AG	AGAGCTACCAGATCCGATGG	GGTTTCGTTTCACTATCCAGG	182	21.8
MSAT5.22	5	MWP19	AT	AGAACAAGTTAGGTGGCT	GGGACAAGAATGGAGT	248	36.4
MSAT5.9	5	MBD2	AG	CGTCATTTTTCGCCGCTCT	CATGGTGGCGCGTAGCTTA	208	49.2
MSAT5.12	5	MXC20	AT	GCATATTGTTGATAGAAAA	AGCCAATGAATCGTT	155	63.9
MSAT5.19	5	MXK3	AT	AACGCATTTGCTGTTTCCCA	ATGGTTATCTCATCTGGTCT	208	75.7

[a] MSAT-labelled markers (26 among the 38) are new markers defined as described in the Results. Other markers are from the TAIR web site (http://www.arabidopsis.org)

[b] BAC and P1 clone in which the marker is situated, from the TAIR web site

[c] Columbia amplification length is estimated from the available sequence

[d] Distance (in cM) from the first marker of the linkage group in the Bay-0 × Shahdara RIL population

'pseudo-molecules') generated by MIPS (The Arabidopsis Genome Initiative 2000), displayed at ftp://ftp.mips.biochem.mpg.de/pub/cress. We used the V211200b version of the 'pseudo-molecules'.

The 420 F6 plants available in the last cycle of SSD (multiplication cycle) were included in the genetic analysis. DNA from the randomly chosen F6 plant was extracted using a high-throughput method as follows: one rosette-leaf (+/– 50 mg of fresh weight) was cut off from each plant approximately 35 days after sowing, and stored in a 96-well Polypropylene storage plate (1-ml well volume). The plate was then frozen in liquid-N_2 and freeze-dried for 2 days. A small glass-bead (4 mm) was added to each well before the plate was placed on a vibrator thus grinding the dry leaves. A CTAB extraction was then performed in the 96-well plate using 300 µl of extraction buffer per well, 30 min at 60 °C with periodic agitation (extraction buffer: CTAB 2%; NaCl 1.4 M; EDTA 20 mM; Tris HCl pH 8 100 mM; β-mercaptoethanol 0.2%). Then, 300 µl of chloroform/isoamyl alcohol (24:1) was added and the plate was placed with its sealing-mat on a small screw-press to ensure hermetic sealing and avoid any cross-contamination between the wells. The plate was then well-shaken and centrifuged, before the aqueous phase was transfered to another 96-well microtitre plate. DNA was precipitated, washed, dissolved in 100 µl of TE, and then diluted 10-times before being used for PCR reactions (5 µl per reaction).

All 38 markers were screened on the 420 RILs, with the same PCR (37 cycles; each cycle: 15 s denaturation at 94 °C, 15 s annealing at 50 °C; 30 s elongation at 72 °C) and electrophoresis (3% agarose gel) conditions.

The segregation of parental alleles for each locus was studied in detail and the recombination fractions were calculated and converted into map order and distances, using a Unix version of MAPMAKER 3.0 (Lander et al. 1987; Lincoln et al. 1993) supplied at http://www-genome.wi.mit.edu/ftp/distribution/mapmaker3/. Unlike previous versions, MAPMAKER 3.0 is designed to construct a genetic map from RIL population data. The Kosambi mapping function (Kosambi 1944) was used to convert recombination fractions into map distances in the Mapmaker program, as it takes into account interference between crossing-over and seems to be well-suited for Arabidopsis linkage analysis (personal observations on distance additivity, data not shown; Koornneef et al. 1998b).

Standard error for the map distance estimation was calculated according to formulas established by Fisher (1937) and Allard (1956), and summarized for RILs in Koornneef et al. (1998b) and Alonso-Blanco et al. (1998b).

Flowering-time QTLs: phenotyping

Flowering precocity was measured on the RILs in two distinct day length conditions: long day (LD) and short day (SD) conditions. The LD experiment was carried out during the last cycle of SSD (F6 to F7), in a greenhouse, with controlled temperature conditions and additional lighting to ensure a 16-h photoperiod (Photosynthetic Photon Flux Density: approximately 100 µmol m^{-2} s^{-1}). The 420 randomly selected F6 plants (the plants from which we obtained F7 seeds and DNA) were observed in these conditions. The SD experiment was carried out between November 2000 and February 2001, in a greenhouse isolated from any interfering artificial light. Sunlight was complemented with approximately 100 µmol m^{-2} s^{-1} (PPFD), ensuring a 8-h photoperiod. The 415 RILs were observed in a randomized complete-block display with two replicates (two plants observed per line).

Flowering time was recorded as the number of days between seed germination and bolting (i.e. the time when the principal bud starts to emerge with the beginning of bolt extension: between 0.5 and 1 cm long). This trait is easily observed and seems to be as robust as total leaf number (Clarke et al. 1995), another method of estimating flowering precocity. This observation was made every other day.

Flowering-time QTLs: data analysis

All QTL analyses were performed using the Unix version of QTL Cartographer 1.14 (Basten et al. 1994, 2000), supplied via http://statgen.ncsu.edu/qtlcart/cartographer.html. First, simple Interval Mapping (Lander and Botstein 1989) was used to determine putative QTLs involved in the variation of the trait. Composite Interval Mapping (CIM) Model 6 of QTL Cartographer 1.14 (Basten et al. 2000) was then performed on the same data: the closest marker to each local LOD Score peak (putative QTL) was used as a cofactor to control the genetic background while testing at a position of the genome. When a cofactor was also a flanking marker of the tested region, it was excluded from the model. The number of cofactors involved in our models was four for SD data and three for LD data. The walking speed chosen for all QTL analysis was 0.1 cM. Additive effects (2a in Table 2) of detected QTLs were estimated from CIM results; 2a represents the mean effect of the replacement of the Shahdara allele by the Bay-0 allele at the studied locus. The LOD significance threshold was calculated by a permutation-test analysis, as suggested by Churchill and Doerge (1994); 1,000 permutations of phenotypic data were analysed using the CIM model with the specific conditions described above, for SD and LD data. The so-called 'experimentwise threshold' obtained is 2.27 and 2.35 (LOD), respectively, for SD and LD data (overall error level: 5%). The running time of these permutation tests on a Sun Sparc Ultra 60 computer running SunOS 5.6, was approximately 15 h. One-LOD support interval of each detected QTL gives information regarding the precision of the estimated position. It should be noted that they represent anti-conservative Confidence Intervals, as shown in simulations by Visscher et al. (1996).

