Heredity (2002) 88, 2–7

www.nature.com/hdy

The genetic variance for multiple linked quantitative trait loci conditional on marker information in a crossed population

H Matsuda[1] and H Iwaisaki[1,2]
[1]*Course of Environmental Management Science, Niigata University, Niigata 950–2181, Japan;* [2]*Department of Animal Science, Niigata University, Niigata 950–2181, Japan*

In the prediction of genetic values and quantitative trait loci (QTLs) mapping via the mixed model method incorporating marker information in animal populations, it is important to model the genetic variance for individuals with an arbitrary pedigree structure. In this study, for a crossed population originated from different genetic groups such as breeds or outbred strains, the variance of additive genetic values for multiple linked QTLs that are contained in a chromosome segment, especially the segregation variance, is investigated assuming the use of marker data. The variance for a finite number of QTLs in one chromosomal segment is first examined for the crossed population with the general pedigree. Then, applying the concept of the expectation of identity-by-descent proportion, an approximation to the mean of the conditional probabilities for the linked QTLs over all loci is obtained, and using it an expression for the variance in the case of an infinite number of linked QTLs marked by flanking markers is derived. It appears that the approach presented can be useful in the segment mapping using, and in the genetic evaluation of, crosses with general pedigrees in the population of concern. The calculation of the segregation variance through the current approach is illustrated numerically, using a small data-set.
Heredity (2002) **88,** 2–7. DOI: 10.1038/sj/hdy/6800011

Keywords: breeding value; crossbred population; IBDP; linked QTLs; marker information; segregation variance

Introduction

Best linear unbiased prediction (BLUP) of breeding values using the mixed model methodology of Henderson (1984) is the predominant approach for genetic evaluation of animals. Evaluation by BLUP requires that genetic means, variances, and covariances are properly modeled. Where only trait phenotype information is used, genetic group theory under additive inheritance was developed to account for differences in genetic means among genetic groups or populations (eg Thompson, 1979). Elzo (1990) further developed a theory to take account of heterogeneity of variances among genetic groups, and later the theory was extended to account for additive variance due to segregation of alleles among populations with different gene frequencies (Lo *et al*, 1993).

For situations where trait phenotype and DNA marker information is available in outbred populations, there are also BLUP methods developed for genetic evaluation (eg, Fernando and Grossman, 1989; Goddard, 1992; Hoeschele, 1993; van Arendonk *et al*, 1994; Saito and Iwaisaki, 1996) and mixed model approaches for quantitative trait loci (QTLs) interval mapping (eg, Grignola *et al*, 1996, 1997; van Arendonk *et al*, 1998; Saito and Iwaisaki, 2000).

Correspondence: H Iwaisaki, Department of Animal Science, Faculty of Agriculture, Niigata University, Igarashi nino-cho 8050, Niigata-shi 950–2181, Japan. E-mail: iwsk@agr.niigata-u.ac.jp
Received 29 December 2000; accepted 1 October 2001

In an outbred population in which there is linkage equilibrium between markers and linked QTLs at the population level the overall genetic mean is not dependent on marker information. The information on marker allele transmission within family is used in genetic evaluation by BLUP. In contrast, in the cases of crossbred populations originated from several genetic groups such as breeds or outbred strains, where linkage disequilibrium between markers and QTLs exists, all the genetic means, variances and covariances depend on marker information. A mixed model for such situations was presented for a single marked QTL, in which a grouping strategy that can take into account crossbreeding and the linkage disequilibrium is included (Goddard, 1992). However, Goddard's model does not properly take account of heterogeneous variances among genetic groups and of the segregation variance. Then, for the similar situation, an alternative approach to model genetic means and (co)variances for genetic evaluation was proposed (Wang *et al*, 1998), combining the covariance theory for a marked QTL (Wang *et al*, 1995) and the covariance theory for a multi-breed population of Lo *et al* (1993).

Very recently, investigating the case where tightly linked multiple QTLs or a cluster of QTLs are contained in a chromosome segment, a mixed model approach was proposed for QTL analysis in F_2 crosses between outbred lines that allows for QTL segregation within lines as well as for differences in mean QTL effects between lines (Perez-Enciso and Varona, 2000). This model was presented as a generalization of the approach by Wang

et al (1998), in which the variance for multiple linked QTLs under the use of marker information was derived. In this paper, we work with a crossed population with the general pedigree, not with F_2 crosses only, derived from crossing between two outbred genetic groups such as breeds and derive an expression for the genetic variance due to a chromosome segment in which linked QTLs are contained, especially for the segregation variance, conditional on flanking marker information.

