Am. J. Hum. Genet. 72:1323–1330, 2003

Report

A Family-Based Test for Correlation between Gene Expression and Trait Values

Peter Kraft,[1] Eric Schadt,[4] Jason Aten,[2] and Steve Horvath[1,3]

Departments of [1]Biostatistics, [2]Biomathematics, and [3]Human Genetics, University of California at Los Angeles, Los Angeles; and [4]Department of Research Genetics, Rosetta Inpharmatics, Seattle

Advances in microarray technology have made it attractive to combine information on clinical traits, marker genotypes, and comprehensive gene expression from family studies to dissect complex disease genetics. Without accounting for family structure, methods that test for association between a trait and gene-expression levels can be misleading. We demonstrate that the standard unstratified test based on Pearson's correlation coefficient can produce spurious results when applied to family data, and we present a stratified family expression association test (FEXAT). We illustrate the utility of the FEXAT via simulation and an application to gene-expression data from lymphoblastoid cell lines from four CEPH families. The FEXAT has a smaller estimated false-discovery rate than the standard test when within-family correlations are of interest, and it detects biologically plausible correlations between beta catenin and genes in the *WNT*-activation pathway in humans that the standard test does not.

Traditional statistical techniques for meiotic mapping, such as linkage analysis and allelic association analysis, have been very successful at mapping Mendelian diseases—that is, rare diseases caused by rare, highly penetrant alleles at a single locus. They have been much less successful at mapping complex diseases, which result from the complex interplay of many loci and environmental factors (Clerget-Darpoux et al. 2001). In part, this is because genes influence disease via their gene products: mRNA and, farther downstream, proteins. These gene products are, in turn, influenced by other genes and environmental factors. This weakens the statistical association between variation in alleles in a functionally relevant gene and variation in disease.

New technologies, such as DNA microarrays, measure the relative abundance of mRNA produced by many genes simultaneously. This information may be one step closer to the relevant biology than classically defined clinical traits and opens up interesting strategies for understanding complex diseases (Horvath and Baur 2000). In an increasingly common application, genes are functionally classified according to their expression across different conditions (Quackenbush 2001). Researchers have also begun to map QTLs underlying gene expression. This has led to the discovery of small regions apparently involved in regulating the expression of a large number of other genes throughout the genome (Brem et al. 2002; Schadt et al. 2003). This information can then be combined with information on clinical-trait QTLs to discern whether a particular locus plays a direct or regulatory role in trait etiology. This approach can produce a relatively small list of candidate genes in an objective fashion (Schadt et al. 2003). Alternatively, the typically large number of candidate genes in regions linked to a clinical trait can be reduced by restricting attention to those the expression levels of which show association with a relevant clinical trait (Wayne and McIntyre 2002). Patterns of gene expression can also be used to define subtypes of clinical trait with distinct genetic components, potentially increasing the power of clinical-trait linkage scans (Schadt et al. 2003).

Studies combining clinical trait linkage mapping and gene-expression data have been successfully completed using model organisms such as *Mus musculus* and *Drosophila melanogaster* (Wayne and McIntyre 2002; Schadt et al. 2003). Similar studies in humans are on the horizon. Family-based studies offer opportunities for multiple analyses: clinical trait linkage mapping, expression QTL mapping, expression profiling, and trait-expression

Received November 19, 2002; accepted for publication February 26, 2003; electronically published April 8, 2003.

