

Detecting Major Genetic Loci Controlling Phenotypic Variability in Experimental Crosses

Lars Rönnegård^{*1} 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 biomolecular 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 cana-

lized 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; QUEITSCH *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; HORNSTEIN 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

Supporting information is available online at: <http://www.genetics.org/cgi/content/full/genetics.111.127068/DC1>.

Available freely online through the author-supported open access option.

¹Corresponding author: Statistics Unit, Dalarna University, Röda vägen 3, SE-781 70 Borlänge, Sweden. E-mail: lm@du.se

TABLE 1
Types of variance contributing to between-group differences in phenotypic variability

Variance group ^a	Sources of phenotypic variability			
	Decanalization (epistasis)	Environmental sensitivity	Temporal fluctuation	Measurement error
Genetically distinct individuals with same allele at a vQTL ^b	•	•	•	•
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.

(*e.g.*, inflammatory diseases) or modern admixture, while yet others are simply too recent in evolutionary history for buffering networks to have evolved (*e.g.*, response to HIV).

Increased variability can also be adaptive. In natural populations disruptive selection favors diversity, with increased “capacitance” (RICE 2008) or “bet-hedging” (BEAUMONT *et al.* 2009) spreading risk over a variable fitness landscape. FEINBERG and IRIZARRY (2010) recently proposed a heritable and selectable mechanism for this based on stochastic epigenetic variation. In controlled populations, variability can be increased through directional selection. For example, in a *Drosophila* selection experiment CLAYTON and ROBERTSON (1957) reported increased bristle number variance, which is consistent with the idea that genotypes associated with higher environmental variance have a greater chance of being selected under directional selection (HILL and ZHANG 2004). Moreover, genetic differences have been observed for phenotypic variability in body weight for chickens (ROWE *et al.* 2006) and snails (ROS *et al.* 2004) and litter size in rabbits (IBANEZ-ESCRICHE *et al.* 2008), sheep (SANCRISTOBAL-GAUDY *et al.* 1998), and pigs (SORENSEN and WAAGEPETERSEN 2003).

In natural populations with stabilizing selection we should expect to find alleles minimizing variance for fitness traits (LANDE 1980; HOULE 1992), whereas directional selection during domestication will favor alleles that increase variance. One may therefore expect to find vQTL in experimental crosses between wild and domestic animals (see ANDERSSON 2001). Nonetheless, genetic buffering that leads to phenotypic robustness need not require an evolutionary explanation to be observed, nor to be useful in medicine and agriculture. Plainly, detecting vQTL and inferring how they arose are separate questions; here we concentrate on the first.

Table 1 lists some sources of phenotypic variability in relation to the genetic groups studied. Unless otherwise qualified, we use “phenotypic variance” to describe the observed marginal variance of the phenotype in the population and distinguish “phenotypic variability” as

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; MARTINEZ 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 reconstruction 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

$$y_i = \mu + \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{q}_i^T \boldsymbol{\alpha} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2), \quad (1)$$

where y_i is the phenotype of individual i , \mathbf{x}_i is the i th row in a design matrix of suitable nongenetic covariates, \mathbf{q}_i represents the genetic state at the QTL, ε_i is the residual with variance σ^2 , and μ , $\boldsymbol{\beta}$, and $\boldsymbol{\alpha}$ are parameters estimated by the model. The QTL genotype g_i is typically unknown but, thanks to information from linked markers M , its underlying haplotype pair is available indirectly as a probability distribution $\mathbf{p}_i = (p_{i1}, p_{i2}, p_{i3})$, 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 \mathbf{q}_i in Equation 1 may therefore be formulated in terms of \mathbf{p}_i . For the additive genetic models considered here, $\mathbf{q}_i =$

(q_{iA}, q_{iB}) , where $q_{iA} = 2p_{i1} + p_{i2}$ and $q_{iB} = 2p_{i3} + p_{i2}$ correspond to the expected “doses” of haplotypes A and B respectively, and $\boldsymbol{\alpha}$ 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 $\mathbf{q}_i = q_{iA}$, leading to a scalar effect $\boldsymbol{\alpha} = \alpha$ that estimates the dosage effect of A relative to B. The predictor \mathbf{q}_i may be alternatively formulated to accommodate more general effects, *e.g.*, as $\mathbf{q}_i = \mathbf{p}_i$, 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, F -statistic 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

$$y_i = \mu + \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{q}_i^T \boldsymbol{\alpha} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2 \exp(\mathbf{z}_i^T \boldsymbol{\gamma} + \mathbf{q}_i^T \boldsymbol{\theta})), \quad (2)$$

