Genet. Res., Camb. (2001), **78**, *pp.* 187–189. *With 2 figures.* © 2001 Cambridge University Press
DOI: 10.1017/S0016672301005262 Printed in the United Kingdom

SHORT PAPER
A comment on Xie and Xu: 'Mapping quantitative trait loci in tetraploid species'

CHRISTINE A. HACKETT*
Biomathematics and Statistics Scotland, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK

(*Received 22 March 2001 and in revised form 17 May 2001*)

Summary

Xie & Xu (2000) present a model for mapping quantitative trait loci in an autotetraploid population. However, one aspect of their model, namely gamete formation, does not properly represent the biological process in autotetraploid species. This paper gives a more realistic formulation for this part of the model, and discusses the consequences for multipoint mapping.

1. Introduction

Although there has been much interest in the theory and practice of mapping quantitative trait loci (QTLs) in diploid species over the last 20 years, there has been very little attention to extending these approaches to autopolyploid species. Xie & Xu (2000) have tackled this problem, and have advanced the theory with a model relating trait values to QTL genotypes in an autotetraploid population of full sibs from a cross between two outbred parents. As for diploid species, it is necessary to infer the QTL genotypes from the genotypes at linked marker loci and this involves modelling the processes of gamete formation and recombination. I believe that Xie & Xu (2000) do not represent this process properly, and present an alternative derivation below.

2. Statistical model

(i) *Modelling recombination between two linked loci*

This paper will consider the formation of gametes in a tetraploid parent with genotype $M_1M_2M_3M_4$ at a marker locus M and $Q_1Q_2Q_3Q_4$ at a linked QTL Q. I assume that M_1 and Q_1 are on the same chromosome, etc. Let the recombination frequency between M and Q be r. Xie & Xu (2000) state that they assume that gametes are produced by random chromosomal segregation. This is defined (e.g. Allard, 1966) by two features:

1. Homologous chromosomes pair with equal frequencies.
2. Either only bivalents are produced, or there is no chiasma between the locus and the centromere in a quadrivalent. These have the consequence that sister chromatids never appear in the same gamete. However, when considering linkage between two loci, one is considering a chiasma between them and hence only the former can give chromosomal segregation at both loci.

Under random chromosomal segregation, a parent will produce six possible gametes at each locus (i.e. M_1M_2, M_1M_3, M_1M_4, M_2M_3, M_2M_4, M_3M_4 at locus M). This order of subscripts will be used in all the matrices below. Xie & Xu (2000) give a matrix $\mathbf{T}$ of the conditional probabilities $\Pr(M = i \mid Q = j)$, where $1 \leqslant i, j \leqslant 6$ represent the above order of gamete subscripts. There is a typing error in their paper at this point, as the probabilities do not sum to 1. Corrected (Xie & Xu, personal communication), their matrix, denoted $\mathbf{T}_X$, is

$$\mathbf{T}_X = \begin{bmatrix} (a+c)/d & (b+c)/d & (b+c)/d & (b+c)/d & (b+c)/d & 2c/d \\ (b+c)/d & (a+c)/d & (b+c)/d & (b+c)/d & 2c/d & (b+c)/d \\ (b+c)/d & (b+c)/d & (a+c)/d & 2c/d & (b+c)/d & (b+c)/d \\ (b+c)/d & (b+c)/d & 2c/d & (a+c)/d & (b+c)/d & (b+c)/d \\ (b+c)/d & 2c/d & (b+c)/d & (b+c)/d & (a+c)/d & (b+c)/d \\ 2c/d & (b+c)/d & (b+c)/d & (b+c)/d & (b+c)/d & (a+c)/d \end{bmatrix}$$

where $a = (1-r)^2$, $b = r(1-r)/3$, $c = r^2/9$ and $d = 1-2r/3+4r^2/9$.