The LD experiment was also analysed using transformed data (log10), to fit a more-normal distribution. Such analyses did not improve the results of QTL mapping and are, therefore, not shown.

The SD experiment was first analysed using the RIL number and block, as classifying factors in an analysis of variance (ANOVA), to determine the RIL effect. No significant block effect was observed, thus this factor was removed from the model. Broad-sense heritability was calculated as the ratio between estimated Genetic Variance and Total Phenotypic Variance (= Genetic Variance + Environmental Variance). Subsequent analyses involved unadjusted mean values from the two observed plants in the SD experiment, because of the absence of a block effect.

The contribution of each identified QTL to the total variance (R^2) was estimated by variance-component analysis, as the estimations obtained seem to be unbiased (Charcosset and Gallais 1996). In each environment (SD and LD), the model involved the genotype at the closest marker to the corresponding detected QTL as random factors in ANOVA (the same markers used as cofactors in CIM mapping). Only homozygous genotypes were included in the ANOVA analysis. Significant QTL × QTL interactions were also added to the linear model via the corresponding marker × marker interactions, and their contribution to the total variance was also estimated. QTL × environment interaction was assessed by a two-factor ANOVA of flowering time, with the corresponding marker genotype and environment as classifying factors. ANOVA analyses were performed using *lm*() and *aov*() functions of the S-PLUS 3.4 statistical package (Statistical Sciences, Inc.).

Results

A new RIL population

The effective cross between Shahdara and Bay-0 was verified on the F1 plant from which seeds were collected to give F2 plants; heterozygosity of markers NGA128 and NGA172 was confirmed (data not shown). F2 plants were, on average, more vigourous than the parents, reflecting a probable effect of the recombination of the parental genomes and the heterozygous status of the plants; 432 of these plants were harvested (F3 seeds) and then underwent three cycles of selfing by single seed descent (SSD) until the F6 generation, where 420 lines were still available for genotyping analysis. This 3% loss was mostly due to plant sterility.

The last cycle of selfing was designed to produce a sufficient amount of F7 seeds for further phenotyping analysis, assuming that the very low expected rate of residual heterozygosity of F6 plants (approximately 3%) will produce only very low residual variability within the descendants of a F6 plant.

Selection of markers

The availability of the complete genome sequence of Arabidopsis prompted us to try to construct a genetic map of the Bay-0 × Shahdara population that should be optimum for QTL mapping. Our principal goals were to ensure an even distribution of the markers, with no major gap between adjacent ones, and to minimize as much as possible the rates of mistyping and missing data, which would then reduce the power of QTL mapping (Hyne et al. 1995). Due to the large number of RILs to be mapped, it was important to minimise the number of markers used, and the genotyping should be possible using standard techniques. For these reasons we elaborated a strategy to design a complete microsatellite-based genetic map.

We selected microsatellite motifs [(AT)*n* or (AG)*n*, where $n > 12$] located on BACs regularly spaced on the five chromosomes. PCR primers were then derived from

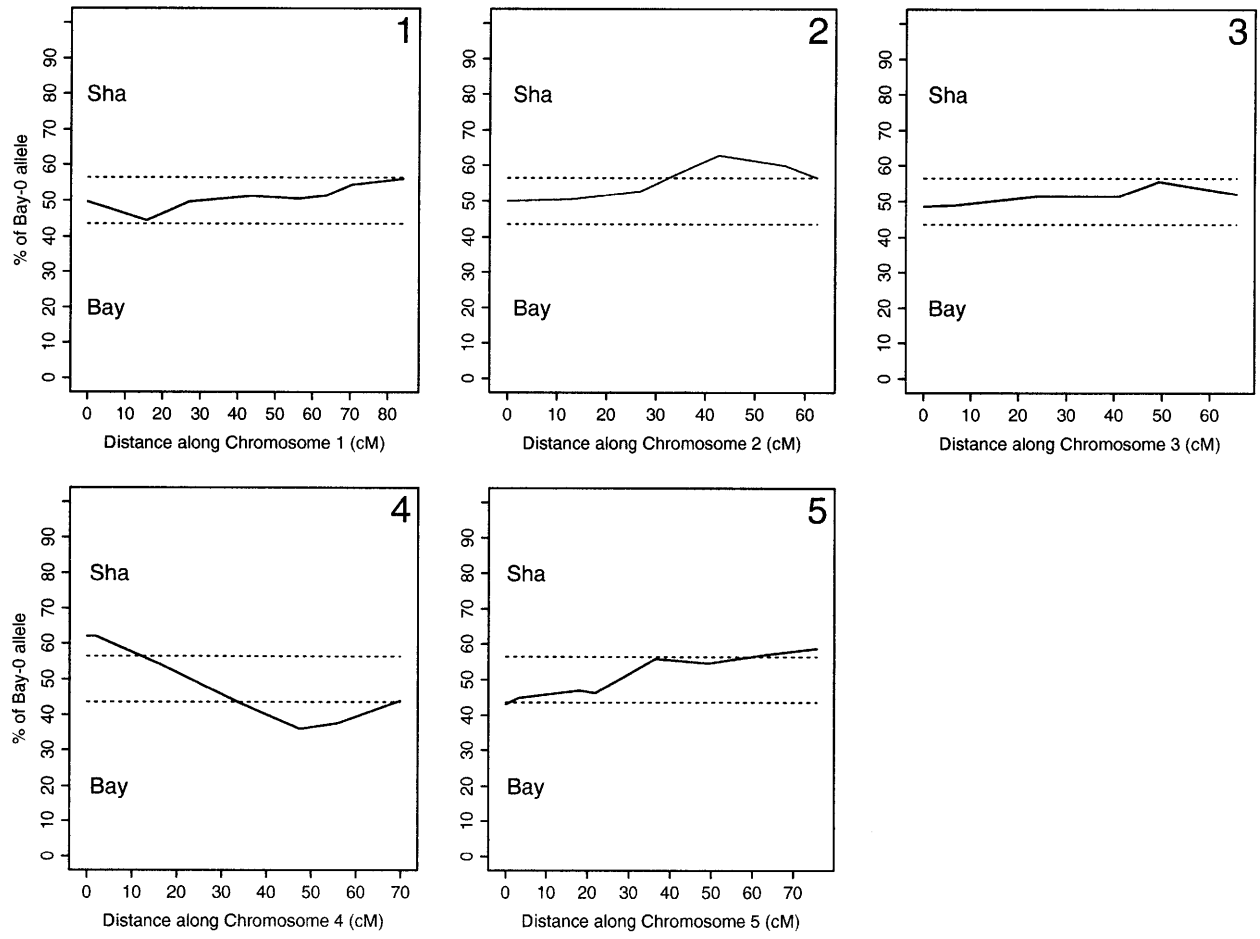

Fig. 1 Segregation between Bay-0 and Shahdara alleles along the five chromosomes. The percentage of the Bay-0 allele is represented along the y axis. The *upper and lower dotted lines* on each graph represent the 56.5% and 43.5% thresholds (at the 1% error level)

