Genet. Res., Camb. (2000), **75**, *pp.* 115–121. *With 2 figures.* Printed in the United Kingdom © 2000 Cambridge University Press 115

A transmission/disequilibrium test approach to screen for quantitative trait loci in two selected lines of Large White pigs

MARCO C. A. M. BINK[1,2]*, MARINUS F. W. TE PAS[1], FRANK L. HARDERS[1] AND LUC L. G. JANSS[1]

[1] *DLO Institute for Animal Science and Health (ID-DLO), PO Box 65, 8200 AB Lelystad, The Netherlands*
[2] *DLO Centre for Plant Breeding and Reproduction Research (CPRO-DLO), 6700 AA Wageningen, The Netherlands*

(*Received 19 February 1999 and in revised form 7 June 1999*)

Summary

Pedigree and marker data from a multiple-generation pig selection experiment have been analysed to screen for loci affecting quantitative traits (QTL). Pigs from a base population were selected either for low backfat thickness at fixed live weight (L-line) or high live weight at fixed age (F-line). Selection was based on single-trait own performance and DNA was available on selected individuals only. Genotypes for three marker loci with known positions on chromosome 4 were available. The transmission/disequilibrium test (TDT) was originally described in human genetics to test for linkage between a genetic marker and a disease-susceptibility locus, in the presence of association. Here, we adapt the TDT to test for linkage between a marker and QTL favoured by selection, and for linkage disequilibrium between them in the base population. The *a priori* unknown distribution of the test statistic under the null hypothesis, no linkage, was obtained via Monte Carlo simulation. Significant TDT statistics were found for markers *AFABP* and *SW818* in the F-line, indicating the presence of a closely linked QTL affecting growth performance. In the L-line, none of the markers studied showed significance. This study emphasizes the potential of the TDT as a quick and simple approach to screen for QTL in situations where marker genotypes are available on selected individuals. The results suggest that previously identified QTL in crosses of genetically diverse breeds may also segregate in commercial selection lines.

1. Introduction

The availability of microsatellite DNA polymorphism has made it possible to identify polymorphic markers in almost any region of the genome. This development has led to numerous studies to test for association between these markers and phenotypes of important quantitative traits such as disease, reproduction and production. Genes affecting these quantitative traits are commonly referred to as quantitative trait loci (QTL). The dissection of traits into their individual Mendelian components allows animal and plant breeders to improve selection decisions in their breeding programme.

So far, QTL mapping has been applied in well-designed experiments. Often crosses of genetically very diverse breeds, e.g. pigs (Andersson *et al.*, 1994), or large paternal half-sib families, e.g. cattle (Georges *et al.*, 1995), are used. In many livestock breeding programmes the identification of large half-sib families is limited and often lacks power to screen for QTL. Including more generations of individuals may improve accuracy and power of QTL detection (Darvasi & Soller, 1995). Recently, in mice, Keightley *et al.* (1996) created a map of QTL for body weight via observing marker frequency divergence after 21 generations in a selection experiment. The feasibility of this type of experiment in livestock in questionable but not infeasible provided that collection and storage of DNA becomes routine in selection experiments. This was recently shown by Ollivier *et al.* (1997) in a pig selection experiment that has been in progress since 1973. Also, there are few or no QTL mapping approaches developed that can use retrospectively available data efficiently. Complications here are the unknown linkage phases in parents with small progeny groups and the continuing selection in the breeding programme.

* Corresponding author. Tel: +31 320 238 265. Fax: +31 320 238 050. e-mail: m.c.a.m.bink@cpro.dlo.nl

In selection experiments family sizes may be small. Human geneticists have developed several statistical tools to utilize phenotypic (disease) and marker data on complex and small pedigrees. Examples of family-based tests for linkage are the haplotype relative risk (HRR; Falk & Rubinstein, 1987), affected-sib-pair (ASP) and the affected family based controls (AFBAC; Thomson, 1988), and the transmission/disequilibrium test (TDT; Spielman *et al*., 1993). Originally, the intended use of the TDT was to test close linkage between a marker and a candidate disease gene. In the case of close linkage, parents that are heterozygous for the marker and disease gene will transmit one of the marker alleles more frequently to its disease-affected offspring (e.g. Spielman *et al*., 1993). Since its introduction, the TDT has also been used as a screening test, i.e. it is applied to data from many markers throughout the genome, without prior evidence of either population association or proximity to candidate genes (Spielman & Ewens, 1996).