* Tel: +44 (0) 1382 562731. Fax: +44 (0) 1382 562426. e-mail: christine@bioss.ac.uk

Table 1. *Gametes produced by an autotetraploid parent with genotype $M_1M_2M_3M_4$ at a marker locus* M *and $Q_1Q_2Q_3Q_4$ at a linked QTL* Q *if chromosomes pair as 1 with 2, 3 with 4. The recombination frequency between* M *and* Q *is* r. *The gamete probabilities are obtained by multiplying the probabilities from each chromosome pair*

Gamete type			Chromosomes 3, 4			
			Parental		Recombinant	
			M_3Q_3 $(1-r)/2$	M_4Q_4 $(1-r)/2$	M_3Q_4 $r/2$	M_4Q_3 $r/2$
Chromosomes 1, 2	Parental	M_1Q_1 $(1-r)/2$	$M_1M_3Q_1Q_3$	$M_1M_4Q_1Q_4$	$M_1M_3Q_1Q_4$	$M_1M_4Q_1Q_3$
		M_2Q_2 $(1-r)/2$	$M_2M_3Q_2Q_3$	$M_2M_4Q_2Q_4$	$M_2M_3Q_2Q_4$	$M_2M_4Q_2Q_3$
	Recombinant	M_1Q_2 $r/2$	$M_1M_3Q_2Q_3$	$M_1M_4Q_2Q_4$	$M_1M_3Q_2Q_4$	$M_1M_4Q_2Q_3$
		M_2Q_1 $r/2$	$M_2M_3Q_1Q_3$	$M_2M_4Q_1Q_4$	$M_2M_3Q_1Q_4$	$M_2M_4Q_1Q_3$

A derivation including the bivalent formation explicitly gives a different form for this matrix. If chromosomes 1 and 2 pair to form one bivalent, and chromosomes 3 and 4 pair to form a second, the possible gametes and their probabilities are shown in Table 1. Two similar tables are obtained if chromosome 1 pairs with chromosome 3, and 2 with 4, or if 1 pairs with 4 and 2 with 3. Under random chromosomal segregation, it is assumed that all these three pairings are equally likely, i.e. probability 1/3.

Combining gamete genotypes from the three possible pairings, there are six genotypes with no recombinations ($M_1M_2Q_1Q_2$, etc., each with probability $(1-r)^2/6$), 24 genotypes with one recombination ($M_1M_3Q_2Q_3$, etc., each with probability $r(1-r)/12$) and six genotypes with two recombinations ($M_1M_3Q_2Q_4$, etc., each with probability $r^2/6$).

The matrix of the conditional probabilities $\Pr(M \mid Q)$ can therefore be written as

$$\mathbf{T} = \begin{bmatrix} a & b & b & b & b & c \\ b & a & b & b & c & b \\ b & b & a & c & b & b \\ b & b & c & a & b & b \\ b & c & b & b & a & b \\ c & b & b & b & b & a \end{bmatrix} \quad (1)$$

where $a = (1-r)^2$, $b = r(1-r)/2$, $c = r^2$. The genotypes with no recombinations are on the leading diagonal, and those with two recombinations are on the other diagonal. The matrix of conditional probabilities for the two parents can then be formed as the Kronecker product $\mathbf{T} \otimes \mathbf{T}$, as in Xie & Xu (2000).

Combining suitable cells of the matrices $\mathbf{T}_X$ or $\mathbf{T}$ gives category probabilities for pairs of dominant markers in different configurations, e.g. simplex coupling. Matrix $\mathbf{T}_X$ does not simplify to the category probabilities for dominant markers given by Wu *et al.* (1992), Hackett *et al.* (1998) and other researchers. Matrix $\mathbf{T}$ does give the same probabilities. The form of matrix $\mathbf{T}$ is used by Luo *et al.* (2001) to construct linkage maps of autotetraploid species using co-dominant markers.

Table 2. *Three-locus parental genotype to illustrate calculation of multipoint linkage*

	Chromosome			
	1	2	3	4
Locus 1	A	B	C	D
Locus 2 (QTL)	A	B	A	C
Locus 3	A	B	C	D

(ii) *Multipoint mapping*

The question of modelling bivalent formation has some important implications for the later section of Xie & Xu (2000) on multipoint mapping. Their multipoint model for m markers and a QTL between markers k and $k+1$ is

$$\Pr(I_M \mid Q = i) = \mathbf{1}^T\mathbf{D}_1\mathbf{T}_{12}\mathbf{D}_2 \ldots \mathbf{D}_k\mathbf{T}_{kQ}\mathbf{D}_{(i)}\mathbf{T}_{Q(k+1)}\mathbf{D}_{k+1} \ldots \mathbf{D}_{m-1}\mathbf{T}_{(m-1)m}\mathbf{D}_m\mathbf{1}$$

where $\mathbf{D}$ is a diagonal matrix with elements equal to 1 for genotypes compatible with the observed phenotypes and 0 elsewhere. Again, I will consider segregation in one parent only for simplicity.

