Theor Appl Genet (1996) 92: 477–482

Wei-Ren Wu · Wei-Ming Li

Model fitting and model testing in the method of joint mapping of quantitative trait loci

Received: 2 May 1995 / Accepted: 22 September 1995

Abstract A previous paper proposed a new method of QTL mapping called joint mapping (JM). Some problems have been found in model fitting and model testing due to the neglect of the correlations among different observations of the dependent variable in this model. The present paper reports a method of solving the problems. The coefficient of correlation between two observations of the dependent variable is derived. A generalized least square (GLS) approach is developed for model fitting and a strategy and procedure of model testing based on a chi-square test is suggested. A simulated example is given. The example shows that the JM method is quite efficient in mapping multiple linked QTLs.

Key words QTLs · Joint mapping · Genetic marker · Linkage

Introduction

Many methods for mapping quantitative trait loci (QTLs) have been proposed since the late 1980s. At present, the most widely used method is the so-called interval mapping (IM) procedure proposed by Lander and Botstein (1989). An important shortcoming of IM, however, is that the method often maps 'ghost' QTLs or gives biased estimates of the QTL's positions when there are multiple QTLs on a chromosome (Martinez and Curnow 1992; Zeng 1994). Some modifications (Lander and Botstein 1989; Haley and Knott 1992; Martinez and Curnow 1992; Jansen 1993; Zeng 1994) have been proposed to overcome these shortcomings. Among them, the so-called composite interval mapping (CIM) proposed by Zeng (1994) seems to be the best. But all of the modified methods still follow the basic idea of IM; namely, using flanking markers to map QTLs.

Communicated by J. W. Snape

W.-R. Wu (✉) · W.-M. Li
Department of Agronomy, Fujian Agricultural University, Fuzhou, Fujian, People's Republic of China

Based on a quite different consideration, we have proposed a method of QTL mapping called joint mapping (JM) (Wu and Li 1994). The basic principle of JM is to utilize the linear functional relationship between the apparent effect of a marker on a quantitative character and the substantial effects of all the related QTLs linked to the marker. The linear coefficients in the function vary with the map distances between the marker and the QTLs. Consequently, each marker on the same chromosome can serve as a node of observation and a multiple regression analysis can be carried out to estimate both the positions and the effects of the QTLs. In principle, any number of linked QTLs can be mapped by JM as long as there are enough markers on a chromosome.

Kearsey and Hyne (1994) have also developed a method of QTL mapping similar to ours. Kearsey (personal communication) made us aware of an important fact we had overlooked in our former paper, namely that different observations of the dependent variable in the model (or the observed apparent effects of different markers) are actually not independent of each other because all the markers are mutually linked. This fact implies that both the conventional least square (LS) method (Kearsey and Hyne 1994) and the conventional weighted least square (WLS) method (Wu and Li 1994) are not suitable for the regression analysis because in such cases, (1) the estimates of parameters (i.e. positions and effects of QTLs) are not of the minimum variance although they are still unbiased (Xiang and Wu 1989); and, particularly, (2) the asymptotic property of either the F statistic used in the LS (Kearsey and Hyne 1994) or the χ^2 statistic used in the WLS (Wu and Li 1994) method is violated (Xiang and Wu 1989; Kearsey and Hyne 1994) so that, strictly speaking, the significance thresholds calculated, based on F distribution or χ^2 distribution, will no longer be correct, although the violation might be small when the sample size is large (Kearsey and Hyne 1994). Kearsey (personal communication) also suggested the use of empirical thresholds determined by Monte Carlo simulation. However, this does not seem to be an ideal solution for the problem because such a simulation is model-dependent and is generally very time con-

suming. In the present paper we report on improvements to the method using another approach.