Address for correspondence and reprints: Dr. Peter Kraft, UCLA School of Public Health, Box 951772, Los Angeles, CA 90095-1772. E-mail: pkraft@ucla.edu

0002-9297/2003/7205-0027$15.00

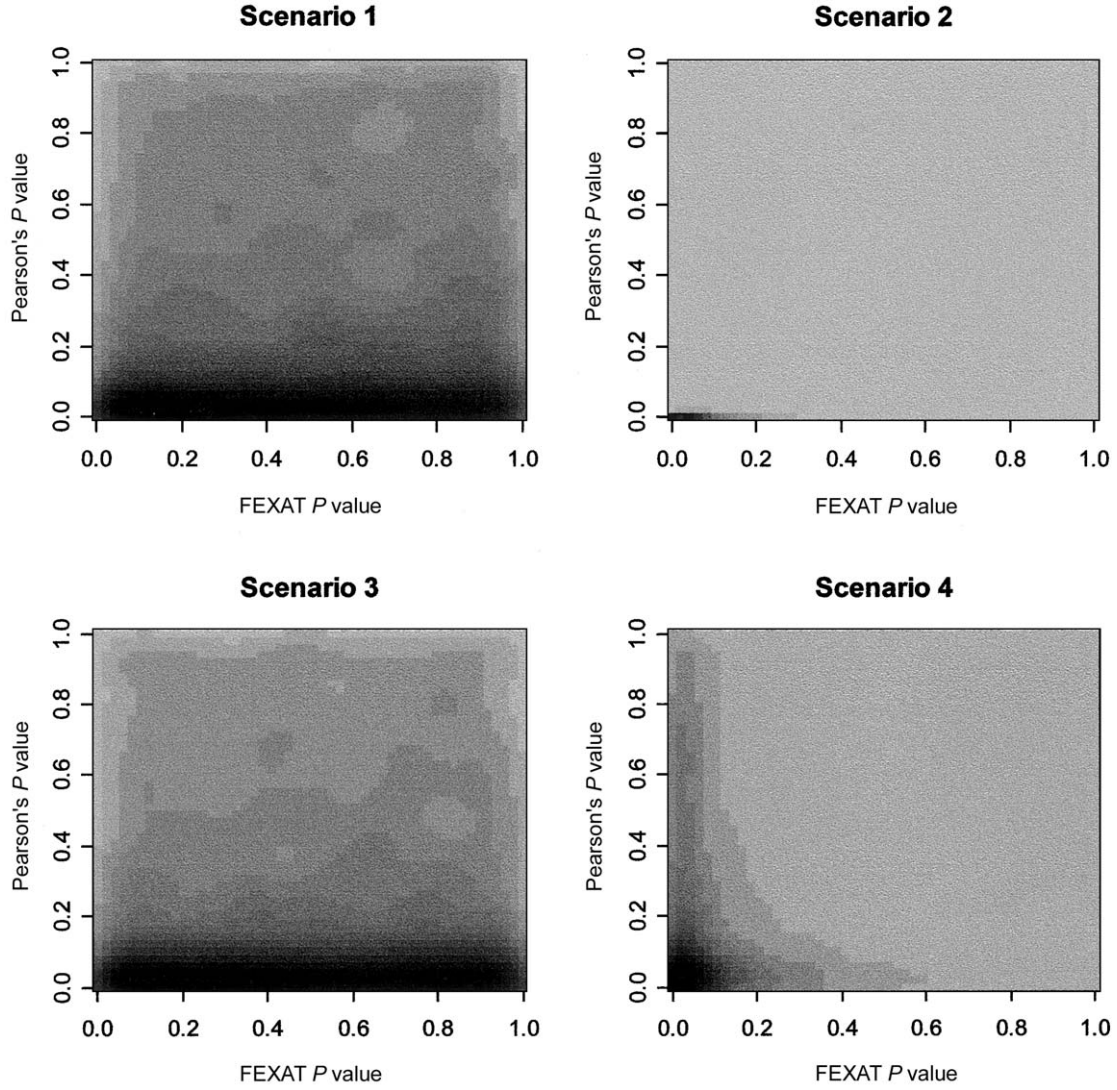

Figure 1 Joint density of Pearson and FEXAT *P* values under different models for the correlation between expression levels and trait values. Darker areas represent areas of higher density, estimated from the *P* values from 10,000 simulated studies of four sibships with 11, 8, 7, and 6 members. The panel for scenario 1 (the null hypothesis) shows that the *P* value for the FEXAT has a roughly uniform distribution, as is appropriate, whereas the distribution for the Pearson's test is skewed toward zero. Under scenario 2 (linear model), both statistics have high power. When the family means for trait and expression are correlated but within-family differences are not (scenario 3), the FEXAT again has a uniform distribution, whereas the Pearson *P* values tend to be small. When the within-family differences are correlated but family means are not (scenario 4), the FEXAT *P* values tend to be much smaller than Pearson *P* values.