where \mathbf{z}_i contains nongenetic covariates affecting the residual variance of the model, $\boldsymbol{\gamma}$ is their corresponding effects vector, and $\boldsymbol{\theta}$ 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 $\varepsilon_i \sim N(0, \sigma_i^2)$ and $\log(\sigma_i^2) = \log(\sigma^2) + \mathbf{z}_i^T \boldsymbol{\gamma} + \mathbf{q}_i^T \boldsymbol{\theta}$, 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 *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 available 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 significance 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 underlying the genotypes at each marker. For every interval 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 $h \times h$ matrix \mathbf{D}_i for each individual i , and the expected line dosages are calculated as the h -vector $\mathbf{q}_i = \mathbf{1}^T(\mathbf{D}_i + \mathbf{D}_i^T)$. Because $\mathbf{q}_i^T \mathbf{1} = 2$ always, \mathbf{q}_i has $h - 1$ degrees of freedom and so we typically omit the h th element during model fitting.

Relaxing Assumption 1: More than two founder lines: For experimental crosses with $h > 2$ founder lines, the predictors \mathbf{q}_i , α , and θ in Equation 2 expand to have $h - 1$ elements each.

Relaxing Assumption 2: Uncertain genotype states: Uncertainty about the QTL genotype is most naturally modeled using a mixture distribution. When modeling QTL affecting the mean only, the marginal likelihood for observation i (omitting covariates) is the mixture $L_i = \sum_{j=1}^3 p_{ij} N(r_j \alpha, \sigma^2)$, where $\mathbf{p}_i = (p_{i1}, p_{i2}, p_{i3})$ is defined as above for two founder lines, $N(\cdot)$ is the normal density, and $\mathbf{r} = (-1, 0, 1)$. This is the likelihood used for maximum-likelihood (ML) estimation in interval mapping (LANDER and BOTSTEIN 1989). The regression approach to interval mapping (Equation 1) treats \mathbf{p}_i as if it were an observed outcome, and as a result overestimates the residual variance for each observation by $v_i = \alpha^2(\mathbf{p}_i^T \mathbf{r})^2 - (\mathbf{p}_i^T \mathbf{r})^2$ (XU 1995; XU 1998; FEENSTRA *et al.* 2006). When modeling mean- and variance-controlling QTL on the basis of Equation 2, the marginal likelihood is

$$L_i = \sum_{j=1}^3 p_{ij} N(r_j \alpha, \exp(r_j \theta) \sigma^2). \quad (3)$$

The regression approach for this model overestimates residuals further, by $v_i^* = v_i + \exp(\theta) (\mathbf{p}_i^T \exp(\mathbf{r}) - \exp(\mathbf{p}_i^T \mathbf{r}))$ (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 marginal likelihood gives biased ML estimates for variances, and the regression approximation confers additional flexibility for modeling different distributions 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 technical problem using the dimension reduction approach described in VALDAR *et al.* (2009), whereby the matrix of line dosage estimates from HAPPY is replaced by its informative eigenvectors.

Quantifying the effects of genotype uncertainty: Regressing on expected dosages incorrectly models uncertainty in the predictors. At a mean-controlling QTL, individuals with less certain \mathbf{p}_i will be less well predicted by the mean part of the model, which could lead to inflated estimates of σ_i^2 and, possibly, a false vQTL signal. We investigated this phenomenon empirically by monitoring the relationship between locus uncertainty 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 $0 \log(0) \equiv 0$. If we represent the states of the F_2 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 $P(g_i = j) = J^{-1}$, $\forall j$. Rescaling as $\text{SIC} = I(M, i) / \log(J)$, the (scaled) selective information content at a locus for individual i ranges from 0, denoting equiprobable diplotypes and minimal certainty, to 1, denoting one diplotype with complete certainty.

Relaxing assumption 3: Multiple QTL: Multiple QTL can be fitted by including additional predictors in Equation 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 independent 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, $Q = 1, 2, \text{ and } 4$ for QTL genotype aa, aA, and AA, respectively. These effects correspond to a QTL with a moderate

effect explaining $\sim 2\%$ of the phenotypic variance. For the scenarios with no vQTL, E was drawn from $N(0, \sigma^2)$ with $\sigma^2 = 100$. For the scenarios with vQTL, $\sigma^2 = 100, 125, 156.25$ 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 for 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 $E \sim N(0, 100)$ throughout. For Scenario E1, $Q = 0, 2, \text{ or } 4$ for QTL genotype aa, aA, or AA. For Scenario E2, the additive QTL effects were calculated as $Q = Q_1 + Q_2$, where $Q_k = 0, 2, \text{ or } 4$ 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 4) for both F₂ and CC, indicating that our DGLM approach produces the appropriate rate of false positives (type I error) when applied to known QTL genotypes.