In this study we apply the TDT to screen for QTL on chromosome 4 in data on a pig selection experiment. Andersson *et al*. (1994) identified regions containing QTL on chromosome 4 for growth and lean meat percentage in crosses of genetically very diverse pig breeds. However, segregation of these QTL in commercial selection lines is of a different nature and as yet unknown. Therefore, evidence from a selection line will be complementary to evidence from data on crosses.

2. Materials and methods

(i) *Selection experiment*

From a commercial Dutch Large White population, two divergent lines were created from one base population and these covered approximately four generations of selection. The F(ast)-line was selected for high growth rate and the L(ean)-line was selected for low backfat thickness. Including the base population, the total number of animals in the pedigree was 4356. The number of sires and dams in the experiment was 242 and 636, respectively; the remainder were final offspring. A maximum of four pigs (two males and two females) were performance-tested per litter and on the basis of performance data candidates were selected as new parents. Selected boars (sires) were mated to selected sows (dams) during a period of 10 weeks, and sows could have a maximum of three litters (= 18 months). Exact details of the trait measurements and mating procedure are given by Sonesson *et al*. (1998). Due to the long use of females, generations became overlapping and generation numbers became fractional (see below).

Table 2. *Number of matings in both selection lines containing genotyped offspring*

	F-line	L-line
No. of matings	309	304
No. of matings with genotyped offspring	197	173
No. of genotyped offspring	324	300

(ii) *Population structure*

Individuals in the base population were mated at random and selection started from their progeny onwards. The generation number of an individual (g_i) was calculated as $g_i = \frac{1}{2}(g_s + g_d) + 1$, where g_s and g_d are the generation numbers of its respective sire and dam. A typical example follows, for members of subsequent litters of a base population sow ($g_d = 0$), $g_i = 1$, $1\frac{1}{2}$ and $1\frac{3}{4}$ when the mates (boars) have generation numbers 0, 1 and $1\frac{1}{2}$, respectively. The mating structure caused the generation intervals to be continuous rather than discrete. Table 1 shows the population structure when using generation classes of 1·0. This table also shows that, in later generations, the selection intensity was somewhat higher in the F-line than in the L-line.

Selection took place immediately after the performance-testing period and selection was both over and within families (or litters). As a result, all, none or some individuals of a particular litter were selected. In Table 2 the numbers of matings for the

Table 1. *Population characteristics of two selection lines: F-line (selection on growth rate) and L-line (selection on low backfat thickness)*

Generation no.	F-line			L-line		
	Tested	Selected	Fraction	Tested	Selected	Fraction
0·0[a]	145	136	0·94	139	124	0·89
1·0	533	182	0·34	425	153	0·39
2·0	606	191	0·32	467	186	0·40
3·0	570	191	0·34	480	185	0·39
Total	1854	700		1511	648	

[a] This class holds all individuals with generation number from 0·0 up to 1·0.

https://doi.org/10.1017/S0016672399004061 Published online by Cambridge University Press

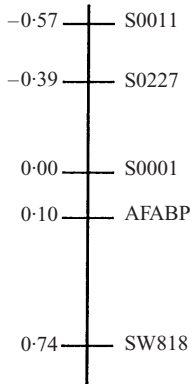

Fig. 1. Genetic markers with their relative positions (morgans) on the linkage map of chromosome 4. (Averaged male/female map; Pigmap October 1998, and F. Gerbens, personal communication.)

two selection lines are given. Not every mating resulted in offspring that were selected after performance testing. In total, 197 and 173 matings resulted in litters with selected offspring in the F- and L-line, respectively. Approximately 1·7 offspring per informative litter were genotyped, where 'informative litter' can be defined as a litter containing at least one genotyped offspring.

