Detecting Major Genetic Loci Controlling Phenotypic Variability in Experimental Crosses

Lars Rönnegård* and William Valdar†

*Statistics Unit, Dalarna University, SE-781 70 Borlänge, Sweden and †Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7265

Manuscript received January 20, 2011
Accepted for publication March 21, 2011

ABSTRACT

Traditional methods for detecting genes that affect complex diseases in humans or animal models, milk production in livestock, or other traits of interest, have asked whether variation in genotype produces a change in that trait’s average value. But focusing on differences in the mean ignores differences in variability about that mean. The robustness, or uniformity, of an individual’s character is not only of great practical importance in medical genetics and food production but is also of scientific and evolutionary interest (e.g., blood pressure in animal models of heart disease, litter size in pigs, flowering time in plants).

We describe a method for detecting major genes controlling the phenotypic variance, referring to these as vQTL. Our method uses a double generalized linear model with linear predictors based on probabilities of line origin. We evaluate our method on simulated F2 and collaborative cross data, and on a real F2 intercross, demonstrating its accuracy and robustness to the presence of ordinary mean-controlling QTL. We also illustrate the connection between vQTL and QTL involved in epistasis, explaining how these concepts overlap. Our method can be applied to a wide range of commonly used experimental crosses and may be extended to genetic association more generally.

QUANTITATIVE trait locus (QTL) analysis has traditionally focused on detection of major genes controlling the expected mean of a phenotype. But there is substantial evidence that not only the mean but also the variance, that is, the stochastic variability of the phenotype about its average value, may itself be under genetic control. The identification of such variance-controlling loci, which we call vQTL, can be helpful in a variety of contexts, including selection of livestock for uniformity, evaluating predictability of response to medical treatment, identification of key biological stabilizers, and assessment of population resilience in ecology and evolution.

One way of interpreting an increase in variability is as a decrease in stability. Waddington (1942) described the concept of canalization, whereby natural selection favors the relative constancy of some attributes, for example, well-formed organs and limbs, and thereby leads to the evolution of heritable architectures that buffer the impact of environmental or background genetic variation that would otherwise cause development to go astray. These architectures create virtual “canals” down which developmental programs flow. For a canalicized phenotype, which modern usage expands to include nondevelopmental traits, the “zone of canalization” is the range of underlying liability over which potentially disruptive variation may be absorbed without serious consequence to the expressed trait value (Lynch and Walsh 1998). A well-studied example of a stabilizing architecture is that provided by heat-shock protein 90 (Hsp90), which buffers genetic and stochastic variation in the development of plants and flies (Rutherford and Lindquist 1998; Quetsch et al. 2002; Sangster et al. 2008).

But in absorbing variation, such stabilizing architectures also hide it from view, and a sensitizing change in the stabilizer that shifts liability outside the zone of canalization can have a dramatic effect on the phenotype. Such shifts release the combined effects of previously “cryptic” genetic variation: now decanalized, the phenotype is more sensitive to internal (including genetic) and external environment, and as a result varies more greatly between individuals (Dworkin 2005; Horstein and Shomron 2006). In this vein, decanalization has been proposed to explain why the genetic architectures of some diseases in human populations seem more amenable than others to genetic dissection through genome-wide association (Gibson and Goldstein 2007). Specifically, whereas some disease phenotypes in homogeneous populations may be heavily canalized and thereby harder to dissect, others may have been decanalized by modern living conditions...
Types of variance contributing to between-group differences in phenotypic variability

| Variance groupa | Sources of phenotypic variability |
|-----------------|-----------------------------------|
| Decanalization (epistasis) | Environmental sensitivity | Temporal fluctuation | Measurement error |
| Genetically distinct individuals with same allele at a vQTLa | • | • | • | • |
| Genetically identical individuals | • | • | • | • |
| Same individual at different times | • | • | • | • |
| Same individual at the same time | • | • | • | • |

aThe group in which variance is assessed, and between which variance is compared.

bThe variance groups compared here.

In this context, we consider the residual variance after controlling for main effects of QTL and other anticipated or manipulated environmental covariates. Phenotypic variability is thus the by-product of unmodeled interactions. Identifying major factors that influence variability requires defining groups between which variances would be contrasted (rows of Table 1). Our goal is to identify loci associated with differences in variance between such groups. For generality we concentrate on groups defined by the first row in Table 1, but note that the groupings defined in the remaining rows allow increasingly specific characterizations of vQTL effect. For instance, experimental crosses having multiple individuals within inbred lines will produce genetically identical individuals and the differences in phenotypic variability within each line are due to both environmental sensitivity and temporal fluctuation, but not epistasis.

Few studies have explicitly looked for vQTL. Among the more recent, ORDAS et al. (2008) studied morphological traits and flowering time in maize. They detected vQTL by contrasting the residual variance between genotypes in replicates of recombinant inbred lines (RILs; see second row, Table 1). The effects were substantial, with alleles associated with a 30–40% increase over the average residual variance. WITTENBURG et al. (2009) examined the sample variance of birth weight within pig litters as a gamma-distributed trait among 3914 sows, estimating a heritability of 0.1 for this trait using a generalized linear mixed model. SANGSTER et al. (2008) used Levene’s test for detection of variance-controlling genes. In that test, the absolute values of the residuals are used as a response in an ANOVA (e.g., FARAWAY 2004). MACKAY and LYMAN (2005) studied Drosophila bristle number and found substantial differences in the coefficient of variation (CV) between inbred lines, comparing CV also using ANOVA. The methods used in these last two studies have the limitation of not being able to model confounding effects in the mean. Using residuals (as in SANGSTER et al. 2008; WITTENBURG et al. 2009) can potentially incorporate covariates but involves conditioning on unknowns. There is thus
considerable utility in a method that simultaneously estimates means and variances, flexibly accommodates covariates, applies to a wide range of experimental crosses, and is robust and fast enough for genome-wide analyses.

Regression-based models (Haley and Knott 1992; Martínez and Curnow 1992) have proven to be fast and powerful at detecting QTL controlling the mean of a complex trait in experimental crosses and flexible since they are straightforwardly extended to include epistatic effects and interactions (Carlborg and Haley 2004). Mott et al. (2000) developed the haplotype re-construction method HAPPY and its associated regression model, which allows for a variable number of strains and may therefore be applied to vQTL mapping in, e.g., heterogeneous stocks (HS; Valdar et al. 2006b) and multiparent advanced generation inbred cross resource populations (MAGIC lines; Cavanagh et al. 2008) such as the collaborative cross (CC; Churchill et al. 2004; Broman 2005; Valdar et al. 2006a; Chesler et al. 2008) and the Arabidopsis recombinant inbred lines of Kover et al. (2009).

Our aim is to develop a regression model for detection of major genes controlling phenotypic variance that can be applied genome wide. The estimation uses double generalized linear models (DGLMs; Smyth 1989) and its parameterization is based on the HAPPY formulation of inferred haplotypes. The method fits ordinary QTL and vQTL simultaneously in the same model. We apply it to simulated data from an F2 and the CC and real data from an F2 intercross of partially inbred lines.

MODELS AND METHODS

The standard regression model for interval mapping of a QTL uses the probability of line origin at a locus to describe its genetic state (Haley et al. 1994). In the simple case of mapping a single QTL in individuals arising from an F2 intercross of inbred founder lines A and B, the model is

\[ y_i = \mu + x_i^T \beta + q_i^T \alpha + e_i, \quad e_i \sim N(0, \sigma^2), \]

where \( y_i \) is the phenotype of individual \( i \), \( x_i \) is the \( i \)th row in a design matrix of suitable nongenetic covariates, \( q_i \) represents the genetic state at the QTL, \( e_i \) is the residual with variance \( \sigma^2 \), and \( \mu \), \( \beta \), and \( \alpha \) are parameters estimated by the model. The QTL genotype \( q_i \) is typically unknown but, thanks to information from linked markers \( M \), its underlying haplotype pair is available indirectly as a probability distribution \( p_i = (p_{i1}, p_{i2}, p_{i3}) \), where \( p_{i1} = P(g_i = AA|M) \), \( p_{i2} = P(g_i \in \{AB, BA\}|M) \), and \( p_{i3} = P(g_i = BB|M) \). The regression predictor \( q_i \) in Equation 1 may therefore be formulated in terms of \( p_i \).