their flanking sequences. When possible, we used already existing microsatellite markers. In most cases, a specific PCR-product was obtained for the five ecotypes tested (Col, Ler, WS, Bay-0, Shahdara), with some exceptions, mostly due to an ecotype-specific inefficient PCR. Among the approximately 100 primer pairs which amplified loci, the global Bay-0/Shahdara polymorphism rate, as determined by migration on agarose gels, was more than 80%. An equivalent result was obtained, using the same loci, between Col and Ler. This confirms the very high level of polymorphism found for microsatellite markers among Arabidopsis ecotypes. However, only 50% of the polymorphic loci were appropriate for high-throughput genotyping performed on agarose gels; the three major sources of rejection were: (1) unbalanced amplification efficiency between Bay-0 and Shahdara alleles; (2) unspecific amplification interfering with Bay-0 or Shahdara PCR products; (3) insufficient size difference between the two alleles. Despite these defects, it was always possible to select additional markers where needed. The numerous markers from the Ler/Col RI map have identified genetic and physical positions. They served as a guide to select microsatellite markers at loci situated every 10–15 cM along the chromosomes. The expected distance between adjacent microsatellite markers chosen should never exceed 20 cM. Table 1 shows detailed information concerning the 38 markers that were chosen to construct our genetic map of the Bay-0 × Shahdara population.

Genetic mapping

DNA extraction was performed on tissue from the F6 plant that was randomly chosen to be selfed during the last cycle of SSD. The high-throughput method used yielded an amount of DNA sufficient for several hundred PCRs, ensuring that the whole genotyping is performed on a unique DNA source. Of 15,960 data points (420 RILs × 38 Markers), only 0.23% are missing (generally due either to the absence of an amplified product or to an uncertainty in interpretation of the gel data); this concerned only six markers. This very low level of missing data ensures a high degree of precision for genetic-distance estimations, and maximises the capabilities of QTL mapping methods; 3.05% of the genotype data showed residual heterozygosity, which is very close to the value expected for F6 Lines (3.12%) when no selection maintaining heterozygous plants has occurred. This also indicates that the DNA extraction method used probably did

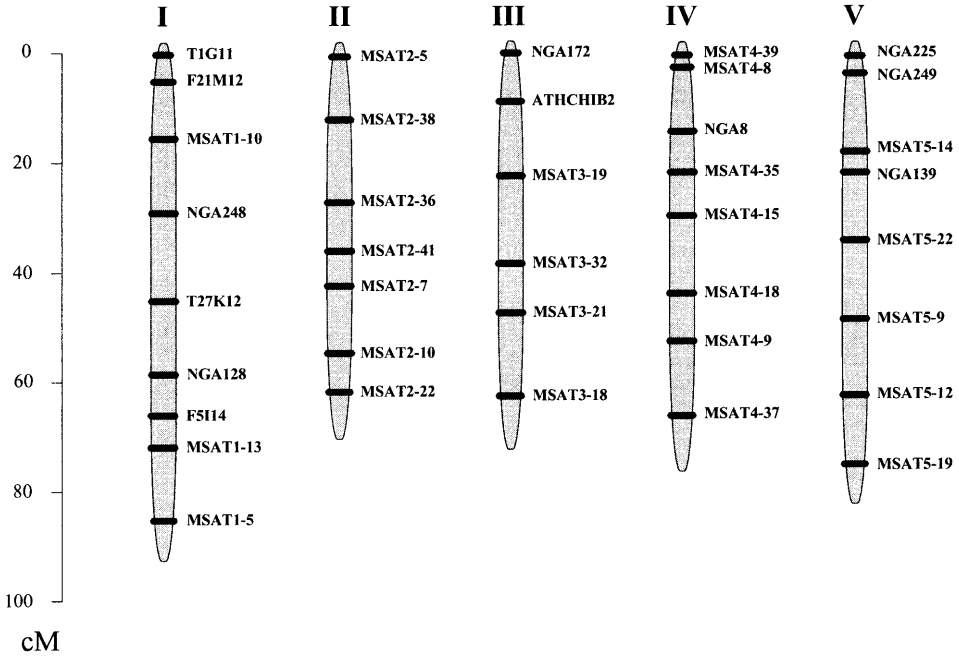

Fig. 2 Genetic map obtained from the Bay-0 × Shahdara RIL population. The 38 microsatellite markers are described in Table 1