Detection of QTL for F₂ and CC with regression on line dosage: For the F₂ simulations, the 5% chromosome-

wide significance thresholds were $\log P = 2.02$ and $\log P = 2.01$ for the mean and variance parts of the model, respectively. Using these thresholds, the proportion of false positives 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 q_i 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 ^a	
	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

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

^a 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.

TABLE 4
Power to detect QTL at a 5% nominal level for regression on QTL genotypes

Simulated effects	F ₂		CC	
	Ordinary QTL	vQTL	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

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 \times 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 ($\hat{\lambda} = 0.55$, SE 0.11), and the extended quasi-likelihood (EQL) procedure of NELDER and PREGIBON (1987), which indicated a linear mean-variance relationship ($\hat{\psi} = 1.2$, 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 5
Power to detect QTL at a 5% chromosome-wide significance level for regression on line dosages

Simulated effects	F ₂		CC	
	Ordinary QTL	vQTL	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

TABLE 6

Proportion of QTL that were detected at a 5% chromosome-wide significance level and whose chromosomal position was estimated accurately^a

Simulated effects	F ₂		CC	
	Ordinary QTL	vQTL	Ordinary QTL	vQTL
Ordinary QTL	0.462	—	0.452	—
vQTL	—	0.318	—	0.262
Ordinary and vQTL	0.442	0.341	0.372	0.272

^aThe chromosomal position was defined to be accurately estimated: (i) for the F₂ cross if the QTL was detected within the correct marker interval, (ii) for the CC if the estimated position was within the correct ± 0.3 cM.

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-wide 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.*, $100 \times (\exp(|\hat{\theta}|) - 1)$. 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

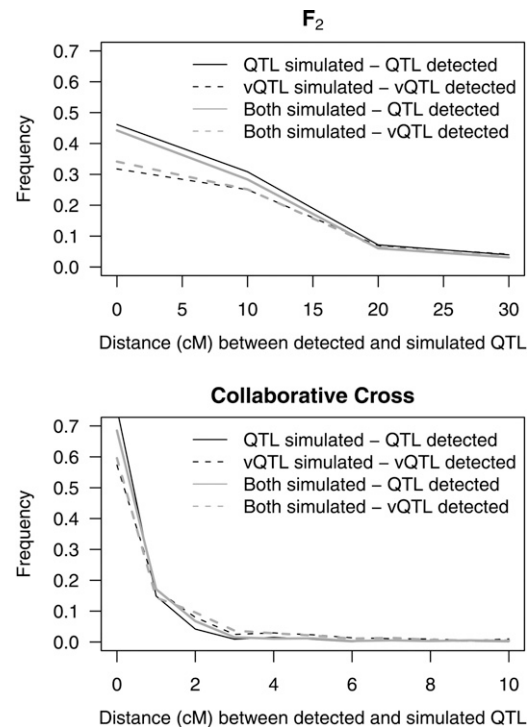


FIGURE 1.—Distance (cM) between simulated and detected QTL for F₂ (10,000 replicates) and the CC (1000 replicates).

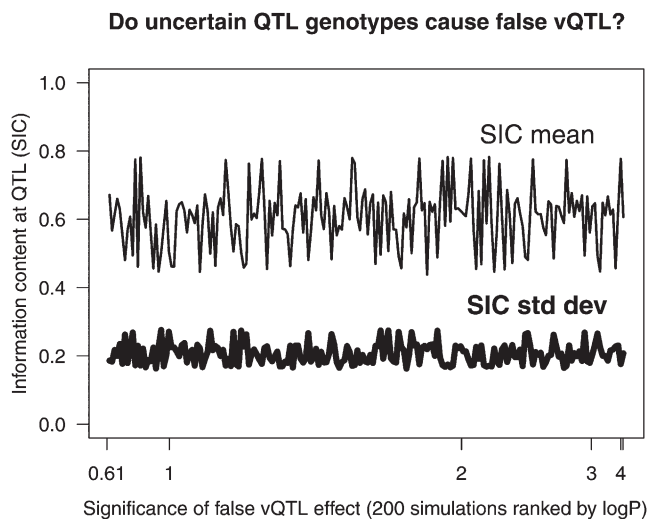


FIGURE 2.—Relationship between log P -values for false vQTL and marker information content (SIC) when simulating a mean-controlling QTL in a CC population. For each of 200 simulations, ordered along the x -axis by their most significant vQTL peak, the plot shows the mean and standard deviation of SIC for 1000 mice. The SIC statistics are stationary, indicating no apparent tendency for marker uncertainty to produce false vQTL signals.