For the additive genetic models considered here, \( q_i = (q_{iA}, q_{iB}) \), where \( q_{iA} = 2p_{i1} + p_{i2} \) and \( q_{iB} = 2p_{i2} + p_{i3} \) correspond to the expected “doses” of haplotypes A and B respectively, and \( \alpha \) corresponds to their estimated “dosage effects” on the phenotypic mean. In practice, to obviate the dependence induced by \( q_{iB} = 2 - q_{iA} \), the regression model is fitted using \( q_i = q_{iA} \), leading to a scalar effect \( \alpha = \alpha \) that estimates the dosage effect of A relative to B. The predictor \( q_i \) may be alternatively formulated to accommodate more general effects, e.g., as \( q_i = p_i \), or represent observed or imputed genotypes of known variants (e.g., Yalcin et al. 2005; Zheng et al. 2011). The QTL scan is usually summarized as a plot of the LOD score, Fstatistic or, \(-\log_{10}(P\text{value})\) (hereafter, \( \log P \)), at each tested position along the genome. Chromosomal regions harboring QTL affecting the trait mean are located by the highest values of the test statistic above a suitable significance threshold (e.g., Broman and Sen 2009).

QTL regression model for detection of major loci controlling phenotypic variability: We consider the regression

\[ y_i = \mu + x_i^T \beta + q_i^T \alpha + e_i, \quad e_i \sim N(0, \sigma^2 \exp(z_i^T \gamma + q_i^T \theta)), \]

where \( z_i \) contains nongenetic covariates affecting the residual variance of the model, \( \gamma \) is their corresponding effects vector, and \( \theta \) is the dosage effect of each line on the residual variance, i.e., the additive vQTL effect. All other variables are defined as in Equation 1. The regression in Equation 2 is thus equivalent to Equation 1 but with \( e_i \sim N(0, \sigma^2 \exp(z_i^T \gamma + q_i^T \theta)) \), describing a model with separate effects for mean and variance.

Regression-based mapping of QTL (including vQTL) using Equation 2 assumes that
1. there are two founder lines,
2. the genetic state of the QTL is predicted accurately by marker data,
3. there is a single major QTL,
4. the QTL is fixed within each founder line,
5. the phenotype is Normally distributed, conditional on the QTL and covariate effects, and
6. the observed values \( y_i \) are exchangeable, conditional on the QTL and covariate effects.

We present a fitting procedure for Equation 2 based on these assumptions and, thereafter, relax the assumptions one at a time to investigate the possibility of using Equation 2 for vQTL detection in empirical studies. In this article we assess Assumptions 1–4 using both simulations on F2 and the CC, and empirical results from a chicken F2 cross. We give theoretical solutions for how to relax Assumptions 5 and 6 and discuss these in the Discussion. Our fitting procedure is based on DGLMs (see Appendix A) and uses the \texttt{dglm} package (Dunn and Smyth 2009) in R (R Development Core Team 2009).
Significance testing: A general method for calculating
P-values applicable to different trait distributions is avail-
able in dglm. The R code to extract the P-values from the
dglm(.) function is given in the supporting information
(File S1). We calculated 5% chromosome-wide signifi-
cance thresholds by simulating chromosomes under
a null model with no mean- and no variance-controlling
QTL effects.

Estimation of line dosages: We estimated line dosages
at each putative QTL using probabilities estimated by
the haplotype reconstruction program HAPPY (Mott
et al. 2000). Given genotype data on individual i and
its h founders, HAPPY uses a hidden Markov model
(HMM) to infer probabilistically the haplotype pair un-
derlying the genotypes at each marker. For every inter-
val between adjacent pairs of markers, it then calculates
the expected diplotype composition: that is, the average
proportion of diplotype AA, AB, etc., that would be
expected across the interval, given the interval’s length
and its descent either side. The diplotype composition
is reported as an

\[ p_{ij} = \frac{1}{1 + \exp(-r_{ij})} \]

and its descent either side. The diplotype composition
is reported as an

\[ p_{ij} = \frac{1}{1 + \exp(-r_{ij})} \]

for observation

\[ r_{ij} = \frac{1}{1 + \exp(-r_{ij})} \]

expected for maximum-likelihood (ML) estimation in interval
mapping (Landerc and Botstein 1989). The regression
approach to interval mapping (Equation 1) treats p, as
if it were an observed outcome, and as a result under-
estimates the residual variance for each observation by

\[ v_i^2 = v_i + \exp(\theta) \left( p_i \exp(r - \exp(p_i r)) \right) \]

(Xu 1995; Xu 1998; Feenstra et al. 2006). When modeling mean- and variance-
controlling QTL on the basis of Equation 2, the mar-
ginal likelihood is

\[ L_i = \sum_{i=1}^{3} p_{ij} N(r_i \alpha, \exp(r_i \theta) \sigma^2). \]  

(3)

The regression approach for this model overesti-
mates residuals further, by

\[ v_i^2 = v_i + \exp(\theta) \left( p_i \exp(r - \exp(p_i r)) \right) \]

(File S1). Although we could obtain ML estimates of the vQTL model using an EM-algorithm
(Appendix B), as is done for interval mapping, we contend that the DGLM regression may be more use-
ful. In particular, the EM-algorithm applied to this
marginal likelihood is computationally slow, a mar-
ginal likelihood gives biased ML estimates for var-
iances, and the regression approximation confers
additional flexibility for modeling different distribu-
tions for the phenotype as well as different mean-variance
relationships.

Relaxing Assumptions 1 and 2: Uncertain genotype
states with more than two founders: Uncertainty about
line origin probability when there are multiple lines can
produce multicollinearity among the genetic predictors
in the regression framework. We overcome this techni-
cal problem using the dimension reduction approach
described in Valdar et al. (2009), whereby the matrix
de of line dosage estimates from HAPPY is replaced by its
informative eigenvectors.

Quantifying the effects of genotype uncertainty:
Regression on expected dosages incorrectly models
uncertainty in the predictors. At a mean-controlling
QTL, individuals with less certain p, will be less well
predicted by the mean part of the model, which could
lead to inflated estimates of \( \sigma^2 \) and, possibly, a false
vQTL signal. We investigated this phenomenon empir-
ically by monitoring the relationship between locus un-
certainty and the proportion of false-positive vQTL in
simulations of the CC. At each marker interval in each
individual we quantified the uncertainty in line origin
using the scaled selective information content (SIC),
defined as follows. If \( P(g_i = j) \) is the prior probability
that individual i is in genetic state j given no marker
information, and \( P(g_i = j | M) \) is the posterior given
marker data M, as estimated by the HAPPY HMM, then
a measure of the information provided by M about the
locus is the Kullback–Leibler divergence,

\[ I(M, i) = \sum_{j=1}^{J} P(g_i = j | M) \log \frac{P(g_i = j | M)}{P(g_i = j)} \]

summed over all J possible states, with \( 0 \log(0) = 0 \). If
we represent the states of the F2 cross as all possible
phased founder diplotypes (i.e., AA, AB, BA, and BB
for the cross of strains A and B) and of the CC as the
set of homozygote diplotypes (i.e., AA, BB, . . . , HH;
denoting founders by A–H), then \( P(g_i = j) = \frac{1}{J} \), \( \forall j \).
Rescaling as SIC = \( I(M, i) / \log(J) \), the (scaled)
selective information content at a locus for individual i ranges
from 0, denoting equiprobable diplotypes and minimal cer-
tainty, to 1, denoting one diplotype with complete certainty.

Relaxing assumption 3: Multiple QTL: Multiple QTL
can be fitted by including additional predictors in Equa-
tion 2. Linked QTL may, however, affect the analysis if
not included in the model. We therefore assessed, by
means of simulations, the influence of additional inde-
pendent and interacting (epistatic) QTL.

Relaxing assumption 4: QTL variation within lines:
Model 2 assumes that both the mean- and variance-
controlling QTL have been fixed within the founder lines, which is reasonable for crosses from highly inbred lines but not necessarily for divergent outbred lines. It is possible that nonfixation of a mean-controlling QTL may be detected as a spurious vQTL. We illustrate this phenomenon for the Growth 2 QTL on chicken chromosome 1, which was found not to have been fixed in a divergent F2 cross (Kerje et al. 2003; Rönnegård et al. 2008).

SIMULATIONS AND DATA

We assessed the performance of our method by applying it to simulated F2 and CC populations, and to a real Red Jungle Fowl × White Leghorn F2 data set. Simulations were generated with software used for Valdar et al. (2006a) and Valdar et al. (2009) (see File S2).

Simulated F2 intercross: Our simulated F2 population included 800 individuals each comprising a single 100-cM chromosome with 10 evenly spaced fully informative SNP markers. QTL, when simulated, were positioned midway between markers, at 45 cM. The simulated population size of 800 was chosen to reflect a typically sized F2 design such as the Red Jungle fowl × White Leghorn F2 cross described further below.