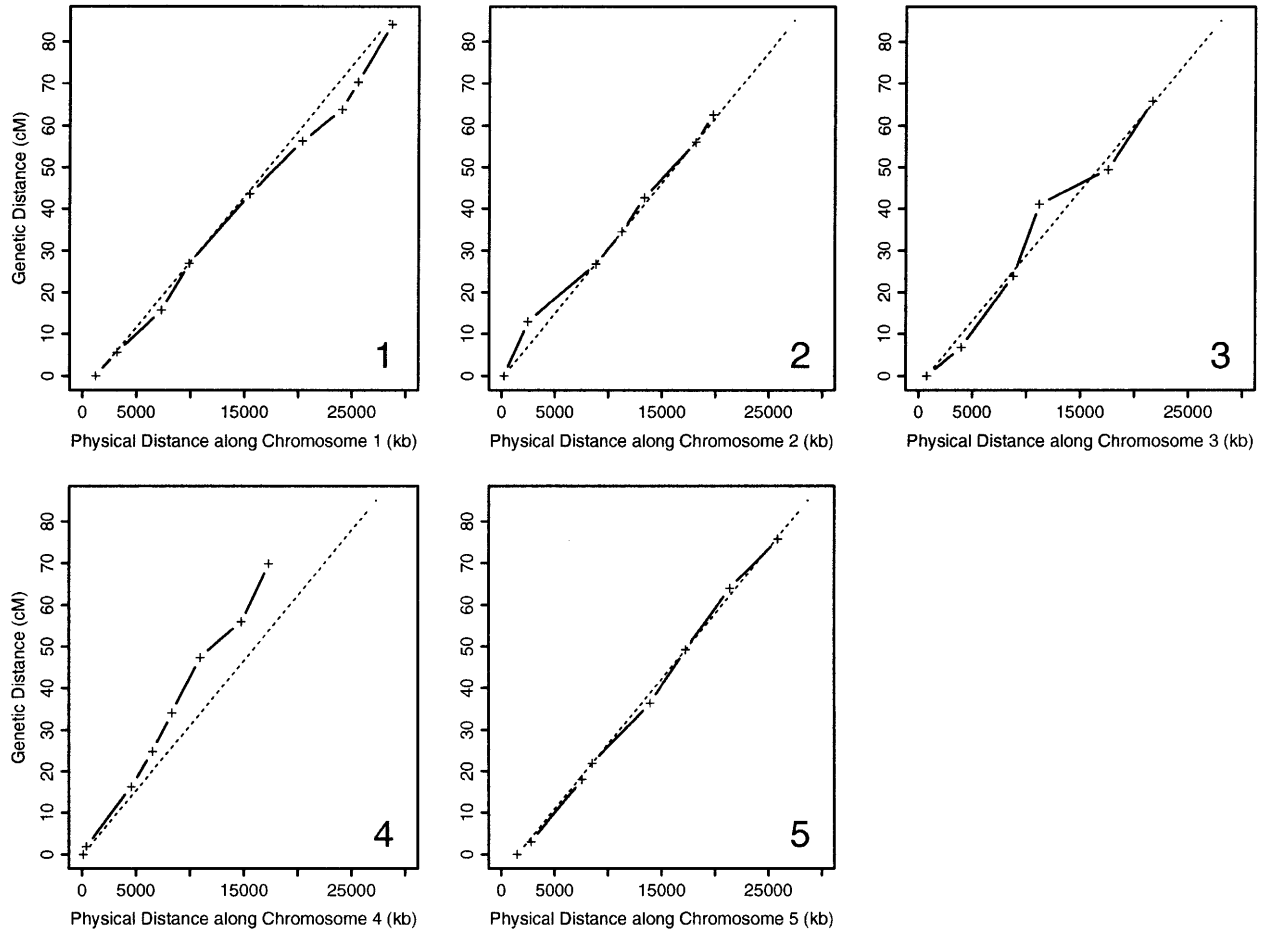

Fig. 3. Relation between genetic and physical length along the five chromosomes in the Bay-0 × Shahdara population. The *dotted line* represents the whole-genome mean ratio (320 kb/cM)

not result in any cross-contamination, which would have significantly increased this percentage. Indeed, no RIL showed more than six heterozygous loci among the 38 tested (data not shown). Among the 38 marker loci studied in the Bay-0 × Shahdara population, residual heterozygosity varied from 1.2% to 5.2%, and was never significantly different from the expected value of 3.12% (as can be tested by a chi-square test at the 1% level). At the population level, the segregation ratio of the two parental alleles was very close to that expected: 51/49% with a slight bias towards Bay-0. Figure 1 shows the variation of the segregation ratio between Bay-0 and Shahdara parental alleles along the five chromosomes. The segregation ratio of a marker should stay between the 43.5% and 56.5% thresholds, if no distortion has occurred around this locus (at the 1% error level). Two distinct regions clearly show a significant deviation from this equilibrium: the bottom of chromosome 2 (between markers MSAT2.41 and MSAT2.10) shows an excess of Bay-0 alleles and the bottom of chromosome 4 (between markers MSAT4.15 and MSAT4.9) shows an excess of Shahdara alleles. In both cases, the magnitude of distortion does not exceed 64%/36% (approximately 1:1.8).

The genetic map obtained with Mapmaker 3.0 is presented in Fig. 2. All markers have been assigned to the desired linkage group, and, without any ambiguity, in the expected order. The five linkage groups of our genetic map represent a total of 358 cM, equivalent to the 373 cM of the original Ler × Col genetic map, constructed with 65 RFLP markers (Lister and Dean 1993). The difference in total genetic length with the Alonso-Blanco et al. (1998c) AFLP-based linkage map of the same Ler × Col population (427 cM), is explained by the addition of new markers at the previously unmapped distal extremities of the chromosomes.

The average genetic distance between two adjacent markers on our map is 10.8 cM, with a minimum of 2.0 cM (MSAT4.39–MSAT4.8 on chromosome 4) and a maximum of 17.2 cM (MSAT3.19–MSAT3.32 on chromosome 3). It appears that 2/3 of the loci in the Bay-0 × Shahdara population are localized at less than 7 cM from a given marker. The precision associated with the estimation of a map distance of 10.8 cM in a 420 RIL population is very high: the assigned mapping standard error (see Materials and methods) is below 0.03 cM. Furthermore, the genotyping of our 420 RILs allows us to observe, on average, 475 effective recombination events per chromosome, which represents one recombination event every 46 kb. Figure 3 details the density of recombination along the five chromosomes: cumulative genetic lengths are plotted against cumulative physical lengths. For chromosomes 1, 2, 3 and 5, we can see that, except in some restricted regions, the recombinations are globally evenly distributed along the chromosomes, following the mean relative ratio of 320 kb per cM. Chromosome 4 also shows a linear relationship between physical and genetic length, althought the slope (density) differs: on average, the physical/genetic ratio for this chromosome is 246 kb per cM. There are proportionally more

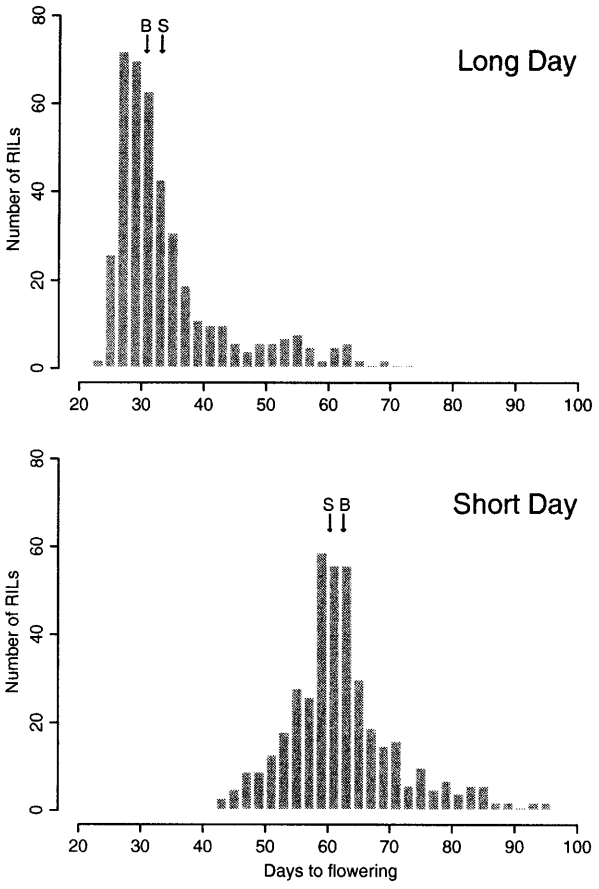

Fig. 4 Distribution of the flowering-time phenotypes of the 420 RILs derived from the Bay-0 × Shahdara cross, in long day (*LD*) and short day (*SD*) experiments. The parent phenotype is represented by an *arrow* (*S* for Shahdara; *B* for Bay-0)