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 line dosages can lead to inflated estimates of the residual variance (see MODELS AND METHODS). In File S1, we describe theoretically the effect of uncertain genotype states on the risk of detecting false vQTL and conclude that detected vQTL can be trusted as long as the marker informativeness is high or, if it is not, that the vQTL is not close to a mean-controlling QTL. Our simulations show that the power of detecting a vQTL in genome scans is largely unaffected by whether the same locus also affects the trait mean (Table 5), and we note that this includes our F_2 simulation with a modest degree of

informativeness from markers spaced 10 cM apart. In the analysis of chicken chromosome 1, we found that a mean-controlling QTL (*Growth 1*) having a very large effect ($\log P = 57.3$) together with moderate marker information (SIC = 0.7) was needed to give any substantial bias in the vQTL estimation (Figure 4). We therefore conclude that the effect of inflated residual variance due to low marker information is not likely to be a major problem for our model.

We apply our model to populations of individuals who are genetically distinct and equally related, phenotyped in studies where genotype assignment is exchangeable across unmodeled environmental variation. A genetic locus whose genotype separates the population into groups with distinct variances is, by our definition, a putative vQTL. Table 1 illustrates how an effect detected in this way could arise from several different sources, each implying a potentially distinct paradigm of vQTL action. For example, it could represent canalizing epistasis, whereby some vQTL genotypes cushion the impact of genetic variation on the phenotype. Our simulations demonstrated that a mean-controlling QTL under epistatic control may indeed be detected as a vQTL with our proposed method, 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

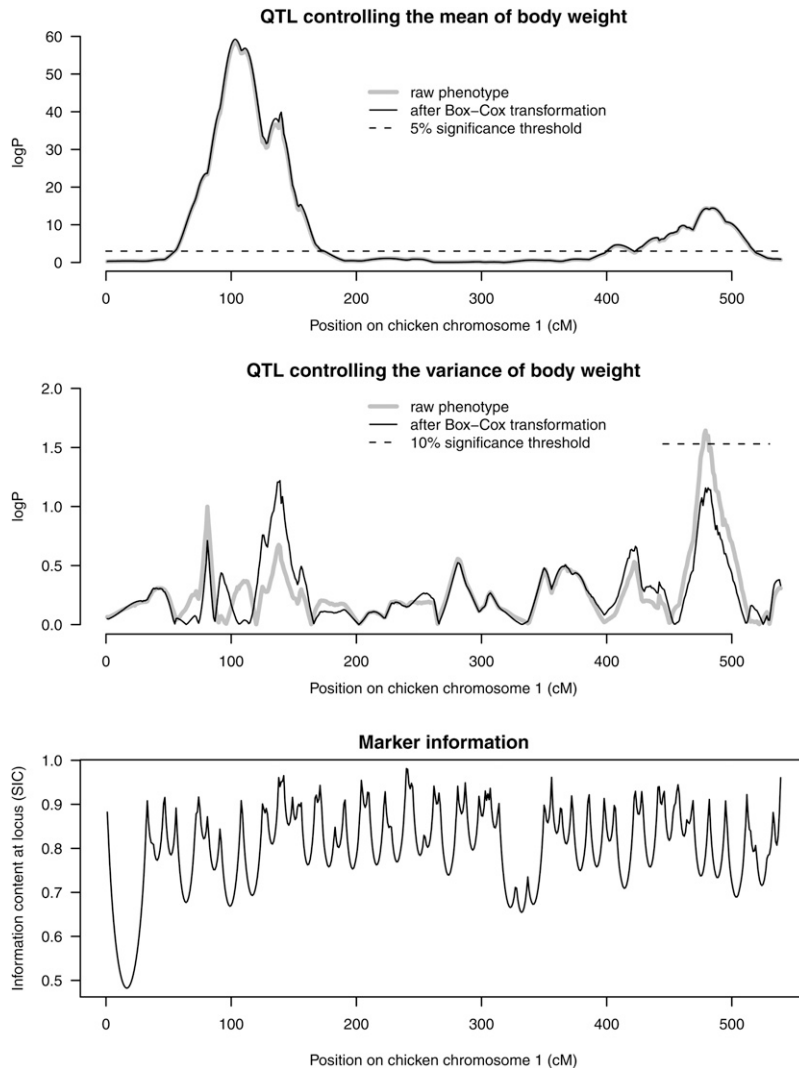


FIGURE 3.—Scan for QTL controlling the mean (top) and the variance (middle) of body weight at 200 days of age on chicken chromosome 1 in an F_2 cross between Red Jungle Fowl and White Leghorn with 756 F_2 offspring. (Bottom) Marker information contents (SIC). Genome-wide significance threshold calculated using 1000 permutations.