association. Methods for expression profiling and testing for trait-expression association should account for family structure. Here, we present a family-based test for correlation between gene expression and trait values. The family-based design also makes this test robust to population stratification bias. We close by briefly discussing family-based approaches to expression profiling.

Standard unstratified measures of the association between expression levels and clinical traits—such as Pearson's correlation coefficient (Quackenbush 2001)—do not take family structure into account. This can lead to spurious results when the number of families is small and the bulk of the variance in the trait and expression levels is due to between-family differences, as opposed to within-family differences. Furthermore, standard unstratified tests may also fail to detect within-family correlation when there is no correlation between gene expression and the trait between families. We illustrate these ideas via simulation below.

To account for family structure, we propose a sibship-stratified family expression association test (FEXAT), which is a continuous-trait version of the Mantel-Haen-

Table 1

Power of Nominal .05-Level Tests for the Pearson's Correlation Test and the FEXAT under Different Models for the Correlation between Expression Levels and Trait Values

Sample Size, Scenario, and Model Parameters	Power for Model and Test					
	$\sigma_\mu^2:\sigma_\varepsilon^2 = 1/2$		$\sigma\mu^2:\sigma_\varepsilon^2 = 1$		$\sigma_\mu^2:\sigma_\varepsilon^2 = 2$	
	Pearson	FEXAT	Pearson	FEXAT	Pearson	FEXAT
Four sibships:						
Scenario 1	.11	.05	.21	.05	.33	.05
Scenario 2:						
$\beta = .333$	.81	.42	.90	.42	.95	.41
$\beta = .667$	1.00	.99	1.00	.98	1.00	.98
Scenario 3:						
$\rho_\mu = .333$	.14	.05	.25	.04	.39	.05
$\rho_\mu = .667$	.21	.05	.38	.05	.55	.05
Scenario 4:						
$\rho_\varepsilon = .333$	.35	.42	.34	.42	.39	.42
$\rho_\varepsilon = .667$	.78	.99	.63	.98	.52	.99
Eight sibships:						
Scenario 1	.09	.05	.14	.05	.21	.05
Scenario 2:						
$\beta = .333$	.89	.36	.99	.34	.96	.35
$\beta = .667$	1.00	.96	1.00	.96	1.00	.96
Scenario 3:						
$\rho_\mu = .333$	.12	.05	.21	.05	.32	.05
$\rho_\mu = .667$	.24	.05	.43	.04	.63	.05
Scenario 4:						
$\rho_\varepsilon = .333$	.30	.35	.25	.34	.26	.34
$\rho_\varepsilon = .667$	.75	.96	.56	.96	.39	.96

Note.—Power calculated via simulation. Scenario 1 corresponds to the null model of no correlation between expression and trait; scenario 2 is a standard linear model; scenario 3 corresponds to a model in which family-specific expression and trait means are correlated but within-family differences are not; and scenario 4 corresponds to a model in which within-family differences from the family mean for expression and trait are correlated but the family means are not. $\sigma_\mu^2:\sigma_\varepsilon^2$ is the ratio of variance in family means to variance in within-family differences; thus, when $\sigma_\mu^2:\sigma_\varepsilon^2 = 2$, two-thirds of the variance in expression (trait) is due to variance in the shared family mean. (See text.)

szel test (Mantel 1963). This is a univariate test (looking for association between *one* gene-expression level and the trait) but can be used in the context of high-dimensional gene-expression data from microarrays as a first-pass analysis to select a subset of mRNAs likely to be associated with the trait, as is now often done using standard methods.

