

Genetics of quantitative traits in *Arabidopsis thaliana*

MJ Kearsey, HS Pooni and NH Syed¹

School of Biosciences, The University of Birmingham, Birmingham, B15 2TT, UK

The genetic control of 22 quantitative traits, including developmental rates and sizes, was examined in generations of *Arabidopsis thaliana* derived from the cross between the ecotypes, Columbia (Col) and Landsberg erecta (Ler). The data were obtained from three sets of families raised in the same trial: the 16 basic generations, that is, parents, F₁'s, F₂'s, backcrosses, recombinant inbred lines (RILs) and a triple test cross (TTC), the latter produced by crossing the RILs to Col, Ler and their F₁. The data were analysed by two approaches. The first (approach A) involved traditional generation mean and variance component analysis and the second (B), based around the RILs and TTC families, involved marker-based QTL analysis.

From (A), genetic differences between Col and Ler were detected for all traits with moderate heritabilities. Height at flowering was the only trait to show heterosis. Dominance was partial to complete for all height traits, and there was no

overdominance but there was strong evidence for directional dominance. For most other traits, dominance was ambidirectional and incomplete, with average dominance ratios of around 80%. Epistasis, particularly of the duplicate type that opposes dominance, was a common feature of all traits. The presence of epistasis must imply multiple QTL for all traits.

The QTL analysis located 38 significant effects in four regions of chromosomes I, II, IV and V, but not III. QTL affecting rosette size and leaf number were identified in all four regions, with days to maturity on chromosomes IV and V. The only QTL for height was located at the expected position of the *erecta* gene (chromosome II; 50 cm), but the additive and dominance effects of this single QTL did not adequately explain the generation means. The possible involvement of other interacting height QTL is discussed.

Heredity (2003) 91, 456–464. doi:10.1038/sj.hdy.6800306

Keywords: dominance; epistasis; heritability; heterosis; QTL analysis; triple test cross

Introduction

In recent years, the analysis of quantitative traits has largely centred round the use of molecular markers to locate and measure the effects of the individual underlying genes, *quantitative trait loci* or QTL (Kearsey and Farquhar, 1998; Tanksley, 1993). In some situations, it has also been possible to examine epistatic interactions and genotype environment interaction involving QTL (Monforte *et al.*, 2001). These approaches have thus provided a very powerful tool for the genetical analysis of quantitative traits and have been a major impetus to quantitative genetical research and breeding.

There are, however, limitations to these approaches. Firstly, the accuracy of chromosomal location of QTL is often low and they focus on those QTL that have large effects. Secondly, the size of the estimated QTL effects will be biased because only those estimates that exceed some given threshold of significance are reported (Utz *et al.*, 2000). Concentration on the proportion of the variation explained by individual QTL tends to encourage acceptance of a low number of genes, and many studies have emphasised how few QTL are necessary to explain the observed genetic variation (Doganlar *et al.*, 2000). However, this can often be misleading. If, for example, a trait is controlled by one QTL with an effect of nine units and 19 QTL each with individual effects of one

unit, the former will account for 81% of the total variation but only 32% of the combined gene effects. Thus, the contribution to the variance is relevant to the response to selection per generation but not to the genetic potential and the ultimate limits to selection.

Conventional biometrical genetical procedures, on the other hand, deal with the combined effects of all the QTL on the means, variances and correlations of traits. These procedures, mainly developed and improved during the second half of the last century, use statistical relations among relatives to infer the nature of the additive, dominance and epistatic effects of the genes. The estimates obtained are automatically weighted by the sizes of the gene effects and their linkage relations. However, they tell us nothing about the location or effects of the individual QTL (Kearsey and Pooni, 1996; Lynch and Walsh, 1998).

The experiments reported in the present paper were designed to be analysed by a combination of conventional biometrical and current molecular methods. The intention was to reconcile and integrate the conclusions from the two approaches in the hope of gaining a more complete picture of the underlying genetical control of several traits. They are based on the well-analysed lines of *Arabidopsis thaliana*, Columbia and Landsberg, and generations derived from them.

Materials and methods

Plant material

The inbred strains of *A. thaliana*, Columbia (Col) and Landsberg (Ler), derived from seed obtained from the

Correspondence: MJ Kearsey, School of Biosciences, The University of Birmingham, Birmingham B15 2TT, UK. E-mail: m.j.kearsey@bham.ac.uk

¹Current address: SCRI, Invergowrie, Dundee DD2 5DA, UK

Received 8 April 2002; accepted 21 April 2003

Nottingham Arabidopsis Stock Centre (NASC, UK), were used as source material. The 16 basic generations (BGs) shown in Table 1, which include both parents and all possible reciprocal F₁'s, F₂'s and first backcrosses (Bc₁ and Bc₂) derived from them, were produced by hand emasculation and pollination. The 101 recombinant inbred lines (RILs) (Lister and Dean, 1993) derived from Ler × Col were each selfed and also crossed as seed parent to Ler, Col and their F₁ in a triple test cross design (Kearsey and Jinks, 1968), but sufficient seeds were obtained for all four families from only 85 of the RILs.

A total of 20 seeds from each basic generation family and five seeds from each RIL or TTC family were sown in separate pots containing JI compost and the 2020 pots were immediately, individually randomised in a single block in an unheated polytunnel in long (16 h) days. A total of 22 traits, representing plant size and development, were scored over the following 2 months and these are listed in Table 2.