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 CV^2 (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

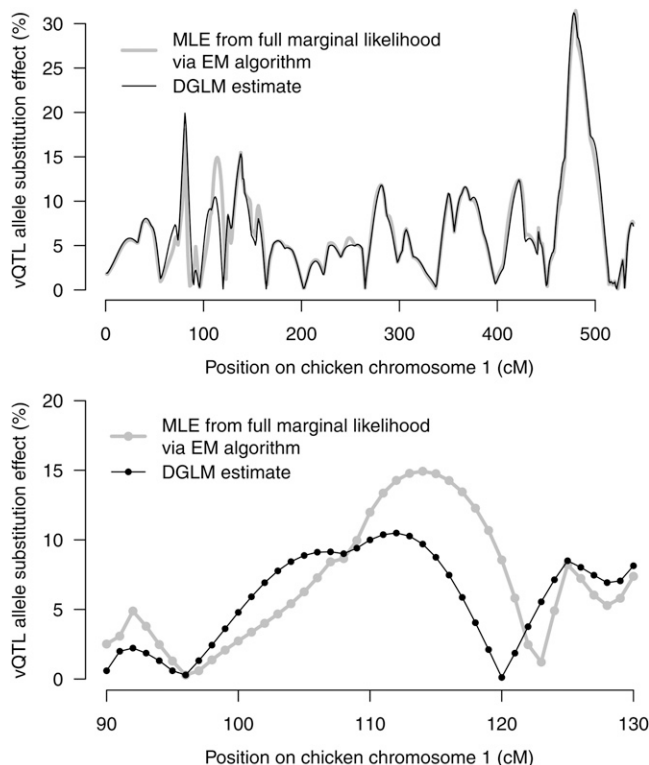


FIGURE 4.—Estimates of vQTL effects given as percentage change in residual variance for allele substitutions for body weight at 200 days of age on chicken chromosome 1 in an F_2 cross between Red Jungle Fowl and White Leghorn. Solid lines: Maximum-likelihood estimates from the full marginal likelihood. Shaded lines: DGLM estimates. Close up for 90–130 cM shown in bottom figure.

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 QTL 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, F_2 , 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 \times 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. NAGARAJ *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. RAINEY, 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.
- BROMAN, K., 2005 The genomes of recombinant inbred lines. *Genetics* **169**: 1133–1146.
- BROMAN, 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.
- CARLBORG, O., and C. S. HALEY, 2004 Epistasis: Too often neglected in complex trait studies? *Nat. Rev. Genet.* **5**: 618–625.
- CAVANAGH, C., M. MORELL, 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., D. C. AIREY, 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, 2009 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. SKOVGAARD and K. W. BROMAN, 2006 Mapping quantitative trait loci by an extension of the Haley–Knott regression method using estimating equations. *Genetics* **173**: 2269–2282.
- 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 variability. *Genet. Res.* **83**: 121–132.
- HOAGLIN, D. C., and R. E. WELSCH, 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.