Regression model

To be consistent with our former paper, a backcross population (BC or $F_1 \times P_1$) will be considered. For the sake of simplicity, epistatic effects among linked QTLs will be neglected. Suppose that there are n markers and m QTLs on a chromosome. In this paper, subscripts i and j will be used to indicate markers, while k and l indicate QTLs. Thus, a multiple regression model can be constructed as follows (Wu and Li 1994):

$$y_i = \sum_{k=1}^{m} b_k x_{ik} + e_i \quad (i = 1,2,\ldots,n) \tag{1}$$

in which y_i is an observed apparent effect, or the difference between a sample mean value of the homozygote genotype and that of the heterozygote genotype, of the i^{th} marker; b_k is the effect, or the difference between the population mean value of the homozygote genotype and that of the heterozygote genotype, of the k^{th} QTL; e_i is the random error; and

$$x_{ik} = 1 - 2r_{ik} \tag{2}$$

where r_{ik} is the recombination frequency between the i^{th} marker and the k^{th} QTL. r_{ik} is a function of the map distance between the i^{th} marker and the k^{th} QTL, i.e.

$$r_{ik} = r(D_{ik}) = r(|p_i - q_k|) \tag{3}$$

where p_i and q_k are the position coordinates on the map of the i^{th} marker and the k^{th} QTL, respectively. Since $p_i (i=1,2,\cdots,n)$ is known, equation (1) is a non-linear model with $2m$ unknown parameters, b_k and q_k $(k=1,2,\ldots,m)$.

Generally speaking, it is difficult to know the exact mathematical form of the function $r(D_{ik})$ in model (1). We have found that it is convenient to assume it to be Haldane's map function, i.e.

$$r = \frac{1}{2}\left(1 - e^{-2D}\right) \tag{4}$$

where the unit of D is in Morgans. In such a case, from (2) we find

$$x_{ik} = e^{-2D_{ik}} = e^{-2|p_i - q_k|}. \tag{5}$$

Model (1) has the following properties:

(1) According to the central limit theorem, asymptotically $y_i \sim N(\sum_k^m b_k x_{ik}, \sigma_i^2)$ or $e_i \sim N(0, \sigma_i^2)$ (Wu and Li 1994).
(2) σ_i^2 is related to r_{ik} $(k=1,2,\cdots,m)$ so that it varies with different marker loci (Wu and Li 1994).
(3) Since all the markers are linked to each other, there must be a correlation between y_i and y_j (or e_i and e_j) $(i,j=1,2,\ldots,n; i \neq j)$. We have proved (see Appendix 1) that theoretically the coefficient of correlation between y_i and y_j is expected to be

$$\rho_{ij} = 1 - 2r_{ij}. \tag{6}$$

So, under the assumption of Haldane's map function, we have

$$\rho_{ij} = e^{-2D_{ij}} = e^{-2|p_i - p_j|}. \tag{7}$$

Model fitting

According to the above three properties, it is obvious that the joint distribution of y_is would be asymptotically a multi-variate normal distribution. So, the likelihood function would be:

$$L = f(Y; X, B, \Sigma) = \frac{1}{(2\pi |\Sigma|)^{n/2}} e^{-\frac{1}{2}(Y-XB)'\Sigma^{-1}(Y-XB)} \tag{8}$$

where Y is an $(n \times 1)$ vector of y_is; X is an $(n \times m)$ matrix of x_{ik}s; B is an $(m \times 1)$ vector of b_ks; Σ is the variance matrix of Y, namely, $\text{var}(Y) = \sigma_e^2 \Sigma = \Sigma$ (here $\sigma_e^2 = 1$).

Σ can be estimated from the sample. In fact, according to the definition of y_i, the variance of y_i (denoted as σ_i^2) will be $= \sigma_{i1}^2/n_{i1} + \sigma_{i2}^2/n_{i2}$, where σ_{i1}^2, n_{i1} and σ_{i2}^2, n_{i2} are the variances and sample sizes of the homozygote and the heterozygote genotypes of the i^{th} marker, respectively, and both σ_{i1}^2 and σ_{i2}^2 can be estimated from the sample. As to the covariance between y_i and y_j (denoted as σ_{ij}^2), it can be estimated either with (A1) or, more conveniently and explicitly, with (A7) (see Appendix 1), in which r_{ij} can be calculated from marker linkage maps constructed in the same experiment or in others. We prefer using formula (A7) to estimate σ_{ij}^2.

