Candidate Gene Analysis for Quantitative Traits Using the Transmission Disequilibrium Test: The Example of the Melanocortin 4-Receptor in Pigs

Jules Hernández-Sánchez,*[,1] Peter Visscher,[†] Graham Plastow[‡] and Chris Haley*

**Roslin Institute, Roslin, Midlothian EH25 9PS, Scotland, United Kingdom, [†]Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JG, Scotland, United Kingdom and [‡]Sygen International, Berkeley, California 94710*

Manuscript received August 15, 2002
Accepted for publication January 23, 2003

ABSTRACT

Population-wide associations between loci due to linkage disequilibrium can be used to map quantitative trait loci (QTL) with high resolution. However, spurious associations between markers and QTL can also arise as a consequence of population stratification. Statistical methods that cannot differentiate between loci associations due to linkage disequilibria from those caused in other ways can render false-positive results. The transmission-disequilibrium test (TDT) is a robust test for detecting QTL. The TDT exploits within-family associations that are not affected by population stratification. However, some TDTs are formulated in a rigid form, with reduced potential applications. In this study we generalize TDT using mixed linear models to allow greater statistical flexibility. Allelic effects are estimated with two independent parameters: one exploiting the robust within-family information and the other the potentially biased between-family information. A significant difference between these two parameters can be used as evidence for spurious association. This methodology was then used to test the effects of the fourth melanocortin receptor (MC4R) on production traits in the pig. The new analyses supported the previously reported results; *i.e.*, the studied polymorphism is either causal or in very strong linkage disequilibrium with the causal mutation, and provided no evidence for spurious association.

POPULATION-wide associations between loci due to linkage disequilibrium can be used in high-resolution mapping of quantitative trait loci (QTL). However, spurious associations between markers and QTL can also arise as a consequence of population stratification, for example due to admixture of two different populations. Associations between a genotype and a trait have been frequently tested, after corrections, with simple one-way ANOVA models, *e.g.*, testing mean genotype differences directly. However, these types of analyses are prone to false-positive results due to confounding effects of population stratification/admixture (*e.g.*, Deng *et al.* 2001). Spielman *et al.* (1993) developed an allele-trait association test called the transmission disequilibrium test (TDT), which is robust to these confounding effects. Different TDTs have since been developed for dichotomous traits (Schaid 1996; Horvath and Laird 1998; Lunetta *et al.* 2000; Martin *et al.* 2000; Zhao *et al.* 2000) and for quantitative traits (Allison 1997; Rabinowitz 1997; Szyda *et al.* 1998). However, most of these TDTs have been formulated in a rather rigid form, hence reducing the scope for further statistical modeling. In addition, previously described TDTs make use of within-family information only.

This study presents a generalized version of TDT to obtain robust estimates of genetic effects within the statistically more flexible mixed-linear model context. This approach allows maximum-likelihood estimates of genetic effects to be obtained via residual maximum likelihood (REML). Additive allele substitution effects were estimated with two independent regression coefficients (b), the within-families coefficient b_{TD} (transmission disequilibrium) and the between-families coefficient b_{PD} (population disequilibrium). Moreover, the rejection of the null hypothesis $b_{TD} = b_{PD}$ provides evidence for stratification/admixture and hence can be used to guard against false-positive results.

The analysis of the fourth melanocortin receptor (MC4R) locus on pig chromosome 1 constitutes a practical demonstration of this method. The interaction between melanocortins and their receptors (MC3R and MC4R) at the hypothalamus is one of the main neuroendocrinological pathways controlling energy balance (Wardlaw 2001). In humans, different allelic variants of both MC4R and MC3R have been associated with obesity (Vaisse *et al.* 1998; Yeo *et al.* 1998; Hinney *et al.* 1999; Li *et al.* 2000). In pigs, the seventh transmembrane region of the MC4R locus contains a mutation at codon 298 that causes a change of aspartic acid for asparagine,

[1]*Corresponding author:* Roslin Institute, Roslin, Midlothian, EH25 9PS, Scotland, United Kingdom.
E-mail: jules.hernandez@bbsrc.ac.uk

i.e., Asp298Asn (Kim *et al.* 1999). This region is highly conserved across all four types of melanocortin receptors in humans (Gantz *et al.* 1993), and it is also very conserved between pigs and humans (Kim *et al.* 2000). The Asp298Asn mutation has been associated with fatter and faster-growing pigs, having significant effects on back fat, days to 110 kg, test daily gain, and daily food intake in a study involving four different commercial pig lines from nucleus breeding farms (Kim *et al.* 1999, 2000). However, the original analyses were performed with methods that are potentially biased in the presence of population stratification. Here we analyze an augmented data set using the new methodology to confirm and extend the original findings.

METHODS

Data: Performance traits were recorded on four different commercial PIC pig lines in the same farm over a 5-year period (1993–1998). The traits were lifetime daily gain (LDG), test daily gain (TDG), daily food intake (DFI), and back fat depth (BF) at the 10th rib. All pigs were performance tested for growth over a fixed period of 12 weeks, during which they were fed *ad lib.* and weighed at the beginning (on-test) and at the end (off-test) of that period. TDG was calculated as off-test weight minus on-test weight divided by the number of days on test. LDG was calculated as off-test weight minus one (assumed the average birth weight) divided by the age of the pig (in days) at off-test. BF was measured ultrasonically in real time at off-test, and it was normalized with the natural log transformation. DFI was electronically recorded for some pigs over the testing period. The sample sizes by line and sex are given in Table 1. In this data set, there were 726 extra records of BF, 574 of TDG, and 44 of DFI, with respect to the data set analyzed by Kim *et al.* (2000). The Asp298Asn substitution mutation is located within a *Taq*I restriction enzyme recognition site (Kim *et al.* 1999), which was used to generate a codominant restriction fragment length polymorphism (RFLP) to distinguish all three genotypic classes (Kim *et al.* 2000).

