Parameterization of Sex-Limited Autosomal Linkage Analysis for Mx

Sarah E. Medland[1,2]

[1] *Queensland Institute of Medical Research, Brisbane, Australia*
[2] *Cognitive Psychophysiology Laboratory, University of Queensland, Australia*

Incorporation of sex-limitation (genotype-sex interaction) effects into a model of quantitative trait loci (QTL) analysis has been shown to increase the power to detect linkage when analyzing traits in which sex limitation is present (Towne et al., 1997). The present note provides a parameterization of the nonscalar sex-limitation ACE model incorporating autosomal sex-limited QTL effects for use with the Mx matrix algebra program (Neale et al., 2002). An example script designed for use with extended sibships that takes advantage of the versatile treatment of covariates within Mx is included.

The purpose of the present note is to provide a parameterization for sex-limited autosomal linkage analyses suitable for extended sibship pedigrees within Mx. Sex limitation of variance, also known as genotype-by-sex (G × S) interaction, may take a number of forms. Nonscalar sex-limitation models are generally used to model sex differences in the magnitude of genetic and environmental effects, and to determine whether the same set of genetic or environmental factors influences a trait in males and females. In the scalar sex-limitation model the proportions of variance explained by genetic and environmental factors do not differ between males and females while the absolute magnitude of these variance components differs by a constant.

Sex limitation has been detected in a wide range of physiological and psychological traits, with recent examples including conduct disorder and alcohol use (Rose et al., 2004), pubertal development (Mustanski et al., 2004), language development (Van Hulle et al., 2004), eating behaviors (Keski-Rahkonen et al., 2004), aggression and hyperactivity (Vierikko et al., 2004) and blood pressure (Zeegers et al., 2004). The presence of sex limitation has important implications for the genetic dissection of complex traits, as the misspecification of genotype-by-environment interactions in quantitative trait loci (QTL) analysis may lead to biased estimates (Purcell & Sham, 2002). The implementation of models where the QTL are allowed to differ between males and females (QTL sex-limited or genotype-by-sex interaction) is relatively straightforward and has been implemented in pedigree analysis programs (e.g., Towne et al., 1997) and programs for structured pedigrees in outbred populations (Seaton et al., 2002).

Within the context of the classical twin design, the parameterization of variance components QTL linkage analysis represents a simple extension of the model through the addition of a QTL-linked component of variance (for a recent description of the methodology see Posthuma et al., 2003). In a similar way the parameterization of a sex-limited nonscalar autosomal QTL model is achieved by extending the variance covariance matrices for a nonscalar sex-limitation script, incorporating the male and female specific QTL parameters. The parameterization adopted in this note uses a *pi-hat* ($\hat{\pi}$) approach, that is, the QTL covariance between relatives is a function of the expected proportion of alleles identical-by-descent (IBD). The extension to a full mixture distribution approach (Eaves et al., 1996) is not complicated for sibships of size two but quickly becomes unwieldy for larger sibships.

Extension of the nonscalar sex-limitation ACE model to include sex-limited autosomal QTL components yields the following for brother pairs:

$$[\Sigma_N]_{ij} = \begin{cases} \sigma^2_{Am} + \sigma^2_{Cm} + \sigma^2_{Qm} + \sigma^2_{Em} & \text{if } i = j \\ 1/2\sigma^2_{Am} + \sigma^2_{Cm} + \hat{\pi}\sigma^2_{Qm} & \text{if } i \neq j \text{ and } i \text{ and } j \text{ are full sibs} \\ \sigma^2_{Am} + \sigma^2_{Cm} + \sigma^2_{Qm} & \text{if } i = j \text{ and } i \text{ and } j \text{ are monozygotic (MZ) twins} \end{cases}$$

sister pairs:

$$[\Sigma_N]_{ij} = \begin{cases} \sigma^2_{Af} + \sigma^2_{Cf} + \sigma^2_{Qf} + \sigma^2_{Ef} & \text{if } i = j \\ 1/2\sigma^2_{Af} + \sigma^2_{Cf} + \hat{\pi}\sigma^2_{Qf} & \text{if } i \neq j \text{ and } i \text{ and } j \text{ are full sibs} \\ \sigma^2_{Af} + \sigma^2_{Cf} + \sigma^2_{Qf} & \text{if } i = j \text{ and } i \text{ and } j \text{ are MZ twins} \end{cases}$$

and brother–sister pairs:

$$[\Sigma_N]_{ij} = \begin{cases} 1/2\, r_A \sigma_{Am}\sigma_{Af} + r_C \sigma_{Cm}\sigma_{Cf} + \hat{\pi}\, r_Q \sigma_{Qm}\sigma_{Qf} \\ \text{if } i \neq j \text{ and } i \text{ and } j \text{ are full sibs} \end{cases}$$

Received 7 April, 2005; accepted 29 September, 2005.