Simulated collaborative cross: The CC is a panel of RILs descended from eight inbred founder strains: A/J, C57BL/6J, 129S1SvImJ, NOD/LtJ, NZO/H1LtJ, CAST/EiJ, PWK/PhJ, and WSB/EiJ (Churchill et al. 2004; Chesler et al. 2008). In keeping with previous simulations of this population (Valdar et al. 2006a), we simulated 1000 RIL individuals each generated from a separate breeding funnel. CC individuals comprised a single 100-cM chromosome with 1000 SNP markers drawn from the Mouse Diversity Genotyping Array (Yang et al. 2009; see http://cgd.jax.org/tools/diversityarray.shtml) and chosen to be roughly equally spaced and informative among the 8 founders. QTL, when simulated, were positioned at 28 and 68 cM, each midway between markers and in regions containing relatively informative markers.

Detection of QTL in F2 and CC with regression on known QTL genotypes: We simulated QTL in the F2 and CC under four scenarios.

Scenario 1: No QTL effects
Scenario 2: Mean-controlling QTL effects
Scenario 3: Variance-controlling QTL effects
Scenario 4: Both mean- and variance-controlling QTL effects

Phenotypic values (Y) were generated as mean-controlling additive QTL effects (Q) plus a residual (E): Y = Q + E. For mean-controlling effects, Q = 1, 2, and 4 for QTL genotype aa, aA, and AA, respectively. These effects correspond to a QTL with a moderate effect explaining ~2% of the phenotypic variance. For the scenarios with no vQTL, E was drawn from N(0, σ²) with σ² = 100. For the scenarios with vQTL, σ² = 100, 125, 156.25 for genotypes aa, aA, and AA, such that the the A allele had additive effects on the log scale and increase variance by 25% for each copy. These effects are moderate in size. Ordas et al. (2008) found variance-controlling QTL in maize where the allele effects resulted in an increase of the residual variance of 30–40%. For each scenario, we generated 10,000 replicates of the F2 simulation and 1000 replicates of the CC (positioning the QTL at 28 cM only). Model fitting was implemented using dglm function in R (R Development Core Team 2009), applying the regression in Equation 2 to the F2 populations and that in Equation 2 to the CC.

Detection of QTL in F2 and CC with regression on line dosage: We repeated the simulations above at each marker interval fitting the DGLM to line dosages inferred by the HAPPY HMM. The simulations from scenario 1 (no effects) were used to obtain 5% chromosome-wise empirical significance levels.

Fitting a single QTL in the presence of linked QTL and epistasis: To study a potential influence of linked QTL on estimated vQTL effects, we simulated three scenarios for the CC. DGLM regression was performed on line dosages as above, with 1,000 replicates simulated for each scenario.

Scenario E1: A single mean-controlling QTL at 28 cM
Scenario E2: Two linked mean-controlling QTL with additive effects at 28 cM and 68 cM
Scenario E3: Two linked mean-controlling QTL with epistatic effects at 28 cM and 68 cM

We generated phenotypes (Y) as Y = Q + E, with constant residual variance E ~ N(0, 100) throughout. For Scenario E1, Q = 0, 2, or 4 for QTL genotype aa, aA, or AA. For Scenario E2, the additive QTL effects were calculated as Q = Q1 + Q2, where Qk = 0, 2, or 4 for aa, aA, or AA at QTL number k. For Scenario E3, QTL with interaction effects were simulated and Q was assigned values according to Table 2. For all three scenarios, we fitted a DGLM with line dosages as predictors for both mean and variance effects at the first QTL position only (28 cM).

Red Jungle fowl × White Leghorn F2 cross: a worst-case scenario for assessing the effects of uncertainty in QTL genotype and nonfixation of QTL within founder lines: In an F2 cross between the chicken lines Red Jungle fowl and White Leghorn, Kerje et al. (2003) had previously detected two QTL affecting body weight at 200 days of age on chromosome 1 around 100 cM (Growth 1) and 490 cM (Growth 2). This trait was chosen because the QTL Growth 1 and Growth 2 have been thoroughly studied previously (Kerje et al. 2003; Rönnegård et al. 2008) where Growth 1 has a very large effect and Growth 2 was not fixed within the founder lines. The
TABLE 2
Simulated epistatic effects for different QTL genotype combinations

| QTL 2 (68 cM) | aa | aA | AA |
|--------------|----|----|----|
| QTL 1 (28 cM) |    |    |    |
| aa           | 0  | 0  | 0  |
| aA           | 0  | 2  | 0  |
| AA           | 0  | 0  | 4  |

cross was composed of four founder individuals (two from each line) and 756 F2 offspring. Although the QTL, Growth 1 is easily shown to have a very strong effect on the mean (explaining 22% of the variance; KERJE et al. 2003), the analysis of Growth 2 is complicated by it being fixed within the Red Jungle fowl founders but variable within the White Leghorn line (RÖNNEGÅRD et al. 2008), leading to additional uncertainty in the underlying QTL genotype. We use these data to explore the differences in estimates from the EM-algorithm (Appendix B), which explicitly takes account of this uncertainty, and the DGLM estimation, which does not, when there is a very strong mean-controlling QTL and moderate marker information. Furthermore, we study the effect of a QTL not being fixed within the founder lines. A sex effect was included as a covariate in both the mean and variance parts of the model. Line dosages were calculated as Haley–Knott probabilities (see File S2) and models were fit as above using the regression in Equation 2.

RESULTS

Detection of QTL for F2 and CC with regression on QTL genotypes: For the F2 simulations the average of the estimated QTL effects was close to those simulated (Table 3). The false-positive rate was close to 0.05 (Table 5) both for Scenario 2 (mean-controlling QTL simulated) and Scenario 3 (vQTL simulated). The power to detect the mean-controlling QTL at a 5% genome-wide significance level was 91.7% (Table 5) and the power decreased slightly when a vQTL was added to the simulations due to the resulting increase in residual variance. The power to detect the vQTL was 77.0% (Table 5) and this was not substantially affected by including a mean-controlling QTL in the simulations (i.e., Scenario 4). For Scenario 4, with both ordinary QTL and vQTL being simulated, most QTL detected were positioned within, or close to, those simulated (Figure 1). The simulations gave similar results for Scenarios 2 and 3. The accuracy of the QTL position does not seem to be substantially affected if either a mean-controlling or a variance-controlling QTL is simulated vs. the Scenario 4 in which both effects are simulated (Table 6).

Marker informativeness was small (SIC around 0.1) for the F2 simulations because markers were spaced 10 cM apart, where these intervals are considerably larger than in most QTL studies today (see, e.g., KERJE et al. 2003). For perfect information about QTL genotype the line dosage predictor qi is 0, 1, or 2, whereas for low information it has an attenuated range and is centered around 1.0. As a result, the regression on line dosage overestimated QTL effect (α and θ) for the F2 simulations. The extent of overestimation depends on the range of the line dosage (i.e., max(q) − min(q)), which is 2 with complete information but was 0.348 in the simulations. In Table 3 we therefore report estimates from the line dosage model after division by 2/0.348 = 5.75.

For the CC simulations the 5% chromosome-wide significance threshold for log P was 3.24 and 3.01 for the mean and variance parts of the model respectively. The power to detect the mean-controlling QTL at a 5% genome-wide significance level was 98.2% (Table 5) and, as above, this decreases slightly when a vQTL is added owing to the increased residual variance. The

TABLE 3
Estimated QTL effects in simulated F2

| Simulated effects | Regression on QTL genotypes | Regression on line dosages* |
|-------------------|-----------------------------|-----------------------------|
|                   | Ordinary QTL | vQTL | Ordinary QTL | vQTL |
| No QTL            | 0.001         | −0.003 | 0.008         | 0.001 |
| Ordinary QTL     | 1.997         | −0.003 | 2.063         | 0.023 |
| vQTL              | −0.010        | 0.2175 | −0.004        | 0.238 |
| Ordinary and vQTL| 1.999         | 0.2173 | 1.913         | 0.247 |

*Simulated value for mean-controlling QTL, 2.0; simulated value for variance-controlling QTL, 0.22.

*Estimates from genome scan with regression on line dosages. Correction for shrunken line dosage estimates in HAPPY due to low marker information contents; corrected values = estimates from genome scan times half the range of line dosages.
power to detect the vQTL was 80.8% (Table 5) and the power was not substantially changed when a mean-controlling QTL was included in the simulations. For the CC simulations, the proportion of QTL detected within the correct ±0.3 cM was highest for a vQTL when simulating both mean- and variance-controlling QTL (Table 6). The estimated QTL positions were well centered around the true simulated QTL position (Figure 1).