Theory and methods

In this paper, we consider a crossed population derived from two breeds, A and B, within each of which there exists some genetic variation. We assume multiple linked QTLs in a chromosome segment, linkage equilibrium in A and B and additive gene action within and between breeds. Here, for convenience, we use similar notation to Perez-Enciso and Varona (2000). For use of flanking marker information, it is supposed that marker linkage phases in founders and the parental origins of marker haplotypes in non-founders are known.

Assuming that all alleles of all loci have equal effects, the distributions of the effects of A and B origin alleles at the kth locus in individual i ($g^h_{Ai,k}$ and $g^h_{Bi,k}$) are written as

$$g^h_{Ai,k} \sim N(\mu_k + \Delta_k/2, \sigma^2_{Ak})$$

and

$$g^h_{Bi,k} \sim N(\mu_k - \Delta_k/2, \sigma^2_{Bk})$$

with

$$\mu_k = \mu/(2q),\ \Delta_k = \Delta/(2q),\ \sigma^2_{Ak} = \sigma^2_A/(2q) \text{ and } \sigma^2_{Bk} = \sigma^2_B/(2q),\ \forall\ h,i,k,$$

where h denote the gametes of paternal and maternal origin (0 and 1), μ is the expected value of the additive effect of linked QTLs in a chromosome segment, q is the number of linked QTLs within the segment, assumed finite, Δ is the difference in the additive effect for linked QTLs in the segment between two breeds, and σ^2_A and σ^2_B are the additive variances for the segment in two breeds. Then, for a cross derived from A and B, the variance for the additive effects due to linked QTLs (g^h_i) for one of the paternal and maternal gametes is given as

$$Var(g^h_i) = Var\left(\sum_{k=1}^{q} g^h_{i,k}\right) = \sum_{k=1}^{q}\sum_{k'=1}^{q} Cov(g^h_{i,k}, g^h_{i,k'}). \tag{1}$$

Following Lo *et al* (1993) and defining $w_{kk'}$ that takes values AA, AB, BA, and BB according to breed origin of each allele at two loci k and k', Perez-Enciso and Varona (2000) obtained:

$$Var(g^h_i) = \sum_{k=1}^{q}\sum_{k'=1}^{q} \{\underset{w}{E}[Cov(g^h_{i,k}, g^h_{i,k'} \mid w_{kk'})] + \underset{w}{Cov}[E(g^h_{i,k} \mid w_{kk'}), E(g^h_{i,k'} \mid w_{kk'})]\}, \tag{2}$$

where E_w represents the expectation for all the combinations of $w_{kk'}$. Note that $E_w[\mathrm{Cov}(g^h_{i,k}, g^h_{i,k'} \mid w_{kk'})]$ is zero, if $k \neq k'$ because we have assumed linkage equilibrium within pure breeds or if $w_{kk'}$ = AB or BA. As a result, we have:

$$\sum_{k=1}^{q}\sum_{k'=1}^{q} \underset{w}{E}[Cov(g^h_{i,k}, g^h_{i,k'} \mid w_{kk'})] = p^h_i\,(\sigma^2_A/2) + (1 - p^h_i)(\sigma^2_B/2), \tag{3}$$

where p^h_i represents the expectation of the fraction of the segment originating from breed A in gamete h for individual i, given breed and marker information.

The second term in equation (2) represents the segregation variance for linked QTLs. Let the probability that loci k and k' are of breed origin b and b', respectively, be $p^h_{i,bb'}$ where $b,b' \in \{A,B\}$. Expressing the expectations of the additive genetic effects in the cases where the locus k is of breed origin A and B by $a_k = \mu_k + \Delta_k/2$ and $\beta_k = \mu_k - \Delta_k/2$, respectively, the second term of equation (2) is written as:

$$\begin{aligned}
&\underset{w}{Cov}[E(g^h_{i,k} \mid w_{kk'}), E(g^h_{i,k'} \mid w_{kk'})] \\
&= \underset{w}{E}[E(g^h_{i,k} \mid w_{kk'})\, E(g^h_{i,k'} \mid w_{kk'})] \\
&- \underset{w}{E}[E(g^h_{i,k} \mid w_{kk'})]\ \underset{w}{E}[E(g^h_{i,k'} \mid w_{kk'})] = p^h_{i,AA}\alpha_k\alpha_{k'} \\
&+ p^h_{i,AB}\alpha_k\beta_{k'} + p^h_{i,BA}\beta_k\alpha_{k'} + p^h_{i,BB}\beta_k\beta_{k'} \\
&- [(p^h_{i,AA} + p^h_{i,AB})\alpha_k + (p^h_{i,BA} + p^h_{i,BB})\,\beta_k] \\
&[(p^h_{i,AA} + p^h_{i,BA})\alpha_{k'} + (p^h_{i,AB} + p^h_{i,BB})\beta_{k'}].
\end{aligned} \tag{4}$$