Since Σ can be estimated in advance, $|\Sigma|$ becomes a constant in (8). Thus, maximizing the value of L in (8) is equivalent to minimizing the value of

$$RSS = (Y - XB)'\Sigma^{-1}(Y - XB). \tag{9}$$

This is actually the generalized least square (GLS) method (Xiang and Wu 1989).

When the positions of putative QTLs, q_k $(k=1,2,\cdots,m)$, are given, (1) becomes a multiple linear regression model. In this case, (9) has an analytical solution (Xiang and Wu 1989):

$$B = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y \tag{10}$$

and the least RSS under the condition of given q_ks can be calculated by substituting (10) back into (9). That means that, as a matter of fact, the function (9) only depends on the variables q_k $(k=1,2,\ldots,m)$. So, a strategy to approach the point of least RSS can be: choose a set of initial values of q_ks; then search for q_k in proper order by iteration until convergence.

Under the assumption of Haldane's map function, the calculation of (10) and (9) will be simplified because, according to (7), ρ_{ij} has the following property

$$\rho_{ij} = \rho_{i,i+1}\,\rho_{i+1,i+2}\cdots\rho_{j-1,j} \qquad (i < j)$$

which makes Σ^{-1} to be a symmetric tri-diagonal matrix, of which the elements of main diagonal and secondary diagonals are

$$\alpha_i = \frac{1-\rho_{i-1,i}^2\,\rho_{i,i+1}^2}{\left(1-\rho_{i-1,i}^2\right)\left(1-\rho_{i,i+1}^2\right)\sigma_i^2} \quad (i=1,2\ldots,n)$$

and

$$\beta_i = \frac{-\rho_{i,i+1}}{\left(1-\rho_{i,i+1}^2\right)\sigma_i\sigma_{i+1}} \quad (i=1,2\ldots,n-1)$$

respectively, where $\rho_{01}=\rho_{n,n+1}=0$.

Model testing

It is known that, for a linear regression model, the RSS of GLS (in the case of $\sigma_e^2=1$) follows asymptotically a chi-square distribution and that the estimates of parameters in the model are of the minimum variance (Xiang and Wu 1989). In this case, model testing will be simple. Unfortunately, model (1) is non-linear, so whether the conclusion from a linear model applies to it needs to be examined. Hence, we conducted a simulation study.

We assumed that on a chromosome there are five markers, evenly spaced with a distance of 5 cM, but no QTLs; the random error $e_i\,(i=1,2,\cdots,5)$ follows a standard normal distribution N(0,1). Two sample sizes, 100 and 1000, were considered; 500 samples were generated for each case. Both null-QTL and one-QTL models were used to analyze the data. The results are listed in Table 1.

These results clearly indicate that for model (1), the RSS of GLS (in the case of $\sigma_e^2=1$) also asymptotically follows a chi-square distribution, of which the degrees of freedom is given by $n-2m$, i.e.

$$RSS(m)\sim\chi^2(n-2m) \tag{11}$$

where n is the number of markers observed and m is the number of QTLs assumed in the model. Consequently, the estimates of parameters must be of minimum variance. When m is equal to, or larger than, the real number of QTLs on the chromosome, the chi-square distribution would be central, otherwise it would be non-central. In addition, according to the additivity of chi-square variables, it can be inferred immediately from (11) that

$$RSS(m)-RSS(m+1)\sim\chi^2(2). \tag{12}$$

Obviously, the chi-square distribution in (12) will be central when m is equal to, or larger than, the real number of QTLs.