Fitting a single QTL in the presence of linked QTL and epistasis: Fitting the DGLM model of Equation 2 to the simulated CC data with a single mean-controlling QTL (Scenario E1) resulted in 5.8% of replicates having P-values less than 0.05 in the variance part of the model. This is consistent with the P-values for the vQTL being robust to the presence of moderate-effect ordinary QTL. Under Scenario E2 (two linked QTL with additive effects) the proportion of replicates having P-values less than 0.05 was 4.0%, and the rate of false positives (type I error) did not seem to be substantially affected by an additional linked QTL acting additively. When there were two linked and interacting mean-controlling QTL (Scenario E3), 17.0% of the replicates had P-values for the variance submodel <0.05, indicating the proportion of false positives is substantially affected by linked QTL that act epistatically. By including interaction effects, between the two interacting loci, in the mean part of the model the false-positive rate was reduced to 5.6%.

By increasing the mean-controlling QTL effect, under Scenario E1, from 2.0 to 20, the empirical type I (false positive) error increased to 12.7%. Hence, vQTL detected close to mean-controlling QTL of large effect should be treated with caution and further analyzed using, e.g., an EM-algorithm where the effects of marker uncertainty are accounted for (see Appendix B).

To investigate the effect of marker informativeness for the CC, we performed further simulations that repositioned the QTL at different locations in a 5-cM-spaced ladder between 25 and 70 cM. For these simulations (200 replicates per QTL location), the SIC varied considerably both within and between the QTL locations, but there was no clear relationship between SIC and the log P-values for false vQTL (Figure 2) when an ordinary QTL was simulated with additive effect of 2.0.

Red Jungle fowl × White Leghorn F₂ cross: a worst-case scenario for assessing the effects of covariate uncertainty and nonfixation of QTL within founder lines: In a preliminary model for chicken body weight without QTL effects, estimation of sex effects in a DGLM gave highly significant (P < 10⁻⁶) estimates of 410.0 and 0.509 for mean and variance predictors, respectively. The estimates being significant and having the same sign suggests a general mean-variance relationship. This was confirmed by applying the Box–Cox procedure (Box and Cox 1964), which suggested a square root transformation (λ = 0.55, SE 0.11), and the extended quasi-likelihood (EQL) procedure of Nelder and Pregibon (1987), which indicated a linear mean-variance relationship (ψ = 1.2, SE 0.2) (see File S1).

A chromosome scan using dglm revealed mean-controlling QTL (Figure 3) that were similar to those estimated by homoscedastic regression (Kerje et al. 2003). There were no large differences between the scans for ordinary QTL with Box–Cox transformed vs. original body weight as response. However, there were substantial differences between the chromosome scans for

### Table 4

| Simulated effects | F₂ Ordinary QTL | vQTL | CC Ordinary QTL | vQTL |
|-------------------|----------------|------|----------------|------|
| No QTL            | 0.052          | 0.052| 0.054          | 0.053|
| Ordinary QTL      | 0.997          | 0.051| 1.000          | 0.043|
| vQTL              | 0.051          | 0.966| 0.057          | 0.998|
| Ordinary and vQTL | 0.984          | 0.963| 1.000          | 0.998|

### Table 5

| Simulated effects | F₂ Ordinary QTL | vQTL | CC Ordinary QTL | vQTL |
|-------------------|----------------|------|----------------|------|
| No QTL            | 0.050          | 0.050| 0.050          | 0.050|
| Ordinary QTL      | 0.917          | 0.055| 0.982          | 0.053|
| vQTL              | 0.059          | 0.770| 0.046          | 0.808|
| Ordinary and vQTL | 0.808          | 0.808| 0.928          | 0.808|
vQTL (Figure 3, middle graph), although no vQTL reached 5% significance.

A moderate-sized peak was detected for the vQTL at approximately the same location as Growth 2 (nominal $P$-value $= 0.02$, chromosome-wise $P$-value $= 0.10$). This peak is likely due to the fact that the QTL alleles of Growth 2 were not fixed within the founder line of domestic White Leghorn hens (Rönnegård et al. 2008). The reason for this is that, if an F2 individual has an allele inherited from the domestic leghorn line, then the residual variance will be greater than if the allele was inherited from the Red Jungle Fowl line since Equation 2 assumes that the QTL alleles are fixed within founder lines. Consequently, the QTL around 490 cM is not a variance-controlling QTL but rather an effect of Growth 2 not being fixed within founder lines.

To assess spurious effects of uncertainty in genotype state, the additive effects for vQTL in Equation 2 were calculated using an EM-algorithm (Appendix B) and were subsequently compared with the DGLM estimates (Figure 4). The estimates are given as percentage change in residual variance for allele substitutions, i.e., $100 \times \exp \left( | \beta | \right) - 1$. The estimates are almost identical except for the positions around 115 cM where there is a strong mean-controlling QTL and moderate marker information (log $P = 57.3$ and SIC $= 0.7$, Figure 3). Hence, no major improvement in the QTL detection was achieved by using the more theoretically correct EM-algorithm.

**DISCUSSION**

We have studied the potential of detecting variance-controlling QTL by fitting a double generalized linear model via a two-stage procedure to simulated phenotype and marker information. The model can detect QTL that affect both the mean and the variance or QTL that affect either the mean or the variance. Because we use line origin probabilities as predictors in the model, as calculated using the HAPPY HMM (Mott et al. 2000), our approach can be applied to a wide range of experimental crosses. There are, however, some important considerations to be emphasized.

Detecting vQTL can be more challenging in theory, practice, and subsequent interpretation than detecting QTL controlling the trait mean. This is unsurprising: at the most basic level, variances are harder to estimate than means, typically requiring five times as many observations to achieve comparable precision (cf. “Tukey’s rule of 5” in Lee and Nelder 2006). Indeed, Visscher and Posthuma (2010) demonstrate analytically that detecting small effect vQTL among unrelated humans using traditional methods would require sample sizes in the 10,000s. More insidiously, inferences about differences in variances can be more sensitive to distributional assumptions than inferences about means. Specifically, it is common for raw phenotype measurements to exhibit a mean-variance relationship that naturally arises through the data generation process: for example, if body length is homoscedastic.
Gaussian then its square (e.g., body area) will not be. In the detection of vQTL, it is therefore especially important to recognize such relationships at an early stage and apply suitable normalizations or explore parametric alternatives to the normal distribution to avoid QTL affecting the mean appearing also to affect the variance. In File S1 we illustrate this problem along with an effective remedy via the Box–Cox transformation. Moreover, we set forth guidelines for how to approach a conservative vQTL analysis of data that are likely to be approximately normal but only after an unknown transformation, and of data that are more suitably modeled by other members of the exponential family, e.g., Poisson for count data, for which a known mean-variance relationship exists. Where it is felt such precautions may be inadequate or are impractical, we suggest resorting to the overconservative strategy of prioritizing “pure vQTL,” that is, strong vQTL with negligible mean effect, subject to all the usual requirements of deconfounding that apply to the detection of ordinary QTL.

Regression on genotype probabilities or expected fractions, and this result is consistent with recent work by Paré et al. (2010) and Struchalin et al. (2010). Alternatively, it could represent a differential sensitivity to environmental variation. Specifically, increased variability under one genotype may manifest as temporally stable phenotypes varying between individuals, phenotypes being highly fluctuant within an individual but on average similar between, little actual variation but a tendency for increased error of measurement, or some combination of these. Dissecting the components of the induced variability could be accomplished by applying our DGLM framework to a more focused experimental design, e.g., incorporating replicates of the CC at different levels (e.g., guided by column 1 of Table 1). In some cases, further statistical exploration on the same data may also be helpful. For example, if the mechanism is epistatic, vQTL detection can be seen as a starting point for modeling the joint action of multiple interacting loci, and this is easily accommodated in our regression framework.

Levene’s test is a popular method for testing equality of variance. Paré et al. (2010) and Struchalin et al. (2010) used it to test for different variances between SNP genotypes. It is capable of modeling group effects and has the advantage of being quite insensitive to non-normality. The DGLM approach, however, is far more flexible: it allows us to model continuous predictors and more complex general relationships of covariates on the variance, which Levene’s test cannot. Levene’s test also does not account for possible imbalance in the data since the estimated residuals mask varying uncertainty among groups. The DGLM approach, by contrast, allows modeling of nongenetic effects in the variance
part of the model, which may be important if for instance the variances are different between batches or contemporary groups in a designed QTL experiment. Moreover, DGLMs can also be extended to include random effects, due to family or polygenic effects, through hierarchical generalized linear models (Lee and Nelder 1996; Lee et al. 2006; Rönnegård et al. 2010b) and double hierarchical generalized linear models (Lee and Nelder 2006; Rönnegård et al. 2010a). This modeling flexibility is necessary in QTL studies and gives a more general approach applicable to a wide range of experimental designs.

