

Proceedings

Open Access

## X chromosome effects and their interactions with mitochondrial effects

Jack W Kent Jr\*, Loren R Lease, Michael C Mahaney, Thomas D Dyer, Laura Almasy and John Blangero

Address: Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX 78245, USA

Email: Jack W Kent\* - [jkent@darwin.sfbr.org](mailto:jkent@darwin.sfbr.org); Loren R Lease - [llease@darwin.sfbr.org](mailto:llease@darwin.sfbr.org); Michael C Mahaney - [mmahaney@darwin.sfbr.org](mailto:mmahaney@darwin.sfbr.org); Thomas D Dyer - [tdyer@darwin.sfbr.org](mailto:tdyer@darwin.sfbr.org); Laura Almasy - [almasy@darwin.sfbr.org](mailto:almasy@darwin.sfbr.org); John Blangero - [john@darwin.sfbr.org](mailto:john@darwin.sfbr.org)

\* Corresponding author

from Genetic Analysis Workshop 14: Microsatellite and single-nucleotide polymorphism Noordwijkerhout, The Netherlands, 7-10 September 2004

Published: 30 December 2005

BMC Genetics 2005, 6(Suppl 1):S157 doi:10.1186/1471-2156-6-S1-S157

### Abstract

We report a simple and rapid method for detecting additive genetic variance due to X-linked loci in the absence of marker data for this chromosome. We examined the interaction of this method with an established method for detecting mitochondrial linkage (another source of sex-asymmetric genetic covariance). When applied to data from the Collaborative Study on the Genetics of Alcoholism, this method found evidence of X-chromosomal linkage for one continuous trait (ntth1) and one discrete trait (SPENT). Evidence of mitochondrial contribution was found for one discrete trait (CRAVING) and three continuous traits (ln(CIGPKYR), ecb21, and tth1). Results for ntth1 suggest that methods that do not also allow for male-female heterogeneity in environmental variance may be overly conservative in detection of X-chromosomal effects.

### Background

Although the X chromosome is one of the largest in the human genome, genetic analysis software packages have been slow to incorporate routines for X-linked loci. This gap is closing: experimental two-point X-linkage analysis routines are now available in MERLIN [1] and SOLAR [2], and Ekstrøm [3] recently proposed an algorithm for multipoint linkage analysis of the X. Even prior to localizing quantitative trait loci within the X, however, it may be useful to determine the additive genetic effect of the X as a whole, both to allow the investigator to decide whether to pursue a more detailed analysis of the chromosome, and to improve the specificity of variance components models for autosomal linkage.

In Ekstrøm [3], in part to avoid assumptions about presence or absence of dosage compensation, both the linkage model and the model for the residual additive genetic effect of the X are estimated in terms of separate variance

components for male-male, female-female, and male-female relative pairs. Here we offer a simplified mechanism for analyzing the 'X effect' as a single parameter (for ease of both computation and interpretation), even when marker data for this chromosome are unavailable. Our formulation derives from Bulmer's [4] dosage compensation model, which is applicable to the majority of X-chromosomal loci in humans and other eutherian mammals. Direct comparison of our simplified approach with that of Ekstrøm [3] is beyond the scope of this paper. However, based on a simulation study that does make this comparison by Kent et al. [5], we believe that the simplified method has equal or greater power to detect dosage compensated loci, and is conservative for non-lyonizing loci.

We used the method as a tool address the more general question of the interaction of sex-related effects in complex traits. Both X-linkage and mitochondrial linkage display asymmetric patterns of allele transmission from

**Table 1: Expected covariances for X-linked traits<sup>a</sup>**

Relationship type	General <sup>b,c</sup>	Dosage compensation
Male-male	$P_1\sigma^2_{Xm}$	$2P_1\sigma^2_{Xf}$
Female-female	$2\phi\sigma^2_{Xf}$	$2\phi\sigma^2_{Xf}$
Male-female	$(1/\sqrt{2})P_1\sigma_{Xm}\sigma_{Xf}$	$P_1\sigma^2_{Xf}$

<sup>a</sup>Expected additive genetic covariances due to the X chromosome without and with assumption of dosage compensation [4].