Genetical analysis

The means and within-family variances of each of the 16 basic generations were calculated for all traits. Models that included standard genetical and cytoplasmic effects (Table 1) were fitted to the means, the parameters being estimated by weighted least squares using the reciprocal of the variance of each family mean as the weight. The genetical parameters were as follows: additive [a], dominance [d] and epistatic additive by additive [aa], additive by dominance [ad] and dominance by dominance [dd] (see Kearsey and Pooni (1996), for definitions and estimation procedures). For consistency, [a] has been defined as 1/2(Ler–Col), and so will be negative when Columbia has the highest score. The nongenetical, cytoplasmic or maternal effects tested were: inbred vs F₁ mothers, Col vs Ler mothers and F₁ vs reciprocal F₁ mothers. The simplest model was accepted for which all the parameters were indivi-

dually significant and the χ^2 , testing the adequacy of the model, was nonsignificant. The genetical and environmental variance parameters, V_G and V_E , were estimated from the within-family variances of the F₂ and nonsegregating generations (Kearsey and Pooni, 1996).

For each trait, the subset of RILs (65) for which the selfed families and the three TTC families had complete progeny sets were chosen, but a few individual plants were inevitably lost due to pests or disease. We refer to the crosses of the *n* RILs to Col, Ler and the F₁ as 'L_{1i}', 'L_{2i}' and 'L_{3i}' (*i* = 1–*n*), respectively. The RIL families were analysed by one-way ANOVA, while the TTC families were subjected to standard ANOVAs (Table 3) to test for additive ($L_{1i}+L_{2i}$), dominance ($L_{2i}-L_{1i}$) and epistatic variation ($L_{1i}+L_{2i}-2L_{3i}$) as described by Kearsey and Pooni (1996) and Kearsey and Jinks (1968). These ANOVAs will be illustrated in the 'Results' section. Phenotypic (V_P), additive (V_A), dominance (V_D) and environmental (V_E) components of variance were estimated from these ANOVAs and used to calculate narrow (V_A/V_P) and broad ($(V_A+V_D)/V_P$) sense heritabilities and dominance ratios, $\sqrt{(2V_D/V_A)}$ (see Kearsey and Pooni, 1996).

QTL analysis was carried out for every trait for which the earlier ANOVAs (RILs and TTC) had indicated significant variation between families. The method of Kearsey and Hyne (1994) (also known as marker difference regression, MDR; Lynch and walsh, 1998) was employed, using the 'QTL Café' (<http://web.bham.ac.uk/g.g.seaton/>). Significance testing and confidence intervals were obtained by simulation. The analyses were carried out on three independent sets of data: (i) the means of the RILs; (ii) the means of the TTC ($L_{1i}+L_{2i}$) values; and (iii) the means of the TTC ($L_{2i}-L_{1i}$) values. The first two identify QTL on the basis of their additive and the third on their dominance genetic effects. The software provides estimates of QTL effects, locations and their respective confidence intervals. A total of 65 markers were used for the QTL analysis. They were

Table 1 Origins of the 16 basic generations from Ler and Col, together with the model for genetical and cytoplasmic effects fitted to the generation means

Generation	Cross Female × male	Genetical parameters					Cytoplasmic parameters			
		m	[a]	[d]	[aa]	[ad]	[dd]	Inbred vs F ₁	Ler vs Col	F ₁ vs RF ₁
P ₁	Ler × Ler	1	1	0	1	0	0	1	1	0
P ₂	Col × Col	1	-1	0	1	0	0	1	-1	0
F ₁	Ler × Col	1	0	1	0	0	1	1	1	0
RF ₁	Col × Ler	1	0	1	0	0	1	1	-1	0
F ₂	F ₁ × F ₁	1	0	0.5	0	0	0.25	-1	0	1
F ₂	F ₁ × RF ₁	1	0	0.5	0	0	0.25	-1	0	1
F ₂	RF ₁ × F ₁	1	0	0.5	0	0	0.25	-1	0	-1
F ₂	RF ₁ × RF ₁	1	0	0.5	0	0	0.25	-1	0	-1
Bc ₁	F ₁ × Ler	1	0.5	0.5	0.25	0.5	0.25	-1	0	1
Bc ₁	RF ₁ × Ler	1	0.5	0.5	0.25	0.5	0.25	-1	0	-1
Bc ₁	Ler × F ₁	1	0.5	0.5	0.25	0.5	0.25	1	1	0
Bc ₁	Ler × RF ₁	1	0.5	0.5	0.25	0.5	0.25	1	1	0
Bc ₂	F ₁ × Col	1	-0.5	0.5	0.25	-0.5	0.25	-1	0	1
Bc ₂	RF ₁ × Col	1	-0.5	0.5	0.25	-0.5	0.25	-1	0	-1
Bc ₂	Col × F ₁	1	-0.5	0.5	0.25	-0.5	0.25	1	-1	0
Bc ₂	Col × RF ₁	1	-0.5	0.5	0.25	-0.5	0.25	1	-1	0

RF₁=reciprocal F₁.