Then, with some arrangement and use of $p^h_{i,AA}(1 - p^h_{i,AA}) = p^h_{i,AA}\,(p^h_{i,AB} + p^h_{i,BA} + p^h_{i,BB})$ and so on, equation (4) can be written as

$$(p^h_{i,AA}p^h_{i,BB} - p^h_{i,AB}p^h_{i,BA})\Delta^2_k.$$

Thus, the additive variance can finally be expressed as

$$Var(g^h_i) = p^h_i(\sigma^2_A/2) + (1 - p^h_i)(\sigma^2_B/2) + \sum_{k=1}^{q}\sum_{k'=1}^{q} (p^h_{i,AA}\,p^h_{i,BB} - p^h_{i,AB}\,p^h_{i,BA})\Delta^2_k, \tag{5}$$

In the case of the model with an infinite number of linked loci considered, assuming that loci are distributed uniformly along the segment, p^h_i can be represented by the expectation of identity-by-descent proportion (IBDP) between the segment of gamete h for individual i and that of an ancestral gamete of breed A. Thus, following the approach of Matsuda and Iwaisaki (1998, 2000) for an outbred population, that uses the concept of IBDP of Guo (1994a,b, 1995) and in which there is assumed to be linkage equilibrium between markers and QTLs, we here consider the current situation where there exist linkage disequilibrium between breeds.

Following Guo's work, let the chromosome region where a cluster of linked QTLs is contained be $0 \leq t \leq l$, and two flanking markers are located at both edges of the region. The length of the region from the origin at which the nth crossover occurred is denoted by S_n. Since, for the two consecutive regions between an even-numbered and the next (odd-numbered) points and between the latter and the next (even-numbered) points where crossover occurred, the offspring receives two homologous chromosomes (paternal and maternal) of its one parent reversely, one chromosome of the offspring can be represented by a time-continuous, two-state Markov chain $c(t)$ at any point t along the segment that is referred to as the gametogenesis process, as follows

$$c(t) = \begin{cases} C & S_{2n} \le t < S_{2n+1} \le l \\ 1 - C & S_{2n+1} \le t < S_{2n+2} \le l, \end{cases} \tag{6}$$

where C is a random variable and takes values 0 and 1 with equal probability, and the time parameter t is the map distance along the chromosome. Since all meiosis are independent, all the gametogenesis processes in a pedigree are also independent and stochastically identical. So, using Haldane's (1919) mapping function, the transition probability matrix for $c(t)$ is given as

$$P(t) = (r_{ij}(t)) = \frac{1}{2}\begin{pmatrix} 1 + e^{-2t} & 1 - e^{-2t} \\ 1 - e^{-2t} & 1 + e^{-2t} \end{pmatrix}. \tag{7}$$

Then, in order to assess the relationships between gametes of base animals of a particular genetic group A and gamete h for a given animal i in a succeeding generation, the m relevant gametogenesis processes $v(t) = (c_1(t), c_2(t), \ldots, c_m(t))$ are considered as the possible pathways of gene transmission from the base animals to gamete h. Given the m processes lying between two gametes, the statement for the joint gametogenesis process that is the possible pathway of gene transmission is common to all the t, and therefore this joint gametogenesis process is a random vector constituting a random walk on an m-dimensional hypercube $Z^m = \{(\eta_1, \eta_2, \ldots, \eta_m): \eta_i = 0 \text{ or } 1\}$. Thus, defining the set of IBD states representing that the gamete h originated from genetic group A by D_A as the collection of vertices on Z^m and denoting the information on breed origin at two flanking marker loci by $v(0) = v_0$ and $v(l) = v_l$, p_i^h, in equations (3) and (5) is obtained as