According to (11) and (12), we can find criteria for testing the model and a corresponding strategy for identifying the best model. Two steps are involved. First, find the models fitted to the data in accordance with the following condition:

$$RSS(m)\leqslant\chi_\alpha^2(n-2m) \tag{13}$$

Table 1 Results of sampling of RSS with 500 simulations (see text). Notes for the symbols: NP=number of parameters in the model; MRSS=mean of RSS; VRSS=variance of RSS; exp=expected by the theoretical (chi-square) distribution; sam=sampled; GF=goodness of fit

Sample size		100		1000	
Model		Null-QTL	One-QTL	Null-CTL	One-QTL
NP		0	2	0	2
MRSS	exp	5	3	5	3
	sam	5.010	3.127	4.940	3.061
VRSS	exp	10	6	10	6
	sam	10.464	6.119	10.320	5.626
GF	χ^2	7.00	3.06	11.06	8.25
	df	12	9	12	9
	$P>$	0.75	0.95	0.50	0.50

where α is the significance level. Then use the following criterion

$$RSS(m)-RSS(m+1)>\chi_\alpha^2(2) \tag{14}$$

to identify the best one among the models which meet condition (13). If inequality (14) is true, then one more QTL needs to be assumed in the model. Otherwise the m-QTL model would be the best one. A series of tests may be carried out successively until no more QTLs can be added to the model.

The determination of significance levels in (13) and (14) is worthy of discussion. When several chromosomes are examined in an experiment, for the test based on condition (13), the probability of refusing a correct model but accepting an alternative model with an additional assumed QTL(s) will increase; and similarly, for the test based on criterion (14), the probability of accepting a false QTL(s) into the model will also increase. Thus, the overall result will be that the risk of mapping false QTLs increases. Therefore, higher significance levels in both (13) and (14) would be required. Considering that different chromosomes are independent, we suggest that the nominal significance level for individual tests be calculated with the following formula

$$\alpha=1-(1-\alpha_0)^{1/c}$$

where α_0 is the overall significance level required and c is the number of chromosomes being analyzed.

In regard to the confidence interval of a QTL's position (q_k), similar to that presented in the former paper (Wu and Li 1994), it can be determined by

$$RSS(m\,|\,q_k)-RSS(m)\leqslant\chi_\alpha^2\ (1)$$

where $RSS(m\,|\,q_k)$ is the minimum RSS depending on q_k.

Example

To illustrate the procedure of JM, a simulated example will be given here. Consider a chromosome, 150 cM long, with

Table 2 Results of JM model fitting

Model	*df*	RSS	$P(\chi^2>\text{RSS})$
Null-QTL	16	55.645	<0.001
One-QTL	14	39.007	<0.001
Two-QTL	12	28.494	<0.005
Three-QTL	10	19.950	<0.05
Four-QTL	8	4.511	>0.75

Table 3 Estimates of parameters obtained by the four-QTL model

QTL no.	Position (cM)	Effect	95% Confidence interval (cM)
1	7.8	0.727	2–17
2	44.7	–0.783	37–54
3	93.2	0.792	84–99
4	123.7	–0.950	119–129

16 evenly spaced markers and 4 QTLs, of which the positions (cM) and effects are 7 and 0.45; 44 and –0.65; 95 and 0.55; 129 and –0.5, respectively. The residual variance $\sigma^2_{res}=1$. A sample with 250 individuals was generated.

A series of models with from 0 to 4 assumed QTLs were fitted (Table 2). The results strongly indicate that there must be 4 QTLs on the chromosome. The estimates of parameters by the 4-QTL model are listed in Table 3, and the curves of $RSS(m|q_k)-RSS(m)\sim q_k$ (k=1,...,4) are given by Fig. 1(a), which intuitively illustrates the estimated positions of the 4 QTLs and their corresponding 95% confidence intervals. From both Table 3 and Fig. 1(a), we see that the estimates of the QTL's positions are satisfactorily precise although the estimates of the QTL's effects are larger than their real values. This example shows that JM is quite efficient in mapping multiple linked QTLs.