crossing-over events on chromosome 4 than on the other chromosomes. This can be partly explained by the fact that chromosome 4 is, physically, the shortest chromosome. If one assumes that each meiosis structurally requires one crossing-over by each bivalent chromosome (Roeder 1997), such an inversely proportional relationship between recombination frequencies and chromosome size is not surprising (Copenhaver et al. 1998).

Flowering-time QTLs

Despite the very limited difference in flowering-time phenotype between Bay-0 and Shahdara in both SD (short day) and LD (long day) conditions, we observed a strong and highly significant ($P < 0.001$ in SD) variability between the RILs (Fig. 4). Most of the lines flower before the most early flowering parent or after the later-flowering one; this transgression in both directions has often been observed, even when the parents did not show any difference for the studied quantitative trait (Alonso-Blanco et al. 1998a; De Vicente and Tanksley 1993). It is classically explained by the fact that a similar phenotype can be the result of two different genotypes, i.e. two

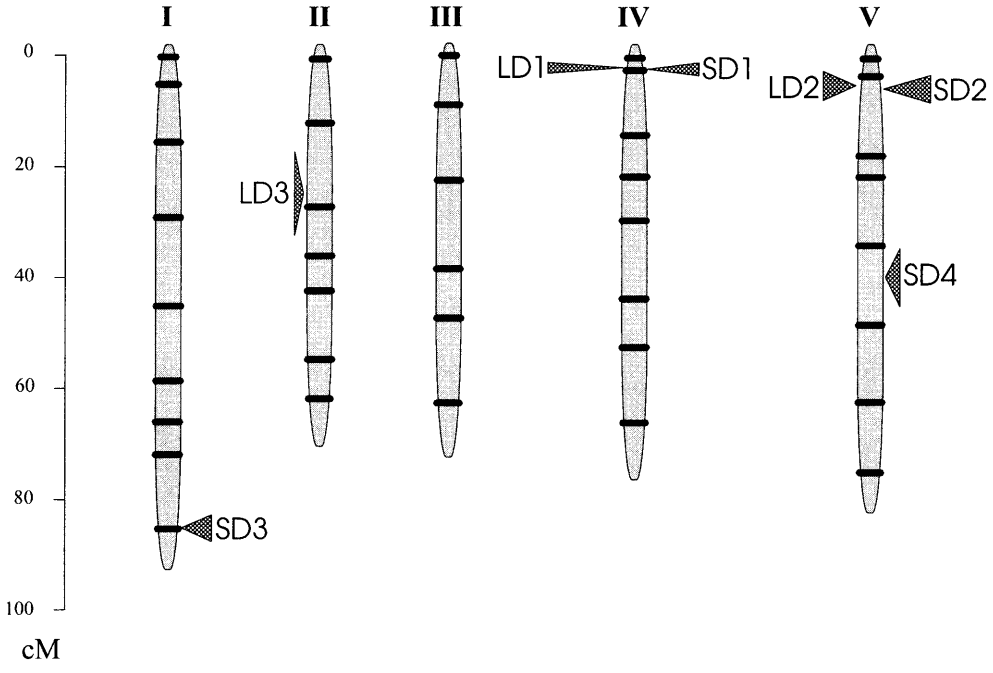

Fig. 5 Detected QTLs explaining flowering-time variability in the Bay-0 × Shahdara RIL population. *LD* represents QTLs detected in the long day environment. *SD* represents QTLs detected in the short day environment. The *horizontal length of arrows* is proportional to the percentage of variance explained by the QTL. The *vertical length of arrows* represents the One-LOD Support Interval. Other characteristics of the QTLs are presented in Table 2

Table 2 Characteristics of the detected QTLs explaining flowering-time variation in the Bay-0 × Shahdara population

QTL[a]	Chromosome – corresponding Marker[b]	Position (cM)[c]	LOD score	% Var [d]	2a (day)[e]	SI (cM)[f]
LD1	Chrom 4 – MSAT4.8	1.3	60.0	35	–14.0	1.5
LD2	Chrom 5 – NGA249	4.5	27.8	17	+8.8	5.2
LD1 × LD2				17		
LD3	Chrom 2 – MSAT2.36	24.9	3.8	2	–3.0	17.7
LD complete model				71%		
SD1	Chrom 4 – MSAT4.8	1.6	58.9	26	–11.4	1.9
SD2	Chrom 5 – NGA249	4.9	35.4	23	+8.7	5.3
SD1 × SD2				8		
SD3	Chrom 1 – MSAT1.5	84.0	20.6	13	+5.8	4.0
SD4	Chrom 5 – MSAT5.22	43.2	9.1	4	+4.2	10.0
SD3 × SD4				1		
SD complete model				75%		

[a] LD represents QTLs detected in the long day environment. SD represents QTLs detected in the short day environment
[b] The corresponding marker is the one used in CIM Model 6, as well as in ANOVA analysis
[c] The position of the QTL is expressed in cM from the first marker of the chromosome
[d] Percentage of variance explained by the QTL or by QTL × QTL interaction, when significant
[e] 2a represents the mean effect (in days to flowering) of the replacement of the Shahdara allele by the Bay-0 allele at the QTL
[f] SI represents the length (in cM) of the one-LOD support interval

different combinations of positive and negative alleles. The earliest LD-grown RIL flowers at 24 days (which is 50 days before the latest one in the same photoperiod conditions). There seem to be strong constraints preventing flowering to happen earlier in our conditions, as shown by the L-shaped distribution of the phenotypes in the LD experiment. In the SD environment, the range of variation was also extended (between 42 and 96 days), but the shape of the distribution is closer to Normal. The population mean is just superior to both parents in SD and LD conditions (respectively 1 and 2 days above the later ecotype). The photoperiod length effect on flowering time is highly significant ($P < 0.001$): the mean delay of flowering in SD in comparison to LD is 27.8 days (ranging from 7 to 50 days). The correlation between flowering time in SD and in LD conditions is highly significant ($P < 0.001$) with a determination coefficient of 56%, indicating that the regulation of both traits can probably be explained in part by the same genetic factors in the Bay-0 × Shahdara population. Nonetheless, these data also show a classical Genotype × Environment interaction, indicating that another part of the regulation of these traits leads to some genotype-specific reactions to environment. The SD experiment allows us to calculate an estimation of trait heritability of 0.86. This value is consistent, for example, with the total percentage of variance explained by the multi-QTL models elaborated by Alonso-Blanco et al. (1998a) for flowering-time traits.