$$\begin{aligned} p_i^h &= \frac{1}{l}\int_0^l P[v(t) \in D_A \mid v(0) = v_0, v(l) = v_l]dt \\ &= \frac{1}{l}\sum_{v_x \in D_A}\int_0^l P[v(t) = v_x \mid v(0) = v_0, v(l) = v_l]dt \\ &= \frac{1}{l[r_{00}(l)]^{m-|v_l - v_0|}\,[r_{01}(l)]^{|v_l - v_0|}} \times \sum_{v_x \in D_A}\int_0^l P[v(t) \\ &= v_x \mid v(0) = v_0]P[v(l) = v_l | v(t) = v_x]dt \\ &= \frac{1}{l[r_{00}(l)]^{m-|v_l - v_0|}\,[r_{01}(l)]^{|v_l - v_0|}} \\ &\times \sum_{v_x \in D_A}\int_0^l \{[r_{00}(t)]^{m-|v_x - v_0|}\,[r_{01}(t)]^{|v_x - v_0|}\,[r_{00}(l \\ &- t)]^{m-|v_x - v_l|}\,[r_{01}(l - t)]^{|v_x - v_l|}\}dt, \end{aligned} \tag{8}$$

where representing two joint gametogenesis processes by $\eta_j = (\eta_{j1}, \ldots, \eta_{ji}, \ldots, \eta_{jm})$ and $\eta_{j'} = (\eta_{j'1}, \ldots, \eta_{j'i}, \ldots, \eta_{j'm})$, $|\bullet|$ stands for $\sum|\eta_{ji} - \eta_{j'i}|$, and denoting the cardinality of D_A by $|D_A|$, events $\{(v(t) = v_x)\}$ for $x = 1, 2, \cdots, |D_A|$ are mutually exclusive. So, for example, the number of joint gametogenesis processes in which no crossover occurred between two flanking markers, with the probability of $r_{00}(l)$, is represented by $m - |v_l - v_0|$. Now, the information on breed composition for founders and parentage for non-founders is referred to as breed information (denoted as $\mathbf{B}$). Then, given $\mathbf{B}$ and flanking marker information (denoted as $\mathbf{M}_F$), the second term in equation (5) can be approximated as

$$\begin{aligned} &\sum_{k=1}^{q}\sum_{k'=1}^{q}(p_{i,AA}^h p_{i,BB}^h \\ &- p_{i,AB}^h\, p_{i,BA}^h)\Delta_k^2 \approx [E(R_{i,AA,BB}^h | \mathbf{BM}_F) \\ &- E(R_{i,AB,BA}^h \mid \mathbf{BM}_F)]\delta^2, \end{aligned} \tag{9}$$

where $\delta^2 = \Delta^2/4$, and $R_{i,AA,BB}^h$ ($R_{i,AB,BA}^h$) is the mean of $p_{i,AA}^h p_{i,BB}^h$ ($p_{i,AB}^h p_{i,BA}^h$) for the whole segment, whose conditional expectation is given as

$$\begin{aligned} &E(R_{i,AA,BB}^h \mid \mathbf{BM}_F) \\ &= \frac{1}{l^2}\int_0^l\int_0^s P[v(t) \in D_{AB}, v(s) \in D_{AB} \mid v(0) = v_0, v(l) \\ &= v_l]dtds = \frac{1}{l^2}\sum_{v_x \in D_{AB}}\sum_{v_y \in D_{AB}}\int_0^l\int_0^s P[v(t) = v_x, v(s) \\ &= v_y \mid v(0) = v_0, v(l) = v_l]dtds \\ &= \frac{1}{l^2[r_{00}(l)]^{m-|v_l - v_0|}\,[r_{01}(l)]^{|v_l - v_0|}} \\ &\times \sum_{v_x \in D_{AB}}\sum_{v_y \in D_{AB}}\int_0^l\int_0^s \{[r_{00}(t)]^{m-|v_x - v_0|}\,[r_{01}(t)]^{|v_x - v_0|} \\ &[r_{00}(s - t)]^{m-|v_x - v_y|}\,[r_{01}(s - t)]^{|v_x - v_y|}\,[r_{00}(l - s)]^{m-|v_y - v_l|} \\ &[r_{01}(l - s)]^{|v_y - v_l|}\}dtds. \end{aligned} \tag{10}$$

which can be obtained using the appropriate IBD sets, D_{AB} and D_{AB}, according to the situations of breed origin at two different loci and denoting two different breeds by b and b'. Note that in the case of $E(R_{i,AB,BA}^h | \mathbf{BM}_F)$, the IBD sets become D_{AB} and D_{BA}.

