J.Natn.Sci.Foundation Sri Lanka 2008 **36** (4):275-280

RESEARCH ARTICLE

Chromosome substitution lines for the analysis of heterosis in *Arabidopsis thaliana*

S.A.C.N. Perera[1*], H.S. Pooni[2] and M.J. Kearsey[2]
[1] *Genetics & Plant Breeding Division, Coconut Research Institute of Sri Lanka, Lunuwila.*
[2] *School of Biosciences, University of Birmingham, Edgbaston, Birmingham, UK.*

Revised: 19 March 2008 ; Accepted: 14 July 2008

Abstract: Heterosis is an important phenomenon used in agricultural crop improvement for several decades. The molecular basis of heterosis is still unclear despite the wide exploitation of heterosis for improving economically important traits of crops. The current experiment investigates the heterosis related to several morphological traits in *Arabidopsis thaliana* on an individual chromosome basis of cr. Whole chromosome substitution lines (CSLs) of chromosomes three, four and five of *Arabidopsis*, recurrent parent Columbia and the three relevant F_1 progenies were grown in a growth chamber in a completely randomised experimental design for the scoring of morphological traits. Data were analysed to detect the presence or absence of heterosis for each trait on per-chromosome basis. The traits rosette width at day 20, flowering time, height at flowering time, rosette and cauline leaf numbers at flowering displayed significant heterosis in different chromosomes. The findings revealed the presence of both positive and negative heterosis for rosette leaf number at flowering on different chromosomes. Genetic components of the means and the potence ratios indicated those parents having more dominant alleles in the cross in each chromosome. The results indicated the appropriateness of CSLs in studying heterosis in a micro scale rather than the whole genome level. The study demonstrates that breeding programmes can be designed to keep the desirable effects and remove the disadvantageous effects of heterosis caused by the genes present in different chromosomal parts in the genome.

Keywords: *Arabidopsis*, chromosome substitution lines, heterosis.

INTRODUCTION

Heterosis is the superiority of an F_1 hybrid over the mean performance of its better-inbred parent. This phenomenon is also known as hybrid vigour and has been widely used in crop and livestock breeding in agriculture and in evolutionary genetics[1,2]. The term heterosis is also used occasionally to describe the F_1 exceeding the mid-parental value of a particular cross or to the hybrid progeny of a top-cross between several inbred lines exceeding the average performance of the parental inbreds. Such heterosis is referred to as mid-parent heterosis while the superiority of the hybrid over the better parent is termed better parent heterosis[3] or heterobeltiosis.

In the absence of gene interactions and other complicating genetic relationships between loci, heterosis is displayed when the average dominance is greater than the degree of gene dispersion. As a result, if the genes are dispersed in the two parents it requires only very little dominance at individual genes to produce quite considerable heterosis. When epistasis, i.e. the interaction between genes is present heterosis can be caused by dominant complementation of genes. This theory indicates that in the presence of complementary epistasis, dominant increasing alleles at each of the two genes have a proportionately greater effect when they occur together than would be expected from their individual effects[3].

Apart from the dominant complementation of genes, yet another hypothesis as to the cause of heterosis states that over-dominance also may contribute to heterosis in some instances. Yet evidence for actual over-dominance remains scarce.

Most of the practical interest in heterosis centres on the breeding of cultivated plants. Heterosis for size, vigour or yield is most evident in outbreeding crops, and over the years plant breeders have exploited the hybrid vigour in producing mostly high yielding and

*Corresponding author

environmentally stable cultivars of crops. Experiments with wheat have provided evidence to support the hypothesis that heterosis in wheat is due to dominant complementation with linkage and interaction of alleles[4]. Maize is a very good example in which heterosis has long been exploited in the production of high yielding F_1 seeds in commercial quantities and the same theory was shown to apply for heterosis in maize in the experiments conducted in 1983[5].

Studies have revealed that over-dominance is not a major cause of heterosis for yield in a cross between the two subspecies of rice, because there was no correlation between most traits and overall genomic heterozygosity [2]. In this experiment heterozygotes were never superior to both homozygotes in analysis of Quantitative Trait (QTL) and some F_8 inbred lines were actually superior to the F_1 for all traits evaluated.

Despite the exploitation of heterosis in improving desirable traits in crop and animal breeding the underlying genetic basis of heterosis is still unclear. The advances of molecular genetics in the recent past, paves the way for the study of the molecular basis of heterosis. A model organism such as *Arabidopsis* to which the most advanced molecular techniques can be applied, is a suitable candidate to be used in such a study. The availability of whole chromosome substitution lines (CSLs) in *Arabidopsis* facilitates the investigation of heterosis on an individual chromosome basis rather than on the whole genome.