Table 2 Traits scored and models fitted to each trait

Trait	Genetic parameters						Adequacy of model		Parental and F ₁ means		
	m	[a]	[d]	[aa]	[ad]	[dd]	χ^2	df	Ler	F ₁	Col
Days to											
First true leaves	12.22	0.71			-1.05		9.4	13	12.90	12.18	11.50
First flower	28.82	0.71			-1.84		10.88	13	29.85	29.05	28.35
Maturity	51.01	1.24	-1.07		-2.39		19.28	11	52.82	50.39	50.10
No. of rosette leaves at											
21 days	8.94	-1.01			1.35		12.02	13	8.00	8.70	10.05
26 days	9.4	-0.72					22.58	14	8.80	9.25	10.55
Flowering	9.59	-1.04			0.52		20.43	13	8.90	9.87	10.90
Rosette size (mm) at											
21 days	27.28	-4.46			7.5		18.2	13	21.30	25.78	30.30
26 days	49.83	-7.13	17.39		11.56	-16.14	6.55	11	42.70	51.12	56.95
Flowering	74.28	-4.34					23.86	14	69.75	70.79	77.65
36 days	100.04	-8.8			6.57		17.11	12	86.45	99.92	112.05
No. of cauline leaves at											
26 days	4.52	1.29					22.39	14	3.65	4.50	3.80
Flowering	16.97						17.06	15	17.60	16.35	19.05
36 days	40.09	-4.58					12.42	14	33.70	43.95	47.25
No. of buds at flowering	9.77						16.67	15	10.00	9.66	10.75
Height (mm) at											
Flowering	43.6		6.56				22.28	14	40.65	46.86	43.65
30 days	67.3	-24.6	49.27		49.14	-36.3	9.63	11	42.70	80.50	91.89
35 days	173.28	-29.93	73.22		57.53	-54.82	7.84	11	143.35	190.78	203.20
40 days	195.29	-42.08	290.37	86.28	54.28	-188.89	8.17	10	239.50	295.98	323.65
45 days	346.6	-44.95	124.28		54.96	-106.35	15.96	11	301.65	365.72	391.55
Maturity	383.08	-45.67	184.86		49.38	-151.02	10.74	11	337.41	415.39	428.75
Leaf size (mm)											
Length at 40 days	53.65	-6.37			3.98		13.49	13	47.16	52.38	59.90
Width at 40 days	22.02				-1.87		19.93	14	22.06	21.18	21.40

All parameters shown are significant and all the χ^2 , testing the adequacy of the models, are non significant. The observed means of Ler, Col and the F₁ are shown in the last three columns. Two significant cytoplasmic effects, F₁ vs RF₁ for days to maturity (0.63) and Ler vs Col for rosette size at 36 days (-4.07) have been taken out to simplify the table.

Table 3 ANOVAs of (i) RILs and (ii) TTC, together with parameter estimates to illustrate methods of analysis: data for 'Days to first leaf'

Source	df	MS	F	P	ems
(i)					
Between RILs	64	2.1855	2.55	***	$\sigma_w^2 + 4.46\sigma_b^2$
Within	246	0.8564			σ_w^2
Estimates					$V_A=0.15$ $V_E=0.86$ $h_n^2=100 \times V_A / (V_A + V_E)$
(ii)					
Additive	64	3.2525	2.15	***	$\sigma_w^2 + 8.92\sigma_A^2$
Dominance	64	2.7761	1.84	***	$\sigma_w^2 + 8.92\sigma_D^2$
Within	471	1.5125			σ_w^2
Estimates					$V_A=0.39$ $V_D=0.14$ $V_E=1.51$ $h_n^2=100 \times V_A / (V_A + V_D + V_E)$ $h_D^2=100 \times (V_A + V_D) / (V_A + V_D + V_E)$ Dominance ratio (D.r.) = $\sqrt{(4V_D / 2V_A)}$ $h_n^2=19.1$ $h_D^2=26.0$ D.r.=0.85
Epistasis [aa]	1	2.5670	1.14	ns	
Epistasis [ad, dd]	64	2.2477	1.69	**	
Within	710	1.3233			

ns=not significant, *0.05>P≥0.01, **0.01>P≥0.001, ***P<0.001.

chosen from the published genotype data at NASC to be evenly spaced across the five chromosomes and markers for which the genotypes for most RILs were known.

Results

A summary of the significant parameters for the 22 traits is presented in Table 2, together with the observed means