Statistical models: The effects of the Asp298Asn mutation on pig production traits were estimated with the models used in Kim *et al.* (2000), which are not robust to population stratification/admixture, and with new robust models. The latter models were a combination of the former models and a TDT (Rabinowitz 1997). The polygenic variance was estimated, fitting sire as a random factor. This model is computationally faster than fitting each animal as a random factor, and the two models were very similar in terms of variance components estimation (results not shown).

ANOVA: The original models included sex, batch, line, and genotype as fixed factors and sire as random effect. BF records were analyzed both in the original scale and in the log-transformed scale. The skewness and kurtosis of the untransformed distribution of BF were 0.88 (±0.049) and 1.96 (±0.099), respectively. After transformation the distribution of BF became more normal (skewness = −0.03 ± 0.049, kurtosis = 0.32 ± 0.099). All two-way interactions between fixed factors were also included in the analyses. Nonsignificant factors were dropped out of the models using a backward elimination procedure. The coefficients for genotypes can be found in column *A* in Table 2. The analyses were performed with the REML procedure in GENSTAT (Payne *et al.* 2001). This method of estimating allele effects is not robust to stratification/admixture (Hernández-Sánchez *et al.* 2002b). We refer to this as the ANOVA method.

Batches as random: There were 54 batch means to estimate when they were fitted as fixed effects in the models. To avoid this unnecessary loss of degrees of freedom, batches were fitted as a random term where direct comparisons were being made with the TDT, and their effect was accounted for with cubic splines. This procedure was feasible because all trait means followed a yearly cycle when plotted against batches. The correction uses up only 2 d.f.: one in fitting a linear regression across all batches and a second in estimating the residual variance around the previous line. The software used to run models with batch as random splines was ASREML (Gilmour *et al.* 2001). This approach was also implemented in the TDT analyses (see below).

TDT: A robust analysis of the Asp298Asn mutation was performed with the same models but substituting genotype for two independent fixed covariates. One of

TABLE 1

Number of pigs (males/females) used in the analyses within each trait, line, and MC4R genotype

MC4R	Line A	Line B	Line C	Line D
		TDG		
11	3/22	27/28	89/349	155/25
12	9/146	38/79	11/177	152/44
22	9/245	12/57	0/32	37/22
		LDG and BF		
11	3/37	27/52	89/504	155/25
12	9/266	38/145	11/250	152/44
22	9/392	12/117	0/50	37/22
		DFI		
11	3/0	27/0	87/0	21/0
12	9/0	38/0	10/0	44/0
22	9/0	12/0	0/0	15/0

TDG, test daily gain; LDG, lifetime daily gain; BF, back fat at 10th rib; DFI, daily food intake. Line A is a Landrace-based population. Line B is a Large White-based population. Line C is a synthetic population based on Duroc and Large White. Line D is a synthetic line based on several different populations including Landrace, Large White, Duroc, and Pietrain. Aspartic acid is coded as 1 and Asparagine as 2.

TABLE 2

Parameterization of covariates *A*, TD, and PD given family genotypes

G_f	G_m	G_o	A	TD	PD
11	11	11	1	0	1
11	12	11	1	½	½
		12	0	−½	½
11	22	12	0	0	0
12	12	11	1	1	0
		12	0	0	0
		22	−1	−1	0
12	22	12	0	½	−½
		22	−1	−½	−½
22	22	22	−1	0	−1

G_f, paternal genotype; G_m, maternal genotype; G_o, offspring genotype. One-way ANOVA: $A = 1, 0, -1$ if $G_o = 11, 12, 22$, respectively. TDT: $TD = H_♂(T_♂ - ½) + H_♀(T_♀ - ½)$, where $H_{♂(♀)} = 1$ if $G_{f(m)} = 12$ and 0 otherwise. $T_{♂(♀)} = 1$ if offspring receives allele 1 from a 12 father (mother) and 0 otherwise. $PD = A - TD$.

these two covariates was based on a TDT (Rabinowitz 1997). Given a biallelic marker, each individual's genotype received a coefficient equal to $H_♂(T_♂ - ½) + H_♀(T_♀ - ½)$; where $H_♂ = 1$ if the individual's sire was heterozygous, or 0 otherwise; $T_♂ = 1$ if the sire had transmitted allele 1 to the individual, or 0 otherwise; and likewise, $H_♀$ and $T_♀$ for the individual's dam. These coefficients can be found in column TD in Table 2. The slope of this covariate, b_{TD}, is a robust estimate of additive substitution effects of alleles at the locus. Allelic effects were also estimated via a second regression coefficient sensitive to the effects of population structure, b_{PD} (L. L. G. Janss, personal communication). The appropriate coefficients to estimate b_{PD} were obtained by subtracting column *A* from column TD in Table 2. We refer to this as the TDT method.

Generating parental genotype data: The TDT method requires parental genotype data, which in the MC4R data set were mostly missing. Missing parental genotypes were generated using Gibbs sampling (*e.g.*, Wang *et al.* 1994; Sorensen and Gianola 2002). Gibbs sampling was equivalent to integrating over all genotype probabilities of parents with missing genotypes. Missing parental genotypes were sampled conditional on genotypes of progeny and other relatives. The sampling algorithm has been reported previously in the literature (Guo and Thompson 1992; Janss *et al.* 1995). Convergence was reached after 10^3 realizations after a burn-in period of 100 realizations. The autocorrelation in Gibbs sampling was minimized by sampling 1 realization every 50 consecutive ones; hence a total of 50×10^4 realizations were generated. The actual integration over missing genotype probabilities was carried out by analyzing the MC4R data set after each sampled realization and averaging results (*i.e.*, *P* values) across all realizations.