- IBANEZ-ESCRICHE, N. D., D. SORENSEN, R. WAAGEPETERSEN and A. BLASCO, 2008 Selection for environmental variation: a statistical analysis and power calculations to detect response. *Genetics* **180**: 2209–2226.
- KERJE, S., O. CARLBORG, L. JACOBSSON, K. SCHUTZ, C. HARTMANN *et al.*, 2003 The twofold difference in adult size between the red jungle fowl and White Leghorn chickens is largely explained by a limited number of QTLs. *Anim. Genet.* **34**: 264–274.
- KOVER, P. X., W. VALDAR, J. TRAKALO, N. SCARCELLI, I. M. EHRENREICH *et al.*, 2009 A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* **5**: e1000551.
- LANDE, R., 1980 Genetic variation and phenotypic evolution during allopatric speciation. *Am. Nat.* **116**: 463–479.
- LANDER, E. S., and D. BOTSTEIN, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185–199.
- LEE, Y., and J. A. NELDER, 1996 Hierarchical generalized linear models (with Discussion). *J. R. Stat. Soc. B* **58**: 619–678.
- LEE, Y., and J. A. NELDER, 2006 Double hierarchical generalized linear models (with discussion). *Appl. Stat.* **55**: 139–185.
- LEE, Y., J. A. NELDER and Y. PAWITAN, 2006 *Generalized Linear Models With Random Effects: Unified Analysis via H-Likelihood*. Chapman & Hall/CRC, London.
- LYNCH, M., and B. WALSH, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.
- MACKAY, T. F., and R. F. LYMAN, 2005 *Drosophila* bristles and the nature of quantitative genetic variation. *Phil. Trans. R. Soc. B.* **360**: 1513–1527.
- MARTINEZ, O., and R. N. CURNOW, 1992 Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. *Theor. Appl. Genet.* **85**: 480–488.
- MCCULLAGH, P., and J. A. NELDER, 1989 *Generalized Linear Models*. Chapman & Hall/CRC, London.
- MOTT, R., C. J. TALBOT, M. G. TURRI, A. C. COLLINS and J. FLINT, 2000 A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc. Natl Acad. Sci. USA* **97**: 12649–12654.
- MULDER, H. A., P. BIJMA and W. G. HILL, 2007 Prediction of breeding values and selection response with genetic heterogeneity of environmental variance. *Genetics* **175**: 1895–1910.
- MULDER, H. A., P. BIJMA and W. G. HILL, 2008 Selection for uniformity in livestock production by exploiting genetic heterogeneity of environmental variance. *Genet. Sel. Evol.* **40**: 37–59.
- NELDER, J. A., and D. PREGIBON, 1987 An extended quasi-likelihood function. *Biometrika* **74**: 221–232.
- ORDAS, B., R. A. MALVAR and W. G. HILL, 2008 Genetic variation and quantitative trait loci associated with developmental stability and the environmental correlation between traits in maize. *Genet. Res.* **90**: 385–395.
- PARATI, G., A. FAINI and M. VALENTINI, 2006 Blood pressure variability: its measurement and significance in hypertension. *Curr. Hypertension Rep.* **8**: 199–204.
- PARÉ, G., N. R. COOK, P. M. RIDKER and D. I. CHASMAN, 2010 On the use of variance per genotype as a tool to identify quantitative trait interaction effects: a report from the women's genome health study. *PLoS Genet.* **6**: e1000981.
- QUEITSCH, C., T. SANGSTER and S. LINDQUIST, 2002 Hsp90 as a capacitor of phenotypic variation. *Nature* **417**: 618–624.
- R DEVELOPMENT CORE TEAM, 2009 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- RICE, S. H., 2008 Theoretical approaches to the evolution of development and genetic architecture. *Ann. NY Acad. Sci.* **1133**: 67–86.
- RÖNNEGÅRD, L., F. BESNIER and O. CARLBORG, 2008 An improved method for QTL detection and identification of within-line segregation in F2 intercross designs. *Genetics* **178**: 2315–2326.
- RÖNNEGÅRD, L., M. FELLEKI, F. FIKSE, H. A. MULDER and E. STRANDBERG, 2010a Genetic heterogeneity of residual variance estimation of variance components using double hierarchical generalized linear models. *Genet. Sel. Evol.* **42**: 8.
- 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. SANCRISTOBAL *et al.*, 2004 Evidence for genetic control of adult weight plasticity in the snail *Helix aspersa*. *Genetics* **168**: 2089–2097.
- ROTHWELL, P. M., S. C. HOWARD, E. DOLAN, E. O'BRIEN, J. E. DOBSON *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.
- SANCRISTOBAL-GAUDY, M., J. M. ELSÉN, 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. SALATHIA, S. UNDURRAGA, R. MILO, 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.
- SMYTH, G. K., 1989 Generalized linear models with varying dispersion. *J. R. Stat. Soc. B* **51**: 47–60.
- SMYTH, G. K., 2002 An efficient algorithm for REML in heteroscedastic regression. *J. Comput. Graph. Stat.* **11**: 836–847.
- SOERENSEN, D., and R. WAAGEPETERSEN, 2003 Normal linear models with genetically structured residual variance heterogeneity: a case study. *Genet. Res.* **82**: 207–222.
- STRUCHALIN, M. V., A. DEHGHAN, J. C. WITTEMAN, C. V. DUIJN 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. GAUGUIER, S. BURNETT, 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.
- VERBYLA, 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**: 728–733.
- WADDINGTON, C., 1942 Canalization of development and the inheritance of acquired characters. *Nature* **150**: 563–565.
- WITTENBURG, D., V. GUIARD, 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. *Heredity* **80**: 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. SZATKIEWICZ, 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. ABECASIS and P. SCHEET, 2011 A comparison of approaches to account for uncertainty in analysis of imputed genotypes. *Genet. Epidemiol.* **35**(2): 102–110.

APPENDIX A: DOUBLE GENERALIZED LINEAR MODEL THEORY

For known genotypes at the QTL, the ML estimates of the effect parameters $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, $\boldsymbol{\theta}$ 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 (ε_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 $\log(E(d_i)) = \log(\sigma^2) + \mathbf{z}_i^T \boldsymbol{\gamma} + \mathbf{q}_i^T \boldsymbol{\theta}$, where the response is now the weighted deviance component $d_i = \varepsilon_i^2 / (1 - h_{ii})$. Here, the h_{ii} is the leverage of observation i , equal to the i th diagonal element of the hat matrix $\mathbf{H} = \mathbf{X}_{\text{all}} (\mathbf{X}_{\text{all}}^T \mathbf{V}^{-1} \mathbf{X}_{\text{all}})^{-1} \mathbf{X}_{\text{all}}^T \mathbf{V}^{-1}$, where \mathbf{X}_{all} is the design matrix for all fixed effects in the mean submodel and $\mathbf{V} = \text{diag}((\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2))$.