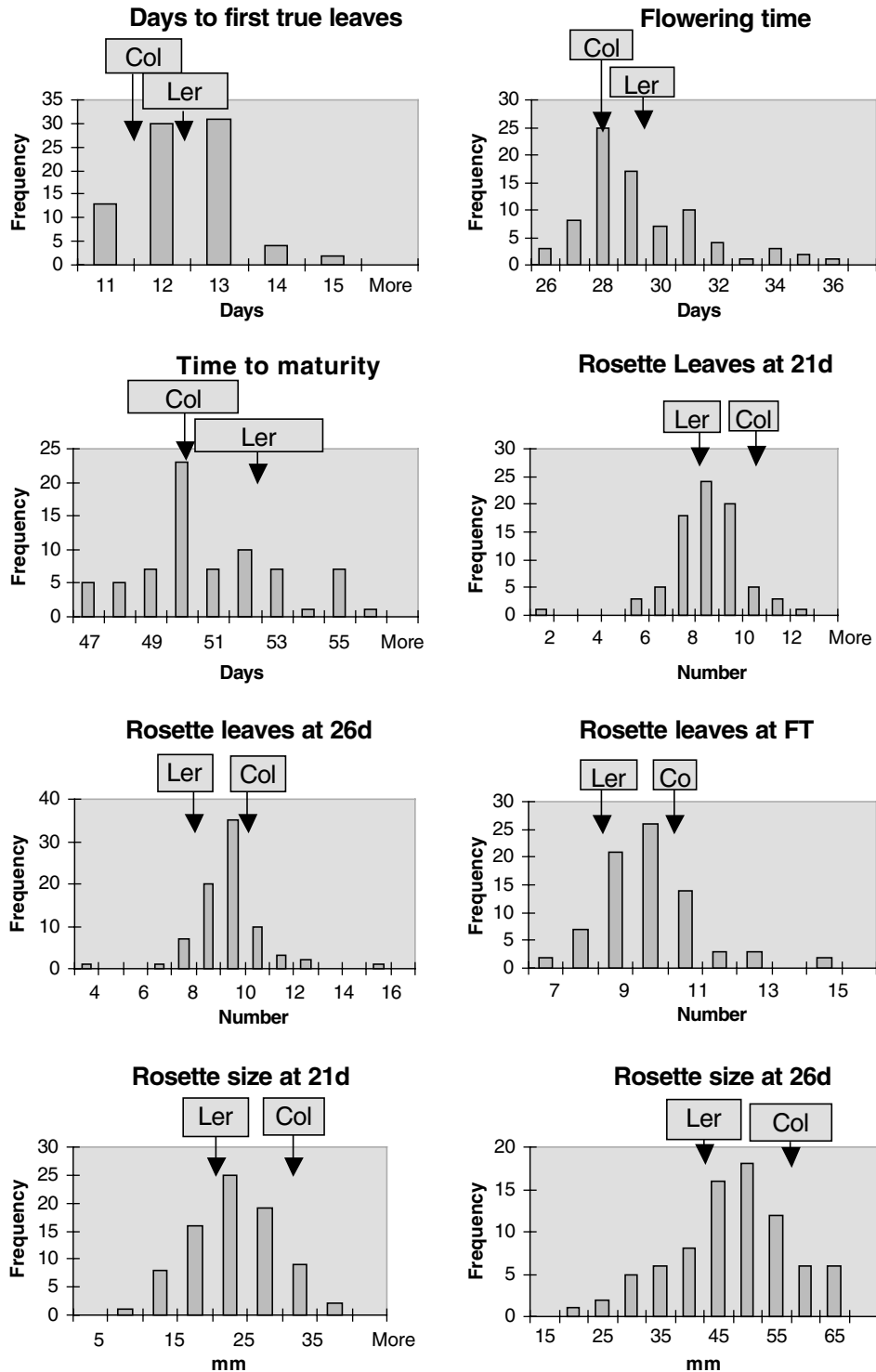


Figure 1 Distribution of RIL means for various traits, together with the means of the parental lines, Col and Ler. The width of the frame containing the line label indicates the 95% confidence interval for that mean.

of the two parental lines and their F_1 . In all cases, fairly simple models could be fitted to the data resulting in a nonsignificant χ^2 , and all but two traits showed significant genetic effects. The two exceptions, both scored at flowering, were cauline leaves and bud number. Additive and/or dominance effects exist for 19 traits and epistasis for 15 traits, so epistasis is a common

feature and must imply two or more QTL for these traits. Where it exists, dominance is for early maturity and greater size. It can be seen that the F_1 mean is generally intermediate to the two parents; better-parent heterosis only occurs for "height at flowering" although it is not significant ($t=0.83$ for 75 df; $P>0.3$). Maternal effects were also a minor component of variation and were

Table 4 Summary of data from RILs and TTC: estimates of variance components and results of tests for epistasis from the TTC

Trait	V_A (RILs)	V_A (TTC)	V_D (TTC)	Epistasis (TTC)	
				$[a \times a]$	$[a \times d], [d \times d]$
Days to					
First true leaves	0.15***	0.39***	0.14***	ns	**
First flower	1.23***	1.69***	0.55***	ns	**
Maturity	1.00***	1.48***	0.43**	ns	**
No. of rosette leaves at					
21 days	0.46***	0.89***	0.29***	ns	***
26 days	0.41***	0.45***	0.10***	ns	**
Flowering	0.06 ns	0.34***	0.06*	ns	ns
Rosette size (mm) at					
21 days	8.16***	26.64***	9.00***	ns	***
26 days	26.32***	75.43***	27.78***	ns	***
Flowering	29.92***	44.25***	11.16***	*	ns
36 days	45.16***	67.24***	19.97***	ns	ns
No. of cauline leaves at					
26 days	1.94***	8.58***	1.85***	ns	***
Flowering	6.62***	5.28***	0.03ns	ns	ns
36 days	26.63***	32.85***	10.04**	ns	ns
No. of buds at flowering	0.46***	0.64***	0.00ns	ns	ns
Height (mm) at:					
Flowering	39.33***	52.75***	39.05***	ns	ns
30 days	281.81***	441.82***	206.63***	ns	**
35 days	946.83***	1213.89***	756.62***	ns	***
40 days	1320.87***	1368.60***	844.27***	ns	*
45 days	1773.02***	1784.41***	1137.26***	ns	*
Maturity	2748.01***	3557.66***	1249.25***	ns	ns
Leaf size (mm)					
Length at 40 days	10.26***	19.80***	4.58**	ns	*
Width at 40 days	0.42 ns	1.39**	0.83**	ns	ns

ns=not significant, * $0.05 > P \geq 0.01$, ** $0.01 > P \geq 0.001$, *** $P < 0.001$.

detected in only two cases; a difference between the progeny of crosses with reciprocal F_1 mothers for "days to maturity" and between crosses with Col *vs* Ler mothers for "rosette size at 36 days". These effects were allowed for as parameters in model fitting and the df for the χ^2 are consequently reduced by 1.