Address for correspondence: Sarah Medland, Queensland Institute of Medical Research, PO Royal Brisbane Hospital, Brisbane 4029, Australia. E-mail: Sarah.Medland@qimr.edu.au

where $\hat{\pi}$ is the proportion of alleles shared IBD at a given locus. The subscripts *A, C, Q* and *E* denote the variance due to additive genetic, common environmental, QTL and unique environmental effects, and the subscripts *m* and *f* refer to male and female parameters. In classical twin designs, the additive genetic correlation between the sexes (r_G) and the correlation between the common environmental effects of males and females (r_C) are confounded, and the latter is usually set to 1. The correlation between the QTL effects in males and females (r_Q) is a locus-specific analogue of r_A. If there are only two QTL alleles segregating in the population, then $|r_Q| = 1$. With more alleles at the QTL it is possible, at least in principle, that some alleles affect only one sex, resulting in a QTL correlation less than 1. Estimation of r_Q is possible but will complicate hypothesis testing because more models can be considered. The problems surrounding estimation and hypothesis testing for sex-limited QTL is analogous to multiple trait QTL mapping. For the remainder of this note it will be assumed that the QTL correlation is set to 1.

The extension of this model to larger sibships in programs in which the variance–covariance structure must be provided by the user, such as Mx, can be accommodated in a number of ways. It is possible to enlarge the variance–covariance model (adding two rows and two columns, one for each sex, for each additional sib) or through stratification of the sample into multiple data groups (with the parameters estimated across all data groups). However, the approach used here is an extension of the parameterization of Medland (2004), which takes advantage of the versatile treatment of covariates within Mx to specify the sex pairings found within each sibship. An example script demonstrating the implementation of this model is given in the Appendix. As with other linkage analyses conducted with Mx, the current parameterization requires prior computation of IBD probabilities using a software package such as Merlin (Abecasis et al., 2002).

In this approach male and female QTL are estimated as separate parameters and a number of possible tests of linkage may result. First, the full model containing male and female QTL estimates may be fitted, and both parameters dropped from the model simultaneously. Minus twice the log-likelihoods of the full model and the nonscalar sex-limitation model are compared using a likelihood ratio chi-square test. This provides a test of sex-limitation linkage with a test statistic distributed as approximately $^1/_4\chi^2_0 : {}^1/_2\chi^2_1 : {}^1/_4\chi^2_2$ (confirmed by simulation of 10,000 replicates of 600 sib-pairs, 100 of each dizygotic [DZ] sex pairing, male–male [MM], female–female [FF], FM and MF, plus 100 MZ pairs of each sex at an unlinked marker). Second, the full model may be compared with a model in which the QTL estimates of males and females have been equated, providing a test of sex limitation or heterogeneity in the QTL estimates at a given locus. When this test is conducted in the presence of linkage, the test statistic is asymptotically distributed as approximately χ^2_1.

A third possible test of linkage comes from comparing the fit of the model in which the QTL parameters of males and females are equal with the fit of a null model of nonscalar sex limitation. In the case of a sex-limited trait this will provide the more familiar test of linkage against the correctly specified polygenic background, with a test statistic distributed as approximately $^1/_2\chi^2_0 : {}^1/_2\chi^2_1$. Hypotheses regarding sex-specific QTL could also be tested by incorporating an additional sex-specific QTL parameter resulting in a general sex-limitation QTL model. The three tests are summarized in Table 1. From a practical point of view, a sex-limited linkage analysis would only be recommended if sex limitation had been found in the data under analysis, and testing for heterogeneity at the QTL would only be recommended in the presence of linkage. Comparison of the results of the sex-limited and non sex-limited linkage analyses, conducted across the whole genome may provide information to aid in the interpretation of linkage findings. Sex limitation of QTL effects may be expected to decrease the size of the QTL detected using a non sex-limitation model. Thus, it is likely that only fitting a sex-limitation model to peaks detected using a non sex-limitation model may result in overlooking the presence of sex-limited QTL in other regions.