Simulation: A simulation study was carried out to investigate the properties of b_{PD} and b_{TD}. Population stratification was generated by sampling from two separated populations with different allele frequencies and analyzing the data jointly. Each population was characterized by: (1) 15 unrelated full-sib families and four offspring per family (60 progeny in total); (2) random mating; (3) a QTL and a linked neutral marker, both biallelic and with allele frequency fixed to 0.9 in one population, and frequencies, at both loci, of 0.9, 0.7, 0.5, 0.3, and 0.1 in the other population; (4) recombination rate between loci of $c = 0$ or ½; (5) standardized linkage disequilibrium in parents ($D' = D/D_{max}$; Lewontin 1988) of either 0 or 1; (6) residual variance of 1 and polygenic variance 0; and (7) QTL explaining 5% (and in some cases 10%) of the total phenotypic variance. There were no interpopulation matings and all analyses were performed at the marker locus.

RESULTS

Properties of b_{PD} and b_{TD} in analyses of simulated data: The power of estimating genetic effects through b_{PD} and b_{TD} in the simulated data is shown in Figure 1, where *F*-ratios are plotted against the level of stratification *S* (allele frequency difference between populations) across four different scenarios. Each dot in Figure 1 is the average of 100 replicates. The significance threshold is based on the tabulated nominal 5% threshold and is shown as a straight line. Figure 1A shows the results after analyzing a marker totally unlinked ($c = ½$) and with no association in the population ($D' = 0$) to a QTL. In this situation, any significant effect is a type I error or due to bias. The F_{TD} (*i.e.*, the *F*-ratio testing whether a significant amount of the total variation is explained by b_{TD}) is approximately 1 across all *S* values, which indicates that the marker did not have a significant effect on the trait. On the contrary, F_{PD} (*i.e.*, the corresponding *F*-ratio test for b_{PD}) appeared positively correlated to the level of *S*. The effect was significant when $S \geq 0.6$. In this case, spurious disequilibrium increases as stratification increases, and b_{PD} cannot distinguish between this sort of disequilibrium and disequilibrium due to linkage.

Figure 1B shows the results after analyzing a marker totally unlinked ($c = 0.5$) but in complete disequlibrium ($D' = 1$) in the parents. Here, F_{PD} is always $> F_{TD}$, and this difference increases with *S*. Moreover, the value of F_{PD} was >1 (~ 2) even without stratification ($S = 0$). This can be explained by considering that, on average, the level of disequilibrium among offspring was one-half, because D' is expected to be halved every generation assuming no linkage and random mating. This feature suggests that b_{PD} could detect an effect given sufficient linkage disequilibrium between a marker and a QTL, even if these two loci

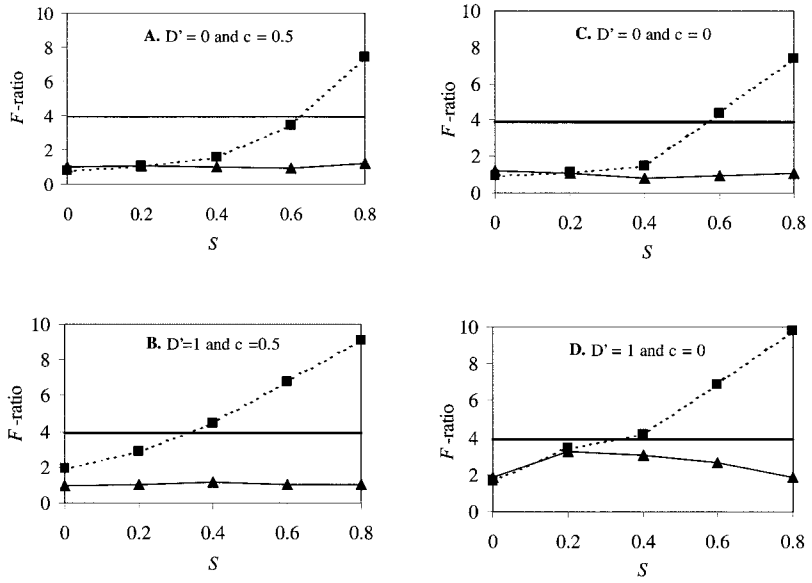

FIGURE 1.—*F*-ratios when testing b_{PD} (squares) and b_{TD} (triangles) given parental linkage disequilibrium (D'), recombination rate (c), and level of stratification (S). Threshold is shown as a straight line.

are totally unlinked, whereas b_{TD} needs the joint occurrence of linkage and linkage disquilibrium to estimate an effect.

Figure 1C shows how F_{PD} rapidly increases for $S > 0.4$, while F_{TD} remains constant and equal to one regardless of S. The marker was totally linked to the QTL locus in this set of simulations ($c = 0$). Despite having simulated no linkage disequilibrium ($D' = 0$) within each population, mixing two populations with different allele frequencies will cause haplotype frequencies to depart from the expected equilibrium frequencies. In this simple scenario, the fact that $F_{TD} = 1$ across all values of S even when $c = 0$ demonstrates the robustness of the test based on b_{TD}.

Finally, Figure 1D shows the effect of effectively analyzing the QTL itself ($D' = 1$ and $c = 0$). The power of estimating b_{PD} increases monotonically with S, a fact consistently observed in all previous graphs. However, the power of estimating b_{TD} reaches its maximum at intermediate levels of S. This is so because intermediate allele frequencies [*e.g.*, at 0.5, represented by $S = 0.9$ (population 1) − 0.5 (population 2) = 0.4 in Figure 1D] are associated with a higher proportion of heterozygous parents, providing better information to estimate b_{TD}.