These expressions enables us to compute the variance for an infinite number of linked QTLs under use of flanking marker data for individuals in general percentages of breed origin A and B. If one of the two flanking markers is uninformative, then transmission can be followed for another informative marker and the variance can be obtained using the information on the informative one with $v(0) = v_0$ or $v(l) = v_l$ (denoted as $\mathbf{M}_s$).

Illustration

In this section, using a crossbred population derived from two breeds A and B as given in Figure 1, we numerically investigate the value of segregation variance for the gamete, originating from individual 3, for individual 4. We compare the three cases of no marker, single marker and flanking marker information available. It is assumed that $\delta^2 = 1.0$. The length of the chromosome segment containing multiple linked QTLs (l) are altered within the range of 0.1–0.5 M. For the gamete of concern, the cases considered were: the case where the allele M1 is transmitted, when information on a single marker is used, and the cases where the non-recombinant haplotype M1N1 or the recombinant haplotype M1N4 is transmitted, when data on flanking marker loci are used.

With $l = 0.1$, the segregation variance for the gamete of concern in the case of no marker information is

$$[E(R_{4,AA,BB}^0 \mid \mathbf{B}) - E(R_{4,AB,BA}^0 \mid \mathbf{B})]\ \delta^2$$

with

$$E(R_{4,AA,BB}^0 \mid \mathbf{B}) = \frac{1}{0.1^2}\,0.25\int_0^{0.1}\int_0^s \{r_{00}^2\,(s - t)\}dtds$$

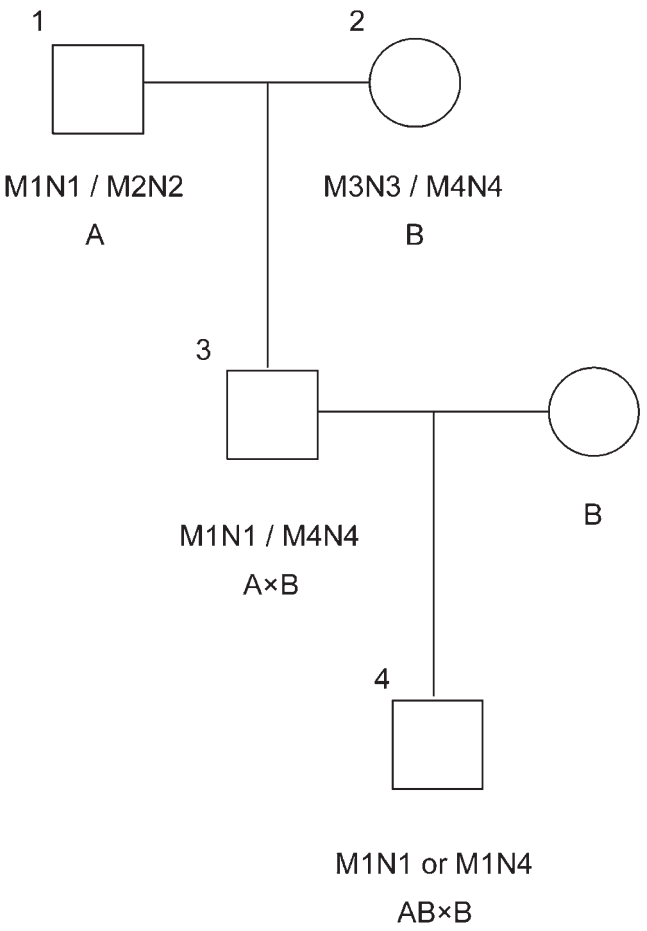

Figure 1 Pedigree plot with breed and marker data for a crossbred population, in which A and B denote two breeds and M and N show flanking markers, representing alleles by figures.

$$= 0.1340$$

and

$$E(R^0_{4,AB,BA} \mid \mathbf{B}) = \frac{1}{0.1^2}\, 0.25 \int_0^{0.1}\int_0^{s} \{r_{01}^2\,(s-t)\}dtds$$
$$= 0.0002,$$

giving the value of 0.1338.

In contrast, the corresponding expression in the case with use of marker information is yielded as

$$[E(R^0_{4,AA,BB} \mid \mathbf{BM}^*) - \mathrm{E}(R^0_{4,AB,BA} \mid \mathbf{BM}^*)]\ \delta^2,$$

where $\mathbf{M}^*$ is $\mathbf{M}_s$ or $\mathbf{M}_F$. For the case of a single marker, we have

$$E(R^0_{4,AA,BB} \mid \mathbf{BM}_S) = \frac{1}{0.1^2}\int_0^{0.1}\int_0^{s} \{r_{00}(t)r_{01}(t)r_{00}^2$$
$$(s-t)\}dtds = 0.0312$$

and

$$E(R^0_{4,AB,BA} \mid \mathbf{BM}_S) = \frac{1}{0.1^2}\int_0^{0.1}\int_0^{s} \{r_{00}(t)r_{01}(t)r_{01}^2$$
$$(s-t)\}dtds = 0.0001,$$

and therefore the segregation variance is given as 0.0311.