<sup>b</sup> $P_1$  and  $P_2$ , the respective probabilities that one and two alleles are shared IBD.

<sup>c</sup> $\phi$ , the autosomal (i.e., diploid) kinship coefficient =  $P_1/4 + P_2/2$ .

males and females, and either may be confounded with genotype × sex or environment × sex interaction effects on traits influenced by autosomal genes. A fully accurate specification of the variance components model for inheritance of a particular trait should distinguish among these possibilities. A companion study [6] examines possible mitochondrial contributions to a set of alcoholism-related phenotypes from the Collaborative Study on the Genetics of Alcoholism (COGA); here we extend this investigation to X-chromosomal effects.

**Methods**

**The COGA data**

We used behavioral and electrophysiological phenotypes in the subset of COGA data generously provided for the Genetic Analysis Workshop 14. The data collection process, including acquisition of written informed consent from all participants and approvals by Institutional Review Boards, are described more fully elsewhere [7].

**Effect of the X chromosome**

We have implemented an extension to the genetic analysis package SOLAR [2] to estimate the additive genetic effect of X-linked loci. For humans and many other placental mammals in which most X-linked loci are subject to random inactivation in females, it should be a reasonable first approximation to employ a simple dosage-compensation model [4]. The additive contribution of an X-linked locus is simply the effect of the single allele in the male, or the single active allele in a homozygous female. As an approximation, the effect of the locus in heterozygous females is just the average of the allele effects, because each female is a mosaic of cells in which one or the other X is inactivated at random. Under these assumptions, the mean X effect is the same in both sexes, while the variance due to the X in males is twice that in females. This linear relationship between the variances allows us to express the additive effect of the X in one sex in terms of the other. For example, in terms of the female variance  $\sigma^2_{Xf}$

$$\Omega = 2\Phi\sigma^2_g + 2\Psi\sigma^2_{Xf} + I_f\sigma^2_{ef} + I_m\sigma^2_{em} \quad (1)$$

where  $\Omega$  is the matrix of phenotypic correlations between pairs of individuals,  $\Phi$  is the matrix of *autosomal* kinship coefficients, and  $\Psi$  (using the notation of Ekström [3]) is a new matrix of X-chromosome-specific coefficients of relationship for male-male, female-female, and male-female pairs (Table 1). Like the elements of  $\Phi$ , the elements of  $\Psi$  can be calculated recursively from pedigree data, as by the algorithm in MINX (MERLIN in X [1,8]).

Inclusion of sex-specific identity matrices and random environmental variances ( $I_f\sigma^2_{ef}$  and  $I_m\sigma^2_{em}$ ) in Equation 1 is desirable because the different  $\sigma^2_{Xf}$  and  $\sigma^2_{Xm}$  may appear as different environmental variances in the restricted

**Table 2: Heritability estimates for selected discrete phenotypes**

	Model test	p-Value
DESIRE	POL vs. sporadic	$3.49 \times 10^{-6c}$
	MIT vs. POL	0.067
	XFX vs. POL	0.050
final model: <sup>a</sup>	$h^2_g: 0.48(0.12) + e^2: 0.52(0.12)$	
CRAVING	POL vs. sporadic	0.0003 <sup>c</sup>
	MIT vs. POL	0.005 <sup>c</sup>
	XFX vs. MIT	0.229
final model: <sup>a</sup>	$h^2_g: 0.00(0.14^b) + e^2: 0.81(0.05) + h^2_{mt}: 0.19(0.05)$	
SPENT	POL vs. sporadic	0.020 <sup>d</sup>
	MIT vs. POL	0.356
	XFX vs. POL	0.034 <sup>d</sup>
final model: <sup>a</sup>	$h^2_g: 0.00(0.05^b) + e^2: 0.94(0.02) + h^2_{Xf}: 0.06(0.02)$	

<sup>a</sup>Estimates (SE) of parameters retained in the final model (the final estimate of  $h^2_g$  is given even when = 0).  $h^2_g$ : autosomal additive genetic heritability;  $h^2_{mt}$ : heritability due to the mitochondrion;  $h^2_{Xf}$ : X-linked heritability in females;  $e^2$ , environmental component.