Analyses of MC4R data: *ANOVA method:* Results were obtained from analyses of the data from the four lines separately and from an overall analysis of the combined data from all lines. Striking differences between genotypes were detected for BF ($P < 0.001$), TDG ($P < 0.001$), and LDG ($P < 0.001$) in the overall (*i.e.*, all lines together) analysis (Table 3). Batches were treated as fixed effects in these analyses to allow us to compare results directly with the findings of KIM *et al.* (2000). Genotypic differences within all pig lines were confirmed for BF (ranging from $P < 0.001$ in lines A and C to $P < 0.02$ in line B) and also within some lines for TDG (ranging from $P < 0.001$ in line C to $P > 0.8$ in line B) and LDG (ranging from $P < 0.01$ in line C to $P > 0.5$ in line D). No significant effect of the Asp298Asn substitution mutation on DFI was found in this analysis, in spite of the significant difference of 0.17 kg between both homozygotes ($P < 0.01$) reported by KIM *et al.* (2000). In this study, the estimated difference in DFI between the two homozygous genotypes, although not significant, was in the same direction as that previously reported at 0.1 kg [standard error of the difference (SED) = 0.063] in the overall analysis.

TDT method: Only results from the overall analyses of the combined data set are shown. The estimation of b_{TD} and b_{PD} via REML was done on the basis of 10^3 replications, where each replicate used a different population of parental genotypes generated with Gibbs sampling. The average P values across these analyses are shown in Table 4. The ASREML software does not perform hypothesis testing for fixed effects in the model although it provides t-values and residual degrees of freedom. The P values associated with genotypic contrasts were obtained from the t-distribution. Two independent null hypotheses were of interest: H_0^1, $b_{PD} = 0$, and H_0^2, $b_{TD} = 0$; both were rejected for all traits except DFI [H_0^1, (BF) $P < 0.0001$, (TDG) $P < 0.01$, (LDG) $P < 0.01$, and (DFI) $P > 0.5$; H_0^2, (BF) $P < 0.0001$, (TDG) $P < 0.001$, (LDG) $P < 0.01$, and (DFI) $P > 0.2$]. For the sake of comparison, Table 4 also incorporates results from the ANOVA method (reported as regression coefficients rather than genotype means as in Table 3). Both the ANOVA and the TDT methods yielded similar results, although the former was a slightly more powerful analysis (the standard errors were generally smaller). Results in Tables 3 and 4 are of similar magnitude.

DISCUSSION

The study demonstrated that the mean additive effect of allele substitution calculated with a one-way ANOVA model can be decomposed into the within- and between-family effects. These two effects can be estimated via a flexible REML analysis (PATTERSON and THOMPSON

TABLE 3

Means by genotype class within and across lines (overall)

Trait	*G*	Line A	Line B	Line C	Line D	Overall
TDG (g/day)	11	960.8	852.1	942.3	895.0	910.5
	12	913.3	859.3	913.6	873.3	888.9
	22	898.5	851.6	889.3	877.8	878.6
	SED	13.88	14.6	14.14	14.61	6.41
	P	0.002	0.819	0.001	0.165	<0.001
LDG (g/day)	11	626.1	660	692.5	696.4	698.1
	12	613.4	656.9	681	677.6	685.4
	22	606.2	651.3	670	676.7	679.5
	SED	7.4	7.3	7.3	9.2	3.4
	P	0.043	0.522	0.005	0.022	<0.001
DFI (kg/day)	11	1.88	1.94	1.89	1.79[a]	1.89
	12	1.81	1.84	1.83	1.8[a]	1.82
	22	2	1.69	NA	1.74[a]	1.79
	SED	0.298	0.108	0.126	0.105[a]	0.063
	P	0.705	0.1	0.655	0.861[a]	0.202
BF (log mm)	11	2.55	2.55	2.55	2.44	2.49
	12	2.46	2.50	2.48	2.36	2.42
	22	2.41	2.46	2.43	2.30	2.37
	SED	0.03	0.028	0.027	0.039	0.013
	P	<0.001	0.017	<0.001	0.002	<0.001
BF (mm)[b]	11	12.8 (12.3–13.2)	12.8 (12.4–13.1)	12.8 (12.4–13.2)	11.5 (11.1–11.9)	12 (11.8–12.2)
	12	11.8 (11.3–12.2)	11.9 (11.6–12.2)	12.2 (11.8–12.5)	10.6 (10.2–11)	11.2 (11–11.4)
	22	11.2 (10.8–11.6)	11.3 (11.1–11.6)	11.7 (11.4–12.1)	9.9 (9.6–10.3)	10.7 (10.5–10.8)

The models were as in Kim *et al.* (2000), although containing only significant terms. *G*, genotypes; 1, Asp and 2, Asn at codon 298 of the porcine MC4R; SED, standard error of the difference between any pair of genotypic means; *P*, *P* values; NA, data not available.

[a] Regarding DFI, only these analyses included sire as a random effect.

[b] The confidence intervals around BF means on the back-transformed scale are asymmetric; thus they are shown within parentheses.

1971) as the regression coefficients b_{TD} and b_{PD}, respectively, using a mixed-linear model that can also incorporate other fixed and random effects.

Parental genotypes are needed to estimate b_{TD} and b_{PD}. This information was not available for the MC4R data analyzed and was generated via Gibbs sampling (1000 realizations). The new methodology was tested via simulation and real data analysis of the effect of the MC4R gene on pig production traits.

The simulation results can be summarized in three main points. First, b_{TD} extracts information from the within- and b_{PD} from the between-family genetic variances, σ^2_{WF} and σ^2_{BF}, respectively. Second, b_{TD} is robust and b_{PD} is biased in the presence of population admixture/stratification. Third, there is generally more power to detect a significant $b_{PD} \neq 0$ than a significant $b_{TD} \neq 0$.