Table 1 Comparison of the values of segregation variance for the four cases

Case	*Segment length (l)*				
	0.1 M	*0.2 M*	*0.3 M*	*0.4 M*	*0.5 M*
No marker	0.1338	0.1435	0.1542	0.1662	0.1796
Single marker	0.0311	0.0586	0.0836	0.1070	0.1296
Flanking marker					
non-recombinant	0.0021	0.0083	0.0183	0.0321	0.0492
recombinant	0.2092	0.2118	0.2162	0.2223	0.2306

For the case where flanking marker information is used, the value becomes 0.0021 with

$$E(R^0_{4,AA,BB} \mid \mathbf{BM}_F) = \frac{1}{0.1^2 r_{00}^2(0.1)}\int_0^{0.1}\int_0^{s} \{r_{00}(t)r_{01}(t)r_{00}^2$$
$$(s-t)\ r_{00}\ (0.1-s)r_{01}\ (0.1-s)\}dtds = 0.0021$$

and

$$E(R^0_{4,AB,BA} \mid \mathbf{BM}_F) = \frac{1}{0.1^2 r_{00}^2(0.1)}\int_0^{0.1}\int_0^{s} \{r_{00}(t)r_{01}(t)r_{01}^2$$
$$(s-t)\ r_{00}\ (0.1-s)r_{01}\ (0.1-s)\}dtds = 0.0000.$$

Table 1 presents all the values of the segregation variance for the considered gamete for the changed values of l that are also depicted in Figure 2. It is revealed that the values for the cases of non-recombinant and recombinant types are reversely and extremely different from those for the case with a single marker used. It is also shown that as l becomes large, the value resulting from use of a single marker becomes considerably near to the corresponding value in the case with no marker information, while the values in the case with use of flanking markers of non-recombinant type essentially remain low. These deviations in patterns from the case of no marker data clearly indicate that use of flanking marker data substan-

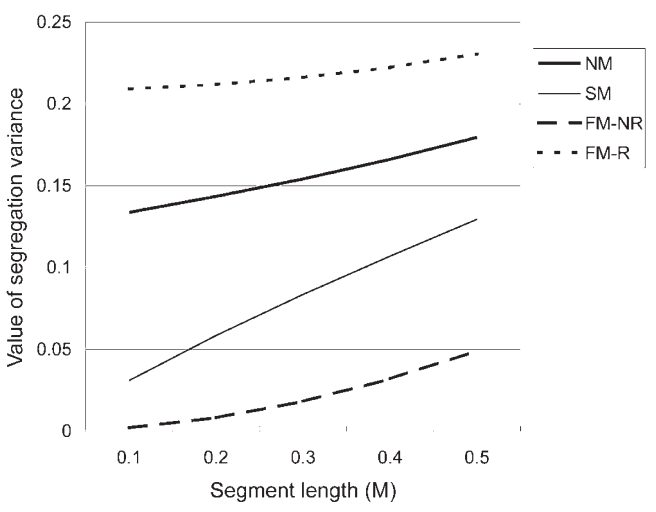

Figure 2 Graphical representation of the values of segregation variance, as listed in Table 1, for the selected lengths of the segment. NM, SM, FM-NR, and FM-R denote the cases with use of no marker, single marker, flanking marker (non-recombinant type), and flanking marker (recombinant type), respectively.

tially increases the information on the segregation variance, as expected.

In a crossed population between breeds or outbred lines, marker information may be partially informative. Therefore, while it is obvious that two informative flanking markers provide more information to account for the segregation variance than a single marker, when one of the two markers is uninformative, only the information on the informative marker can be used. In such situations, however, use of different informative flanking markers out of multiple linked markers on the chromosome would be a reasonable approach.