(iii) *Marker data*

Three microsatellite marker loci on chromosome 4 were considered, i.e. *S0001*, *AFABP* (adipocyte fatty acid-binding protein locus) and *SW818* (Fig. 1). The map positions of *S0001* and *SW818* loci were near the significant peaks of test-statistics for backfat thickness and growth rate, respectively, reported by Andersson *et al.* (1994). The data on marker *AFABP* were available since it was a candidate gene in another study (Gerbens *et al.*, 1998). The allelic frequencies in the Large White base population are unknown since base individuals were not blood sampled. In principle, blood samples were taken from all selected individuals and DNA was available from approximately 1050 individuals. When pedigree and marker data conflicted, marker data were not used. This was the case for 2%, 11% and 2% of genotypes for markers *S0001*, *AFABP* and *SW818*, respectively. Polymorphism of the marker loci was moderate, i.e. 3, 4 and 4 alleles, and allele frequencies were not very equal (see below).

(iv) *Transmission/disequilibrium test* (*TDT*)

The TDT is used to check jointly for linkage and linkage disequilibrium by testing whether alleles at a particular marker locus segregate randomly from parents to a specific subset of their offspring. In the original TDT, this subset contained the disease-affected offspring (children in human genetics). Here, the subset comprises those offspring that have been selected for a quantitative trait. Trait selection probably favours certain alleles at important QTL and consequently also affects the segregation of closely linked marker alleles. So, instead of 'affected offspring', we now consider 'selected offspring'.

(a) *Two-allele marker locus*. Following Spielman & Ewens (1996), consider a marker locus M, with two alleles M_1 and M_2, and obtain genotypes for affected individuals and their parents. For heterozygous parents, data to be analysed are numbers of 'transmissions', that is the number of times that the M_1 allele or the M_2 allele was transmitted to an affected offspring. Spielman *et al.* (1993) denote these counts as follows:

$$\left.\begin{aligned} b &= \text{number of times that } M_1M_2 \\ &\quad \text{transmits } M_1 \text{ to selected offspring,} \\ c &= \text{number of times that } M_1M_2 \\ &\quad \text{transmits } M_2 \text{ to selected offspring.} \end{aligned}\right\} \quad (1)$$

The counts may come from families that are simplex (i.e. data are from only one selected offspring), multiplex (data are from two or more selected sibs) or multigenerational; and the population may exhibit structure. The TDT statistic is

$$(b-c)^2/(b+c), \quad (2)$$

and it tests for equal numbers of transmissions of M_1 and M_2 from heterozygous parents to selected offspring. If marker locus M is linked to a QTL, b and c will tend to differ in value if there is linkage disequilibrium in the population or the analysis is done within a large family.

(b) *Multiple-allele marker locus*. We now give two possible ways to extend the TDT to marker loci with multiple alleles. As with the two-allele TDT we consider only parents heterozygous for marker alleles. The first and very straightforward approach is to consider one of the alleles at the marker locus as unique and group all other alleles into one composite allele. After this, a simple two-allele TDT can be performed for each allele at the marker locus. The second approach follows the same idea but is more general. Let k be the number of alleles at marker M. Then, a statistic can be used that compares for each heterozygous parent M_i, M_j ($1 \leqslant j \leqslant k$ and $j \neq i$) the number of times that M_i is transmitted to affected offspring with the number of times that M_j is transmitted to such offspring. Extension of (1) to a multiple allele situation can be

$$\left.\begin{aligned} n_{i\bullet} &= \text{number of times that } M_iM_j \\ &\quad \text{transmits } M_j \text{ to selected offspring,} \\ n_{\bullet i} &= \text{number of times that } M_iM_j \\ &\quad \text{transmits } M_j \text{ to selected offspring.} \end{aligned}\right\} \quad (3)$$

Note that transmitting M_j is identical to 'not-transmitting' M_i to selected offspring. And a

https://doi.org/10.1017/S0016672399004061 Published online by Cambridge University Press

