genes in which the breakpoint lies between the two primers used in the LightCycler PCR.

MARKUS FELDKÖTTER,[1] VERENA SCHWARZER,[1] RADU WIRTH,[2] THOMAS F. WIENKER,[3] AND BRUNHILDE WIRTH[1]

[1]*Institute of Human Genetics,* [2]*Department of Surgery, and* [3]*Institute for Medical Biometry, Informatics and Epidemiology, University Clinic, Bonn*

References

Andreassi C, Jarecki J, Zhou J, Coovert DD, Monani UR, Chen X, Whitney M, Pollok B, Zhang M, Androphy E, Burghes AH (2001) Aclarubicin treatment restores SMN levels to cells derived from type I spinal muscular atrophy patients. Hum Mol Genet 10:2841–2849

Chang JG, Hsieh-Li HM, Jong YJ, Wang NM, Tsai CH, Li H (2001) Treatment of spinal muscular atrophy by sodium butyrate. Proc Natl Acad Sci USA 98:9808–9813

Feldkötter M, Wirth R, Schwarzer V, Wirth B (2002) Quantitative analysis of SMN1 and SMN2 based on real-time LightCycler PCR: fast and highly reliable heterozygosity testing and prediction of severity of spinal muscular atrophy. Am J Hum Genet 70:358–368

Hahnen E, Schönling J, Rudnik-Schöneborn S, Zerres K, Wirth B (1996) Hybrid survival motor neuron genes in patients with autosomal recessive spinal muscular atrophy: new insights into molecular mechanisms responsible for the disease. Am J Hum Genet 59:1057–1065

Ogino S, Wilson RB (2002) *SMN* dosage analysis and risk assessment for spinal muscular atrophy. Am J Hum Genet 70:1596–1598 (in this issue)

Wirth B, Herz M, Wetter A, Moskau S, Hahnen E, Rudnik-Schöneborn S, Wienker T, Zerres K (1999) Quantitative analysis of survival motor neuron copies: identification of subtle *SMN1* mutations in patients with spinal muscular atrophy, genotype-phenotype correlation, and implication for genetic counseling. Am J Hum Genet 64:1340–1356

Address for correspondence and reprints: Dr. Brunhilde Wirth, Institute of Human Genetics, University Clinic, Wilhelmstrasse 31, D-53111 Bonn, Germany. E-mail: bwirth@uni-bonn.de

0002-9297/2002/7006-0025$15.00

Am. J. Hum. Genet. 70:1599–1602, 2002

The Power of Multivariate Quantitative-Trait Loci Linkage Analysis Is Influenced by the Correlation between Variables

To the Editor:

In a recent article, Sham et al. (2000) investigated the power of variance-components linkage analysis by deriving an analytic expression for the noncentrality parameter (NCP) of the linkage test. The authors demonstrated that the NCP—and, hence, the power of the test to detect linkage—was determined primarily by the square of the additive and dominance genetic components of variance due to the quantitative-trait locus (QTL) and by the residual correlation between siblings. However, Sham et al. presented calculations for the univariate case only. Recently, it has been demonstrated that the power of QTL linkage analysis may be increased by use of multivariate techniques that analyze the pleiotropic action of the QTL on several variables (Boomsma 1996; Martin et al. 1997). In particular, the power of multivariate linkage analysis is strongly influenced by the correlation between the variables, being greatest when the QTL induces covariation between the variables in the direction opposite to the residual correlation (Allison et al. 1998; Amos et al. 2001). Here, I follow the methodology of Sham et al., to demonstrate analytically, for the first time, how the power of a bivariate variance-components linkage analysis depends not only on the magnitude and direction of the correlation between variables but also on the source of this correlation.

The relationship between two observable variables is parameterized in terms of the path model displayed in figure 1. Observed variables for each sib pair (*square boxes*) are due to the combined action of several latent variables (*circles*), including a pleiotropic QTL (Q), poly-