Population stratification/admixture increases σ^2_{BF} and not σ^2_{WF}. As a consequence, estimates of b_{PD} may be biased in the presence of stratification/admixture, whereas estimates of b_{TD} are not as the simulation results demonstrate. Tests such as the TDT are robust because they exploit only σ^2_{WF}. However, the one-way ANOVA model uses both σ^2_{WF} and σ^2_{BF}, from which a pooled estimate between b_{PD} and b_{TD} is obtained, and because of the latter component of variance, this pooled estimate of allelic effects may be biased if there is population admixture/stratification. In spite of this potential bias producing false-positive results (*e.g.*, Figure 1, A and B), robust methods such as the TDT have seldom been used in animal breeding (some exceptions are Bink *et al.* 2000 and Hernández-Sánchez *et al.* 2002a). This may be because TDT is viewed as a not very powerful method (*e.g.*, Figure 1D, with severe limitations in some circumstances, *e.g.*, low frequency of heterozygous parents).

The fact that b_{TD} and b_{PD} exploit different sources of information can be intuitively appreciated by inspecting the coefficients in Table 2. First, the coefficients required in the estimation of b_{TD} are weights given to all individuals with records (*e.g.*, offspring) according to their genotypes and to the genotypes of their parents (*i.e.*, family type). Hence, b_{TD} is the slope of the regression of phenotypes onto explanatory variables that combine both offspring's genotypes and family type. Second, the coefficients required in the estimation of b_{PD} can

TABLE 4

Average effect of allele substitution for each trait using TDT (PD, TD) and ANOVA

Trait	Estimation	b	SE (b)	T	d.f.	P value
TDG	PD	16.1	5.84	2.76	2430	<0.01
	TD	14.1	4.45	3.17		<0.01
	ANOVA	14.8	3.66	4.05		<0.0001
LDG	PD	9.1	3.31	2.75	2423	<0.01
	TD	8.8	2.38	3.69		<0.001
	ANOVA	8.9	2.01	4.42		<0.0001
BF	PD	0.07	0.013	5.14	2423	<0.0001
	TD	0.06	0.009	6.7		<0.0001
	ANOVA	0.06	0.009	8.11		<0.0001
DFI	PD	0.03	0.047	0.64	272	>0.5
	TD	0.08	0.061	1.26		>0.2
	ANOVA	0.05	0.037	1.28		>0.2

Cycle was treated as random. PD, population disequilibrium (between-families gene effect); TD, transmission disequilibrium (within-families gene effect); ANOVA, genotype-substituted covariates PD and TD; b, gene effect; SE (b), standard error of gene effect; T, t-statistic from testing $b = 0$ *vs.* b $\neq$ 0; d.f., nominal residual degrees of freedom. P value: $N(0, 1)$ was used as an approximation to the t-distribution for TDG, LDG, and BF.

alternatively be obtained as $\Sigma_{j=1}^{2} G_{ij}$, where $G_{ij} = \frac{1}{2}$, 0, or $-\frac{1}{2}$ if the genotype of the jth parent in the ith family is 11, 12, or 22, respectively. Therefore b_{PD} is the slope of the regression of phenotypes onto family type.

More explicitly, let us model y_{ij}, the phenotype of the jth individual having the ith QTL genotype, as $y_{ij} = \mu + g_i + a_{ij} + e_{ij}$, where μ is the population mean, g_i is the effect of the ith QTL genotype on the jth offspring, and a_{ij} and e_{ij} are the polygenic and residual random terms drawn from two independent normal distributions with zero means and variances σ_A^2 and σ_e^2, respectively. Let there be random mating, no population stratification, and only additive genetic effects at the QTL locus. Under these circumstances, the total additive genetic variance splits equally between and within families; therefore $E[\sigma_{BF}^2] = \sigma_G^2/2$ and $E[\sigma_{WF}^2] = \sigma_e^2 + \sigma_G^2/2$, where $\sigma_G^2 = \sigma_A^2 + \sigma_Q^2$, and σ_Q^2 is the variance due to the QTL. If a statistical model explains all σ_Q^2, then the additive genetic effects of the QTL will be fully accounted for. However, if the model estimates additive effects only through either b_{TD} or b_{PD}, then σ_Q^2 will be only partially explained. For example, if b_{TD} is the only estimator of QTL effect, then $E[\sigma_{WF}^2] = \sigma_e^2 + \sigma_A^2/2$ and $E[\sigma_{BF}^2] = \sigma_G^2/2$, as only within-family variation can be used to estimate b_{TD}. If, on the other hand, b_{PD} is the only estimator of QTL effect, then $E[\sigma_{WF}^2] = \sigma_e^2 + \sigma_G^2/2$ and $E[\sigma_{BF}^2] = \sigma_A^2/2$, as only between-family variation can be used to estimate b_{PD}. These changes in the within- and between-family variances due to estimation of b_{TD} or b_{PD} were validated via computer simulations (data not shown).

The situation is more complex when there is stratification within a population. The APPENDIX shows how the proportion of the total phenotypic variation explained, fitting a linear regression, differed when either b_{TD} or b_{PD} was estimated in the presence of population stratification.

Simulation results showed that testing the null hypothesis $b_{PD} = 0$ rather than $b_{TD} = 0$ tends to produce higher F-ratios when it is not true (Figure 1). Nevertheless, the analysis of MC4R data showed the opposite effect (*i.e.*, t-tests for b_{TD} were always higher than those for b_{PD} in Table 4). This result is possible when there is no admixture/stratification, because b_{TD} is expected to be equivalent to b_{PD}. Moreover, in simulations, the only between-family component was one-half of the total QTL variance, whereas in reality other factors (*e.g.*, litter and sow effects) may also increase σ_{BF}^2. The analysis of real data showed that b_{TD} was very similar to b_{PD} across all traits and, furthermore, that both estimates were also similar to allelic effects obtained with ANOVA (Table 4). This suggests that there was no significant stratification/admixture in the population and that there were equivalent amounts of genetic information between and within families.