Arabidopsis has a haploid karyotype of five chromosomes and accordingly five whole CSLs can be produced. Previous studies[6] reported the production of CSLs in *Arabidopsis* for the genetic analysis of quantitative traits using a marker assisted breeding programme. *Arabidopsis* accessions Columbia (Col) and Landsberg (L*er*) have been used as the recurrent parent and donor parent respectively in this breeding programme because of the extensive existing genetic analysis information and the sequence information on these lines.

When using CSLs in studying heterosis, the relevant CSL and the recurrent parent (Columbia) are considered as the inbred parents. A cross between these two parents will generate an F_1 progeny in which the particular chromosome is entirely heterozygous in a pure homozygous genetic background of the recurrent parent. The genetic analysis of desired traits in these three populations (two inbred parents and the F_1) facilitates the studying of heterosis on an individual chromosome basis. The objective of the current study was to study the chromosomal basis of heterosis for some morphological traits in *Arabidopsis* chromosome substitution Lines.

METHODS AND MATERIALS

Plant material: Seeds of CSLs of *Arabidopsis* chromosomes 3, 4 and 5, (CSL3, CSL4 and CSL5), the relevant F_1 progeny of CSLs (obtained by backcrossing each CSL to the recurrent parent Columbia (CSL3-F_1, CSL4-F_1 and CSL5-F_1) along with Columbia (Col) (the recurrent parent in substitution lines) were used in the experiment.

Twenty-five plants of each line (a total of 175 plants) were grown in a completely randomised experimental design surrounded by a peripheral guard row under controlled environmental conditions. Each plant was grown in a 5-inch pot filled with compost mixture. Three seeds were placed in each pot at sowing in order to allow for non-germination, and ten days after germination the plants were thinned down to one plant per pot. The environment in the growth chamber was controlled, and maintained at 16 h photoperiod and 24^0C temperature.

Traits scored: The traits germination time (GT), bolting time (BT-days from germination to bolting), flowering time (FT-days from germination to opening of the first flower) leaf number and the rosette width at day 20 from sowing (LN-20 and RW 20), height, rosette and cauline leaf number at flowering time (HF, RLNF and CLNF) were scored for each plant.

Data Analysis: Descriptive statistics were calculated for all the traits for the lines Col, CSL3, CSL4, CSL5, CSL3-F_1, CSL4-F_1 and CSL5-F_1. Analysis of variance was performed among the 3 lines (two parents and the F_1) for each chromosome for each trait. The means of Col, each CSL and the relevant F_1 were compared to detect whether the F_1 means were either lower or higher than the means of the parents. When the mean value of the F_1 was either below or above the means of both the parents, Student's t test was performed to compare the means of the F_1 and the parent that has a mean closer to the F_1 value. If it was significant, the trait was identified to display hybrid vigour. When germination-time was shown to have an effect on the observed variation among the lines (as displayed by ANOVA using germination time as a co-variate) the adjusted means with the covariate were used in calculating the heterosis.

In addition, homogeneity tests (Bartlett's test) for the variance between the three populations in each chromosome were performed for each trait to confirm the uniformity of environmental variation for each trait.

The mid parental value m, the additive and dominance components (a_A and d_A respectively) were calculated for each trait for chromosomes 3, 4 and 5 using the following formulae.

$$m = (P_1 + P_2) / 2$$
$$a_A = (P_1 - P_2) / 2$$
$$d_A = F_1 - (P_1 + P_2) / 2$$

P_1 = mean of Col, P_2 = mean of CSL, F_1 = mean of F_1 The potence ratio ($d_A / |a_A|$) was then calculated for each trait in each chromosome.

RESULTS

Analysis of variance

Analysis of variance revealed statistically significant differences for GT ($p>0.0001$), FT ($p=0.013$), CLNF ($p<0.0001$) and RWF ($p=0.001$) between Col (parent 1) and CSL3-F_1. Traits BT ($p>0.0001$), FT ($p=0.0009$), LN-20 ($p=0.003$) and RW-20 ($p=0.004$) were significantly different between CSL3 (parent 2) and CSL3-F_1.

For chromosome 4, FT ($p=0.04$) was different between Col (parent 1) and CSL4-F_1, while the traits GT ($p=0.013$), BT ($p<0.0001$), FT ($p>0.0001$), HF ($p=0.008$), RLNF ($p>0.0001$), CLNF ($p>0.0001$) and RWF ($p=0.0008$) were significantly different between CSL4 (parent 2) and CSL4-F_1.