There are several possible extensions of our model. Dominance can be included similarly as in other QTL regression models (Haley et al. 1994), as can multiple loci (e.g., Valdar et al. 2009). Because DGLM allows for any response distribution from the exponential family (Smyth 2002), our model is straightforwardly extended to binomial, Poisson, or gamma-distributed traits. In particular, several studies have focused on QTL controlling the CV rather than the variance (e.g., Mackay and Lyman 2005; Ansel et al. 2008). Traits with a constant CV, given the explanatory variables, are naturally modeled using a gamma distribution (McCullagh and Nelder 1989), and the DGLM method can be adapted to model such traits by setting the dispersion to CV² (see File S1). If we do not know whether to search for variance-controlling or CV-controlling QTL (i.e., whether the trait should be normal or gamma distributed), we can use the EQL (Nelder and Pregibon 1987) to compare model fits (see File S1). Ideally, distributional assumptions should always be checked for a detected QTL using a QQ-plot for the deviance residuals (McCullagh and Nelder 1989) from the mean part of the DGLM.

We anticipate identification of vQTL will be of interest in a wide range of genetics studies and applications. A clear application is in breeding systems, where vQTL detection could help selection for more robust livestock production (Mulder et al. 2007; Mulder et al. 2008). We believe there is also substantial scope for application in the study of animal models of human

Figure 3.—Scan for QTL controlling the mean (top) and the variance (middle) of body weight at 200 days of age on chicken chromosome 1 in an F₂ cross between Red Jungle Fowl and White Leghorn with 756 F₂ offspring. (Bottom) Marker information contents (SIC). Genome-wide significance threshold calculated using 1000 permutations.
disease. For example, medical diagnosis of hypertension has traditionally focused on achieving reliable estimates of the underlying mean blood pressure. However, increasing evidence points to the dangers of temporal fluctuation about the mean (Parati et al. 2006; Brunelli et al. 2008; Rothwell et al. 2010). Our framework could be used to detect vQTL affecting such variability in animal models and could be adapted to use in suitably controlled human studies.

In conclusion, we have developed a regression model for detection of major loci controlling phenotypic variance, which can be applied on a wide range of experimental crosses such as backcross, F₂, and MAGIC lines such as the CC. We have studied the robustness of the model to varying marker information, misspecification of the response distribution, linked QTL, and epistasis, proposed recommendations for its use, and discussed the meaning of detected vQTL and how they might be further dissected. We expect detection of vQTL will have wide application in genetics.

We acknowledge Leif Andersson for providing data from the Red Jungle fowl × White Leghorn cross. We also acknowledge Gary Churchill for helpful discussions regarding use of the EM algorithm. L.R. recognizes financial support by the Swedish Research Council Formas. W.V. acknowledges partial support by National Institute of Mental Health/National Human Genome Research Institute grant P50 MH090338 and by a Career Development Award from the Medical Research Council, UK.

LITERATURE CITED

ANDERSSON, L., 2001 Genetic dissection of phenotypic diversity in farm animals. Nat. Rev. Genet. 2: 1–11.

ANSEL, J., H. BOTTIN, C. RODRIGUEZ-BELTRAN, C. DAMON, M. NAGARAJU et al., 2008 Cell-to-cell stochastic variation in gene expression is a complex trait. PLoS Genet. 4: e1000049.

BEAUMONT, H. J., J. GALLIE, C. KOST, G. C. FERGUSON and P. B. RAINLEY, 2009 Experimental evolution of bet hedging. Nature 462: 90–93.

BOX, G. E. P., and D. R. COX, 1964 An analysis of transformations (with discussion). J. R. Stat. Soc. B 26: 211–252.

BRONMAN, K., 2005 The genomes of recombinant inbred lines. Genetics 169: 1135–1146.

BRONMAN, K. W., and S. SEN, 2009 A Guide to QTL Mapping with R/qtl. Springer, New York.

BRUNELLI, S., R. I. THADHANI, K. E. LYNCH, E. D. ANKERS, M. M. JOFFE et al., 2008 Association between long-term blood pressure variability and mortality among incident hemodialysis patients. Am. J. Kidney Dis. 52: 716–726.

CARLBOERG, O., and C. S. HALEY, 2004 Epistasis: Too often neglected in complex trait studies? Nat. Rev. Genet. 5: 618–625.

Cavanagh, C., M. Morelli, I. Mackay and W. Powell, 2008 From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. Curr. Opin. Plant Biol. 11: 215–221.

Chesler, E. J., D. R. MILLER, L. R. BRANSTETTER, L. D. GALLOWAY, B. L. JACKSON et al., 2008 The Collaborative Cross at Oak Ridge National Laboratory: developing a powerful resource for systems genetics. Mamm. Genome 19(6): 382–389.

Churchill, G. A., A. D. ABBEY, H. ALLAYEE, J. M. ANGEL, A. D. ATTIE et al., 2004 The Collaborative Cross, a community resource for the genetic analysis of complex traits. Nat. Genet. 36: 1133–1137.

Clayton, G., and A. Robertson, 1957 An experimental check on quantitative genetical theory. II. The long-term effects of selection. J. Genet. 55: 152–170.

Dunn, P. K., and G. K. Smyth, 2008 dglm: double generalized linear models. R package version 1.6.1.

Dworkin, I., 2005 Canalization, cryptic variation, and developmental buffering: a critical examination and analytical perspective, pp. 131–158 in Variation: A Central Concept in Biology, edited by B. Hallgrimsson and B. K. Hall. Elsevier Academic Press, New York.

Faraway, J. J., 2004 Linear Models with R. Chapman & Hall/CRC, London.

Feenstra, B., I. M. SROVGAARD and K. W. BROMAN, 2006 Mapping quantitative trait loci by an extension of the Haley–Knott regression method using estimating equations. Genetics 173: 2290–2292.

Feinberg, A. P., and R. A. IRIZARRY, 2010 Evolution in health and medicine Sackler colloquium: stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. Proc. Natl. Acad. Sci. USA 107(Suppl.): 1757–1764.

Gibson, G., and D. B. GOLDSTEIN, 2007 Human genetics: the hidden text of genome-wide associations. Curr. Biol. 17: 929–932.

Haley, C. S., and S. A. Knott, 1992 A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69: 315–324.

Haley, C. S., S. A. Knott and J. M. ELSEN, 1994 Mapping quantitative trait loci in crosses between outbred lines using least squares. Genetics 136: 1195–1207.

Hill, W. G., and X. S. ZHANG, 2004 Effects on phenotypic variability of directional selection arising through genetic differences in residual variance. Genet. Res. 83: 121–132.

Hoaglin, D. C., and R. E. WELCH, 1978 The hat matrix in regression and ANOVA. Am. Stat. 32: 17–22.

Hornstein, E., and N. SHOMRON, 2006 Canalization of development by microRNAs. Nat. Genet. 38: S20–S24.

Houle, D., 1992 Comparing evolvability and variability of quantitative traits. Genetics 130: 195–204.
RÖNNEGÅRD, L., X. SHEN and M. ALAM, 2010b – hglm: A package for fitting hierarchical generalized linear models. R J. 2(2): 20–28.

ROS, M., D. SORENSEN, R. WAAGEPETERSEN, M. DUPONT-NIVET, M. SANCRISTORAL et al., 2004 – Evidence for genetic control of adult weight plasticity in the snail Helix aspersa. Genetics 168: 2089–2097.

ROTGERELLI, P. M., S. G. HOWARD, E. DOLAN, E. O’BRIEN, J. E. DORBON et al. 2010 – Prognostic significance of visit-to-visit variability, maximum systolic blood pressure, and episodic hypertension. Lancet 375: 895–905.

ROWE, S. J., I. M. S. WHITE, S. AVENDANO and W. G. HILL, 2006 – Genetic heterogeneity of residual variance in broiler chickens. Genet. Sel. Evol. 38: 617–635.

RUTHERFORD, S., and S. LINDQUIST, 1998 – Hsp90 as a capacitor for morphological evolution. Nature 396: 336–342.

SANCRISTORAL-GAUDY, M. J., M. ELSEN, L. BODIN and C. CHEVALET, 1998 – Prediction of the response to a selection for canalisation of a continuous trait in animal breeding. Genet. Sel. Evol. 30: 423–451.