<sup>b</sup>One-tailed empirical SEs estimated for point estimates at the boundary.

<sup>c</sup> $p < 0.01$

<sup>d</sup> $p < 0.05$

**Table 4: Nested models for traits LNCIGP and nthl**

Trait	Log-likelihood	$\sigma^2_g$	$\sigma^2_e$	$\sigma^2_{ef}$	$\sigma^2_{em}$	$\sigma^2_{mt}$	$\sigma^2_{xf}$
<b>LNCIGP</b>							
A-polygenic	-1048.50	0.54	1.53	--	--	--	--
B-A+ diff $\sigma^2_e$	-1047.28	0.52	-- <sup>a</sup>	1.66	1.40	--	--
C-all parameters	-1045.51	0.35	--	1.72	1.44	0.15	0.00
D-eq. $\sigma^2_e$	-1046.91	0.36	1.58	--	--	0.15	0.00
E- $\sigma^2_{mt} = 0$	-1047.24	0.49	--	1.67	1.38	--	0.03
F- $\sigma^2_{xf} = 0$	-1045.51	0.34	--	1.72	1.44	0.15	--
<b>nthl</b>							
A-polygenic	224.36	0.07	0.15	--	--	--	--
B-A+ diff $\sigma^2_e$	227.06	0.07	--	0.17	0.13	--	--
C-all parameters	229.40	0.02	--	0.19	0.12	0.01	0.03
D-eq. $\sigma^2_e$	225.33	0.05	0.16	--	--	0.02	0.00
E- $\sigma^2_{mt} = 0$	229.14	0.02	--	0.19	0.11	--	0.03
F- $\sigma^2_{xf} = 0$	228.27	0.04	--	0.18	0.13	0.02	--
G-A + $\sigma^2_{mt}$	225.33	0.05	0.16	--	--	0.02	--
H-A + $\sigma^2_{xf}$	224.37	0.07	0.15	--	--	--	0.00

<sup>a</sup>--, parameter not estimated.

model where  $\sigma^2_{xf} = \sigma^2_{xm} = 0$ , and also to account for genotype  $\times$  sex or environment  $\times$  sex interactions, if any. An important exception is that discrete traits (e.g., CRAVING) are analyzed using heritability parameterization with the phenotypic variance in both sexes fixed at 1 (because the variance-components model for discrete traits assumes an underlying liability with distribution  $N(0, 1)$ ).

The significance test is straightforward, testing the ratio of the likelihood of the model with the single added parameter  $\sigma^2_{xf}$  to that of the restricted model. Because  $\sigma^2_{xf}$  is tested on its lower bound, the likelihood ratio statistic is distributed as a 1/2:1/2 mixture of  $\chi^2$  with one degree of freedom, and a point mass at zero [9].

**Mitochondrial effect**

We used the mitochondrial identity-by-descent (IBD) calculation implemented in SOLAR [10] (and see discussion in Lease et al. [6]). The additive genetic variance due to the mitochondrion can be added to the basic polygenic model as

$$\Omega = 2\Phi\sigma^2_g + M\sigma^2_{mt} + I_f\sigma^2_{ef} + I_m\sigma^2_{em} \quad (2)$$

where **M** is the mitochondrial IBD matrix. For comparison to the X effects, we retain separate environmental variance terms for males and females (except for the discrete traits). The significance of the model that includes the mitochondrial variance is tested against the restricted model ( $\sigma^2_{mt} = 0$ ) by a likelihood ratio test.

**Results and Discussion**

In the following, evidence of mitochondrial or X effect is listed as significant at  $p < 0.05$ , without correction for multiple testing (the traits are genetically and phenotypically

correlated [6] and measured in the same individuals). Note, however, that a modified Bonferroni correction that accommodates these genetic correlations, described by Lease et al. [6], yields a more stringent experiment wise type I error rate = 0.01.