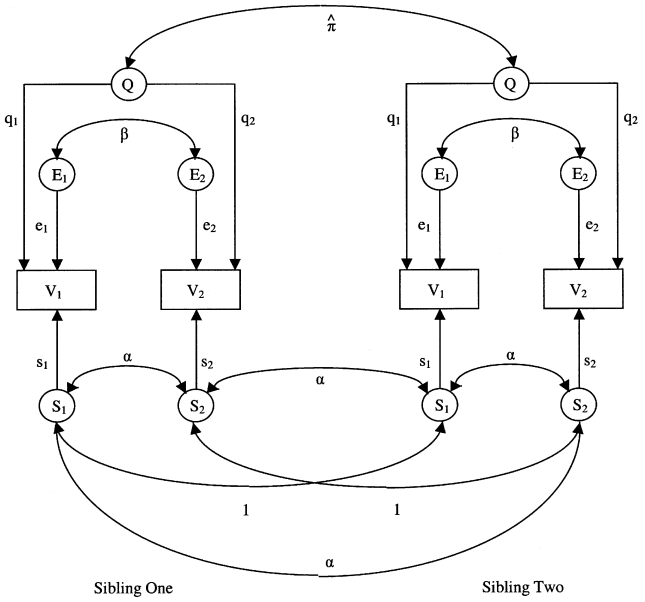

Figure 1 Path diagram showing the relationship between two observed variables (V_1 and V_2) for a pair of siblings. Covariation between the phenotypes is due to the QTL (Q), genetic and environmental sources that are shared among siblings (S_1 and S_2), and nonshared sources of variation (E_1 and E_2).

genic and environmental effects common to each member of the sib pair (S_1 and S_2), and unique environmental influences specific to each sibling (E_1 and E_2). Causal paths between variables are represented by unidirectional arrows, whereas correlations between variables are represented by bidirectional arrows. The strength of association between each variable is measured by a path coefficient (equivalent to a partial regression coefficient), in the case of a causal path, or a correlation coefficient, in the case of a bidirectional path. The correlation between siblings for the common QTL is $\hat{\pi}$, the estimated proportion of genes shared identical by descent at the trait locus, whereas the correlation between siblings for shared polygenic and environmental sources (i.e., S_1 and S_2) is 1. Correlations between phenotypes arise because of the pleiotropic action of the QTL (represented by the product of the path coefficients q_1 and q_2), from polygenic and environmental effects shared between siblings (represented by the product of α, s_1, and s_2) and from nonshared residual effects (represented by the product of β, e_1, and e_2). It is assumed that each variable is standardized to unit variance. The test for linkage is computed as twice the difference in log-likelihood between a model where q_1 and q_2 are estimated and a model where q_1 and q_2 are constrained to 0. Since q_1 (or, alternatively, q_2) is constrained to be positive, whereas q_2 has no such constraint (to allow for the possibility of a negative correlation between the observed variables), the test statistic is distributed asymptotically as a 50:50 mixture of χ^2_1 and χ^2_2 (Self and Liang 1987).

Under the null hypothesis of no linkage (N), the asymptotic parameter estimates for the covariance matrix, implied by figure 1, of the ith sib pair are

$$\Sigma_{i_N} = \begin{matrix} 1 & & & \\ q_1q_2 + \alpha s_1s_2 + \beta e_1e_2 & 1 & & \\ \frac{q_1^2}{2} + s_1^2 & \frac{q_1q_2}{2} + \alpha s_1s_2 & 1 & \\ \frac{q_1q_2}{2} + \alpha s_1s_2 & \frac{q_2^2}{2} + s_2^2 & q_1q_2 + \alpha s_1s_2 + \beta e_1e_2 & 1 \end{matrix}$$

(only lower elements of the matrix are shown). Under the alternative hypothesis of linkage (L), the asymptotic parameter estimates are given by:

$$\Sigma_{i_L} = \begin{matrix} 1 & & & \\ q_1q_2 + \alpha s_1s_2 + \beta e_1e_2 & 1 & & \\ \hat{\pi}_i q_1^2 + s_1^2 & \hat{\pi}_i q_1q_2 + \alpha s_1s_2 & 1 & \\ \hat{\pi}_i q_1q_2 + \alpha s_1s_2 & \hat{\pi}_i q_2^2 + s_2^2 & q_1q_2 + \alpha s_1s_2 + \beta e_1e_2 & 1 \end{matrix}$$

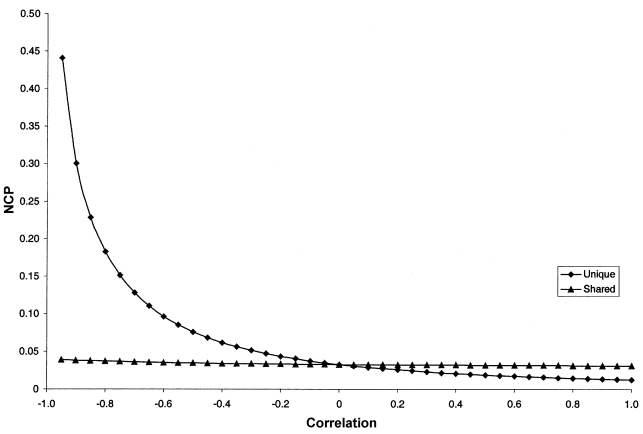

Figure 2 NCP as a function of either the correlation between unique sources of variation (*lines with diamonds*) or the correlation between shared sources of variation (*lines with triangles*).