Let $i = 1,\ldots,I$ index the sibships in the study (extended pedigrees may contribute multiple sibships) and $j = 1,\ldots,J_i$ index the subjects in a sibship. Further, let X_{ij} denote the expression level for subject j in family i and Y_{ij} the trait value. X_{ij} could be measured as the (log of the) fold change of the expression for the gene under study in the subject's RNA relative to a reference sample, as in cDNA arrays (Hedge et al. 2000) or as match-mismatch score, as in oligonucleotide arrays (Dalma-Weiszhaus et al. 2002). Then the FEXAT is:

$$\frac{\left[\sum_i \sum_j X_{ij}\left(Y_{ij} - \bar{Y}_{i\cdot}\right)\right]^2}{\sum_i \frac{1}{J_i - 1}\sum_j \left(X_{ij} - \bar{X}_{i\cdot}\right)^2 \sum_j \left(Y_{ij} - \bar{Y}_{i\cdot}\right)^2}.$$

This statistic can be compared with its asymptotic χ_1^2 distribution or an empirical permutation distribution (described below) to calculate P values. We motivate this statistic by noting that the terms $\sum_j X_{ij}(Y_{ij} - \bar{Y}_{i\cdot})^2$ in the numerator have expectation zero under the null, and the terms

$$\frac{1}{J_i - 1}\sum_j (X_{ij} - X_{i\cdot})^2 \sum_j \left(Y_{ij} - \bar{Y}_{i\cdot}\right)^2$$

in the denominator are the permutation variances of the numerator terms. The FEXAT makes no explicit distributional assumptions. It requires only that, conditional on the observed order statistics, all $(J_i!)^2$ combinations within a sibship are equally likely (see, e.g., Cox and Hinkley 1974, pp. 184–186).

Simulation studies show that the FEXAT on the basis of the asymptotic critical value has a conservative type I error–rate size even when the number of families is small and that it has greater power to detect within-family correlation than Pearson's correlation test. We simulated two small studies with 32 observations. The first con-

Table 2

Power of Nominal .001-Level Tests for the Pearson's Correlation Test and the FEXAT under Different Models for the Correlation between Expression Levels and Trait Values

Sample size, scenario, and Model Parameters	Power for Model and Test					
	$\sigma^2_\mu:\sigma^2_\varepsilon = 1/2$		$\sigma^2_\mu:\sigma^2_\varepsilon = 1$		$\sigma^2_\mu:\sigma^2_\varepsilon = 2$	
	Pearson	FEXAT	Pearson	FEXAT	Pearson	FEXAT
Four sibships:						
Scenario 1	.008	.0003	.028	.0003	.090	.0001
Scenario 2:						
$\beta = .333$	.387	.022	.600	.023	.789	.021
$\beta = .667$	.962	.593	.980	.607	.991	.594
Scenario 3:						
$\rho_\mu = .333$	.013	.0001	.048	.0003	.136	.0000
$\rho_\mu = .667$	.034	.0002	.118	.0002	.266	.0001
Scenario 4:						
$\rho_\varepsilon = .333$	.070	.024	.086	.023	.138	.023
$\rho_\varepsilon = .667$	.440	.594	.330	.600	.274	.594
Eight sibships:						
Scenario 1	.003	.0002	.011	.0002	.027	.0002
Scenario 2:						
$\beta = .333$	.465	.006	.735	.006	.926	.006
$\beta = .667$	.979	.257	.993	.739	.999	.257
Scenario 3:						
$\rho_\mu = .333$	.007	.0002	.023	.0002	.073	.0001
$\rho_\mu = .667$	.028	.0002	.100	.0001	.260	.0002
Scenario 4:						
$\rho_\varepsilon = .333$	.041	.007	.042	.007	.055	.007
$\rho_\varepsilon = .667$	.328	.257	.200	.261	.130	.255