Where estimates of b_{TD} and b_{PD} differ, a t-test could be used to test the null hypothesis $b_{TD} = b_{PD}$. If no evidence to reject the null hypothesis were found, *i.e.*, no evidence for admixture/stratification in the population, then a more powerful one-way ANOVA could be safely implemented within the mixed-linear model to estimate allele effects. Otherwise, robust approaches such as the one proposed here, *i.e.*, estimating b_{TD} or other TDTs, should be considered as the only reliable methods of analysis.

A complication of this approach applied to many real data sets, including the MC4R data used as an example, is generating missing parental genotypes. A Gibbs sampling technique to generate 10^3 replicates was readily implemented (and we know they were accurate enough because 10 times more replicates did not affect the outcome) in this simple scenario: *i.e.*, a single biallelic marker, a maximum of three generations, and few missing data. Other robust tests that do not require parental genotypes, *e.g.*, sib-TDTs (*e.g.*, SCHAID and ROWLAND 1999; SPIELMAN and EWENS 1998), were found to be less powerful than the method outlined here in additional analyses of simulated data [*e.g.*, $F_{sib} = 5.5$ *vs.* $F_{TD\text{-}PD} = 7.7$ (based on 1000 replicates)]. Analyzing more complex or unbalanced data sets will probably demand more realizations of the Gibbs sampler and it will present an additional problem: testing fixed effects in REML, because the asymptotic properties of the t-test or the Wald test could not be guaranteed (see KENWARD and ROGER 1997; WELHAM and THOMPSON 1997; ELSTON 1998). If few fixed effects are fitted in the model, then maximum likelihood rather than REML can be used to test allele effects via likelihood-ratio tests.

There is strong evidence from both this study and KIM *et al.* (2000) that the substitution-mutation Asp298Asn in

the MC4R gene affects production traits in the pig, *e.g.*, BF and growth (TDG, LDG). These effects were reestimated with extra records with respect to the original work. Significant effects were found for BF, TDG, and LDG across all lines ($P < 0.001$), although not for all pig lines (ranging from $P < 0.001$ to $P > 0.8$).

DFI was the only trait not significant in this study ($P > 0.2$) that was significant in the study of Kim *et al.* (2000), even after using their statistical models. The lack of effect on DFI was unexpected given previously reported results ($P < 0.01$) and other reported associations of MCR with appetite and feeding behavior in macaques (Koegler *et al.* 2001), mice (Butler *et al.* 2000), and layer chicks (Tachibana *et al.* 2001). One possible explanation for not detecting an effect on DFI is the lack of power due to a small data set ($n = 275$). Although there were less data ($n = 231$) in the study of Kim *et al.* (2000) than in this study, the difference is small and additional statistical noise may have been introduced through factors such as data structure and/or chance. An independent study that used a larger data set ($n = 619$) found a significant ($P < 0.05$) additive gene effect of 0.075 kg/day on DFI (G. Plastow, personal communication).

Further experiments are needed to ascertain whether the Asp298Asn substitution mutation at the MC4R locus is causative. For example, one could test the effect of this and other polymorphisms on the same exon/gene in different segregating populations. Final proof will require molecular experiments, *e.g.*, testing whether different molecular pathways are activated in independent cell cultures having different Asp298Asn variants and/or creating mouse models for testing this polymorphism on identical genetic backgrounds.

This study demonstrates that TDT can be implemented within the REML framework. As a guideline, to maximize the power in an association analysis, one should test associations with ANOVA (within a REML or maximum-likelihood framework) only after having checked that b_{TD} and b_{PD} are not significantly different from each other, *i.e.*, making sure no false-positive results are being caused by population stratification.

We thank Dr. R. Pong-Wong for his advice on Gibbs sampling; Dr. O. Southwood, D. Waddington, and anonymous referees for helpful comments that improved earlier versions of this manuscript; and especially Dr. L. L. G. Janss for sharing his ideas with us. We acknowledge Kwan Suk Kim and Max Rothschild (Iowa State University) for providing the MC4R data set and the Biotechnology and Biological Sciences Research Council and PIC International Group for funding this project.

LITERATURE CITED

Allison, D. B., 1997 Transmission-disequilibrium tests for quantitative traits. Am. J. Hum. Genet. **60:** 676–690.

Bink, M. C., M. F. Te Pas, F. L. Hardens and L. L. Janss, 2000 A transmission/disequilibrium test approach to screen for quantitative trait loci in two selected lines of large white pigs. Genet. Res. **75:** 115–121.

Butler, A. A., R. A. Kesterson, K. Khong, M. J. Cullen, M. A. Pelleymounter, *et al.*, 2000 A unique metabolic syndrome causes obesity in the melanocortin-3 receptor-deficient mouse. Endocrinology **141:** 3518–3521.

Deng, H. W., W. M. Chen and R. R. Recker, 2001 Population admixture: detection by Hardy-Weinberg test and its quantitative effects on linkage-disequilibrium methods for localizing genes underlying complex traits. Genetics **157:** 885–897.

Elston, D. A., 1998 Estimation of denominator degrees of freedom of F-distributions for assessing Wald statistics for fixed-effect factors in unbalanced mixed models. Biometrics **54:** 1085–1096.

Gantz, I., H. Miwa, Y. Konda, Y. Shimoto, T. Tashiro *et al.*, 1993 Molecular cloning, expression and gene localization of a 4th melanocortin receptor. J. Biol. Chem. **268:** 15174–15179.