**Mitochondrial and X effects in discrete traits**

Phenotypes include 12 discrete behavioral traits related to alcohol consumption [7]. Two traits (BINGE and MORNING) did not show significant heritability in the polygenic model and were not considered further. Of the remaining traits (SMOKER, DESIRE, CRAVING, SPENT, NARROW, GAVEUP, BLACKOUT, WITHDRAWAL, HEALTH, and PSYCHO), data are shown only for those whose model likelihoods were increased by the addition of mitochondrial or X-effect terms (DESIRE, which showed a marginally significant X effect, is also shown). The sequence of analysis was: estimation of a polygenic model (POL); addition of a mitochondrial heritability term (MIT); addition of a heritability term for the additive genetic effect of the X (XFX). At each step the more restricted model was retained unless the model with the added parameter had a likelihood that was significantly greater at  $p < 0.05$ . Table 2 reports the  $p$ -values for the successive models and point estimates of the parameters retained in the final model for each trait. Trait SPENT showed a significant effect ( $p = 0.034$ ) of the X chromosome, but the size of this effect (6% of total variance) is small. CRAVING showed a significant mitochondrial effect ( $p = 0.005$ ), as reported [6], but no significant effect of the X chromosome. In each case the new parameter accounts for the entire additive genetic effect (i.e.,  $h^2_g = 0$ ) in the final model.

**Table 3: Heritability estimates for selected discrete phenotypes retained in the final model**

	Model test	p-Value
LNMXDK	POL vs. sporadic	1.37 × 10 <sup>-17c</sup>
	EMF vs. POL	0.008 <sup>c</sup>
	MIT vs. EMF	0.418
	XFX vs. EMF	0.500
final model: <sup>a</sup> ttdt2	$\sigma_g^2: 0.22(0.00^b) + \sigma_{ef}^2: 0.32(0.00^b) + \sigma_{em}^2: 0.42(0.00^\dagger)$	
	POL vs. sporadic	6.48 × 10 <sup>-12c</sup>
	EMF vs. POL	0.000 <sup>c</sup>
	MIT vs. EMF	0.136
final model: <sup>a</sup> LNCIGP	$\sigma_g^2: 0.28(0.00^b) + \sigma_{ef}^2: 0.70(0.00^b) + \sigma_{em}^2: 0.40(0.00^\dagger)$	
	POL vs. sporadic	5.39 × 10 <sup>-14c</sup>
	EMF vs. POL	0.059
	MIT vs. POL	0.037 <sup>d</sup>
final model: <sup>a</sup> ecb21	$\sigma_g^2: 17.55(9.24) + \sigma_e^2: 333.79(0.29) + \sigma_{mt}^2: 55.06(2.46)$	
	POL vs. sporadic	7.34 × 10 <sup>-35c</sup>
	EMF vs. POL	0.002 <sup>b</sup>
	MIT vs. EMF	0.012 <sup>d</sup>
final model: <sup>a</sup> ttth1	$\sigma_g^2: 10.76(0.21) + \sigma_{ef}^2: 17.22(0.07) + \sigma_{em}^2: 10.50(0.07) + \sigma_{mt}^2: 3.24(0.21)$	
	POL vs. sporadic	6.25 × 10 <sup>-11c</sup>
	EMF vs. POL	0.149
	MIT vs. POL	0.030 <sup>d</sup>
final model: <sup>a</sup> ntth1	$\sigma_g^2: 0.08(0.01) + \sigma_e^2: 0.28(0.00^b) + \sigma_{mt}^2: 0.04(0.00^b)$	
	POL vs. sporadic	5.93 × 10 <sup>-9c</sup>
	EMF vs. POL	0.010 <sup>d</sup>
	MIT vs. EMF	0.060
final model: <sup>a</sup>	$\sigma_g^2: 0.02(0.01) + \sigma_{ef}^2: 0.19(0.00^b) + \sigma_{em}^2: 0.11(0.00^b) + \sigma_{xf}^2: 0.03(0.00^b)$	0.021 <sup>d</sup>

<sup>a</sup>Estimates (SE) of parameters retained in the final model (the final estimate of  $h_g^2$  is given even when = 0).  $\sigma_g^2$ , autosomal additive genetic variance;  $\sigma_{mt}^2$ , additive genetic variance due to the mitochondrion;  $\sigma_{xf}^2$ , X-linked additive genetic variance in females;  $\sigma_{ef}^2$ , random environmental variance in females;  $\sigma_{em}^2$ , random environmental variance in males.