Figure 1 shows the distribution of the RIL means for a sample of traits scored either at a fixed time or at a set physiological age, namely at flowering. It can be seen, by comparing these distributions with the confidence intervals of the parental lines, that there is considerable transgressive segregation in all cases, indicating that increasing alleles are dispersed among the parents. This, again, implies two or more QTL for such traits. Where there are differences between the parental lines, Col is always the faster developer and has the larger size.

Table 3 illustrates the basic ANOVAs of the RILs and TTC families. Given the completely randomised design, the RILs analysis involves a simple one-way ANOVA with the expected mean squares and parameter estimates as shown. The coefficients of the expected mean squares are not whole numbers because of missing data. They have been calculated, following Sokal and Rohlf (1981), as $n_0 = \{1/(a-1)\} \{\sum n_i - (\sum n_i^2 / \sum n_i)\}$, where n_i is the size of the i th family and a is the number of families. The TTC families were analysed as two separate ANOVAs. The first follows the standard NCIII design (Kearsey and Pooni, 1996) to detect additive and dominance effects from the crosses of the RILs to the two parents. The

second is based on an analysis of the comparison $(L_{1i} + L_{2i} - 2L_{3i})$, where L_{1i} , L_{2i} and L_{3i} represent the means of the families derived from crossing RIL $_i$ to Col, Ler and the F_1 , respectively. This comparison is solely a function of epistasis and should be zero for all ' i ' in the absence of epistasis. The correction term in the ANOVA tests for additive-by-additive epistasis, while the variation among RILs is due to additive by dominance and dominance-by-dominance epistasis (Kearsey and Jinks, 1968; Kearsey and Pooni, 1996). In the example shown in Table 3, only the latter is significant. When epistasis is present, estimates of the additive (V_A) and dominance (V_D) components from the NCIII ANOVA are inflated, but it is not easy to correct for this. The estimates given for V_A and V_D , therefore, must be accepted as being biased.

The estimates of V_A and V_D and tests for epistasis from the analyses illustrated in Table 3 are shown for all traits in Table 4. V_A and/or V_D are significant for all traits while epistasis was detected for 14 traits. Using these estimates plus those from the basic generations, heritabilities and average dominance ratios, $\sqrt{(2V_D/V_A)}$, were calculated and they are summarised in Table 5. The conclusions concerning epistasis from the BGs and TTC are generally highly consistent, indicating the almost complete absence of [aa] but the general presence of [ad] epistasis. Of the seven traits showing no epistasis in the BGs, four also showed no epistasis in the TTC.

QTL analysis was carried out on 84 RILs and 74 TTC families for which both L_1 and L_2 data were available.

Table 5 Estimates of heritabilities (h_{ni}^2 , h_b^2), dominance ratio and number (k) of QTLs controlling each trait

Trait	h_{ni}^2		h_b^2		Dominance ratio	k
	RILs	TTC	TTC	BG		
Days to						
First true leaves	14.8	19.1	26.0	23.6	0.85	5
First flower	27.1	27.6	36.7	12.9	0.81	7
Maturity	19.4	21.3	27.5	—	0.76	7
No. of rosette leaves at						
21 days	18.8	29.8	39.5	—	0.80	7
26 days	19.1	27.1	33.0	46.3	0.66	10
Flowering	0.3	21.3	25.1	—	0.59	13
Rosette size (mm) at						
21 days	19.1	33.4	44.7	9.3	0.82	4
26 days	20.5	32.2	44.0	1.2	0.86	3
Flowering	28.1	27.6	34.5	32.1	0.71	10
36 days	24.8	26.4	34.2	53.6	0.77	10
No. of cauline leaves at						
26 days	11.7	32.1	39.0	—	0.66	5
Flowering	21.7	15.7	15.7	49.1	0.11	21
36 days	16.0	17.9	23.3	—	0.78	19
No. of buds at flowering	15.7	18.5	17.6	13.0	0.00	28
Height (mm) at						
Flowering	21.5	17.6	30.6	10.9	1.22	15
30 days	24.4	21.0	30.9	34.1	0.97	5
35 days	36.7	28.9	46.8	30.5	1.12	6
40 days	37.8	31.3	50.6	51.2	1.11	10
45 days	42.5	34.6	56.6	55.9	1.13	9
Maturity	54.0	54.3	73.4	59.8	0.84	6
Leaf size (mm)						
Length at 40 days	18.9	23.2	28.5	27.5	0.68	8
Width at 40 days	1.8	10.4	16.6	3.4	1.10	9

Values in bold are significant at $P \leq 0.05$.

The RILs and L_1+L_2 data provide locations and additive effects (a) while L_2-L_1 provides locations and dominance effects (d). Such small numbers of genotypes are far from ideal and reduce the power of the analysis considerably. Of the 330 analyses performed (22 traits \times 5 chromosomes \times 3 sets {RILs, L_1+L_2 and L_2-L_1 }), 38 were significant at $P < 1\%$ on a whole chromosome basis (compared to those expected by chance alone of $330/100 = 3.3$) giving an expected false-positive rate of about 8% (3.3/38). However, consistency of locations and effects across the three independent sets of data strengthens belief in the QTL identified. The estimated QTL locations and effects are shown in Table 6 for cases where $P < 1\%$, together with the additive and dominance effects of the same traits estimated from the basic generations, for comparison. The locations are also shown in Figure 2.