Note.—Power calculated via simulation. Scenario 1 corresponds to the null model of no correlation between expression and trait; scenario 2 is a standard linear model; scenario 3 corresponds to a model where family-specific expression and trait means are correlated but within-family differences are not; and scenario 4 corresponds to a model where within-family differences from the family mean for expression and trait are correlated but the family means are not. $\sigma^2_\mu:\sigma^2_\varepsilon$ is the ratio of variance in family means to variance in within-family differences; thus, when $\sigma^2_\mu:\sigma^2_\varepsilon = 2$, two-thirds of the variance in expression (trait) is due to variance in the shared family mean.

sisted of four sibships of sizes 11, 8, 7, and 6 (the same family structure as the CEPH data analyzed below); the second consisted of eight sibships of size 4. We drew gene expression levels X_{ij} and trait values Y_{ij} from random effects model

$$\begin{pmatrix} X_{ij} \\ Y_{ij} \end{pmatrix} = \begin{pmatrix} \mu_{xi} \\ \mu_{yi} \end{pmatrix} + \begin{pmatrix} \varepsilon_{xij} \\ \varepsilon_{xij} \end{pmatrix} ,$$

where

$$\begin{pmatrix} \mu_{xi} \\ \mu_{yi} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{\mu_x} & \rho_\mu \sigma_{\mu_x} \sigma_{\mu_y} \\ \rho_\mu \sigma_{\mu_x} \sigma_{\mu_y} & \sigma^2_{\mu_y} \end{pmatrix}\right] \text{ and } \begin{pmatrix} \varepsilon_{xij} \\ \varepsilon_{yij} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{\varepsilon_x} & \rho_\varepsilon \sigma_{\varepsilon_x} \sigma_{\varepsilon_y} \\ \rho_\varepsilon \sigma_{\varepsilon_x} \sigma_{\varepsilon_y} & \sigma^2_{\varepsilon y} \end{pmatrix}\right] .$$

The family mean vectors $(\mu_x\ \mu_y)^T$ and the within-family differences $(\varepsilon_x\ \varepsilon_y)^T$ were assumed to be independent. The family-mean vectors model within-family similarities in expression and trait values owing to either shared genes or environment; the correlation in mean vectors might be due to confounding factors such as population stratification. We considered four scenarios: (1) no correlation between X_{ij} and Y_{ij}; (2) a linear relationship between Y_{ij} and X_{ij}, $Y_{ij} = \beta X_{ij}$ + individual-level noise; (3) correlation in family expression and trait means but not within-family differences (i.e., $\rho_\mu \neq 0$ and $\rho_\varepsilon = 0$); and (4) correlation in within-family expression and trait differences but not family means (i.e., $\rho_\mu = 0$ and $\rho_\varepsilon \neq 0$). We fixed $\sigma^2_\varepsilon = \sigma^2_{\varepsilon x} = \sigma^2_{\varepsilon y}$ at 1 and varied $\sigma^2_\mu = \sigma^2_{\mu x} = \sigma^2_{\mu y}$ to investigate the relative effects of variation in family means and within-family differences. (Simulations not shown

Table 3

***P* Values for the Pearson's Correlation Test and FEXAT for Association between Beta-Catenin and Genes Known To Be Involved in the *WNT*-Activation Process in Lymphoblastoid Cell Lines from Four CEPH Families**

Gene	*P* Value	
	Pearson	FEXAT
TCF4	.34	.04
LEF1	.08	.07
CTBP2	.21	.03
WNT11	.44	.0002
WISP2	.002	.02
MAP3K[a]	.36	.01

[a] Represents the mean *P* value over 4 of 11 mitogen-activated protein 3-kinases represented on the microarray.