Figure 5 and Table 2 present a summary of the QTLs found (seven in total) for LD and SD flowering time. Three and four QTLs were mapped respectively in LD and SD environments. Detected QTLs explain from 2% (LD3) to 35% (LD1) of the total phenotypic variation. Most of them explain more than 15% of the variation. The two main QTLs in the LD environment have opposite effects, the Shahdara allele at LD1 delays flowering with respect to the Bay-0 allele, whereas the Shahdara allele at LD2 promotes flowering earlier than the Bay-O allele. SD1 and SD2 QTLs show the same pattern of opposite effects. Nevertheless, when tested through marker MSAT4.8, LD1/SD1 locus shows a significant ($P < 0.01$) interaction with the environment; this is not the case for the NGA249-linked locus (LD2/SD2). As well as single-QTL additive effects, we detected strong epistatic relationships between these QTLs: the interaction between LD1 and LD2, as well as between SD1 and SD2, explained a substantial part of the phenotypic variation (respectively 17% and 8%). This epistasis reveals reciprocal conditional effects of LD1 (SD1) and LD2 (SD2) loci: the strong additive effect of LD1 (SD1) on flowering time was apparent essentially when LD2 (SD2) carries the Bay-0 allele; reciprocally, the strong additive effect of LD2 (SD2) on flowering time was expressed essentially when LD1 (SD1) carries the Shahdara allele. For example, the additive effect of LD2 was +19.5 days when the Shahdara allele was present at LD1, while it was only +2.0 when the Bay-0 allele was present at LD1. A much smaller interaction was also observed between SD3 and SD4 (Table 2). The complete model established by the CIM study of the LD (SD) flowering time, contains 3 (4) QTLs and 1 (2) epistatic relations between them; this explains as much as 71% (75%) of the total phenotypic variation (Table 2). Concerning the estimated heritability of the SD flowering time trait (0.86), it can be concluded that almost 9/10 of the genotypic variation has been dissected in single QTL factors and interactions between them.

Discussion

The Bay-0 × Shahdara population is a new RIL set designed for QTL mapping studies. Its genetic map was constructed using microsatellite markers defined from the available sequence; therefore, they are all physically anchored on the genome, and have proven to be very efficient to draw an optimized genetic map for the population. It is important that such a map should be as precise as possible, with very low levels of mistyping and missing-data. Only robust markers were used to genotype the whole population, thus ensuring a very high level of precision to the assigned genetic positions of markers. The comparison between our genetic map and other maps, such as the Ler × Col RI map, is not easy as the only consistent bridge between them is the physical position of markers. Nevertheless, it appears that there are strong variations between the populations in the level of recombination, which are localized to specific regions. Intervals NGA225–NGA249 and MSAT5.14–NGA139, for example, seem to show a higher degree of recombination in the Ler × Col population, as the corresponding genetic distances are higher than expected from the physical length of these intervals (data not shown). Such variations in the density of recombination have also been observed when comparing the Ler × Col and Ler × Cvi genetic maps (Alonso-Blanco et al. 1998c), with up to 3-fold differences. Despite these unpredictable variations in recombination density, our genetic map is composed of markers which are regularly distributed along the chromosomes (Fig. 2). This is an important factor, since Zeng et al. (1999) showed that, in some cases (CIM), the QTL mapping statistic test can be affected by an uneven distribution of markers in the genome. The average genetic distance between two adjacent markers (10.8 cM) has been targeted to optimize the efficiency of QTL mapping. Indeed, simulation studies have shown that the advantages of increasing marker density beyond one marker every 10 cM are less significant than those obtained when increasing the size of the population (Darvasi and Soller 1994; Charmet 2000). Finally, microsatellite markers, which are robust and polymorphic, should also help to connect different genetic maps using a common set of markers.

In addition to the unanimously accepted augmentation in the power and accuracy of QTL detection (Darvasi et al. 1993; Charmet 2000), a large RIL population (> 300 individuals) is also advantageous for the identification of recombination between tightly linked markers. Fine-resolution mapping of a major locus would therefore be possible without having to develop another population. The quality of these RILs is a fundamental factor for subsequent QTL analysis over the entire genome. Detailed study of the residual heterozygosity and segregation ratio along the chromosomes, permits the analysis of the possible consequences of genetic distortions. In the Bay-0 × Shahdara population, residual heterozygosity was never significantly higher than the expected value of 3.12%. A classical error that increases heterozygosity (at least in some loci) is by unwittingly imposing a 'selection', for example by size, on the plants selfed during SSD cycles. The display used to randomly designate the plant selected at each SSD cycle allowed us to avoid this type of distortion in our population. The bias in segregation ratio observed between parental alleles essentially concerned two limited regions of chromosomes 2 and 4. Moreover, this bias is counterbalanced by the size of the population: for the locus most affected (i.e. MSAT4.18) the Bay-0 allele is still represented in the population by 151 lines (36% of 420 RILs). The incidence of an equivalent segregation distortion would certainly be more dramatic in a 150-individual population, where the minority allele would be represented by only 54 lines, which could strongly compromise the QTL analysis in this region. Finally, figure 1 clearly illustrates that 80% of the total genetic length in the Bay-0 × Shahdara population is free of distortions. Similar and often more extended

distortions have been described in Arabidopsis for all populations already available (Lister and Dean 1993; Alonso-Blanco et al. 1998c). Despite the care taken in the design of SSD display, such distortions could simply be explained by the existence of unintentional selection pressure. For example, an epistatic relation between two loci might lead to a decreased germination rate of the seeds containing a certain combination of alleles. These alleles would be, at least partly, counter-selected and, by linkage, the markers around the loci would reflect this distortion.