Those traits representing height at various ages locate a single major QTL at ~ 50 cM on chromosome II, the site of the *erecta* mutation (Table 6). This location, the size and direction of the genetic effects are consistent over time and across RILs and TTC families. Thus, the additive effect of the QTL, a , is negative (ie Col is larger than Ler) and the dominance effect, d , is positive as before, that is Col alleles are dominant. However, there are inconsistencies. The additive effect of the '*erecta*' QTL is considerably less than the additive effect [a] at 30–40 days but is larger by 45 days and beyond. This small initial effect, which increases greatly at maturity, is entirely consistent with the known effects of *erecta*. The size of d is considerably smaller than [d] as estimated

from the basic generations by weighted least squares (eg 68.9 as opposed to 184.9 for height at maturity). Conversely, d is always much larger than we would expect from the F_1 mean (eg 68.9 opposed to 32.3 for height at maturity). These discrepancies between the effects from the basic generation and QTL analyses are consistent with our previous conclusion that more than one QTL is involved, but the other QTLs have failed to achieve significance. Similar effects associated with *erecta* are found for rosette size at all ages, except at 26 days, and for leaf size at 40 days. However, there is also evidence for another QTL for rosette size on chromosome IV.

A further QTL at ~ 60 cM on chromosome IV affects time to maturity, rosette size at 26 days and number of cauline leaves at 26 days. There is some indication of QTL on chromosomes I (cauline leaves at 26 days and at flowering) and V (rosette leaves at 26 days and at flowering, and cauline leaves at flowering), but the locations are inconsistent. No QTL was detected on chromosome III.

Discussion

All traits show significant genetical variation that, apart from two exceptions (buds and cauline leaves at flowering), is consistently detectable from the basic generations (means and variances), and from the RILs and TTC ANOVAs. Heritabilities are typically 20–40% and, in the case of the height traits, increase consistently with age.

Table 6 QTL locations and effects compared to genetic effects from basic generations

Trait	Basic Gens. (WLS)		(Ler–Col)/2	F_1-m	RILs	L_1+L_2	L_2-L_1	Chrom	Loc. (cM)
	[a]	[d]							
Days to									
First true leaves	0.71	–	–	–	–	–	–	–	–
First flower	0.71	–	0.75	–0.05	–	–	–	–	–
Maturity	1.24	–1.07	1.36	–1.07	0.46*	0.51	0.79**	4	50;52;78
No. of rosette leaves at:									
21 days	–1.01	–	–1.03	–0.33	2.9& –2.9**	–	–	2	30&48;-;-
26 days	–0.72	–	–0.88	–0.43	–0.42**	–	–	5	28;-;-
Flowering	–1.04	–	–1.00	–0.03	–0.75**	–0.49**	–	2	44;28;-
Rosette size (mm) at									
21 days	–4.46	–	–4.50	–0.03	–	–	3.12**	2	-;-;56
26 days	–7.13	17.39	–7.13	1.30	–2.92**	–	–	4	62;-;-
Flowering	–4.34	–	–3.95	–2.91	–5.7**	–7.28**	5.71**	2	50;42;64
36 days	–8.8	–	–12.80	0.67	–7.95**	–8.11**	6.15*	2	52;36;50
No. of cauline leaves at									
26 days	1.29	–	–0.08	0.78	–0.98**	–1.61*	–	1	82;92;-
					0.96**	1.38	–	2	24;14;-
					–0.94**	–0.96	–	4	58;56;-
Flowering			–0.73	–1.98	–	2.56***	1.52*	1	-;44;70
					–2.27**	–	–	5	26;-;-
36 days	–4.58	–	–6.78	3.48	–	–	–	–	–
No. of buds at flowering	–	–	–0.38	–0.71	–	–	–	–	–
Height (mm) at									
flowering		6.56	–1.50	4.71	–5.7**	–7.28**	5.71**	2	50;42;64
30 days	–24.6	49.27	–24.60	13.20	–5.89	–15.15**	25.32**	2	52;64;50
35 days	–29.9	73.22	–29.93	17.50	–16.37**	–27.2**	44.68**	2	52;52;54
40 days	–42.1	290.4	–42.08	14.40	–30.49**	–36**	52.87**	2	50;48;52
45 days	–45	124.3	–44.95	19.13	–44.19**	–40.83*	63.07**	2	50;48;52
Maturity	–45.7	184.9	–45.67	32.31	–61.04**	–56.63**	68.91***	2	50;48;50
Leaf size (mm)									
Length at 40 days	–6.37	–	–5.92	–0.70	–4.39**	–4.61**	2.67**	2	52;44;50
Width at 40 days	–	–	0.33	–0.54	–	–	–	–	-;-;-

Columns 2 and 3, [a] and [d] from BGs; cols. 4 and 5, [a] and [d] from Col, Ler and F_1 ; cols. 6 and 7 additive QTL effects from RILs and TTC (L_1+L_2); col. 8, dominance QTL effects from TTC (L_2-L_1); cols. 9 and 10, chromosome and location of QTL. (* $0.05 \geq P > 0.01$, ** $0.01 \geq P > 0.001$, *** $P < .001$).

This trend is not observed for the other traits but is not uncommon for quantitative traits (Jayasekera *et al*, 1994). There is good agreement between the heritability estimates from the Basic Generation, RILs and the TTC families which points to a consistent expression of alleles under the varying genotypic backgrounds.

Given the repeated measurements over time and the nature of the traits scored, one would expect them to be correlated *inter se*. Ler develops more slowly than Col (Table 2) and hence, not surprisingly, it is smaller at any given chronological time from sowing. Thus it has fewer leaves, smaller rosettes and is shorter. However, this is also true when these same traits are scored at a fixed physiological time, that is, at flowering.