According to Sham et al. (2000), the NCP for linkage (λ_L) is equal to twice the difference in expected log-likelihoods between the alternative and null hypotheses:

$$\begin{aligned}\lambda_L &= E(2\ln L_L) - E(2\ln L_N) \\ &= \ln|\Sigma_N| - \frac{1}{4}\ln|\Sigma_{\pi=0}| - \frac{1}{2}\ln|\Sigma_{\pi=0.5}| - \frac{1}{4}\ln|\Sigma_{\pi=1}| \ .\end{aligned}$$

To evaluate this expression, note that the determinant of a matrix of order n is a sum of $n!$ signed products, each involving n elements of the matrix. The computation is made easier, in the present case, because the variables are standardized and, therefore, the diagonal terms of the matrix are equal to 1:

$$\begin{aligned}|\Sigma| &= 1 + 2r_{21}r_{31}r_{32} + 2r_{21}r_{41}r_{42} + 2r_{31}r_{41}r_{43} \\ &\quad + 2r_{32}r_{42}r_{43} + r_{21}^2r_{43}^2 + r_{32}^2r_{41}^2 + r_{31}^2r_{42}^2 \\ &\quad - r_{21}^2 - r_{31}^2 - r_{32}^2 - r_{41}^2 - r_{42}^2 - r_{43}^2 \\ &\quad - 2r_{21}r_{32}r_{41}r_{43} - 2r_{21}r_{31}r_{42}r_{43} \\ &\quad - 2r_{31}r_{32}r_{41}r_{42} \ ,\end{aligned}$$

where r_{ij} is the element corresponding to the ith row and jth column of Σ. If we denote the right half of this equation as "$1 + x$" and note that the first-order Taylor-series expansion of $\ln(1 + x) \approx x$, then the NCP may be approximated as

$$\lambda_L \approx x_S - \frac{1}{4}x_2 - \frac{1}{2}x_1 - \frac{1}{4}x_0 \ ,$$

where x_S, x_2, x_1, and x_0 are the first-order Taylor-series approximations for the null hypothesis and the alternative

hypotheses of sharing two, one, or zero alleles identical by descent at the trait locus.

Evaluation of this expression in terms of the parameters in figure 1 yields

$$\begin{aligned}\lambda_L \approx & \frac{q_1^4}{8} + \frac{q_2^4}{8} + \frac{q_1^2 q_2^2}{4} + \frac{q_1^4 q_2^4}{2} \\ & - \frac{q_1^2 q_2^4}{2} - \frac{q_1^4 q_2^2}{2} - \frac{q_1^4 s_2^4}{8} \\ & - \frac{q_2^4 s_1^4}{8} - \frac{q_1^2 q_2^2 s_1^2 s_2^2}{4} \\ & + \frac{\alpha q_1 q_2^3 s_1 s_2 (q_1^2 + s_1^2 - 1)}{2} \\ & + \frac{\alpha q_1^3 q_2 s_1 s_2 (q_2^2 + s_2^2 - 1)}{2} \\ & + \frac{\beta q_1 q_2^3 e_1 e_2 (q_1^2 - 1)}{2} + \frac{\beta q_1^3 q_2 e_1 e_2 (q_2^2 - 1)}{2} \\ & + \alpha\beta q_1^2 q_2^2 s_1 s_2 e_1 e_2 + \frac{\beta^2 q_1^2 q_2^2 e_1^2 e_2^2}{2} \quad .\end{aligned}$$

Note particularly that the second part of the equation (i.e., the last four lines) contains terms involving the correlation between shared polygenic and environmental effects (α) and the correlation between unique environmental effects (β). The sign of these correlations contributes to the magnitude of the NCP. Consider first the terms containing the correlation between shared polygenic and environmental effects (i.e. the terms containing α). It is apparent that the parts of the expression inside parentheses must be negative. Therefore, if the QTL and shared polygenic and environmental influences produce correlations in the same direction, the terms will be negative, and therefore the NCP and the power to detect linkage will decrease. In contrast, when the QTL and shared influences induce correlations in opposite directions, the terms will become positive increasing the NCP and power. The power to detect linkage increases as the correlation between shared sources decreases (i.e., becomes more negative). A similar argument also applies to terms containing the QTL and unique sources of variation (i.e., the terms that include α), although the increase in power is more dramatic because the terms inside the parentheses are greater in magnitude and because there is an additional term containing β that is always positive (i.e., $1/2\beta^2 q_1^2 q_2^2 e_1^2 e_2^2$). The last term in the equation (i.e., $\alpha\beta q_1^2 q_2^2 s_1 s_2 e_1 e_2$) suggests that the increase in power will be greatest when both shared and nonshared influences induce covariance in the direction opposite to the QTL.