To provide a preliminary comparison of the power of QTL mapping, the methods of IM and CIM were also used to analyze the data. The results are shown in Fig. 1(b). For convenience of comparison, the likelihood ratio (LR) rather than LOD was used as the statistic in IM. The LR threshold of 5% overall significance level for IM is ≈7.00, which was calculated by conversion from the equivalent empirical LOD threshold (≈1.52) obtained by Lander and Botstein (1989) for the case of mapping a single 100-cM-long chromosome on which markers are spaced evenly with a distance of 10 cM. Since the chromosome in this example is longer than 100 cM, and thus there are more marker intervals to be analyzed, the appropriate LR threshold may be a little larger than 7.00. The LR threshold of 5% overall significance level for CIM is ≈8.50, which was calculated based on the assumptions (Zeng 1994) that the LR statistic follows a chi-square distribution with 1 degree of freedom and that the tests on individual intervals be approximately independent.

We see from Fig. 1(b) that only one (the 4th) QTL was detected by both IM and CIM (the corresponding estimates of position and effect of the QTL were 127 cM and –0.57 by IM and 124 cM and –0.978 by CIM, respectively). The results show that both IM and CIM are not as efficient as JM in QTL mapping in this example. Certainly, this does not mean that JM would be better than IM and CIM in every case. But the example at least indicates that the power of JM should be comparable to those of currently widely used QTL mapping methods based on flanking-marker analysis. More comparative studies are needed to provide further knowledge on this point.

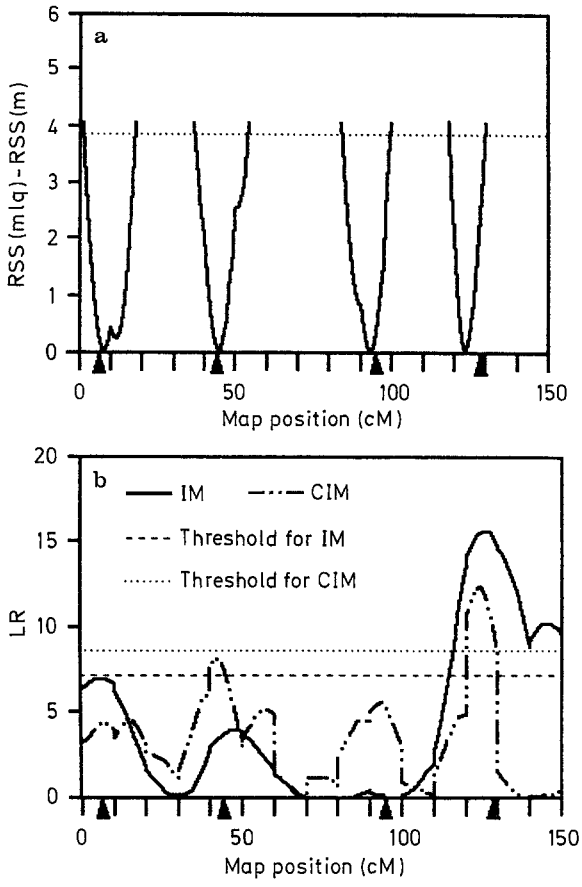

Fig. 1a, b Results of QTL mapping of the simulated example (see text) with JM, IM and CIM. The *vertical short bars* on the bottom indicate the positions of markers. The *black triangles* indicate the real positions of QTLs. **a** JM: curves of $RSS(m|q_k)$–$RSS(m)\sim q_k$ (k=1,···,4). The lowest point of each curve indicates the most probable position of a QTL. The *horizontal dotted line* shows the significance threshold $\chi^2_{0.05}(1)=3.84$, below which the range of each curve defines a 95% confidence interval of the corresponding QTL's position. **b** IM and CIM: LR scores varying along the chromosome