Gilmour, A. R., B. R. Cullis, S. J. Welham and R. Thompson, 2001 *ASREML Reference Manual.* ftp.res.bbsrc.ac.uk in pub/aar.

Guo, S. W., and E. A. Thompson, 1992 A Monte Carlo method for combined segregation and linkage analysis. Am. J. Hum. Genet. **51:** 1111–1126.

Hernández-Sánchez, J., D. Waddington, P. Wiener, C. S. Haley and J. L. Williams, 2002a Genome-wide search for markers associated with bovine spongiform encephalopathy. Mamm. Genome **13:** 164–168.

Hernández-Sánchez, J., C. S. Haley and P. M. Visscher, 2002b Power of association and transmission disequilibrium tests. World Congress of Genetics Applied to Livestock Production, Communication 21–24, Montpellier, France.

Hinney, A., A. Schmidt, K. Nottebom, O. Heibult, I. Becker *et al.*, 1999 Several mutations in the melanocortin-4 receptor gene including a nonsense and a frameshift mutation associated with dominantly inherited obesity in humans. J. Clin. Endocrinol. Metab. **84:** 1483–1486.

Horvath, S., and N. M. Laird, 1998 A discordant-sibship test for disequilibrium and linkage: no need for parental data. Am. J. Hum. Genet. **63:** 1886–1897.

Janss, L. L. G., R. Thompson and J. A. M. van Arendonk, 1995 Application of Gibbs sampling for inference in a mixed major gene-polygenic inheritance model in animal populations. Theor. Appl. Genet. **91:** 1137–1147.

Kenward, M. G., and J. H. Roger, 1997 Small sample inference for fixed effects from restricted maximum likelihood. Biometrics **53:** 983–997.

Kim, K. S., H. J. Larsen and M. F. Rothschild, 1999 Rapid communication: linkage and physical mapping of the porcine melanocortin-4 receptor (MC4R) gene. J. Anim. Sci. **78:** 791.

Kim, K. S., N. Larsen, T. Short, G. Plastow and M. F. Rothschild, 2000 A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth and feed intake traits. Mamm. Genome **11:** 131–135.

Koegler, F. H., K. L. Grove, A. Shiffmacher, M. S. Smith and J. L. Cameron, 2001 Central melanocortin receptors mediate changes in food intake in the rhesus macaque. Endocrinology **142:** 2586–2592.

Lewontin, R. C., 1988 On measures of gametic disequilibrium. Genetics **120:** 849–852.

Li, W. D., E. J. Joo, E. B. Furlong, M. Galvin, K. Abel *et al.*, 2000 Melanocortin 3 receptor (MC3R) gene variants in extremely obese women. Int. J. Obes. Metab. Dis. **24:** 206–210.

Lunetta, K. L., S. V. Faraone, J. Biederman and N. M. Laird, 2000 Family-based tests of association and linkage that use unaffected sibs, covariates, and interactions. Am. J. Hum. Genet. **66:** 605–614.

Martin, E. R., S. A. Monks, L. L. Warren and N. L. Kaplan, 2000 A test for linkage and association in general pedigrees: the pedigree disequilibrium test. Am. J. Hum. Genet. **67:** 146–154.

Patterson, H. D., and R. Thompson, 1971 Recovery of inter-block information when block sizes are unequal. Biometrika **58:** 545–554.

Payne, R. W., P. W. Lane, P. G. N. Digby, S. A. Harding, P. K. Leech *et al.*, 2001 *Reference Manual.* GENSTAT 5 Release 4.2, Oxford University Press, Oxford.

Rabinowitz, D., 1997 A transmission disequilibrium test for quantitative trait loci. Hum. Hered. **47:** 342–350.

Schaid, D. J., 1996 General score tests for associations of genetic markers with disease using cases and their parents. Genet. Epidemiol. **13:** 423–449.

Schaid, D. J., and C. M. Rowland, 1999 Quantitative trait transmission disequilibrium test: allowance for missing parents. Genet. Epidemiol. **17:** S307–S312.

Sokal, R. R., and R. J. Rohlf, 1995 *Biometry*. W. H. Freeman, New York.

Sorensen, D., and D. Gianola, 2002 *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics*. Springer Verlag, New York.

Spielman, R. S., and W. J. Ewens, 1998 A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. Am. J. Hum. Genet. **62:** 450–458.

Spielman, R. S., R. E. McGinnis and W. J. Ewens, 1993 Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am. J. Hum. Genet. **52:** 506–516.

Szyda, J., Z. Liu and V. Wild, 1998 Application of the transmission-disequilibrium test to detection of major genes, 49th Annual Meeting of the European Association for Animal Production, Warsaw.

Tachibana, T., K. Sugahara, A. Ohgushi, R. Ando, S. Kawakami *et al.*, 2001 Intracerebroventricular injection of agouti-related protein attenuates the anorexigenic effect of alpha-melanocyte stimulating hormone in neonatal chicks. Neurosci. Lett. **305:** 131–134.

Vaisse, C., K. Clement, B. Guy-Grand and P. Froguel, 1998 A frameshift mutation in human MC4R is associated with a dominant form of obesity. Nat. Genet. **20:** 113–114.

Wang, C. S., J. J. Rutledge and D. Gianola, 1994 Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs. Genet. Sel. Evol. **26:** 91–115.

Wardlaw, S. L., 2001 Obesity as a neuroendocrine disease: lessons to be learned from proopiomelanocortin and melanocortin receptor mutations in mice and men. J. Clin. Endocrinol. Metab. **86:** 1442–1446.

Welham, S. J., and R. Thompson, 1997 Likelihood ratio tests for fixed model terms using residual maximum likelihood. J. R. Stat. Soc. Ser. B **59:** 701–714.