Dominance effects among the means were detected consistently for the height traits, but among the other traits, only rosette size at 26 days and days to maturity had significant dominance. Where it exists, dominance was for faster development and greater height. 'Better parent heterosis' does not occur for any trait except 'height at flowering', although the F_1 is not significantly taller than tallest parent, Col. Conversely, the TTC ANOVAs consistently detect dominance variation, except for 'buds at flowering', and the dominance ratios,

$\sqrt{(2V_D/V_A)}$, indicate partial dominance for most traits through to complete dominance for height. There is no evidence for significant overdominance. The fact that V_D occurs when [d] does not, implies ambidirectional dominance for such traits, because the former is a function of Σd^2 and the latter Σd . Similarly, the very large ratios of [d]/[a] (from 2 to 7) for heights contrast with the dominance ratios of around 1, indicating directional dominance and dispersed increasing alleles in the parents.

Model fitting confirms the above interpretation to a large extent. While additive differences [a] are significant for 18 of the 22 traits under study, dominance and dominance \times dominance epistasis are present for only eight traits, six of which are heights. In all cases where both [d] and [dd] are significant, they take opposing signs indicating the presence of duplicate epistasis; that is, heterozygosity at several loci has less heterotic effect than would be suggested by their individual effects. The only trait for which [d] is negative and significant is days to maturity, indicating that genes conferring early flowering are dominant to those responsible for late flowering. Another interesting feature of the results in Table 2 is the consistency with which [ad] is detected in

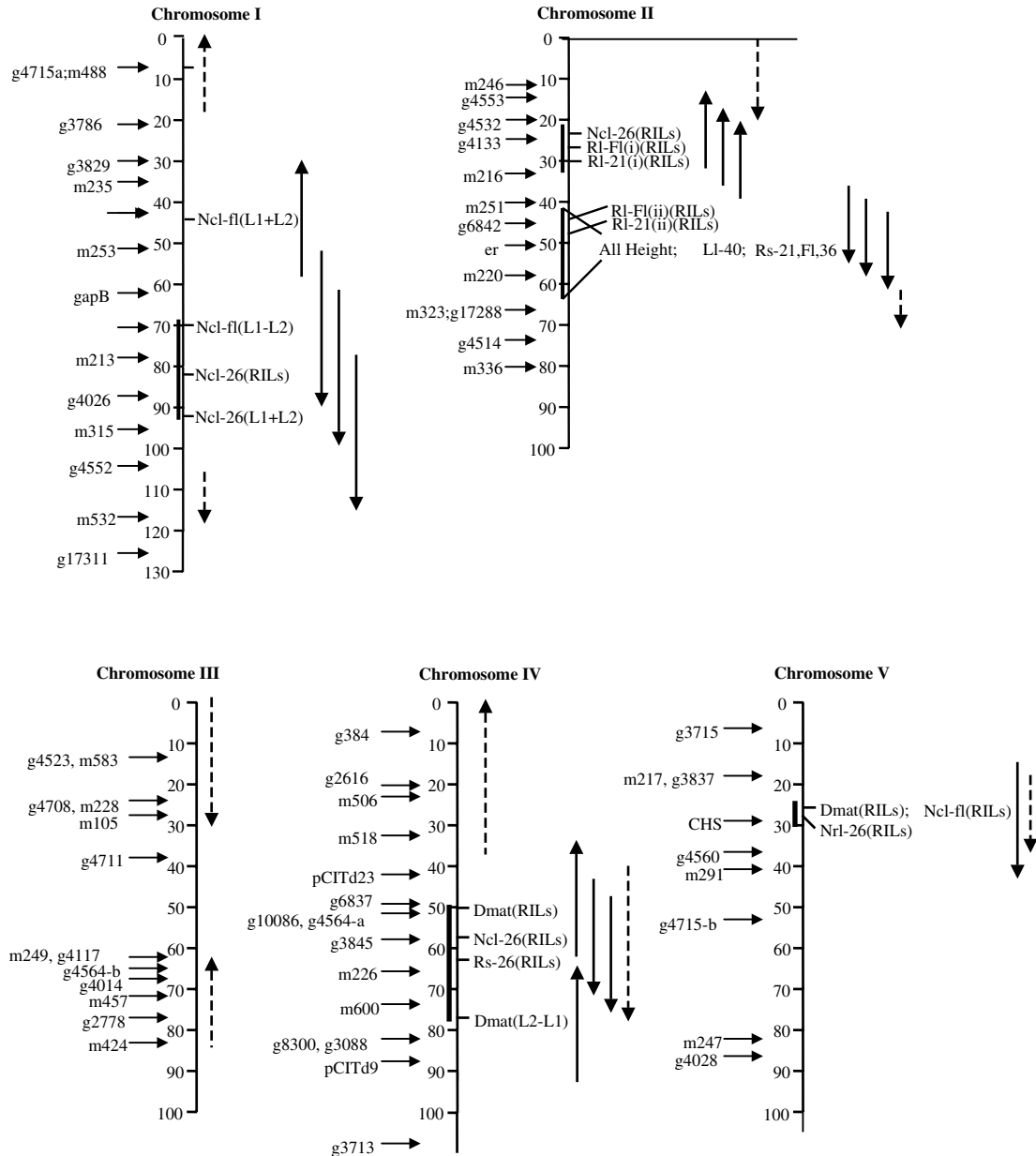


Figure 2 Chromosome maps indicating locations of markers and QTLs. Arrows indicate confidence interval of QTL. Solid arrows are QTL located in present study, dashed arrows indicate leaf-related QTLs identified by Jansen *et al* (1995). Down-arrow indicates Col allele increases score; up-arrow indicates Ler allele increases score. Vertical bars indicate QTL clusters likely to indicate the same QTL.

this study. This component is significant for 15 traits and it takes a negative sign in every case. Detection of [ad] with such a high frequency and the consistency of its sign suggests that Ler and Col have an excess of genes in coupling for most traits, particularly for plant height, leaf size, rosette size, rosette leaves and maturity. There is little evidence for additive-by-additive [aa] epistasis either in the BGs or in the TTC.