Towne et al. (1997) explored the increase in power associated with modeling a genotype-by-sex interaction, within the context of analyzing data from the Genetic Analysis Workshop 10. They concluded that considering sex limitation when modeling linkage 'increases power to detect linkage … even when there is only a relatively modest amount of G × S interaction' (Towne et al., 1997, p. 1057). In addition, the increased information regarding QTL action may inform the design of subsequent fine-mapping and

Table 1

Likelihood Ratio Tests for Autosomal Sex-QTL Interaction

Test	Null hypothesis	Alternative hypothesis	Asymptotic distribution of test statistic
Sex-limited QTL	$\sigma_{Qf} = \sigma_{Qm} = 0$	$\sigma_{Qf} > 0; \sigma_{Qm} > 0$	$^1/_4\chi^2_0 : {}^1/_2\chi^2_1 : {}^1/_4\chi^2_2$
Heterogeneity at QTL	$\sigma_{Qf} = \sigma_{Qm} \neq 0$	$\sigma_{Qf} \neq \sigma_{Qm}$	χ^2_1
Non sex-limited QTL	$\sigma_{Qf} = \sigma_{Qm} = 0$	$\sigma_{Qf} = \sigma_{Qm} > 0$	$^1/_2\chi^2_0 : {}^1/_2\chi^2_1$

https://doi.org/10.1375/twin.8.6.569 Published online by Cambridge University Press

association studies. From a practical point of view, these analyses are of course more computationally intensive than an analysis that does not include sex limitation. However, in our experience we have found that the results obtained using this method can help to elucidate the underlying biology of the phenotype, both in cases where sex-limited QTLs are identified (Morley et al., in press) and when no such QTLs are detected (Cornes et al., 2005) in the analyses of traits shown to exhibit sex limitation.

Acknowledgments

The author would like to thank Katherine Morley, Belinda Cornes, Peter Visscher, Nick Martin, David Duffy, Stuart McGregor and two anonymous reviewers for their useful comments.

The author was supported in part by a Thenie Baddams Grant from the Australian Federation of University Women — SA Inc.

References

Abecasis, G. R., Cherny, S. S., Cookson, W. O., & Cardon, L. R. (2002). Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genetics, 30*, 97–101.

Cornes, B. K., Medland, S. E., Ferreira, M. A. R., Morley, K. I., Duffy, D. L., Heijmans, B., Montgomery, G. W., & Martin, N. G. (2005). Sex-limited genome-wide linkage scan for body mass index in an unselected sample of 933 Australian twin families. *Twin Research and Human Genetics, 8,* 616–632.

Eaves, L. J., Neale, M. C., & Maes, H. (1996). Multivariate multipoint linkage analysis of quantitative trait loci. *Behavior Genetics, 26*, 519–525.

Keski-Rahkonen, A., Viken, R. J., Kaprio, J., Rissanen, A., & Rose, R. J. (2004). Genetic and environmental factors in breakfast eating patterns. *Behavior Genetics, 34*, 503–514.

Medland, S. E. (2004). Alternate parameterization for scalar and non-scalar sex-limitation models in Mx. *Twin Research, 7*, 299–305.

Morley, K., Medland, S., Ferreira, M., Lynskey, M., Montgomery, G., Heath, A., Madden, P., & Martin, N. (in press). A smoking susceptibility locus on chromosome 11p13: Evidence from sex-limitation linkage analyses in a sample of Australian twin families. *Behavior Genetics*.

Mustanski, B. S., Viken, R. J., Kaprio, J., Pulkkinen, L., & Rose, R. J. (2004). Genetic and environmental influences on pubertal development: Longitudinal data from Finnish twins at ages 11 and 14. *Developmental Psychology, 40*, 1188–1198.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2002). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Posthuma, D., Beem, A. L., de Geus, E. J., van Baal, G. C., von Hjelmborg, J. B., Iachine, I., & Boomsma, D. I. (2003). Theory and practice in quantitative genetics. *Twin Research, 6*, 361–376.

Purcell, S., & Sham, P. (2002). Variance components models for gene–environment interaction in quantitative trait locus linkage analysis. *Twin Research, 5*, 572–576.

Rose, R. J., Dick, D. M., Viken, R. J., Pulkkinen, L., & Kaprio, J. (2004). Genetic and environmental effects on conduct disorder and alcohol dependence symptoms and their covariation at age 14. *Alcoholism Clinical and Experimental Research, 28*, 1541–1548.

Seaton, G., Haley, C. S., Knott, S. A., Kearsey, M., & Visscher, P. M. (2002). QTL Express: Mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics, 18,* 339–340.

Towne, B., Siervogel, R. M., & Blangero, J. (1997). Effects of genotype-by-sex interaction on quantitative trait linkage analysis. *Genetic Epidemiology, 14*, 1053–1058.

Van Hulle, C. A., Goldsmith, H. H., & Lemery, K. S. (2004). Genetic, environmental, and gender effects on individual differences in toddler expressive language. *Journal of Speech Language and Hearing Research, 47*, 904–912.

Vierikko, E., Pulkkinen, L., Kaprio, J., & Rose, R. J. (2004). Genetic and environmental influences on the relationship between aggression and hyperactivity-impulsivity as rated by teachers and parents. *Twin Research, 7*, 261–274.

Zeegers, M. P., Rijsdijk, F., Sham, P., Fagard, R., Gielen, M., De Leeuw, P. W., & Vlietinck, R. (2004). The contribution of risk factors to blood pressure heritability estimates in young adults: The East Flanders prospective twin study. *Twin Research, 7*, 245–253.