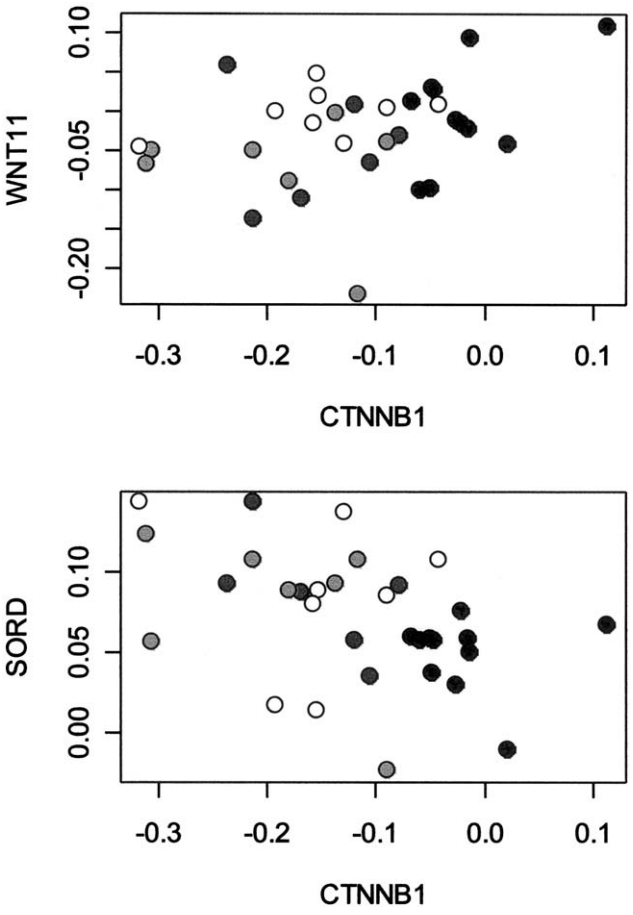

Figure 2 *WNT11* and *SORD* mean log expression ratios plotted against the *CTNNB1* expression in four CEPH sibships. Each sibship is colored a different shade of gray. The FEXAT found *WNT11* and *CTNNB1* to be significantly correlated ($P = .0002$), whereas the standard test did not ($P = .44$). The situation was reversed for *SORD* and *CTNNB1* (FEXAT $P = .76$; Pearson $P = .02$).

varied the ratios $\sigma^2_{\varepsilon x}:\sigma^2_{\varepsilon y}$ and $\sigma^2_{\mu x}:\sigma^2_{\mu y}$. Results did not differ qualitatively from those presented here.) For each scenario, we generated 10,000 replicate studies and calculated Pearson and FEXAT correlation tests on the basis of χ_1^2 critical values for $\alpha = .05$ and $\alpha = .001$. Simulations and analysis were conducted with the freely available R software (Ihaka and Gentleman 1996). An R function to calculate the FEXAT correlation and test is available at the first author's Web site.

Simulation results are presented infigure 1 andtables 1 and 2. Under scenario 1, the FEXAT has a conservative type I error rate, whereas the standard correlation test is anticonservative, with the type I error rate increasing as the ratio of between-family variance σ_μ^2 to within-family variance σ_ε^2 increased. Intuitively, when σ_ε^2 is small relative to σ_μ^2, the relevant sample size for the standard correlation test is the number of families. When this sample size is small, asymptotic properties fail. In particular, the four family means $(\mu_{xi}\ \mu_{yi})^T$ may, by chance, fall on a line, suggesting a linear relationship when, in fact, there is none. As the number of sibships studied increased, the type I error associated with the standard correlation test decreased, although it was still appreciable (>10% for nominal .05-level test) for eight sibships.

Under scenarios 2 and 3, the standard correlation test has greater power to detect association than the FEXAT. This is not surprising, since the standard test is optimal when there is a linear relationship between X_{ij} and Y_{ij}, and the FEXAT, by construction, is not sensitive to correlations in family-mean expression and trait values.