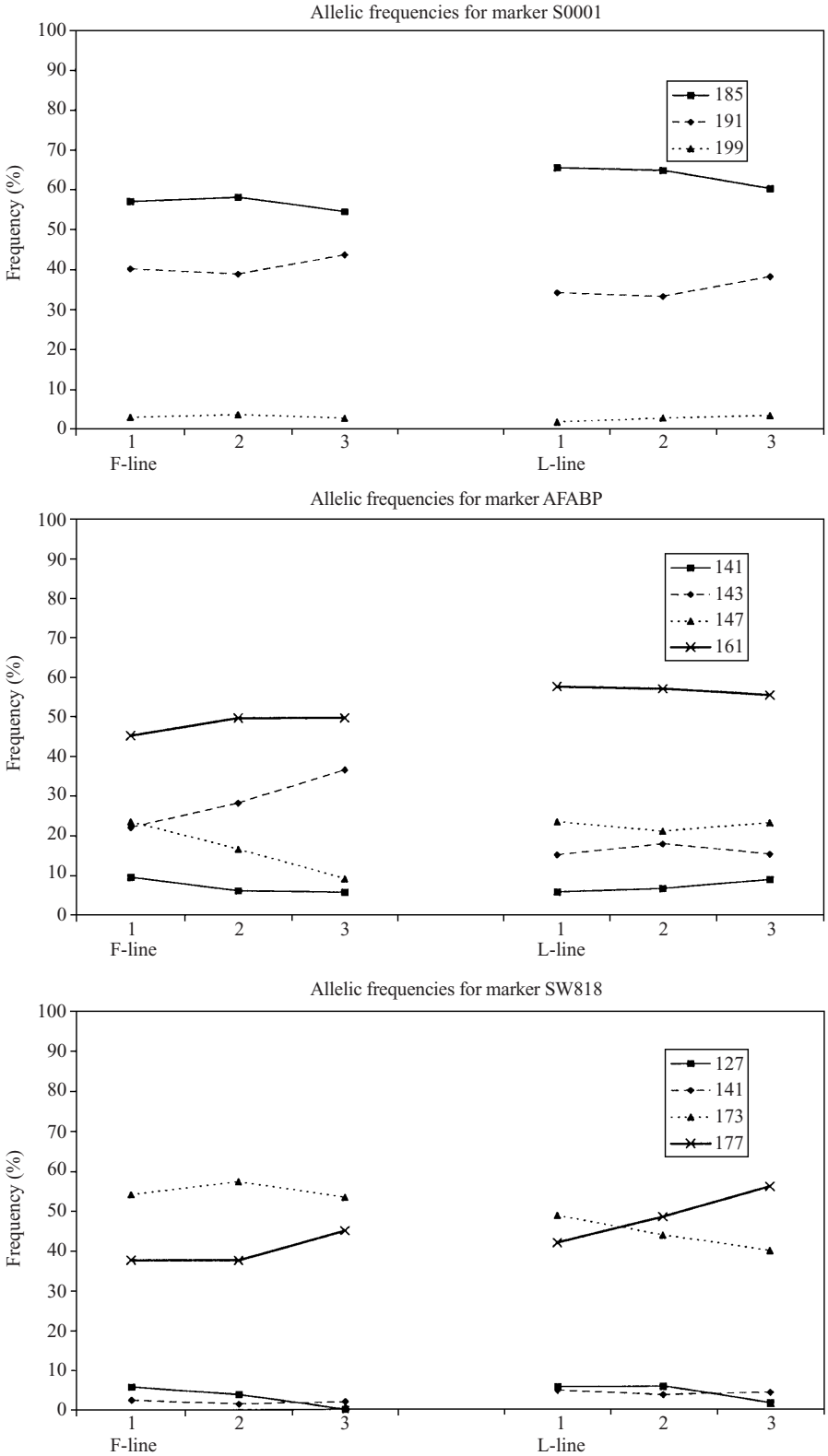

Fig. 2. Allelic frequencies at three marker loci on chromosome 4, over subsequent generations of two selection lines, i.e. selection on high growth rate (F) and selection on low backfat thickness (L).

https://doi.org/10.1017/S0016672399004061 Published online by Cambridge University Press

generalization of the two-allele TDT test statistic, as given by Spielman & Ewens (1996), is then

$$\frac{k-1}{k} \times \sum_{i=1}^{k} \frac{(n_{i\bullet} - n_{\bullet i})^2}{n_{i\bullet} + n_{\bullet i}}. \quad (4)$$

(v) *Distribution of the TDT statistic under the null hypothesis: Monte Carlo simulation*

For nuclear families with a single affected offspring, the test statistic (4) has very nearly a χ^2 distribution with $k-1$ degrees of freedom when the null hypothesis of no linkage is true (Kaplan *et al.*, 1997). However, in our selection experiment families are strongly related and multiple sibs per family (= litter) may be selected. Consequently, we can not use the tabulated threshold values to determine significance of effects.