Consider a parent with three loci and genotypes given in Table 2. A, B, C and D will be used to represent different alleles, to avoid confusion with the chromosome number. Loci 1 and 3 have all four alleles distinguishable, but for locus 2 the alleles on chromosomes 1 and 3 are identical.

Say, for example, that we want to calculate the probability of a gamete with phenotype Locus 1: AC, Locus 2: AC, Locus 3: AD. Let the recombination frequency be r between loci 1 and 2, and s between loci

https://doi.org/10.1017/S0016672301005262 Published online by Cambridge University Press

2 and 3. The multipoint formula above would give the probability of this gamete as

$$\mathbf{1}^T\mathbf{D}_1\mathbf{T}_{12}\mathbf{D}_2\mathbf{T}_{23}\mathbf{D}_3\mathbf{1}$$

where $\mathbf{D}_1$ is diagonal with a 1 in the second place (i.e. from chromosomes 1 and 3), $\mathbf{D}_3$ is diagonal with a 1 in the third place, and $\mathbf{D}_2$ is diagonal with 1s in the third and sixth places, as this gamete could have received alleles from chromosomes 1 and 4, or 3 and 4.

The multipoint formula is evaluated (using matrix **T** from equation 1) as

$$\frac{r(1-r)}{2}\left[\frac{s(1-s)}{2}+(1-s)^2\right].$$

The first part of this expression represents the probability that the individual has inherited alleles from chromosomes 1 and 3 for locus 1, chromosomes 3 and 4 for locus 2 and chromosomes 1 and 4 for locus 3. The second part represents the probability that the individual has inherited chromosomes 1 and 3 for locus 1 and chromosomes 1 and 4 for loci 2 and 3.

However, the fully informative loci 1 and 3 show that to obtain this gamete there has been a crossover between chromosomes 3 and 4 and so the bivalent pairing is chromosome 1 pairing with 2 and chromosome 3 pairing with 4, and crossovers occur within these pairs. Because of this, the individual cannot have alleles from chromosomes 3 and 4 expressed together for locus 2, and its AC phenotype must come from chromosomes 1 and 4. The correct expression for the probability of this individual is

$$\frac{r(1-r)}{2}\left[(1-s)^2\right].$$

A consistent method for a multipoint analysis would be to calculate the probability of each sequence of markers conditional on each possible bivalent pairing separately, and then to sum over the probability of each bivalent pairing.

3. Conclusions

The version of matrix **T** derived here can easily be incorporated into the model in Xie & Xu (2000) to give a more realistic representation of the process of gamete formation in autotetraploids. Their model can be modified to give a multipoint analysis taking the process of bivalent formation into account. With these useful tools, the process of QTL mapping in autotetraploid species may soon become as practical as it is in diploids.

This research was supported by a research grant from the UK Biotechnology and Biological Sciences Research Council, and by the Scottish Executive Rural Affairs Department. I am grateful to Dr John Bradshaw and Dr Shizhong Xu for useful discussions.

References

Allard, R. W. (1966). *Principles of Plant Breeding*. New York: Wiley.

Hackett, C. A., Bradshaw, J. E., Meyer, R. C., McNicol, J. W., Milbourne, D. & Waugh, R. (1998). Linkage analysis in tetraploid species: a simulation study. *Genetical Research* **71**, 143–154.

Luo, Z. W., Hackett, C. A., Bradshaw, J. E., McNicol, J. W. & Milbourne, D. (2001). Construction of a genetic linkage map in tetraploid species using molecular markers. *Genetics* **157**, 1369–1385.

Wu, K. K., Burnquist, W., Sorrells, M. E., Tew, T. L., Moore, P. H. & Tanksley, S. D. (1992). The detection and estimation of linkage in polyploids using single-dose restriction fragments. *Theoretical and Applied Genetics* **83**, 294–300.

Xie, C. & Xu, S. (2000). Mapping quantitative trait loci in tetraploid populations. *Genetical Research* **76**, 105–115.