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

$$L_i = \sum_{j=1}^3 p_{ij} N(\boldsymbol{\mu} + \mathbf{x}_i^T \boldsymbol{\beta} + r_j \boldsymbol{\alpha}, \sigma^2 \exp(\mathbf{z}_i^T \boldsymbol{\gamma} + r_j \boldsymbol{\theta})) = \sum_{j=1}^3 p_{ij} f_{ij}$$

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

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

where $(\partial L_i / \partial \boldsymbol{\Psi}) = \sum_{j=1}^3 p_{ij} (\partial f_{ij} / \partial \boldsymbol{\Psi})$ and

$$\begin{aligned} \frac{\partial L_i}{\partial \boldsymbol{\Psi}} &= \sum_{j=1}^3 p_{ij} \frac{\partial}{\partial \boldsymbol{\Psi}} f_{ij} = \sum_{j=1}^3 p_{ij} f_{ij} \frac{\partial g_{ij}}{\partial \boldsymbol{\Psi}} \\ &\Rightarrow \frac{\partial l_i}{\partial \boldsymbol{\Psi}} = \frac{1}{\sum_{j=1}^3 p_{ij} f_{ij}} \sum_{j=1}^3 p_{ij} f_{ij} \frac{\partial g_{ij}}{\partial \boldsymbol{\Psi}}. \end{aligned}$$

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 $\boldsymbol{\Psi}$ are obtained by iterating the following EM steps until convergence:

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

GENETICS

Supporting Information

<http://www.genetics.org/cgi/content/full/genetics.111.127068/DC1>

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

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`.

```
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.

```
##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:

```
$P.mean
```

[1] 0.2156352

\$P.disp

[1] 0.262968

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 ε_i^2 (SMYTH 2002; LEE *et al.* 2006). In the more general GLM setting, the dispersion term ϕ 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):

$$Var(y) = \phi V(\mu)$$

where $V(\mu)$ is the “variance function”, which gives the relationship between the variance of y and the mean μ . 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, $\phi \equiv \sigma_\varepsilon^2$. For the gamma distribution $V(\mu) = \mu^2$ and since $Var(y)/V(\mu) = \phi$ 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 *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

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

or with \mathbf{x}_i representing an indicator vector for the genotype and $\boldsymbol{\alpha}$ the genotype effects,

$$E(y_i|\mathbf{x}_i) = \mathbf{x}_i^T \boldsymbol{\alpha} \quad \text{and} \quad \text{Var}(y_i|\mathbf{x}_i) = \sigma^2$$

The constant variance model with uncertain genotype When genotype is uncertain, such that $p(x_i = g_j|\text{Data}) = p_{ij}$ and $\mathbf{p}_i = [p_{i1}, p_{i2}, \dots]$, then the sampling distribution conditional on \mathbf{p}_i is the mixture

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

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

$$\begin{aligned} \text{Var}(y_i|\mathbf{p}_i) &= \text{Var}_g(E(y_i|g_j)|\mathbf{p}_i) + E_g(\text{Var}(y_i|g_j)|\mathbf{p}_i) \\ &= [E_g(E(y_i^2|g_j)) - E_g^2(E(y_i|g_j))] + E_g(\text{Var}(y_i|g_j)) \\ &= \left[\sum_j p_{ij} \alpha_{ij}^2 - \left(\sum_j p_{ij} \alpha_{ij} \right)^2 \right] + \sigma^2 \\ &= [\mathbf{p}_i^T (\boldsymbol{\alpha}^2) - (\mathbf{p}_i^T \boldsymbol{\alpha})^2] + \sigma^2 \\ v_i &= v_i^{\text{extra}} + \sigma^2. \end{aligned}$$

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

$$p_{\text{HK}}(y_i|\mathbf{p}_i) = \text{N}(\mathbf{p}_i^T \boldsymbol{\alpha}^2, \sigma^2)$$

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

$$p_{\text{EHK}}(y_i|\mathbf{p}_i) = \text{N}(\mathbf{p}_i^T \boldsymbol{\alpha}^2, v_i)$$

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 α_j and θ_j are specific to genotype g_j . The expectation and variance of y_i are

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

The vQTL model with uncertain genotype. With uncertain genotype, the mean is the same as the constant variance model, $E(y_i|\mathbf{p}_i) = \mathbf{p}_i^T \boldsymbol{\alpha}$, but the variance is

$$\begin{aligned} Var(y_i|\mathbf{p}_i) &= Var_g(E(y_i|g_j)|\mathbf{p}_i) + E_g(Var(y_i|g_j)|\mathbf{p}_i) \\ &= [\mathbf{p}_i^T (\boldsymbol{\alpha}^2) - (\mathbf{p}_i^T \boldsymbol{\alpha})^2] + \mathbf{p}_i^T \exp(\boldsymbol{\theta}) \\ v_i &= v_i^{\text{extra}} + v_i^{\text{vQTL}} \end{aligned}$$