ANOVA revealed significant differences for GT ($p=0.003$), LN-20 ($p=0.049$), HF ($p=0.007$), RLNF ($p=0.009$) and RWF ($p= 0.006$) between Col (parent 1) and CSL5-F_1 while the traits LN-20 ($p>0.0001$), RW-20 ($p>0.0001$), RLNF ($p=0.0123$) and RWF ($p>0.0001$) were significantly different between CSL5 (parent 2) and CSL5-F_1.

Detecting heterosis

The means for the traits, and traits showing heterosis are given in Table 1. Out of all the traits scored for each of the chromosome, GT (chr. 3), LN-20 (chr.3, chr.5), RW-20 (chr.3, chr.5), HF (chr.3), RLNF (chr.3, chr.5) and CLNF (chr.3, chr.5) showed F_1 means exceeding the means of parent. The traits RW-20 (chr.4), BT (chr.4), HF (chr.4), RLNF (chr.4) and CLNF (chr.4) showed means less than the means of each parent.

The results of the t-tests conducted to determine whether the observed mean differences were significantly different indicated the presence or absence of heterosis of the above listed traits on the relevant chromosomes. Accordingly, significant positive heterosis (better parent) was observed for the traits LN-20 (chr.5), RW-20 (chr.3),

Table 1: Means and the standard errors of the means (in parenthesis) of each line for each of the traits scored. The means of the F_1s that exceed the parents, in either direction, are underlined while those showing significant heterosis are underlined and given in bold letters.

Line	GT	LN-20	RW-20	BT	FT	HF	RLNF	CLNF
Col	3.200 (0.100)	13.420 (0.447)	47.490 (2.053)	16.600 (0.252)	21.170 (0.310)	49.080 (4.630)	22.080 (1.230)	7 .000 (0.510)
CSL3	4.750 (0.296)	12.880 (0.445)	44.640 (2.045)	19.545 (0.314)	23.890 (0.453)	40.550 (2.780)	23.730 (1.470)	8.364 (0.509)
CSL3-F_1	4.864 (0.249)	14.580 (0.463)	**55.660** (2.128)	17.318 (0.304)	21.900 (0.367)	**61.710** (3.920)	26.120 (1.110)	**9.941** (0.433)
CSL4	3.545 (0.157)	13.110 (0.441)	47.370 (2.028)	18.522 (0.416)	24.37 (0.378)	67.000 (5.670)	24.000 (1.310)	11.250 (0.413)
CSL4-F_1	3.208 (0.104)	13.390 (0.445)	46.200 (2.946)	15.640 (0.416)	21.900 (0.367)	(48.080) (3.930)	**16.380** (1.150)	6.610 (0.610)
CSL5	4.240 (0.210)	12.680 (0.432)	41.560 (1.983)	15.440 (0.332)	20.220 (0.321)	72.950 (4.410)	22.136 (0.825)	7.000 (0.354)
CSL5-F_1	4.000 (0.225)	**15.890** (0.440)	**60.310** (2.022)	15.826 (0.241)	20.290 (0.311)	98.090 (5.540)	**26.170** (1.130)	8.261 (0.480)

GT= germination time, LN-20 = leaf number at day 20, RW-20 = rosette width at day 20, BT = bolting time, FT = flowering time, HF = height at flowering time, RLNF = rosette leaf number at flowering time and CLNF = cauline leaf number at flowering time

RW-20 (chr.5), HF (chr.3), RLNF (chr.5) and CLNF (chr.3) while significant negative heterosis was observed for RLNF (chr.4).

Bartlett's homogeneity test for variance

Test for equal variance in chromosome three showed that the variation among the three populations (parent 01–Col, parent 02-CSL and F_1) differed for germination time ($p>0.0001$) while for all the other traits the observed variation among the populations was uniform. The same was observed for the three populations in chromosome 4 (p for germination time =0.041). In chromosome 5, in addition to germination time ($p=0.001$), the two traits leaf number ($p>0.0001$) and rosette width at day 20 ($p>0.012$) were observed to be significantly different while the variance in other traits was uniform.

Genetic components of means and potence ratios

The calculated genetic components of means and the potence ratios based on the three population means for each chromosome for each trait are presented in Table 2. The dominant genetic components of the means were greater than the additive genetic components of the mean for several traits indicating (pseudo) over-dominance when each chromosome is considered as a whole.