Discussion

In the above discussion, only a backcross population was involved. But the results can be directly applied to a dou-

bled haploid (DH) population. It is also applicable to a population of recombinant inbred lines (RIL) except that some modifications are needed because, in a RIL population, the proportion of recombinants (R) of two linked loci does not equal the corresponding recombination frequency (r) but equals to $2r/(1+2r)$ (Haldane and Waddington 1931). Therefore, for a RIL population, the parameters r in model (1) must be replaced by R and the coefficient of correlation between y_i and y_j becomes $\rho_{ij}=1-2R_{ij}$. Interestingly, if we assume Kosambi's map function (Kosambi 1944)

$$r=\frac{1}{2}\frac{e^{2D}-e^{-2D}}{e^{2D}+e^{-2D}}$$

and let $D^*=2D$, then we find

$$R=\frac{1}{2}\left(1-e^{-2D^*}\right)$$

which has the same form as (4). In such a case, therefore, the calculation of (10) and (9) can also be simplified.

In the above discussion, epistatic effects among linked QTLs were ignored. But in principle, the method of JM applies to the situation when there are epistatic effects. Following the digenic interaction system of polygenes described by Mather and Jinks (1982), we have found that, for DH and RIL populations, when epistatic effects between two linked QTLs are taken into account, the form of model (1) will not be influenced, whereas for a BC population two new terms related to epistatic effects will appear in model (1). The two terms are:

$$\sum_{k<i=2}^{m}(1-r_{ik}-r_{il})(aa_{kl}-dd_{kl}) \tag{15}$$

and

$$\sum_{k<l=2}^{m}(r_{ik}-r_{il})(ad_{lk}-ad_{kl}) \tag{16}$$

where subscripts i, k and l denote the i^{th} marker and the k^{th} and the l^{th} QTLs, respectively; a denotes the additive effect, d denotes the dominance effect, aa, ad and dd denote epistatic effects of $a\times a$, $a\times d$ and $d\times d$, respectively. If it is assumed that $ad_{lk}=ad_{kl}$, then term (16) will be eliminated so that there will be only one epistatic effect term in the model. This is what we have presented in the former paper (Wu and Li 1994).

Although epistatic effects can be included in the model, both model fitting and model testing will then become more complicated, and the precision of mapping may be reduced, too. In fact, for the sake of simplicity, epistatic effects are generally assumed to be negligible in most of the current methods of QTL mapping. However, if epistatic effects actually exist and are significant, then they should be considered. Since our first purpose is to map QTLs rather than estimate the interactive effects among the QTLs, an ideal situation would be that epistatic effects can be omitted from the model of QTL mapping whether or not they exist. From this point of view, the DH and RIL populations are better than that of BC for QTL mapping.

Appendix 1

Assume that the i^{th} and the j^{th} markers are linked to each other with a recombination frequency r_{ij}. Let n denote the number of individuals, $\bar{\mu}$ denote the sample mean and σ^2 denote the variance of the sample mean of a genotype; subscripts 1 and 2 indicate the homozygote genotype and the heterozygote genotype of a marker, respectively; N denotes the total number of individuals of the sample. Then, in accordance with the result of Appendix 2, we have

$$\begin{aligned}COV(y_i,y_j)&=COV(\bar{\mu}_{i1}-\bar{\mu}_{i2},\bar{\mu}_{j1}-\bar{\mu}_{j2})\\&=COV(\bar{\mu}_{i1},\bar{\mu}_{j1})-COV(\bar{\mu}_{i1},\bar{\mu}_{j2})\\&\quad-COV(\bar{\mu}_{i2},\bar{\mu}_{j1})+COV(\bar{\mu}_{i2},\bar{\mu}_{j2})\\&=\frac{n_{i1j1}}{\sqrt{n_{i1}n_{j1}}}\sigma_{i1}\sigma_{j1}-\frac{n_{i1j2}}{\sqrt{n_{i1}n_{j2}}}\sigma_{i1}\sigma_{j2}\\&\quad-\frac{n_{i2j1}}{\sqrt{n_{i2}n_{j1}}}\sigma_{i2}\sigma_{j1}+\frac{n_{i2j2}}{\sqrt{n_{i2}n_{j2}}}\sigma_{i2}\sigma_{j2}\end{aligned} \tag{A1}$$