SANGSTER, T. A., N. SALATHA, S. UNDERRAGA, R. ILLO, K. SCHELLENBERG et al., 2008 – HSP90 affects the expression of genetic variation and developmental stability in quantitative traits. Proc. Natl. Acad. Sci. USA 105: 2963–2968.

SMITH, G. K., 1989 – Generalized linear models with varying dispersion. J. R. Stat. Soc. B 51: 47–60.

SMITH, G. K., 2002 – An efficient algorithm for REMI in heteroscedastic regression. J. Comput. Graph. Stat. 11: 836–847.

SORENSEN, D., and R. WAAGEPETERSEN, 2003 – Normal linear models with genetically structured residual variance heterogeneity: a case study. Genet. Res. 82: 207–222.

STRUDLICH, M. V., A. DEIGHAN, J. C. WITTEMAN, C. V. DUSJN and Y. S. AULCHENKO, 2010 – Variance heterogeneity analysis for detection of potentially interacting genetic loci: method and its limitations. BMC Genet. 11(1): 92.

VALDAR, W., J. FLINT and R. MOTT, 2006a – Simulating the Collaborative Cross: power of quantitative trait loci detection and mapping resolution in large sets of recombinant inbred strains of mice. Genetics 172: 1783–1797.

VALDAR, W., L. C. SOLBERG, D. GAUGUER, S. BURNET, P. KLENERMAN et al., 2006b – Genome-wide genetic association of complex traits in heterogeneous stock mice. Nature Genetics 38: 879–887.

VALDAR, W., C. C. HOLMES, R. MOTT and J. FLINT, 2009 – Mapping in structured populations by resample model averaging. Genetics 182: 1263–1277.

VERRYLA, A. P., 1993 – Modelling variance heterogeneity: residual maximum likelihood and diagnostics. J. R. Stat. Soc. B 55: 493–508.

Visscher, P. M., and D. POSTHuma, 2010 – Statistical power to detect genetic loci affecting environmental sensitivity. Behav. Genet. 40: 729–733.

Waddington, C., 1942 – Canalization of development and the inheritance of acquired characters. Nature 150: 563–565.

WITTENBURG, D., V. GUAID, F. TEUSCHER and N. REINSCH, 2009 – Comparison of statistical models to analyse the genetic effect on within-litter variance in pigs. Animal 2: 1559–1568.

XU, S., 1995 – A comment on the simple regression method for interval mapping. Genetics 141: 1657–1659.

XU, S., 1998 – Further investigation on the regression method of mapping quantitative trait loci. Hereditas 130: 364–373.

YALCIN, B., J. FLINT and R. MOTT, 2005 – Using progenitor strain information to identify quantitative trait nucleotides in outbred mice. Genetics 171(2): 673–681.

YANG, H., Y. DING, L. HUTCHINS, J. SZATKEWICZ, T. BELL et al., 2009 – A customized and versatile high-density genotyping array for the mouse. Nat. Meth. 6: 663–666.

ZHENG, J., Y. LI, G. R. ABECASSIS and P. SCHEET, 2011 – A comparison of approaches to account for uncertainty in analysis of imputed genotypes. Genet. Epidemiol. 35(2): 102–110.

Communicating editor: K. M. NICHOLS
APPENDIX A: DOUBLE GENERALIZED LINEAR MODEL THEORY

For known genotypes at the QTL, the ML estimates of the effect parameters $\beta$, $\alpha$, $\gamma$, $\theta$ in Equation 2 can be estimated using Fisher scoring. Smyth (1989) showed this to be equivalent to predicting the mean effects using a linear model (Equation 1) and the squared residuals ($e_i^2$) using a generalized linear model (GLM) with a gamma-distributed response and the log link.

ML gives biased variance component estimates and the estimates are also sensitive to influential observations with high leverages (Verbyla 1993). A restricted maximum likelihood (REML) approach, implemented in the dglm package, resolves these issues by incorporating leverage into the estimation (Smyth 2002; Lee et al. 2006). Specifically, the gamma GLM is used to predict $\log(E(d_i)) = \log(\sigma^2) + z_i^T \gamma + q_i \theta$, where the response is now the weighted deviance component $d_i = \frac{e_i^2}{1 - h_{ii}}$. Here, the $h_{ii}$ is the leverage of observation $i$, equal to the $i$th diagonal element of the hat matrix $H = X_{all}(X_{all}^T V^{-1} X_{all})^{-1} X_{all}^T V^{-1}$, where $X_{all}$ is the design matrix for all fixed effects in the mean submodel and $V = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$.

Both the hat matrix and the associated leverages play an important role in ordinary regression theory (Hoaglin and Welsch 1978; Faraway 2004) and are used extensively in DGLM theory (Smyth 1989).

APPENDIX B: AN EM-ALGORITHM FOR ESTIMATING vQTL FOR NORMALLY DISTRIBUTED TRAITS IN F2 CROSSES

For an observation $i$, the full likelihood for vQTL with genotype uncertainty is

$$L_i = \sum_{j=1}^3 p_j N(\mu + x_i^T \beta + r_j \alpha, \sigma^2 \exp(z_i^T \gamma + r_j \theta)) = \sum_{j=1}^3 p_j f_{ij}$$

with variables other than $f_{ij}$ defined as for Equation 3. Let $l_i = \log(L_i)$, $g_{ij} = \log(f_{ij})$, and $\Psi$ be the parameter vector, and assume observations are independent given the parameters such that $L = \sum_i l_i$. Then,

$$\frac{\partial l_i}{\partial \Psi} = \frac{1}{L_i} \frac{\partial L_i}{\partial \Psi}$$

where $(\partial L_i/\partial \Psi) = \sum_{j=1}^3 p_j (\partial g_{ij}/\partial \Psi)$ and

$$\frac{\partial L_i}{\partial \Psi} = \sum_{j=1}^3 p_j \frac{\partial}{\partial \Psi} f_{ij} = \sum_{j=1}^3 p_j f_{ij} \frac{\partial g_{ij}}{\partial \Psi}$$

$$\Rightarrow \frac{\partial l_i}{\partial \Psi} = \frac{1}{\sum_{j=1}^3 p_j f_{ij}} \sum_{j=1}^3 p_j f_{ij} \frac{\partial g_{ij}}{\partial \Psi}$$

Here $g_{ij}$ is the log-likelihood given the genotypes, so the gradient and Hessian of $g_{ij}$ are those obtained from the logarithm of the normal density function $f_{ij}$. ML estimates of $\Psi$ are obtained by iterating the following EM steps until convergence:

1. Calculate the “weights” given the parameter estimates: $w_{ij} = p_j f_{ij} / (\sum_j p_j f_{ij})$
2. Calculate the parameter estimates given $w_{ij}$ from: $\partial l/\partial \Psi = \sum_{i=1}^n \sum_{j=1}^3 w_{ij} (\partial g_{ij}/\partial \Psi) = 0$, which can be calculated using a Newton iterative method since we have the gradient and Hessian of $g_{ij}$.
Detecting Major Genetic Loci Controlling Phenotypic Variability in Experimental Crosses

Lars Rönnegård and William Valdar

Copyright © 2011 by the Genetics Society of America
DOI: 10.1534/genetics.111.127068
1. FUNCTION TO EXTRACT P-VALUES FROM DGLM

Here we give an R function `dglm.Pvalues(.)` that extracts the P-values for a fitted model `dglm.fit`.

```r
# Example
set.seed(123)
require(dglm)
n = 200
a <- rbinom(n, 1, 0.5)  # Covariate for additive genetic effects
sex <- rbinom(n, 1, 0.5)  # Covariate for a non-genetic effect
add.effect = 0
sex.effect = 1
res.var = exp( a*add.effect )  # Residual variance
y <- 10 + a*add.effect + sex*sex.effect + rnorm(n, 0, sqrt(res.var))
d.fit <- dglm( formula = y ~ a + sex, dformula = ~ a )
P.values <- dglm.Pvalues( d.fit )
print( P.values )
```

This code gives the following output in R:

```r
$P.mean
```

---

**FILE S1**

**Supporting Theory**

1. FUNCTION TO EXTRACT P-VALUES FROM DGLM

Here we give an R function `dglm.Pvalues(.)` that extracts the P-values for a fitted model `dglm.fit`.

```r
dglm.Pvalues <- function(dglm.fit){
  P.disp = anova.dglm(dglm.fit)$Adj.P[2]
  P.mean = summary(dglm.fit)$coef[2,4]
  list(P.mean=P.mean, P.disp=P.disp)
}
```