The fact that there is epistasis, dispersion and ambidirectional dominance indicates that these traits must be controlled by at least two QTL. Following Wright (1934), it is possible to estimate the number of genes, k , from $(\Sigma a)^2 / \Sigma a^2$; Σa is best estimated as half the difference between the extreme RILs, while Σa^2 can be estimated from $2V_A$. These estimates of k are shown in

Table 5 and they suggest that there are between three and 15 genes for each trait. It is well known that such estimates are imprecise and tend to be minimal estimates (Mayo and Hopkins, 1985; Kearsley and Pooni, 1996), but they are all consistent in suggesting that, despite their close origins, Col and Ler now differ by several genes controlling most quantitative traits.

The QTL analyses are consistent with the biometrical approach in many respects (Table 6). Thus, the direction and size of the additive effects for height, rosette size and rosette leaf number agree, with Col alleles giving larger plants: the positive direction of the dominance effects for height and rosette size and the absence of evidence for dominance for rosette leaf number. However, there are also significant inconsistencies. The basic generations

and TTC indicated significant gene interaction for most traits implying two or more QTL, although QTL analysis seldom revealed more than one QTL for any trait. This may reflect the low power of QTL detection given the restricted sets of genotypes. However, the estimated sizes of the dominance effects of the QTL for height were much larger than the corresponding additive effects and also larger than expected given the observed F_1 values. Both these effects predict overdominance at the QTL and heterosis, neither of which was detected by the more sensitive biometrical analyses. The estimates of additive and dominance effects from the TTC have equal precision, so this apparent exaggeration of dominance cannot be due to estimation bias. The squared additive effects of the QTL from RILs and TTC in Table 6 equal or slightly exceed the estimated additive genetic variances from these generations, suggesting that all the variance has been explained by these QTL. The squared dominance effects, however, are considerably larger than the estimated dominance variation. These effects strongly suggest an upward bias of estimates as shown by Melchinger *et al* (1998) and Utz *et al* (2000). We have good evidence from other work (Koumproglou *et al*, 2002) that there is another QTL, located at the top of chromosome III, at which the Col allele also increases height. Jansen *et al* (1995) also located a QTL affecting leaf number in this location. It is almost certainly flowering related and decreases in effect with time. The combined effects of the two QTL at 30 days to maturity would probably match the additive effect [a] from the basic generations. This other gene may be interacting with *erecta* in the heterozygous condition causing the discrepancies in the amount of dominance discussed above.

There are potentially other analyses that may be carried out on these data, such as the use of principal components analysis to identify key components for QTL and generation mean analysis and these are ongoing.

References

- Doganlar S, Frary A, Tanksley SD (2000). The genetic basis of seed weight variation: tomato as a model system. *Genetics* **100**: 1267–1273.
- 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.
- Jayasekera NEM, Karunasekera KB, Kearsey MJ (1994). Genetics of production traits in *Hevea brasiliensis* (rubber). I. Changes in genetical control with age. *Heredity* **73**: 650–656.
- Kearsey MJ, Farquhar AJ (1998). QTL analysis in plants: where are we now? *Heredity* **80**: 137–142.
- Kearsey MJ, Hyne V (1994). QTL analysis: a simple 'marker-regression' approach. *Theor Appl Genet* **89**: 698–702.
- Kearsey MJ, Jinks JL (1968). A general method of detecting additive dominance and epistatic variation for metrical traits. I. Theory. *Heredity* **23**: 403–409.
- Kearsey MJ, Pooni HS (1996). *The Genetical Analysis of Quantitative Traits*. Chapman and Hall: London.
- Koumproglou R, Wilkes TM, Townson P, Wang XY, Beynon J, Pooni HS *et al* (2002). STAIRS: a new genetic resource for functional genomic studies of *Arabidopsis*. *Plant J* **31**: 355–364.
- Lister C, Dean C (1993). Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. *Plant J* **4**: 745–750.
- Lynch M, Walsh B (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer Associates: Sunderland, MA.
- Mayo O, Hopkins AM (1985). Problems of estimating the minimum number of genes contributing to quantitative variation. *Biom J* **27**: 181–187.
- Melchinger AE, Utz HF, Schon CC (1998). QTL mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* **149**: 383–403.
- Monforte AJ, Friedman E, Zamir D, Tanksley SD (2001). Comparison of a set of QTL NILs for chromosome 2 of tomato: deductions about natural variation and implications for germplasm utilization. *Theor Appl Genet* **102**: 572–590.
- Sokal RR, Rohlf FJ (1981). *Biometry*. WH Freeman and Co.: New York.
- Tanksley SD (1993). Mapping polygenes. *Annu Rev Genet* **27**: 205–233.
- Utz HF, Melchinger AE, Schön CC (2000). Bias and sampling error of the estimated proportion of genetic variance explained by QTL determined from experimental data in maize using cross validation and validation with independent samples. *Genetics* **154**: 1839–1849.
- Wright S (1934). The results of crosses between inbred strains of guinea pigs differing in number of digits. *Genetics* **19**: 537–551.