Yeo, G. S., I. S. Farooqui, S. Aminian, D. J. Halsall, R. G. Stanhope *et al.*, 1998 A frameshift mutation in MC4R associated with dominantly inherited human obesity. Nat. Genet. **20:** 111–112.

Zhao, H., S. Zhang, K. R. Merikangas, M. Trixler, D. B. Wildenauer *et al.*, 2000 Transmission/disequilibrium tests using multiple tightly linked markers. Am. J. Hum. Genet. **67:** 936–946.

Communicating editor: M. A. F. Noor

APPENDIX: IMPACT OF STRATIFICATION ON b_{TD} AND b_{PD}

Population stratification affects both the estimation of the effect via b_{PD} and the power of that estimation (SE of b_{PD}). We assume the simplest scenario where a population is divided into two subpopulations of equal size, where mating is at random within each subpopulation, and there is no mating across subpopulations.

The expected mean square of a linear model that regresses phenotypes onto a single explanatory variable X is $E[\text{MSR}] = \sigma_e^2 + B^2\,\Sigma(X - \overline{X})^2$, where B is the expected regression parameter and σ_e^2 is the residual variance (Sokal and Rohlf 1995). When n phenotypes are simulated with gene effect (B) and $\sigma_e^2 = 1$, then $E[\text{MSR}] = 1 + E[\Sigma(X - \overline{X})^2] = 1 + n\{E[X^2] - (E[X])^2\}$. Let us assume that p_i is the frequency of allele i at the trait locus in subpopulation 1 and q_i is the equivalent frequency in subpopulation 2. Furthermore, let us assume that $p_{11} = (p_1)^2$ is the frequency of genotype 11 in subpopulation 1 and q_{11} is the frequency of the equivalent genotype in subpopulation 2. It can be shown that $E[\text{PD}] = (p_1 - p_2 + q_1 - q_2)/2$ and that $E[\text{PD}^2] = 0.5(p_{11}^2 + p_{22}^2 + q_{11}^2 + q_{22}^2) + 0.25(p_{11}(1 - p_{11}) + q_{11}(1 - q_{11}))$. The same process is followed to develop the expected mean squares when using TD; thus $E[\text{TD}^2] = (p_1^2 + q_1^2)/4$, and $E[\text{TD}] = 0$. These two predictions were very similar to simulation results (not shown). We can see the effect of stratification on $E[\text{MSR}]_{TD}$ and $E[\text{MSR}]_{PD}$ for $n = 100$ in Figure A1, A and B. Figure A1A takes a hill-type shape (seen from above) where the highest point is at the center and gradually decays in all directions away from the center. Figure A1B takes a valley-type shape (seen from above) where the lowest points are on the diagonal (in fact, Figures A1A and A1B are identical on the diagonal) and quickly rising away from the diagonal. If there is no stratification, *e.g.*, on the diagonal passing through points $p_1 = q_1$, then $E[\text{MSR}]_{TD} = E[\text{MSR}]_{PD}$, and thus both regression lines are identical. However, when there is stratification (*e.g.*, $p_1 \neq q_1$), $E[\text{MSR}]_{PD}$ increases and $E[\text{MSR}]_{TD}$ decreases. The effect on $E[\text{MSR}]_{TD}$ is not due to stratification itself but rather to a reduction of the information content due to a decrease in the frequency of heterozygous genotypes. Although we have shown the effect of analyzing the trait locus itself, this effect is transferred to a marker as a function of linkage disequilibrium between both loci.

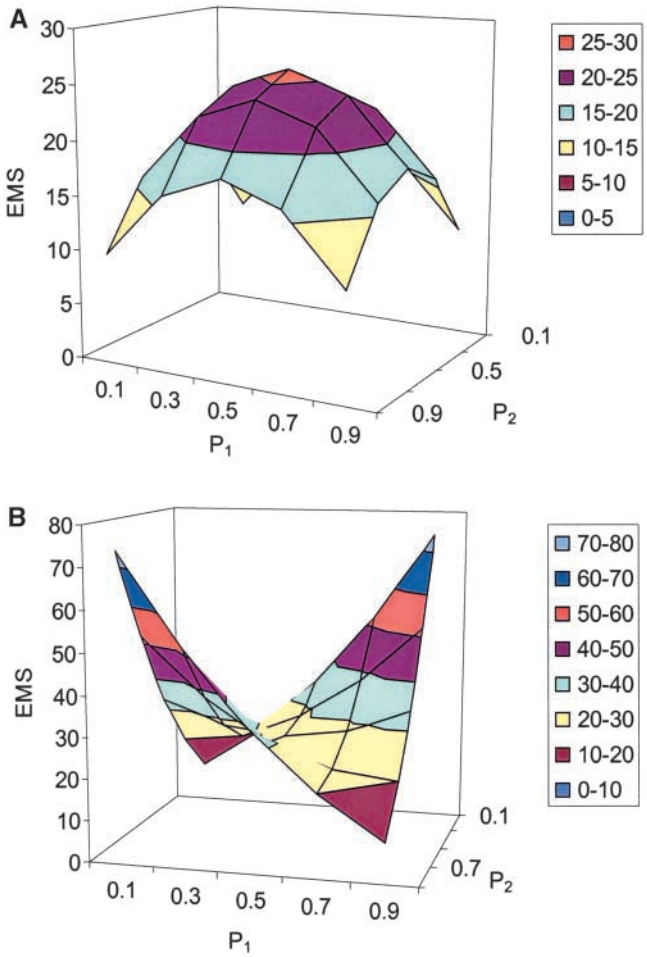

Figure A1.—(A) Expected mean squares (EMS) for b_{TD}, estimated with a simple linear regression (see text), when frequencies of allele A in two different populations are P_1 and P_2. (B) EMS for b_{PD}.