where COV denotes co-variance. (A1) is a general formula for calculating $COV(y_i,y_j)$. But theoretically, for a BC population, it is expected that

$$n_{i1}=n_{i2}=n_{j1}=n_{j2}=\frac{1}{2}N \tag{A2}$$

$$n_{o1j1}=n_{i2j2}=\frac{1}{2}(1-r_{ij})N \tag{A3}$$

$$n_{i1j2}=n_{i2j1}=\frac{1}{2}r_{ij}N \tag{A4}$$

$$\sigma_{i1}^2=\sigma_{i2}^2=\frac{1}{2}\sigma_i^2 \tag{A5}$$

$$\sigma_{j1}^2=\sigma_{j2}^2=\frac{1}{2}\sigma_j^2 \tag{A6}$$

where σ_i^2 and σ_j^2 have the same meaning as in the text. So, by substituting (A2)–(A6) into (A1) we have

$$\begin{aligned}COV(y_i,y_j)&=\frac{1}{2}(1-r_{ij})\sigma_i\sigma_j-\frac{1}{2}y_{ij}\sigma_i\sigma_j\\&=\frac{1}{2}r_{ij}\sigma_i\sigma_j+\frac{1}{2}(1-r_{ij})\sigma_i\sigma_j\\&=(1-2r_{ij})\sigma_i\sigma_j.\end{aligned} \tag{A7}$$

Hence, we find that the expected coefficient of correlation between y_i and y_j is

$$\rho_{ij}=1-2r_{ij}.$$

Appendix 2

Assume that there are two random variable sets X and Y. There is a random sample with size n obtained from set union $X\cup Y$, in which the numbers of variables sampled from X, Y and set intersection $X\cap Y$ are n_x, n_y and n_{xy}, respec-

tively. Let $\bar{x}$ and $\bar{y}$ denote the mean values of variables sampled from X and Y. Then, we have

$$COV(\bar{x},\bar{y}) = COV\left(\frac{1}{n_x}\sum_{i=1}^{n_x} x_i, \frac{1}{n_y}\sum_{j=1}^{n_y} y_j\right)$$

$$= \frac{1}{n_x n_y}\sum_{i=1}^{n_x}\sum_{j=1}^{n_y} COV(x_i, y_j)$$

$$= \frac{1}{n_x n_y}\sum_{i=1}^{n_x}\sum_{j=1}^{n_y} \rho_{ij}\,\sigma_i\,\sigma_j .$$

Because

$$\rho_{ij} = \begin{cases} 0 & \text{when } x_i \text{ and } y_j \text{ are the same variable} \\ 1 & \text{when } x_i \text{ and } y_j \text{ are different variables} \end{cases}$$

we get

$$COV(\bar{x}, y) = \frac{n_{xy}}{\sqrt{n_x n_y}}\,\sigma_{\bar{x}}\,\sigma_{\bar{y}}$$

and

$$\rho_{x,y} = \frac{n_{xy}}{\sqrt{n_x n_y}} .$$

References

Haldane JBS, Waddington CH (1931) Inbreeding and linkage. Genetics 16:357–374

Haley CS, Knott SA (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315–324

Jansen RC (1993) Interval mapping of multiple quantitative trait loci. Genetics 135:205–211

Kearsey MJ, Hyne V (1994) QTL analysis: a simple 'marker-regression' approach. Theor Appl Genet 89:698–702

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Martinez O, Curnow RN (1992) Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. Theor Appl Genet 85:480–488

Mather K, Jinks JL (1982) Biometrical genetics (3rd edn). University Press, Cambridge

Wu WR, Li WM (1994) A new approach for mapping quantitative trait loci using complete genetic marker linkage maps. Theor Appl Genet 89:535–539

Xiang KF, Wu QG (1989) Experiment design and data analysis (in Chinese). Shanghai Science and Technology Press, Shanghai, PRC

Zeng Z-B (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468