The dominant components of means were greater than the additive components of the means (in either direction) for the traits GT, LN-20, RW-20, HF, RLNF and CLNF in chromosome 3. This was true with the traits RW-20, BT, HF, RLNF, and CLNF in chromosome 4 and LN-20, RW-20, RLNF and CLNF in chromosome 5 but only significantly so for those underlined and bold.

DISCUSSION AND CONCLUSION

Traits showing heterosis

Significant positive heterosis was observed for trait LN-20 in chromosome 5, RW-20 in chromosomes 3 and 5, HF in chromosome 3, RLNF in chromosome 5 and CLNF in chromosome 3. Negative heterosis (F_1 mean lower than the mean of lower parent) was observed for RLNF in chromosome 4.

Due to the non homogeneity of population variances for certain traits, heterosis has been confirmed only for RW-20, HF and CLNF in chromosome 3, RLNF in chromosome 4 and RLNF in chromosome 5.

Identification of both positive and negative effects of heterosis on the same trait (RLNF) in different chromosomes was facilitated by the use of CSLs. If conventional hybrids were used, only the net effect of the different effects would have been noticed.

Inequality of population variances

The three populations P_1, P_2 and F_1 are all non segregating populations. Thus, there is no genetic parameter that contributes to the observed variation. Because all three generations are grown under the same environment the variability that is observed among the populations should

Table 2: The mid parental values, additive (a_A) and dominant (d_A) genetic components of means, and the potence ratio (d_A/a_A) for each trait in each chromosome

	GT	LN-20	RW-20	BT	FT	HF	RLNF	CLNF
Chr.3								
m	3.985	13.150	46.065	18.075	22.530	44.815	22.905	7.682
a_A	-0.775	0.270	1.425	-1.473	-1.360	4.265	-0.825	-0.682
d_A	0.889	1.430	9.595	-0.754	-0.630	16.895	3.215	2.259
d_A/a_A	1.147	5.296	**6.733**	-0.512	-0.460	**3.960**	3.897	**3.312**
Chr.4								
m	3.373	13.265	47.430	17.561	22.770	58.040	23.040	9.125
a_A	-0.172	0.155	0.060	-0.961	-1.600	-8.960	-9.960	-2.125
d_A	-0.165	0.125	-1.230	-1.921	-1.325	-9.960	-6.660	-6.465
d_A/a_A	-0.954	0.806	-20.50	-1.999	-0.828	-1.112	**-6938**	-3.042
Chr.5								
m	3.720	13.050	44.525	16.020	20.695	61.015	22.108	7.000
a_A	-0.520	0.370	3.165	0.580	0.475	-11.94	-0.028	0.000
d_A	0.280	2.840	15.785	-0.194	-0.405	7.075	4.062	1.261
d_A/a_A	0.538	**7.676**	**4.987**	-4.653	-0.853	0.593	**145.07**	n/a

GT= germination time, LN-20 = leaf number at day 20, RW-20 = rosette width at day 20, BT = bolting time, FT = flowering time, HF = height at flowering time, RLNF = rosette leaf number at flowering time and CLNF = cauline leaf number at flowering time

be equal provided that there is no interaction between the genotypes and the environment. The non homogeneity of variance in germination time can be attributed to maternal effects rather than genotype environment interaction because germination time has been shown to have considerable maternal effects in *Arabidopsis*[7]. Inequality in variances has also been observed in populations of chromosome 5 for LN-20, RW-20 as well. Although no definitive conclusion can be drawn as to the reasons for this, genotype–environment interaction can be suggested as a reason. It is also noteworthy that this inequality in variances for these two additional traits has been observed only in chromosome 5.

Genetic components of the means and dominance/ potence ratios

The additive and dominant components of the means help determining the presence or absence of dominance for the traits tested. Potence/dominance ratio gives an indication about the parent having the most dominant alleles and is, therefore, more potent in the cross. This ratio cannot be equated with the true dominance ratio that can be obtained only when dominance is uni-directional and there is complete association of alleles in the parents. Genetic variance based calculations would have been better in understanding the true genetic basis including epistasis. However, this study is limited to non segregating populations, thus preventing the deployment of such calculations. Consequently, the stated observations have been made on the observed values of potence ratio with increasing and decreasing effects of alleles. When considering the increasing and decreasing alleles also, this discussion is based on the net increase or decrease of all the responsible alleles present within the chromosome concerned.