The null hypothesis is no linkage between the marker and QTL. Here, rather than some approximate χ^2 test, we prefer to derive the distribution of our test statistic under the null hypothesis by Monte Carlo simulation. Let genotypic founders be all genotyped animals whose parents were both ungenotyped. For each Monte Carlo simulation the observed marker alleles of genotypic founders were used since allelic frequencies in the base population were unknown. The transmission of the observed marker alleles in genotypic founders to their offspring occurred at random and also transmission of alleles in subsequent generations occurred fully at random. After each simulation, the TDT statistics were calculated as described previously.

3. Results

(i) *Change in marker allele frequencies*

The two selection lines were selected for almost 3·5 generations. Pigs were grouped into four generation classes, i.e. 0, 1, 2 and 3, and comparison of allele frequencies was made for generations 1 to 3 (Fig. 2). For marker *S0001*, changes in allelic frequencies are small and show similar trends in both lines, a slight decrease in frequency of allele 185 and an increase in allele 199. For the *AFABP* marker, clear changes in allele frequencies were found, especially in the F-line where allele 147 decreased dramatically, and allele 143 increased in occurrence. For marker *SW818*, changes in allele frequencies occurred in both selection lines, where the changes in the L-line showed a clear trend, i.e. a steady increase of allele 177 and a decrease in allele 173. In the F-line, the frequency of allele 177 also showed an increase in the last generation. The small number of generations under selection and the rather continuous distribution seriously prevented testing the significance of changes in allele frequencies.

(ii) *TDT statistic under the null hypothesis*

For the multi-allele situation, two approaches to compute TDT statistics were suggested (see above). First, consider each allele at a marker locus as the allele of interest and merge the other alleles into one composite allele. Subsequently, all alleles can be analysed jointly. Table 3 shows the significance threshold values of the TDT statistics under the null hypothesis. These values were obtained via a Monte Carlo simulation of 65000 replicates (this number was arbitrarily taken to effectively reduce the Monte Carlo errors). For alleles with a moderate to high frequency of occurrence (see Fig. 2), the 0·1% significance threshold levels ranged from 10·29 up to 15·75. The cumulative threshold, i.e. alleles jointly, was clearly higher for all marker loci. The threshold values for allele 141 for marker *SW818* were all equal, due to its very low frequency in the population.

Table 3. *Threshold values of distribution of the transmission/disequilibrium test (TDT) statistic. Values are based on Monte Carlo gene-drop simulation with 65000 iterations*

	S0001				*AFABP*					*SW818*				
	185[a]	191	199	Cum[b]	141	143	147	161	Cum	127	141	173	177	Cum
F-line														
0·1%	13·30	13·59	10·67	19·23	11·76	11·05	13·83	13·56	20·41	11·57	3·00	11·17	12·26	16·62
1·0%	8·53	8·31	6·37	12·15	7·36	6·86	8·66	8·65	14·14	7·25	3·00	7·02	7·36	11·76
5·0%	4·92	4·86	3·85	7·44	4·26	3·92	5·17	5·12	9·63	4·31	3·00	4·17	4·31	7·98
10·0%	3·48	3·45	2·67	5·54	3·10	2·79	3·71	3·66	7·70	3·07	3·00	2·91	3·06	6·36
L-line														
0·1%	12·60	13·39	11·00	18·17	15·36	15·75	12·12	19·00	29·11	10·29	8·00	10·89	13·30	17·83
1·0%	8·10	8·40	7·20	11·68	10·53	10·59	7·22	13·07	20·44	6·37	6·00	6·82	8·33	12·14
5·0%	4·78	4·90	4·26	7·30	6·81	6·87	4·25	8·53	14·40	3·86	3·57	3·94	4·92	8·19
10·0%	3·33	3·46	3·00	5·53	5·12	5·23	3·00	6·55	11·66	2·67	2·67	2·81	3·53	6·52

[a] Pseudo two-allele TDT: this allele is unique, other alleles are grouped into one other allele.
[b] Multiple-allele TDT using equation (4).

https://doi.org/10.1017/S0016672399004061 Published online by Cambridge University Press

Table 4. *Transmission of alleles from heterozygous parents to selected offspring for three marker loci at chromosome 4, and values for the transmission/disequilibrium test (TDT), where P values < 0·15 are given in parentheses*