Figure 2 displays the effect that varying the correlation between shared and unique sources of variation has on the NCP for a plausible biological model. In this model, the QTL accounts for 20% of the variance of each trait (i.e., $q_1^2 = q_2^2 = 0.2$), and induces a positive correlation between the variables (i.e., q_1 and q_2 are both positive). Both shared and unique effects account for forty percent of the variance for both traits (i.e., $s_1^2 = s_2^2 = 0.4$; $e_1^2 = e_2^2 = 0.4$). The correlation between unique sources of variation is varied, while the shared correlation is fixed at 0 (*lines with diamonds*), and the correlation between shared factors is varied, whereas the unique environmental correlation is fixed at 0 (*lines with triangles*). Note that the graph is based on exact values for the NCP and not on the Taylor-series approximation.

In both cases, the NCP increases as the correlation between the latent sources of variation decreases. However, although the increase in NCP is small and linear for the shared case, the increase is dramatic and exponential as the correlation between the unique sources of variation decreases. Thus, the power of bivariate QTL linkage analysis depends not only on the phenotypic correlation between variables but also on the source of this correlation.

In conclusion, these results imply that, in a bivariate linkage analysis, one is most likely to detect a QTL that produces a correlation between variables opposite in direction to the background correlation. In particular, power is dramatically affected by the correlation between the unique environmental sources of variation. This combination of latent sources would tend to produce variables that have low or moderate phenotypic correlations, a fact that should be kept in mind when deciding which variables to include in a bivariate linkage analysis.

Acknowledgments

I would like to thank Dr. David Duffy for fruitful discussions and Dr. Nick Martin for helpful comments on the manuscript.

DAVID M. EVANS
Queensland Institute of Medical Research and Joint Genetics Program
University of Queensland
Brisbane

References

Allison DB, Thiel B, St Jean P, Elston RC, Infante MC, Schork NJ (1998) Multiple phenotype modeling in gene-mapping studies of quantitative traits: power advantages. Am J Hum Genet 63:1190–201

Amos C, de Andrade M, Zhu D (2001) Comparison of multivariate tests for genetic linkage. Hum Hered 51:133–144

Boomsma DI (1996) Using multivariate genetic modeling to detect pleiotropic quantitative trait loci. Behav Genet 26: 161–166

Martin N, Boomsma D, Machin G (1997) A twin-pronged attack on complex traits. Nat Genet 17:387–392

Self SG, Liang K (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J Am Stat Assoc 82:605–610

Sham PC, Cherny SS, Purcell S, Hewitt JK (2000) Power of linkage versus association analysis of quantitative traits, by use of variance-components models, for sibship data. Am J Hum Genet 66:1616–1630

Address for correspondence and reprints: Dr. David M. Evans, Queensland Institute of Medical Research, PO Royal Brisbane Hospital, Brisbane 4029, Australia. E-mail: davidE@qimr.edu.au

0002-9297/2002/7006-0026$15.00

Am. J. Hum. Genet. 70:1602, 2002

The National Institutes of Health Announces Online Availability of "Points to Consider When Planning a Genetic Study That Involves Members of Named Populations"

To the Editor:

The National Institutes of Health (NIH) has developed a guide for researchers, called "Points to Consider When Planning a Genetic Study That Involves Members of Named Populations." The NIH supports and encourages the concept and process of community consultation in many research areas and believes that investigators who are planning genetic-research projects involving members of named populations should consider whether and how those communities should be consulted. The new "Points to Consider" document describes what is meant by "community consultation"; presents situations in which community consultation should be considered; identifies potential benefits, both for researchers and for communities, that engagement in this process offers; and provides practical examples of how to plan a community consultation. The "Points to Consider" document is posted on the NIH Web site, at the URL given below.

It is increasingly important for researchers to realize that nonscientists may not be well versed in the scientific benefits resulting from genetics research. Individuals and the communities to which they belong may fear that participation in genetic studies involving named populations may end up stereotyping that particular named population, potentially putting the entire community at risk of discrimination by insurers or other third parties. In creating the "Points to Consider" document, the NIH aims to assist scientists in the design of studies that operate in variable social and cultural contexts and that yield meaningful data while they work with communities.

JUDITH H. GREENBERG
National Institute of General Medical Sciences
National Institutes of Health
Bethesda, MD

Electronic-Database Information

The URL for data in this letter is as follows:

Points to Consider When Planning a Genetic Study That Involves Members of Named Populations, http://www.nih.gov/sigs/bioethics/named_populations.html

Address for correspondence and reprints: Dr. Judith H. Greenberg, National Institutes of Health, Division of Genetics and Developmental Biology, Building 45, Room 2As25, 45 Center Drive, MSC-6200, Bethesda, MD 20892-6200. E-mail: greenbej@nigms.nih.gov

0002-9297/2002/7006-0027