Below we illustrate the use of the function with a simple example where the simulated genetic effects are set to 0.

```r
## An example using dglm.Pvalues(.) with 200 simulated observations
set.seed(123)
require(dglm)
n = 200
a <- rbinom(n, 1, 0.5)  # Covariate for additive genetic effects
sex <- rbinom(n, 1, 0.5)  # Covariate for a non-genetic effect
add.effect = 0
sex.effect = 1
res.var = exp( a*add.effect )  # Residual variance
y <- 10 + a*add.effect + sex*sex.effect + rnorm(n, 0, sqrt(res.var))
## The additive genetic effect must be given first in the following formula
d.fit <- dglm( formula = y ~ a + sex, dformula = ~ a )
P.values <- dglm.Pvalues( d.fit )
print( P.values )
```

This code gives the following output in R:

```r
$P.mean
```
2. DOUBLE GLM FOR NON-NORMAL TRAITS

For non-normal distributed traits a GLM can be used to model the mean-controlling QTL and the deviance components from this GLM are subsequently used in eq 3 (in the paper) instead of the squared raw residuals \( \varepsilon_i^2 \) (SMYTH 2002; LEE et al. 2006). In the more general GLM setting, the dispersion term \( \phi \) is modeled in the second submodel of the double GLM. For the exponential family of distributions the variance of \( y \) can be written as (MCCULLAGH and NELDER 1989):

\[
\text{Var}(y) = \phi V(\mu)
\]

where \( V(\mu) \) is the “variance function”, which gives the relationship between the variance of \( y \) and the mean \( \mu \). For the normal distribution \( V(\mu) = 1 \) (i.e. the variance does not vary with the mean) and the dispersion term is the residual variance, \( \phi \equiv \sigma^2 \). For the gamma distribution \( V(\mu) = \mu^2 \) and since \( \text{Var}(y)/V(\mu) = \phi \) we have that \( \phi \equiv CV^2 \). Consequently, we can model the coefficient of variation using a gamma GLM (see chapter 8 in (MCCULLAGH and NELDER 1989)). This gives a method to fit CV-controlling QTL. ANSEL et al. (2008) referred to such QTL in a gene expression study as “noise QTL”. For these calculations the \textit{dglm} package fits a digamma distribution (SMYTH 1989).

3. THE EFFECT OF UNCERTAINTY IN GENOTYPE STATE ON THE INFERENCE OF VQTLS

In a constant variance model, the sampling distribution of phenotype \( y_i \) of animal \( i \) given the genotype \( x_i = g_j \) at the QTL locus is \( p(y_i|x_i = g_j) = N(\alpha, \sigma^2) \) with

\[
E(y_i|x_i = g_j) = \alpha_j \quad \text{and} \quad \text{Var}(y_i|x_i = g_j) = \sigma^2
\]
or with $x_i$ representing an indicator vector for the genotype and $\alpha$ the genotype effects,

$$E(y_i|x_i) = x_i^T \alpha \quad \text{and} \quad Var(y_i|x_i) = \sigma^2$$

### The constant variance model with uncertain genotype

When genotype is uncertain, such that $p(x_i = g_j|\text{Data}) = p_{ij}$ and $p_i = [p_{i1}, p_{i2}, ...]$, then the sampling distribution conditional on $p_i$ is the mixture

$$p(y_i|p_i) = \sum_j p_{ij} p(y_i|x_i = g_j)$$

with $p(y_i|x_i = g_j)$ defined as above. Conditional on $p_i$ the expectation is $E(y_i|p_i) = p_i^T \alpha$, much as before, but the variance is now

$$Var(y_i|p_i) = Var_g(E(y_i|g_j)|p_i) + E_g(Var(y_i|g_j)|p_i)$$

$$= \left[ E_g(E(y_i^2|g_j)) - E_g^2(E(y_i|g_j)) \right] + E_g(Var(y_i|g_j))$$

$$= \left[ \sum_j p_{ij} \alpha_i^2 \left( \sum_j p_{ij} \alpha_i \right)^2 \right] + \sigma^2$$

$$= \left[ p_i^T (\alpha^2) - (p_i^T \alpha)^2 \right] + \sigma^2$$

$$v_i = v_i^{\text{extra}} + \sigma^2.$$

A quick alternative to fitting the mixture by maximum likelihood is the regression model used for Haley-Knott or HAPPY where it is assumed

$$p_{HK}(y_i|p_i) = N(p_i^T \alpha^2, \sigma^2)$$

or, as described by BROMAN and SEN (2009), an “extended Haley-Knott” regression fitted by ML to provide a closer approximation to the mixture model

$$p_{EHK}(y_i|p_i) = N(p_i^T \alpha^2, v_i)$$

which is slower than the HK model but faster than a full ML fit.
The vQTL model with known genotype. In the double GLM model, the sampling distribution of phenotype \( y_i \) of animal \( i \) given the genotype \( x_i = g_j \) at the QTL locus is

\[
p(y_i|g_j) = N(\alpha_j, \exp(\theta_j)),
\]

where \( \alpha_j \) and \( \theta_j \) are specific to genotype \( g_j \). The expectation and variance of \( y_i \) are

\[
E(y_i|x_i = g_j) = \alpha_j \quad \text{and} \quad \text{Var}(y_i|x_i = g_j) = \exp(\theta_j)
\]

The vQTL model with uncertain genotype. With uncertain genotype, the mean is the same as the constant variance model, \( E(y_i|p_i) = p_i^T\alpha \), but the variance is

\[
\text{Var}(y_i|p_i) = \text{Var}_a(E(y_i|g_j)|p_i) + \text{E}_a(\text{Var}(y_i|g_j)|p_i)
\]

\[
= [p_i^T(\alpha^2) - (p_i^T\alpha)^2] + p_i^T\exp(\theta)
\]

\[
v_i = v_i^{\text{extra}} + v_i^{\text{vQTL}}
\]

Since we are interested in how \( x_i \) affects the variance, our only measure of \( x_i \) is \( p_i \) and since \( v_i^{\text{extra}} = f(p_i, \alpha) \) such that \( v_i \) can potentially change with \( x_i \) in the absence of a vQTL effect, there is the potential for genotype uncertainty to mimic the effect of a vQTL. For low informative markers the \text{dglm} \ approach captures only part of the vQTL effect, as \( \text{Var}(y_i|p_i) = \exp(p_i^T\theta) \), so that

\[
v_i = v_i^{\text{vQTL}} + v_i^{\text{extra}} + (p_i^T\exp(\theta) - \exp(p_i^T\theta))
\]

A Taylor expansion shows that the last term is small for small to moderate vQTL effects \( (\theta << 1) \). This term does not depend on the effects in the mean model (i.e. \( \alpha \)). Its contribution to \( v_i \) should only reduce the power of the model, whereas the term \( v_i^{\text{extra}} \) may increase the Type I error.
4. TRANSFORMATION AND ASSESSMENT OF MEAN-VARIANCE RELATIONSHIP

Maximum likelihood can be used to estimate the parameter $\lambda$ in a Box-Cox transformation (Box and Cox 1964) for the linear model $z = X\beta + \varepsilon$ where the response is

$$z = \begin{cases} \frac{y^{\lambda} - 1}{\lambda} & \lambda \neq 0 \\ \log(y) & \lambda = 0 \end{cases}$$

and the residuals $\varepsilon$ are assumed normal i.i.d. The function boxcox(.) in the MASS library in R can be used for this estimation. We can thereby obtain a response $z$ that is approximately normal. See Pawitan (2001, page 178) for a description of the method. For the case of a pleiotropic gene controlling both the mean and the variance of a trait, the Box-Cox transformation of the raw data might be over-conservative by removing large parts of the vQTL effects. A potential remedy to this problem could be to include the Box-Cox parameter in the likelihood and to estimate it simultaneously with all other parameters in our model. A similar solution for mixed effect models has previously been developed using Bayesian techniques (Yang 2010) but we do not investigate this possibility in our current paper.

For GLM distributions in general, the concept of extended quasi-likelihood (Nelder and Pregibon 1987) can be used to assess the mean-variance relationship. Using e.g. the eql(.) function in the EQL package in R, the parameter $\psi$ for the relationship $\text{Var}(y) = \phi \mu^\psi$ with $\phi$ being a constant dispersion parameter. So that for a normal distribution $\psi = 0$, for a Poisson distribution $\psi = 1$, and for a gamma distribution $\psi = 2$. See Pawitan 2001 for a recent and clear description of the method.