<sup>b</sup>SEs reported as (0.00) are nonzero but <0.01.

<sup>c</sup> $p < 0.01$

<sup>d</sup> $p < 0.05$

**Mitochondrial and X effects in continuous traits**

We examined two behavioral traits, MAXDRINKS and CIGPKYR. Both measures were leptokurtic and were log<sub>e</sub>-transformed: κ before(after) transformation: MAXDRINKS 9.18 (0.06); CIGPKYR 9.69 (-1.46). We also examined 13 electrophysiology traits (ecb21, ntth1-4, ttdt1-4, ttth1-4). All showed a significant additive genetic component; data are shown only for those traits whose model likelihoods were increased by addition of more parameters. The sequence of analysis was: estimation of a (model POL); estimation of a polygenic model with separate environmental variance terms for the two sexes (EMF); addition of a mitochondrial variance term (MIT); addition of a variance term for the additive genetic effect of the X (XFX). At each step the more restricted model was retained unless the model with the added parameter had a significantly greater likelihood at  $p < 0.05$ . Table 3 reports the  $p$ -values for the successive models and the esti-

mates of the parameters retained in the final model for each trait. Two traits (LNMXDK and ttdt2) had significantly different environmental variances in males and females, although neither showed significant mitochondrial or X effects. Three traits (LNCIGP, ecb21, and ttth1) showed evidence for mitochondrial effect. One trait (ntth1) showed a significant effect ( $p = 0.021$ ) of the X.

**Interaction of sex-related effects**

The initial screen of traits followed a step-wise procedure in which each additional parameter was retained in subsequent analyses only if it significantly increased the likelihood of the model. We were curious if estimating other combinations of parameters might alter the conclusions from this analysis. Table 4 presents results from a series of nested models for continuous traits LNCIGP and ntth1. (In the following, models A, B, etc. for each trait refer to their listing in Table 4.)

**Table 5: Variance estimates for selected continuous phenotypes in the final model**

Trait	Estimate (SE) of each parameter retained in the final model <sup>a</sup>					
	POL		EMF		MIT	XFX
	$\sigma^2_g$	$\sigma^2_e$	$\sigma^2_{ef}$	$\sigma^2_{em}$	$\sigma^2_{mt}$	$\sigma^2_{Xf}$
LNMXDK	0.22 (0.00) <sup>b</sup>	-- <sup>c</sup>	0.32 (0.00)	0.42 (0.00)	--	--
ttd2	0.28 (0.00)	--	0.70 (0.00)	0.40 (0.00)	--	--
LNCIGP	17.55 (9.24)	333.79 (0.29)	--	--	55.06 (2.46)	--
ecb2l	10.76 (0.21)	--	17.22 (0.07)	10.50 (0.07)	3.24 (0.21)	--
ttth1	0.08 (0.01)	0.28 (0.00)	--	--	0.04 (0.00)	--
ntth1	0.02 (0.01)	--	0.19 (0.00)	0.11 (0.00)	--	0.03 (0.00)

<sup>a</sup> $\sigma^2_g$ , the autosomal additive genetic variance;  $\sigma^2_{mt}$ , the additive genetic variance due to the mitochondrion; and  $\sigma^2_{Xf}$ , the X-linked additive genetic variance in females;  $\sigma^2_{ef}$ , random environmental variances in females;  $\sigma^2_{em}$ , the random environmental variances in males.

<sup>b</sup>SEs reported as (0.00) are nonzero but <0.01.

<sup>c</sup>--, parameter not estimated.

LNCIGP showed a significant mitochondrial effect in the initial screen (Table 3). This result is robust; the effect is detectable whether male and female environmental variances are constrained to be the same (model D vs. A,  $p = 0.037$ ) or allowed to diverge (F vs. B,  $p = 0.030$ ). In model E ( $\sigma^2_{mt} = 0$ ) most of the variance actually due to the mitochondrion appears as residual additive genetic variance, with only minor misallocation of variance to the X.