Discussion

In this paper, we first considered the situation in which a finite number of linked QTLs are located on a particular chromosome region and then presented an expression for the variance due to the QTLs in a cross originated from two breeds. Using the concept of IBDP given by Guo (1994a,b, 1995), we further investigated the variance in the mean of the conditional probabilities for an infinite number of linked QTLs in a chromosome segment. The variance for the linked QTLs conditional on flanking marker information was modeled. The results obtained can straightforwardly be used in genetic evaluation by BLUP using marker and trait information (Matsuda and Iwaisaki, 2001a). Moreover, in the segment mapping of Perez-Enciso and Varona (2000), an approach presented here could makes it possible to utilize information not only on F_2 crosses but also on any crossbred individual in the population under study, taking into account the segregation variance appropriately. Also in the segment mapping, as indicated by Grignola *et al* (1997) for the usual interval mapping, it would be possible to detect a plural of clusters of QTLs by conducting the intersection-union test. The effects for QTLs positioned outside the segment or the marker interval can be handled as the effects of the remaining chromosome, as in equation (5) of Perez-Enciso and Varona (2000).

We have allowed the use of flanking markers. Indeed, it is known that in QTL mapping utilizing flanking marker information is more effective than use of only single marker information (eg, Mackinnon *et al*, 1996). As demonstrated in the numerical investigation, data on flanking marker loci are also informative in taking account of the segregation variance, relative to those on a single marker locus.

Wang *et al* (1998) presented the theory to properly model means and (co)variances in a multi-breed population, given single marker information, in the presence of gametic disequilibrium between the marker locus and a linked QTL. Compared with the segregation variance in the single QTL situation modeled by them, we note that the current expression for the segregation variance for multiple linked QTLs is of a different form, as would be expected. Our expression for the segregation variance requires the assumption that multiple linked QTLs are in coupling phase or that the differences in the expectations for linked QTLs between two breeds are in the same sign. The variance component due to segregation depends on the differences. Therefore, it is expected that the segregation variance becomes highest and lowest with the coupling and repulsion phases, respectively. When the QTLs are not in coupling phase, the relationship between twice the segregation variance δ^2 defined by Lande (1981) and Lo *et al* (1993) and the difference in the expectation Δ is no longer represented by

$$\delta^2 = \sum_{k=1}^{q} \sum_{k'=1}^{q} [\Delta/(2nl)]^2.$$

Previously, the variation in genetic composition for a finite number of multiple linked genes or a chromosome segment was theoretically investigated by eg, Hill (1993) and Visscher (1996). The theory of this kind was applied to determine whether the detected entity in QTL mapping using inbred lines is a single QTL of relatively large effect or a cluster of multiple QTLs with smaller effects (Visscher and Haley, 1996) and to investigate the power of a chromosome test for detecting the genetic variation on a single chromosome (Visscher and Haley, 1998). Thus, for the crossbred population with general pedigree derived from pure original breeds, the current approach would also be useful to provide information on determining if a genomic region under study contains a single QTL or multiple linked QTLs.

In this study, we assumed that parental linkage phases of flanking markers are known. When phase information is unknown, one way would be to compute the expectation of IBDP for any possible phase, and then to add up the expected IBDPs weighted by the corresponding phase probabilities. Generally, however, such calculations would be tremendous, since it is likely that there are a number of possible linkage phases and parental origins. Hence, Markov chain Monte Carlo approximation would be a method of choice, as used by Grignola *et al* (1996), Perez-Enciso and Varona (2000) and Perez-Enciso *et al* (2000). This approach requires one set of assumed linkage phases for markers in a process of the sampling. Therefore, it seems that use of the current approach assuming known marker linkage phases in each Gibbs iteration could be useful in efficiently computing the mean of the conditional probabilities required. Finally, for populations with complex pedigrees of current type, it would be not easy to construct the IBD sets necessitated in equations (8) and (10). On this point, we comment that the application of a computing procedure to systematically calculate the genetic covariance matrix whose elements are the required (co)variances and its inverse by recursively constructing the IBD sets (Matsuda and Iwaisaki, 2001b) can be useful.

References

Elzo MA (1990). Recursive procedures to compute the inverse of the multiple trait additive genetic covariance matrix in inbred and noninbred multibreed population. *J Anim Sci* **68**: 1215–1228.

Fernando RL, Grossman M (1989). Marker assisted selection using best linear unbiased prediction. *Genet Sel Evol* **21**: 467–477.

Goddard ME (1992). A mixed model for analyses of data on multiple genetic markers. *Theor Appl Genet* **83**: 878–886.

Grignola FE, Hoeschele I, Tier B (1996). Mapping quantitative trait loci in outcross populations via residual maximum likelihood. I. Methodology. *Genet Sel Evol* **28**: 479–490.