Flowering-time QTL mapping in the Bay-0 × Shahdara RIL population illustrates the possible use of this tool to analyse the genetic basis of a quantitative trait. Nine tenths of the total genotypic variability of this character has been decomposed essentially in major QTLs. This result has often been obtained in previous flowering-time QTL studies (Kowalski et al. 1994; Clarke et al. 1995; Jansen et al. 1995; Mitchell-Olds 1996; Kuittinen et al. 1997; Alonso-Blanco et al. 1998a): it seems that Arabidopsis natural variation in flowering time is controlled mostly by a small number of large-effect loci (Alonso-Blanco et al. 1998a; Koornneef et al. 1998a). Three of the detected QTLs are specific to the photoperiod environment, LD or SD (Fig. 5). This accounts for the fact that some of the flowering-pathway genes polymorphic between Bay-0 and Shahdara only control flowering time in specific day length conditions. However, these environment-specific loci are often small-effect QTLs (Table 2). In contrast, large-effect QTLs seem to correspond to central-role genes, whose polymorphism broadly influences flowering-time phenotype.

Some of the QTLs that we found in the Bay-0 × Shahdara population colocalize with previously published QTLs detected in other populations (Kowalski et al. 1994; Clarke et al. 1995; Jansen et al. 1995; Kuittinen et al. 1997; Alonso-Blanco et al. 1998a). Moreover, the epistatic interaction between RLN3 and RLN4 described by Clarke et al. (1995) in the Ler × H51 cross resembles the LD1 × LD2 and SD1 × SD2 interactions. In fact, only LD3 maps in a region of chromosome 2 where, to our knowledge, no other QTLs had been found previously.

Flowering time represents a 'model' character for QTL analysis, in that many of the genes involved in the regulation of this phenomenon have already been cloned (see Koornneef et al. 1998a for a review; see Sheldon et al. 2000 for an updated regulation model), like *FRIGIDA* (*FRI*) by Johanson et al. (2000) and *FLOWERING LOCUS C* (*FLC*) (Lee et al. 1994; Sheldon et al. 1999). This is advantageous for the assignment of good candidate genes to explain our QTLs. The very close estimation of LD1/SD1 and LD2/SD2 positions and effects (additive and epistatic relationship between them) strongly suggests that they correspond to the same loci (Fig. 5). If one assumes a linear relationship between genetic and physical distances in the MSAT4.39–MSAT4.8 interval (which is 2.1 cM in length), the LD1 one-LOD Support Interval (SI) covers the BAC clone F6N23, and the end of the neighbouring BACs. The best estimation of LD1 position (LOD Score peak) within this SI corresponds to the middle of the BAC F6N23, which contains the *FRI* gene. Although the LD2 SI is much-more extended (5-cM length) and covers approximately 20 BACs, the best estimated position of LD2 falls very close to the *FLC* gene. Moreover, the strong LD1 × LD2 interaction is reminiscent of the interaction described between the *FRI* and *FLC* genes (Koornneef et al. 1994; Lee et al. 1994; Michaels and Amasino 1999). At the molecular level, allelic variation at the *FRI* locus has been extensively studied by Johanson et al. (2000). Using their PCR markers, we find that Bay-0 seems to carry the same recessive allele as the Ler group (data not shown), and verified that the Shahdara *FRI* allele seems to be fully functional. We conclude that a polymorphism exists between Shahdara and Bay-0 at the *FRI* gene, and that this polymorphism is responsible for LD1 (and SD1) QTLs. Moreover, Johanson et al. (2000) have shown that the Shahdara *FLC* allele was defective, which is another element indicating that *FLC* could be a good candidate to explain LD2 (and SD2).

In this study, large-effect as well as small-effect QTLs have been detected simultaneously, though not with the same precision. The size of the population involved in the analysis should explain the high power of QTL detection, although the number of observed plants per RIL was reduced to a minimum. This is confirmation that a large RIL population of good quality is of particular interest when the position of the QTL needs to be precisely estimated, for example to ensure further molecular cloning of the gene responsible for the variation. This had been demonstrated before by simulation studies only. Moreover, new methods and models, like Multiple Interval Mapping (Kao et al. 1999; Zeng et al. 1999), will certainly greatly improve the dissection of the whole genetic architecture of quantitative traits, as computer calculation capacity becomes less restrictive.

With the advent of map-based cloning of major QTLs on model species like Arabidopsis, we can predict that the coming years will see a great increase in QTL mapping studies, concerning all types of traits. An important element in such studies will be the development of segregating populations specifically adapted and dedicated to QTL mapping, high-throughput genetic-mapping methods and more powerful QTL-mapping statistical tools. The Bay-0 × Shahdara RIL population is a step towards this achievement. Being the product of a cross between genetically distant ecotypes, adapted to diverging constraints of their specific habitat, ensures the segregation of particularly informative traits in the same population. Studies are in progress using the Bay-0 × Shahdara population, to map QTLs explaining diverse traits such as nitrogen-use efficiency, root architecture, seed germination, drought tolerance and virus resistance. Furthermore, Bay-0 and Shahdara have also shown phenotypic differences for several other characters of interest such as soil acidity tolerance (personal observations), powdery mildew resistance (Adam et al. 1999) or seed oil fatty-acyl composition (Millar and Kunst 1999).

Acknowledgements We are grateful to F. Gosse and B. Auclair for taking care of the plants. We thank C. Dean and U. Johanson for information concerning *FRIGIDA* PCR markers, and H. North for her careful reading of the manuscript. The Bay-0 × Shahdara RILs set and the marker data files will be made accessible through deposition in public Stock Centers.