	S0001				*AFABP*					*SW818*				
	185[a]	191	199	Cum[b]	141	143	147	161	Cum	127	141	173	177	Cum
F-line														
Transmitted	110	98	15		11	61	28	101		20	4	101	113	
Not-transmitted	101	100	21		22	47	54	78		25	3	122	8	
TDT	0·38	0·02	1·00	0·93	3·67 (0·07)	1·81	8·24 (0·01)	2·96 (0·14)	12·51 (0·02)	0·56	0·14	1·98	5·28 (0·03)	5·97 (0·12)
L-line														
Transmitted	122	89	8		24	30	54	101		27	6	104	124	
Not-transmitted	97	112	10		23	38	71	77		22	4	104	128	
TDT	2·85 (0·13)	2·63 (0·15)	0·22	3·81	0·22	0·94	2·31 (0·15)	3·24	4·88	0·51	0·40	0·00	0·06	0·73

[a] Pseudo two-allele TDT: this allele is unique, other alleles are grouped into one other allele.
[b] Multiple-allele TDT using equation (4).

(iii) *TDT statistic*

Transmission of alleles from a heterozygous parent to its selected offspring was not always obvious, i.e. when the offspring was heterozygous for the same alleles. In these cases, transmission of alleles might be inferred by including information from the other parent. We used transmission only events where it was clear which allele was inherited. Table 4 shows the numbers of alleles transmitted by heterozygous parents and the resulting TDT statistics for all marker alleles. For marker *S0001* none of the TDT statistics was significant, all exceeding the 10% level. In the F-line, the test statistics for several alleles for marker *AFABP* were significant, especially for allele 147 (1% level). Apparently, selection did not favour this allele since the number of transmissions to selected offspring was much lower than its companion allele in the heterozygous parent. The cumulative test statistic for this locus was also highly significant and clearly suggests the presence of a closely linked, segregating QTL for growth rate in this selection line. In the F-line, the test statistic for allele 177 (marker *SW818*) was highly significant (3% level) and here selection favoured this allele. The cumulative test statistic was moderately significant, however, i.e. only at a 12% level. In the L-line, the test statistics were not significant in general, indicating absence of segregating QTL for backfat thickness near the markers studied.

4. Discussion

In this study data on pig selection lines were successfully analysed by using the transmission/disequilibrium test (TDT) (Spielman *et al.*, 1993). Instead of subsets of diseased offspring, we studied subsets of offspring selected due to superiority for a quantitative trait. The procedure of collecting blood samples that was followed in the experiment proved to be well suited for the TDT since DNA was available on parents and their selected offspring. Note that in a multi-generation experiment parents were the selected offspring in the previous generation of selection. The absence of DNA on the 'not-selected' (= culled) sibs does not affect the efficiency of the TDT but does prohibit approaches that are based on sib-pair analysis, e.g. ASP and traditional QTL mapping approaches. Also, the complex pedigree structure and limited marker data prevented the use of traditional QTL mapping approaches, e.g. regression or maximum likelihood analysis. Two other complications, next to population structure, in applying the original TDT were the multiplex families and the multi-allelic marker loci. These modifications prohibited the use of tabulated distribution threshold values for the TDT statistic under the null hypothesis. This distribution was now obtained by Monte Carlo simulation and the

https://doi.org/10.1017/S0016672399004061 Published online by Cambridge University Press

resulting threshold values clearly differed from the tabulated values of the χ^2 distribution. These differences might be due to population structure or to unequal marker allele frequencies. Results on changes in allele frequencies over generations were also presented and were consistent with the TDT results. The small number of generations and the continuous distribution of generation numbers hampered significance testing of changes in allele frequencies. Therefore, the TDT appears to be more flexible and useful to screen and test for QTL in data from selection programmes with overlapping generations.

Results showed significant test statistics in the F-line for markers *AFABP* and *SW818*. In the L-line, however, no significance for segregating QTL closely linked to one of the three markers was observed. This difference between lines may be explained by the success of selection in both lines. First, the selection intensity in the L-line was somewhat lower than in the F-line (Table 1). Secondly, the estimated genetic trend in the quantitative trait under selection was higher in the F-line than in the L-line, i.e. $0{\cdot}55\sigma_a$/generation and $0{\cdot}30\sigma_a$/generation, respectively (table 4 in Sonesson *et al*., 1998). The low test statistics in the L-line might also be the result of the relatively low information content of the markers studied, or due to absence of strong linkage disequilibrium at the population level.