Grignola FE, Zhang Q, Hoeschele I (1997). Mapping linked quantitative trait loci via residual maximum likelihood. *Genet Sel Evol* **29**: 529–544.

Guo SW (1994a). Computation of identity-by-descent proportions shared by two siblings. *Am J Hum Genet* **54**: 1104–1109.

Guo SW (1994b). Proportion of genes survived in offspring conditional on inheritance of flanking markers. *Genetics* **138**: 953–962.
Guo SW (1995). Proportion of genome shared identical by descent by relatives: concept, computation, and applications. *Am J Hum Genet* **56**: 1468–1476.
Haldane JBS (1919). The combination of linkage values and the calculation of distance between the loci of linked factors. *J Genet* **8**: 299–309.
Henderson CR (1984). *Applications of Linear Models in Animal Breeding*, 1st edn. University of Guelph Press: Guelph, Ont.
Hill WG (1993). Variation in genetic composition in backcross programs. *J Hered* **84**: 212–213.
Hoeschele I (1993). Elimination of quantitative trait loci equations in an animal model incorporating genetic marker data. *J Dairy Sci* **76**: 1693–1713.
Lande R (1981). The minimum number of genes contributing to quantitative variation between and within populations. *Genetics* **99**: 541–553.
Lo LL, Fernando RL, Grossman M (1993). Covariance between relatives in multibreed populations: additive model. *Theor Appl Genet* **87**: 423–430.
Mackinnon MJ, van der Beek S, Kinghorn BP (1996). Use of deterministic sampling for exploring likelihoods in linkage analysis for quantitative traits. *Theor Appl Genet* **92**: 130–139.
Matsuda H, Iwaisaki H (1998). A mixed linear model for best linear unbiased prediction of additive and dominance effects using information of markers flanking a cluster of quantitative trait loci. In: Proceedings of the 8th World Conference on Animal Production, vol. II, Seoul National University, Seoul. pp 594–595.
Matsuda H, Iwaisaki H (2000). Best linear unbiased prediction of QTLs-cluster effects using flanking and upstream marker information in outbred populations. *Jpn J Biometrics* **21**: 39–49.
Matsuda H, Iwaisaki H (2001a). A mixed model method to predict QTL-cluster effects using trait and marker information in a multi-group population. *Genes Genet Syst* **76**: 81–88.
Matsuda H, Iwaisaki H (2001b). Calculating gametic relationship matrix and its inverse in the model containing effects of QTL-cluster marked by flanking markers. *Biometrical J* submitted.
Perez-Enciso M, Varona L (2000). Quantitative trait loci mapping in F_2 crosses between outbred lines. *Genetics* **155**: 391–405.
Perez-Enciso M, Varona L, Rothschild MF (2000). Computation of identity by descent probabilities conditional on DNA markers via Monte Carlo Markov chain method. *Genet Sel Evol* **32**: 467–482.
Saito S, Iwaisaki H (1996). A reduced animal model with elimination of quantitative trait loci equations for marker-assisted selection. *Genet Sel Evol* **28**: 465–477.
Saito S, Iwaisaki H (2000). A procedure for QTL interval mapping and variance component estimation using REML. In: Proceedings of the 9th AAAP Congress, vol. C, p. 260. University of New South Wales, Sydney.
Thompson R (1979). Sire evaluation. *Biometrics* **35**: 339–353.
van Arendonk JAM, Tier B, Kinghorn BP (1994). Use of multiple genetic markers in prediction of breeding values. *Genetics* **137**: 319–329.
van Arendonk JAM, Tier B, Bink MCAM, Bovenhuis H (1998). Restricted maximum likelihood analysis of linkage between genetic markers and quantitative trait loci for a granddaughter design. *J Dairy Sci* **81** (Suppl 2): 76–84.
Visscher PM (1996). Proportion of the variation in genetic composition in backcrossing programs explained by genetic markers. *J Hered* **87**: 136–138.
Visscher PM, Haley CS (1996). Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models. *Theor Appl Genet* **93**: 691–702.
Visscher PM, Haley CS (1998). Power of a chromosomal test to detect genetic variation using genetic markers. *Heredity* **81**: 317–326.
Wamg T, Fernando RL, van der Beek S, Grossman M, van Arendonk JAM (1995). Covariance between reratives for a marked quantitative trait locus. *Genet Sel Evol* **27**: 251–274.
Wang T, Fernando RL, Grossman M (1998). Genetic evaluation by best linear unbiased prediction using marker and trait information in a multibreed population. *Genetics*, **148**: 507–515.