Appendix

An example Mx script for use with continuous data. A detailed description of a non-QTL sex-limitation script using this style of parameterization is given in Medland (2004).

Nonscalar Sex-Limitation Model

```
#define nsib 4     !size of largest sib-ship — in this case 4
! Zygosity information is included in the data file as z1 and z2 and contains the kinship
! coefficients for the additive genetic effects
! z1 contains 1 if the family is MZ and .5 if the family is DZ
! z2 contains .5 for all families
! The four siblings are referred to as a b c d with traita denoting the phenotype of sibling
! a etc.

!G1: Data from the twins and their siblings
Data NInput_vars = 21 Ngroups = 1
missing = -9.00
Rectangular File = marker1
Labels
family z1 z2 sexa sexb sexc sexd AGEa AGEb AGEc AGEd
traita traitb traitc traitd pihat_ab pihat_ac pihat_bc pihat_ad pihat_bd pihat_cd
Select fam z1 z2 sexa sexb sexc sexd agea ageb agec aged traita traitb traitc traitd pihat_ab pihat_ac pihat_bc
pihat_ad pihat_bd pihat_cd ;
Definition fam z1 z2 sexa sexb sexc sexd agea ageb agec aged pihat_ab pihat_ac pihat_bc pihat_ad pihat_bd
pihat_cd ;
Begin Matrices ;
X Lower 1 1 Free          ! male genetic structure
Y Lower 1 1 Free          ! male common environmental structure
Z Lower 1 1 Free          ! male specific environmental structure
K Lower 1 1 Free          ! male specific genetic structure
G Full 1 1 Free           ! male qtl
U Lower 1 1 Free          ! female genetic structure
V Lower 1 1 Free          ! female common environmental structure
W Lower 1 1 Free          ! female specific environmental structure
F Full 1 1 Free           ! female qtl
M Full nsib 1 Free        ! means
S Full nsib 1             ! sex definition variables
I Full 1 nsib             ! age definition variables
B Full 1 2 Free           ! sex correction for means model
J Unit nsib 1             ! Matrix of 1s to create sex correction
H Stand nsib nsib         ! specifies the kinship between sibs
R Stand nsib nsib         ! pihats
End matrices                  ;
Specify S sexa sexb sexc sexd ;
Specify H z1 z2 z2 z2 z2 z2;
Specify I agea ageb agec aged ;
Specify R pihat_ab pihat_ac pihat_bc pihat_ad pihat_bd pihat_cd ;
Begin Algebra;
A = (U*U')+(V*V')+(W*W')+(F*F');                ! FEMALE VARIANCE
C = (X*X')+(Y*Y')+(Z*Z')+(G*G')+(K*K');         ! MALE VARIANCE
P = J-S;
O = \v2d(S');
T = \v2d(P');
D = A| (U*U')%A|(V*V')%A|(W*W')%A|(F*F')%A| C| (X*X')%C|(Y*Y')%C|(Z*Z')%C|(G*G')%C
|(K*K')%C ;
E = A| (U*U')|(V*V')|(W*W')|(F*F')|C| (X*X')|(Y*Y')|(Z*Z')|(G*G')|(K*K') ;
End Algebra ;
Means M +(B*(S'_I))' ;
Covariances       (((S@X+(P@U))* (S@X+(P@U))').H)
                  +(((S@K)* (S@K)').H)
```


https://doi.org/10.1375/twin.8.6.569 Published online by Cambridge University Press

```
                +((S@Y+(P@V))* (S@Y+(P@V))')
                +(((S@G+(P@F))* (S@G+(P@F))').R)
                + O@(Z*Z')+T@(W*W') ;
Start 130 M 1 1 — M nsib 1
Start 4 Y 1 1 V 1 1
Start 6 X 1 1 U 1 1 K 1 1
Start 4 Z 1 1 W 1 1
Bo 0 20 Y 1 1 V 1 1 X 1 1 U 1 1 K 1 1 Z 1 1 W 1 1
Labels Column D F_Var F_A F_C F_E F_Q M_Var M_A M_C M_E M_Q M_R
Labels Column E F_Var F_A F_C F_E F_Q M_Var M_A M_C M_E M_Q M_R
Labels Row D Standardised_estimates
Labels Row E Unstandardised_estimates
Options Multiple issat
END
Equate G 1 1 1 F 1 1 1
! Equate Male and Female QTL
End
Drop G 1 1 1 F 1 1 1
! Null model to which:
! Model 1 can be compared to test for sex-limited linkage
! Model 2 can be compared to test for non sex-limited linkage
End
```


https://doi.org/10.1375/twin.8.6.569 Published online by Cambridge University Press