A different pattern appears for ntth1, which showed a significant X effect in the initial screen. Interestingly, a significant X effect is detectable only if male and female environmental variances are allowed to differ (model E vs. B,  $p = 0.02$ ): the addition of  $\sigma^2_{Xf}$  to the polygenic model barely increases the model likelihood (model H vs. A,  $p = 0.44$ ). Different estimates of male and female environmental variance are retained in models that include the X effect, which may suggest genotype  $\times$  sex interactions. Exclusion of  $\sigma^2_{Xf}$  does not yield a falsely significant estimate of  $\sigma^2_{mt}$  (model F vs. model B,  $p = 0.06$ ).

**Conclusion**

In this study we have applied a new method for detecting additive genetic variance due to the X chromosome to a representative set of real data. (Simulation studies to validate the method will be reported elsewhere.) We find that real mitochondrial effects (estimated by an established method) are not confounded with X effects. However, it appears that a true X effect can appear as heterogeneous environmental variance in males and females; indeed, in one case (ntth1), detection of an X effect *depended* on modelling this heterogeneity. This suggests that tests of X-chromosomal effect that constrain males and females to have equal environmental variances (as, for example, analyses of discrete traits) may be overly conservative.

**Abbreviations**

COGA: Collaborative Study on the Genetics of Alcoholism

EMF: Model estimating separate environmental variances by sex

IBD: Identity-by-descent

MIT: Model estimating mitochondrial variance component

POL: polygenic model

XFX: Model estimating X-linked variance component

**Authors' contributions**

JWK, TDD, and JB developed the methodology for analyzing the effect of the X chromosome. JWK conceived the study, performed the X-effect analyses, and drafted the manuscript. LRL and MCM performed the mitochondrial analyses. MCM and LA provided critical reviews of the methodology. TDD and LA coordinated preparation of the COGA data for investigators at SFBR. All authors have read and approved this manuscript.

**Acknowledgements**

This work was supported in part by NIH/NIMH grant MH059490-06 (JB, principal investigator). We appreciate comments on the first draft by two anonymous reviewers.

**References**

1. Abecasis GR, Cherny SS, Cookson WO, Cardon LR: **Merlin-rapid analysis of dense genetic maps using sparse gene flow trees.** *Nat Genet* 2002, **30**:97-101.
2. Almasy L, Blangero J: **Multipoint quantitative-trait linkage analysis in general pedigrees.** *Am J Hum Genet* 1998, **62**:1198-1211.
3. Ekström CT: **Multipoint linkage analysis of quantitative traits on sex-chromosomes.** *Genet Epidemiol* 2004, **26**:218-230.

4. Bulmer MG: *The Mathematical Theory of Quantitative Genetics* New York: Oxford University Press; 1985.
5. Kent JW Jr, Dyer TD, Blangero J: **Estimating the additive genetic effect of the X chromosome.** *Genet Epidemiol* in press.
6. Lease LR, Winnier DA, Williams JT, Dyer TD, Almasy L, Mahaney MC: **Mitochondrial genetic effects on latent class variables associated with susceptibility to alcoholism.** *BMC Genet* 2005, **6(Suppl 1)**:S158.
7. Begleiter H, Reich T, Nurnberger J Jr, Li TK, Conneally PM, Edenberg H, Crowe R, Kuperman S, Schuckit M, Bloom F, Hesselbrock V, Porjesz B, Cloninger CR, Rice J, Goate A: **Description of the Genetics Analysis Workshop II Collaborative Study on the Genetics of Alcoholism.** *Genet Epidemiol* 1999, **17(Suppl 1)**:S25-S30.
8. **MINX: Chromosome X Analyses** [<http://www.sph.umich.edu/csg/abecasis/Merlin/reference.html>]. *The University of Michigan Center for Statistical Genetics MERLIN Reference Sheet*
9. Self SG, Liang K-Y: **Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions.** *J Am Stat Assoc* 1987, **82**:605-610.
10. Czerwinski SA, Williams JT, Demerath EV, Towne B, Siervogel RM, Blangero J: **Does accounting for mitochondrial genetic variation improve the fit of genetic models?** *Genet Epidemiol* 2001, **21(Suppl 1)**:S779-S782.

Publish with **BioMed Central** and every scientist can read your work free of charge

*"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."*

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