The use of marker haplotypes, a set of alleles at linked marker loci, in the TDT seems a powerful extension to a TDT on single marker loci. However, this extension is hampered when the linkage phases of parental alleles at linked markers are unknown. Linkage phases may be inferred from marker genotypes on progeny or (grand) parents. Since the number of genotyped offspring per parent was very small, the grandparental genotypes, when observed, are most useful to decipher linkage phase in parents. However, the relatively low polymorphism of the three markers and their large distances did not allow an easy reconstruction of linkage phases.

The significant TDT results justify further efforts to disentangle segregating QTL on chromosome 4. Genotyping more markers with high polymorphism may enable multi-loci TDT and also multi-point linkage analysis. However, the use of a multi-point linkage analysis requires additional statistical efforts to accommodate the complex pedigree structure and the ungenotyped individuals in the pedigree. Phenotypes on both selected and culled animals should be included to avoid bias and improve accuracy in estimation of genetic parameters (e.g. Johnson *et al*., 1999; Bink *et al*., 2000). Joint analysis of the two selection lines seems worthwhile due to the common base population and due to reversed genetic trends in quantitative traits observed in both lines (Sonesson *et al*., 1998). Selection was on different traits and a multi-trait analysis is therefore desired. A multi-trait analysis would also allow detection of QTL for other traits measured in the experiment, such as fertility and meat quality traits (Sonesson *et al*., 1998).

The authors wish to thank Frans Verburg, Frans Gerbens and Karel de Greef for data preparation and for valuable discussions.

References

Andersson, L., Haley, C. S., Ellegren, H., Knott, S. A. & Johansson, M., *et al*. (1994). Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science* **263**, 1771–1774.

Bink, M. C. A. M., Janss, L. L. G. & Quaas, R. L. (2000). Markov chain Monte Carlo for mapping a quantitative trait locus in outbred populations. *Genetical Research* (in press).

Darvasi, A. & Soller, M. (1995). Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics* **141**, 1199–1207.

Falk, C. T. & Rubinstein, P. (1987). Haplotype relative risks: an easy reliable way to construct a control sample for risk calculations. *Annals of Human Genetics* **51**, 227–233.

Georges, M., Nielsen, D., Mackinnon, M., Mishra, A., Okimoto, R., *et al*. (1995). Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics* **139**, 907–920.

Gerbens, F., Jansen, A., Van Erp, A. J. M., Harders, F., Meuwissen, T. H. E., *et al*. (1998). The adipocyte fatty acid-binding protein locus: characterization and association with intramuscular fat content in pigs. *Mammalian Genome* **12**, 1022–1026.

Haley, C. S., Knott, S. A. & Elsen, J.-M. (1994). Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**, 1195–1207.

Johnson, D. L., Jansen, R. C. & Van Arendonk, J. A. M. (1999). Mapping quantitative trait loci in a selectively genotyped outbred population using a mixture model approach. *Genetical Research* **73**, 75–84.

Kaplan, N. L., Martin, E. R. & Weir, B. S. (1997). Power studies for the Transmission/Disequilibrium Tests with multiple alleles. *American Journal of Human Genetics* **60**, 691–702.

Keightley, P. D., Hardge, T., May, L. & Bulfield, G. (1996). A genetic map of quantitative trait loci for body weight in mice. *Genetics* **142**, 227–237.

Ollivier, L., Messer, L. A., Rothschild, M. F. & Legault, C. (1997). The use of selection experiments for detecting quantitative trait loci. *Genetical Research* **69**, 227–232.

Sonesson, A. K., De Greef, K. H. & Meuwissen, T. H. E. (1998). Genetic parameters and trends of meat quality, carcass composition and performance traits in two selected lines of large white pigs. *Livestock Production Science* **57**, 23–32.

Spielman, R. S. & Ewens, W. J. (1996). The TDT and other family-based tests for linkage disequilibrium and association. *American Journal of Human Genetics* **59**, 983–989.

Spielman, R. S., McGinnis, R. E. & Ewens, W. J. (1993). Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *American Journal of Human Genetics* **52**, 506–516.

Thomson, G. (1988). Mapping disease genes: family-based association studies. *American Journal of Human Genetics* **57**, 487–498.