Below we investigate the use of Box-Cox transformation to remove false vQTL from non-normal data and thereafter specify guidelines to avoid detecting false vQTL due to scale effects.
Using Box-Cox transformation to remove false vQTL from non-normal data. We simulated data with a homoscedastic (ie, constant) variance as

\[ y \sim N(\mu + \alpha q, 1) \]

setting \( \mu = 1 \), \( \alpha = 10 \) and fixing \( q \) as a vector of 50 zeros and 50 ones to give 100 observations in total. For each condition described below we simulated 20 replicates. For each simulated \( y \) we obtained four versions of an “observed phenotype” \( w \). We obtained these versions by applying four alternative transformations that set \( w \) to \( y \) (ie, the identity transformation), \( y^2 \), \( \sqrt{y} \), or \( \frac{1}{y} \) (as listed in Table A1 below).

The Box-Cox parameter \( \lambda \) was estimated by means of MLE (see, eg, page 178 in PAWITAN 2001) for the model

\[ \frac{w_i^\lambda - 1}{\lambda} = \mu + \alpha q_i + \varepsilon_i \]

where residuals are normally distributed and \( \sigma^2 \) is constant (i.e. for a model with homoscedastic residual variance). Two models were fitted using \textit{dglm} (DUNN and SMYTH 2009):

\[ w_i = \mu + \alpha q_i + \varepsilon_i \]

\[ \log(\sigma_i^2) = m + \theta q_i \]

and

\[ \frac{w_i^\lambda - 1}{\lambda} = \mu + \alpha q_i + \varepsilon_i \]

\[ \log(\sigma_i^2) = m + \theta q_i \]

where \( \mu \) and \( \alpha \) is an intercept term and regression coefficient for the mean model, and \( m \) and \( \theta \) is an intercept term and regression coefficient for the variance model.
Table A1: Estimated Box-Cox parameter $\lambda$ and estimated regression parameter $\theta$ in the variance part of the model (standard errors within brackets)

| $w$ | $\hat{\lambda}$ | $\hat{\theta}$ with response $w$ | $\hat{\theta}$ with response $\frac{w^{\lambda-1}}{\lambda}$ |
|-----|----------------|---------------------------------|---------------------------------|
| $y$ | 0.96 (0.015)   | 0.113 (0.047)                   | -0.006 (0.018)                  |
| $y^2$| 0.48 (0.007)   | 2.92 (0.074)                    | -0.006 (0.018)                  |
| $\sqrt{y}$| 1.92 (0.031) | -1.39 (0.039)                   | -0.006 (0.018)                  |
| $\frac{1}{y}$| -0.96 (0.015) | -6.47 (0.045)                   | -0.006 (0.018)                  |

For $y$, generated from normal random numbers with a homogeneous variance, the average estimate of $\theta$ over the 20 replicates was close to 0 as expected ($\hat{\theta} = 0.113$). By chance the samples of $y$ were slightly skewed and the average estimate of $\lambda$ was 0.96 with $\hat{\theta} = -0.006$ after Box-Cox transformation. Note that the average estimate of $\theta$ was -0.006 for all four cases and that the risk of detecting false vQTL due to scale effects therefore can be avoided using Box-Cox transformation. Note also that all three transformations gave estimates of $\theta$ significantly different from 0, implying that we can expect non-normal data to give false vQTL.

To assess the risk of Box-Cox transformation removing true vQTL, we simulated data with heteroscedastic variances:

$$y \sim N(\mu + \alpha q, \log(m + \theta q_d))$$

where $\mu = 1$, $\alpha = 10$, $m = 0$, and $\theta$ was set to either 1 or -1. As above, $q$ was a vector of 50 zeros and 50 ones. We studied two cases for each simulated value of $\theta$:

- Case 1: $q_d = q$
- Case 2: $q_d$ and $q$ are uncorrelated

The Box-Cox parameter $\lambda$ was estimated by means of MLE as described above, and two
models were fitted using dglm

\[ y_i = \mu + \alpha q_i + \varepsilon_i \]

\[ \frac{y_i^\lambda - 1}{\lambda} = \mu + \alpha q_i + \varepsilon \]

\[ \log(\sigma_i^2) = m + \theta_{q_{d,i}} \]

From the results for Case 1 \((q_d = q)\) in Table A2, we can see that Box-Cox transformation removes the vQTL effects if the covariates in the mean and variance parts of the model are highly correlated, i.e. if the vQTL is located at, or very close to, an ordinary QTL. From Case 2, we can see that Box-Cox transformation does not remove the vQTL effect if the vQTL and the ordinary QTL are not linked.

Table A2: Estimated Box-Cox parameter \(\hat{\lambda}\) and estimated regression parameter \(\hat{\theta}\) in the variance part of the model (standard errors within brackets)

| Simulated \(\theta\) | Simulated case | \(\hat{\lambda}\) | \(\hat{\theta}\) for response \(y\) | \(\hat{\theta}\) for response \(y^{\frac{\lambda-1}{\lambda}}\) |
|-------------------|----------------|----------------|----------------|----------------|
| 1                 | \(q_d = q\)   | 0.65 (0.013)   | 1.11 (0.047)   | 0.070 (0.018)  |
| -1                | \(q_d = q\)   | 1.28 (0.018)   | -0.89 (0.047)  | -0.073 (0.020) |
| 1                 | \(q_d \neq q\) | 0.98 (0.023)   | 1.03 (0.060)   | 1.00 (0.065)   |
| -1                | \(q_d \neq q\) | 0.95 (0.022)   | -1.00 (0.061)  | -0.95 (0.059)  |

Guidelines for fitting vQTL when the data generation process is not known or expected to be approximately normal.

1. Fit ordinary QTL with a homoscedastic linear model and include ordinary QTL with large effects in this preliminary linear model.

2. Do a QQplot of the residuals and estimate the Box-Cox transformation parameter \(\lambda\) for the preliminary linear model.
3. If the QQplot is satisfactory and \( \lambda \) is close to 1 then perform dglm to detect both ordinary QTL and vQTL.

4. If the QQplot is not satisfactory or \( \lambda \) is not close to 1. Use the Box-Cox transformed response in a dglm to detect both QTL and vQTL.

Guidelines for fitting vQTL when the data is expected to have been generated by a non-normal GLM, e.g. binomial, Poisson or gamma.

1. Fit ordinary QTL with a GLM and include QTL with large effects in a preliminary linear model.

2. Perform GLM diagnostics plots, including QQplot of for the standardized deviance residuals, and estimate the EQL parameter \( \psi \).

3. If the diagnostics plots are satisfactory and the estimate of \( \psi \) is close to the assumed variance function for \( \mu^\psi \) then use dglm to detect both ordinary and vQTL.

4. If the diagnostics plots are not satisfactory, or the EQL estimate of \( \psi \) is not close to the expected value, then compare different distributions and link functions prior to the DGLM scan.

The interpretation of vQTL varies with different GLM distributions. Generally, vQTL are effects explaining differences in the dispersion parameter of the GLM, such that for a gamma distribution a vQTL explains differences in CV (coefficient of variation), whereas for a Poisson distribution a vQTL explains differences in the ratio of variance/mean.

LITERATURE CITED

Ansel, J., H. Bottin, C. Rodriguez-Beltran, C. Damon, M. Nagaraj, S. Fehrmann, J. Francois, and G. Yvert, 2008 Cell-to-cell stochastic variation in gene expression is a complex trait. PLoS Genetics 4: e1000049.
Box, G. E. P. and D. R. Cox, 1964 An analysis of transformations (with discussion). Journal of the Royal Statistical Society B 26: 211–252.

Broman, K. W. and S. Sen, 2009 A Guide to QTL Mapping with R/qtl. Springer.

Dunn, P. K. and G. K. Smyth, 2009 dglm: Double generalized linear models. R package version 1.6.1.

Lee, Y., J. A. Nelder, and Y. Pawitan, 2006 Generalized linear models with random effects - unified analysis via h-likelihood. Chapman & Hall/CRC.

McCullagh, P. and J. A. Nelder, 1989 Generalized linear models. Chapman & Hall/CRC.

Nelder, J. A. and D. Pregibon, 1987 An extended quasi-likelihood function. Biometrika 74: 221–232.

Pawitan, Y., 2001 In All Likelihood: statistical modeling and inference using likelihood. Clarendon Press, Oxford, U.K.

Smyth, G. K., 1989 Generalized linear models with varying dispersion. Journal of the Royal Statistical Society B 51: 47–60.

Smyth, G. K., 2002 An efficient algorithm for REML in heteroscedastic regression. Journal of Computational and Graphical Statistics 11: 836–847.

Yang, Y., 2010 The genetics of environmental variation. Ph. D. thesis, Faculty of Agricultural Sciences, Aarhus University.
FILE S2

Supporting Data and Simulation Program

File S2 is available for download as a compressed (.zip) folder at http://www.genetics.org/cgi/content/full/genetics.111.127068/DC1.