Since we are interested in how \mathbf{x}_i affects the variance, our only measure of \mathbf{x}_i is \mathbf{p}_i and since $v_i^{\text{extra}} = f(\mathbf{p}_i, \boldsymbol{\alpha})$ such that v_i can potentially change with \mathbf{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 *dglm* approach captures only part of the vQTL effect, as $Var(y_i|\mathbf{p}_i) = \exp(\mathbf{p}_i^T \boldsymbol{\theta})$, so that

$$v_i = v_i^{\text{vQTL}} + v_i^{\text{extra}} + (\mathbf{p}_i^T \exp(\boldsymbol{\theta}) - \exp(\mathbf{p}_i^T \boldsymbol{\theta}))$$

A Taylor expansion shows that the last term is small for small to moderate vQTL effects ($\boldsymbol{\theta} \ll 1$). This term does not depend on the effects in the mean model (i.e. $\boldsymbol{\alpha}$). Its contribution to v_i should only reduce the power of the model, whereas the term v_i^{extra} may increase the Type I error.

4. TRANSFORMATION AND ASSESSMENT OF MEAN-VARIANCE RELATIONSHIP

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

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

and the residuals ε 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 ψ for the relationship $Var(y) = \phi\mu^\psi$ with ϕ 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

$$\mathbf{y} \sim N(\mu + \alpha \mathbf{q}, 1)$$

setting $\mu = 1$, $\alpha = 10$ and fixing \mathbf{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 \mathbf{y} we obtained four versions of an “observed phenotype” \mathbf{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 λ 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 σ^2 is constant (i.e. for a model with homoscedastic residual variance). Two models were fitted using *dglm* (DUNN and SMYTH 2009):

$$w_i = \mu + \alpha q_i + \varepsilon_i$$

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

and

$$\frac{w_i^{\hat{\lambda}} - 1}{\hat{\lambda}} = \mu + \alpha q_i + \varepsilon_i$$

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

where μ and α is an intercept term and regression coefficient for the mean model, and m and θ is an intercept term and regression coefficient for the variance model.

Table A1: Estimated Box-Cox parameter λ and estimated regression parameter θ in the variance part of the model (standard errors within brackets)

w	$\hat{\lambda}$	$\hat{\theta}$ with response w	$\hat{\theta}$ with response $\frac{w^{\hat{\lambda}}-1}{\hat{\lambda}}$
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 θ 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 λ was 0.96 with $\hat{\theta} = -0.006$ after Box-Cox transformation. Note that the average estimate of θ 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 θ 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:

$$\mathbf{y} \sim N(\mu + \alpha \mathbf{q}, \log(m + \theta \mathbf{q}_d))$$

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

- Case 1: $\mathbf{q}_d = \mathbf{q}$
- Case 2: \mathbf{q}_d and \mathbf{q} are uncorrelated

The Box-Cox parameter λ was estimated by means of MLE as described above, and two

models were fitted using *dglm*

$$y_i = \mu + \alpha q_i + \varepsilon_i$$

$$\frac{y_i^{\hat{\lambda}} - 1}{\hat{\lambda}} = \mu + \alpha q_i + \varepsilon$$

$$\log(\sigma_i^2) = m + \theta q_{d,i}$$

From the results for Case 1 ($\mathbf{q}_d = \mathbf{q}$) 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 λ and estimated regression parameter θ in the variance part of the model (standard errors within brackets)

Simulated θ	Simulated case	$\hat{\lambda}$	$\hat{\theta}$ for response y	$\hat{\theta}$ for response $\frac{y^{\hat{\lambda}} - 1}{\hat{\lambda}}$
1	$\mathbf{q}_d = \mathbf{q}$	0.65 (0.013)	1.11 (0.047)	0.070 (0.018)
-1	$\mathbf{q}_d = \mathbf{q}$	1.28 (0.018)	-0.89 (0.047)	-0.073 (0.020)
1	$\mathbf{q}_d \neq \mathbf{q}$	0.98 (0.023)	1.03 (0.060)	1.00 (0.065)
-1	$\mathbf{q}_d \neq \mathbf{q}$	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 λ for the preliminary linear model.

3. If the QQplot is satisfactory and λ is close to 1 then perform dglm to detect both ordinary QTL and vQTL.
4. If the QQplot is not satisfactory or λ 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 ψ .
3. If the diagnostics plots are satisfactory and the estimate of ψ is close to the assumed variance function for μ^ψ then use dglm to detect both ordinary and vQTL.
4. If the diagnostics plots are not satisfactory, or the EQL estimate of ψ 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 Agricul-
tural 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>.