References

Adam L, Ellwood S, Wilson I, Saenz G, Xiao S, Oliver RP, Turner JG, Somerville S (1999) Comparison of *Erysiphe cichoracearum* and *E. cruciferarum* and a survey of 360 *Arabidopsis thaliana* accessions for resistance to these two powdery mildew pathogens. Mol Plant Microbe Interact 12:1031–1043

Allard RW (1956) Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia 24:235–278

Alonso-Blanco C, Koornneef M (2000) Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. Trends Plant Sci 5:22–29

Alonso-Blanco C, El-Assal SED, Coupland G, Koornneef M (1998a) 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

Alonso-Blanco C, Koornneef M, Stam P (1998b) The use of recombinant inbred lines (RILs) for genetic mapping, in Arabidopsis protocols. In: Martinez-Zapater JM, Salinas J (eds.) Humana Press Inc, Totowa, pp 137–146

Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MTR (1998c) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. Plant J 14:259–271

Basten CJ, Weir BS, Zeng Z-B (1994) Zmap – a QTL cartographer. In: Smith C, Gavora JS, Burnside EB (eds.) 5th Computing strategies and software: World Congress on Genetics Applied to Livestock Production. Guelph, Ontario, Canada, pp 65–66

Basten CJ, Weir BS, Zeng Z-B (2000) QTL cartographer, version 1.14. North Carolina State University, Department of Statistics, Raleigh, North Carolina

Breyne P, Rombaut D, van Gysel A, van Montagu M, Gerats T (1999) AFLP analysis of genetic diversity within and between *Arabidopsis thaliana* ecotypes. Mol Gen Genet 261:627–634

Buell CR, Somerville SC (1997) Use of Arabidopsis recombinant inbred lines reveals a monogenic and a novel digenic resistance mechanism to *Xanthomonas campestris* pv campestris. Plant J 12:21–29

Casacuberta E, Puigdomenech P, Monfort A (2000) Distribution of microsatellites in relation to coding sequences within the *Arabidopsis thaliana* genome. Plant Sci 157:97–104

Chang C, Bowman AW, Lander ES, Meyerowitz EW (1988) Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. Proc Natl Acad Sci USA 85:6856–6860

Charcosset A, Gallais A (1996) Estimation of the contribution of quantitative trait loci (QTLs) to the variance of a quantitative trait by means of genetic markers. Theor Appl Genet 93:1193–1201

Charmet G (2000) Power and accuracy of QTL detection: simulation studies of one-QTL models. Agronomie 20:309–323

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

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

Copenhaver GP, Browne WE, Preuss D (1998) Assaying genome-wide recombination and centromere functions with Arabidopsis tetrads. Proc Natl Acad Sci USA 95:247–252

Darvasi A, Soller M (1994) Optimum spacing of genetic markers for determining linkage between marker loci and quantitative trait loci. Theor Appl Genet 89:351–357

Darvasi A, Weinreb A, Minke V, Weller JI, Soller M (1993) Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. Genetics 134:943–951

De Vicente MC, Tanksley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134:585–596

Edwards MD, Stuber CW, Wendel JF (1987) Molecular-marker facilitated investigations of QTL in maize. I. Number, genomic distribution and types of gene action. Genetics 116:113–125

Erschadi S, Haberer G, Schoniger M, Torres-Ruiz RA (2000) Estimating genetic diversity of *Arabidopsis thaliana* ecotypes with amplified fragment length polymorphisms (AFLP). Theor Appl Genet 100:633–640

Fisher RA (1937) The design of experiments. Oliver and Boyd, Edinburgh, London

Hyne V, Kearsey MJ, Pike DJ, Snape JW (1995) QTL analysis: unreliability and bias in estimation procedures. Mol Breed 1:273–282

Innan H, Terauchi R, Miyashita NT (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics 146:1441–1452

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

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

Kao CH, Zeng ZB, Teasdale RD (1999) Multiple interval mapping for quantitative trait loci. Genetics 152:1203–1216

Koornneef M, van Eden J, Hanhart CJ, Stam P, Braaksma FJ, Feenstra WJ (1983) The linkage map of *Arabidopsis thaliana* (L.) Heynh. J Hered 74:265–272

Koornneef M, Alonso-Blanco C, Peeters AJM, Soppe W (1998a) Genetic control of flowering time in Arabidopsis. Annu Rev Plant Physiol Plant Mol Biol 49:345–370

Koornneef M, Alonso-Blanco C, Stam P (1998b) Genetic analysis. In: Martinez-Zapater JM, Salinas J (eds) Arabidopsis protocols. Humana Press Inc, Totowa, pp 105–117

Koornneef M, Blankestijn-de Vries H, Hanhart CJ, 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 distances from recombination values. Ann Eugen 12:172–175

Kowalski SP, Lan TH, Feldmann KA, Paterson AH (1994) QTL Mapping of naturally occurring variation in flowering time of *Arabidopsis thaliana*. Mol Gen Genet 245:548–555

Kuittinen H, Sillanpaa MJ, Savolainen O (1997) Genetic basis of adaptation: flowering time in *Arabidopsis thaliana*. Theor Appl Genet 95:573–583

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

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, 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, Michaels SD, Masshardt AS, Amasino RM (1994) The late-flowering phenotype of *FRIGIDA* and mutations in *LUMINIDEPENDENS* is suppressed in the Landsberg *erecta* strain of Arabidopsis. Plant J 6:903–909

Lincoln SE, Daly MJ, Lander ES (1993) Constructing genetic linkage maps with MAPMAKER/EXP version 3.0: a tutorial and reference manual. Whitehead Institute.

Lister C, Dean C (1993) Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. Plant J 4:745–750

Michaels SD, Amasino RM (1999) *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell 11:949–956

Millar AA, Kunst L (1999) The natural genetic variation of the fatty acyl composition of seed oils in different ecotypes of *Arabidopsis thaliana*. Phytochemistry 52:1029–1033

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

Nam HG, Giraudat J, den Boer B, Moonan F, Loos WDB, Hauge BM, Goodman HM (1989) Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. Plant Cell 1:699–705

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726

Reiter RS, Williams J, Feldman K, Rafalski JA, Tingey SV, Scolnik PA (1992) Global and local genome mapping in *Arabidopsis thaliana* recombinant inbred lines and random amplified polymorphic DNAs. Proc Natl Acad Sci USA, 89:1477–1481

Roeder GS (1997) Meiotic chromosomes: it takes two to tango. Genes Dev 11:2600–2621

Sharbel TF, Haubold B, Mitchell-Olds T (2000) Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. Mol Ecol 9:2109–2118

Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ, Dennis ES (1999) The *FLF* MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. Plant Cell 11:445–458

Sheldon CC, Finnegan EJ, Rouse DT, Tadege M, Bagnall DJ, Helliwell CA, Peacock WJ, Dennis ES (2000) The control of flowering by vernalization. Curr Opin Plant Biol 3:418–422

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Vision TJ, Brown DG, Shmoys DB, Durrett RT, Tanksley SD (2000) Selective mapping: a strategy for optimizing the construction of high-density linkage maps. Genetics 155:407–420

Visscher PM, Thompson R, Haley CS (1996) Confidence intervals in QTL mapping by bootstrapping. Genetics 143:1013–1020

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Fritjers A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Zeng Z-B, Kao CH, Basten CJ (1999) Estimating the genetic architecture of quantitative traits. Genet Res 74:279–289