Table 2 shows that when considering the increasing and decreasing effects of alleles, Col carried increasing alleles for LN20, RW-20 and HF while L*er* alleles had increasing effects on the traits BT and flowering related traits RLNF and CLNF in chromosome 3. Similarly the observation of the genetic components of the means and the potence ratio revealed that Col contained more dominant alleles in chromosome 3 for the traits LN-20. RW-20 and HF while, L*er* contained more dominant alleles for RLNF and CLNF.

In chromosome 4, Col alleles had increasing effect on traits LN-20, and RW-20 while L*er* alleles had increasing effects on traits BT, HF, RLNF and CLNF. The very highly negative potence ratio of RW-20 in chromosome 4 is attributed to the very narrow difference between the parental means. In addition, the potence ratio revealed that L*er* contained more dominant alleles for the traits HF, RLNF and CLNF.

With regard to chromosome 5, Col alleles were increasing the trait in LN-20, RW-20, BT and FT and the potence ratio indicated that Col alleles were more dominant in traits LN-20, RW-20, BT. L*er* alleles were observed to be increasing the traits HF and RLNF in chromosome 05. The very high potence ratio for the trait RLNF in chromosome 05 was due to the very narrow difference in parental means.

Narrow parental means results in high potence ratios as shown in Table 2 for RLNF in chromosome 5. With respect to heterosis such instances mainly imply higher degree of gene dispersion in the parents. This points out the possibility of producing lines which outperform F_1 by incorporating several increasing alleles dispersed in the parents, into lines developed by further cycles of breeding.

General comments

A study including the populations for all the five chromosomes in *Arabidopsis* and more basic generations including segregating populations such as F_2 and backcrosses would facilitate a comprehensive analysis on heterosis on each chromosome level which could later be added for the whole genome scale. The unavailability of any further basic generations beyond F_1 prevented tests for determining the presence or absence of complicating factors such as epistasis. Hybrid vigour has been reported in *Arabidopsis* for height at flowering in a far more detailed study using sixteen basic generations of the same ecotypes Col and L*er*[8]. They had scored more or less the same traits but had reported the findings on a whole genome basis so that the results are not strictly comparable with the current study. In the current experiment we have observed heterosis for more traits but on an individual chromosome basis with a fewer number of generations. But, as the results of the current study indicate, there are positive and negative effects of different chromosomes on the same trait. The whole genome analysis will reveal the net effect of all the effects. This stresses the advantage of studying heterosis at a more micro level so that the breeding programmes can be designed to keep the desirable effects and eliminate the disadvantageous effects caused by the genes present in different chromosomal parts of the genome. The use of CSL in the study of heterosis helps investigating the phenomenon of hybrid vigour at a more micro level rather than the whole genome level. This would help in resolving so far unanswered questions about the genetic basis of heterosis.

Acknowledgement

The authors wish to thank Andrew Bennet for recording morphological data.

References

1. Stuber C.W., Lincoln S.E., Wolff D.W., Helentjaris T. & Lander E.S. (1992). Identification of genetic-factors contributing to heterosis in a hybrid from 2 elite maize inbred lines using molecular markers. *Genetics* **132**(3):823-839.
2. Xiao J.H., Li J.M., Yuan L.P. & Tanksley S.D. (1995). Dominance is the major genetic-basis of heterosis in rice as revealed by QTL analysis using molecular markers. *Genetics* **140**(2):745-754.
3. Kearsey M.J. & Pooni H.S. (1996). Applications. In: *The Genetical Analysis of Quantitative Traits*. pp. 302-334. Chapman & Hall, UK.
4. Pickett A.A. & Galwey N.W. (1997). A further evaluation of hybrid wheat. *Plant Varieties and Seeds* **10**(1):15-32.
5. Sprague G. F. (1983). Heterosis in maize: theory and practice. In: *Heterosis: Reappraisal of Theory and Practice.* (Ed. R. Franken) pp. 47-70 Monographs on Theoretical and Applied Genetics 6. Springer, Germany.
6. Koumproglou R., Wilkes T.M., Townson P., Wang X.Y., Beynon J., Pooni H.S., Newbury H.J. & Kearsey M.J. (2002). STAIRS: a new genetic resource for functional genomic studies of *Arabidopsis. Plant Journal* **31**(3): 355-364.
7. Perera S.A.C.N. (2004). Fine Mapping of QTL and Microarray Gene Expression Studies in *Arabidopsis* using STAIRS. *Ph. D. thesis*. The University of Birmingham.
8. Kearsey M.J., Pooni H.S. & Syed N.H. (2003). Genetics of quantitative traits in *Arabidopsis thaliana. Heredity* **91**(5):456-464.