Under scenario 4, when variation in trait values is associated with variation in within-family differences in expression, as opposed to across-family differences, the FEXAT has more power to detect the association between trait values and gene-expression levels (except when the nominal test size is $\alpha = .001$ and within-family variation accounts for the majority of variance in trait and expression). The relative efficiency increases as the proportion of variance in traits and expression levels owing to between-family differences increases. This implies that the FEXAT may have more power than standard correlation tests when both the trait and expression level under study have high heritabilities or are both sensitive to differences in shared family environment.

The observed type I error for the FEXAT is slightly conservative when the nominal rate is $\alpha = .05$ but is much smaller than the nominal rate when $\alpha = .001$. This suggests that the tails of the FEXAT distribution are truncated relative to the asymptotic distribution for small sample sizes. Alternatively, *P* values can be calculated from a permutation distribution, as discussed below.

To further examine whether the FEXAT statistic offers advantages over more classical association-based statistics in the gene expression setting, we considered *WNT* signaling in lymphoblastoid cell lines from 32 individuals from four large CEPH sibships (in CEPH/Utah pedigrees 1362, 1375, 1377, and 1408). Details concerning gene-expression measurement are described more fully elsewhere (Schadt et al. 2003). The *WNT* pathway (*WNT* signaling transduction) is involved in many different developmental processes. This pathway serves to regulate *WNT*-responsive genes and has been shown to regulate B lymphocyte proliferation (Reya et al. 2000). Therefore, we expected the *WNT* pathway to be activated in the CEPH lymphoblastoid cell lines, given that these cell lines had been cultured and maintained in the log phase of cell growth for at least 2 d before harvest.

Central to activation of the *WNT* pathway is the formation of a beta-catenin/*TCF* complex that forms in the presence of *WNT* expression, which results in the activation of *WNT*-responsive genes (Seidensticker et al. 2000). In the absence of *WNT,* the beta-catenin/*TCF* complex does not form; instead, another complex forms,

the beta-catenin deconstruction complex, which promotes the phosphorylation of beta-catenin by *Glycogen synthase kinase 3b* (*GSK3B*). The phosphorylated beta-catenin becomes multiubiquitinated and is subsequently degraded in proteasomes (Seidensticker et al. 2000). Therefore, in the presence of *WNT,* we would expect beta-catenin levels to be associated with *TCF/LEF1* levels and related factors that can lead to the formation and stabilization of the beta-catenin/*TCF* complex. In the absence of *WNT,* we would expect levels of beta-catenin to be associated with genes making up the beta-catenin deconstruction complex. Treating beta-catenin as a clinical trait and testing associations with genes represented in each complex, we applied the FEXAT and the standard correlation test to determine whether the activation or deconstruction complexes were active in the CEPH lymphoblastoid cell lines.

The heritability of beta-catenin (*CTNNB1* [GenBank accession number NM_001904]) over the four CEPH families was found to be very significant, with a *P* value of .00001 (Schadt et al. 2003). We sought to test whether the high degree of heritability detected for this trait in such a small number of families translated into FEXAT statistics that were more significant than standard Pearson correlation statistics computed between *CTNNB1* and genes known to associate with beta-catenin as part of beta-catenin deconstruction or the *WNT* activation pathway. We computed the Pearson correlation coefficient and FEXAT statistic between *CTNNB1* and seven genes associated with the beta-catenin deconstruction complex: *GSK3B* (GenBank accession number NM_002093), axin (*AXIN1* [GenBank accession number AF009674]), conductin (*AXIN2* [GenBank accession number NM_004655]), *APC* (GenBank accession number NM_000038), and the dishevelled genes *DVL1–3* (GenBank accession numbers NM_004421–3). Neither test statistic was significant at the 0.1 level for any of these genes, with the exception of *APC,* which had *P* values of .04 and .03 for the FEXAT and Pearson test statistics, respectively. It has been established in the literature that

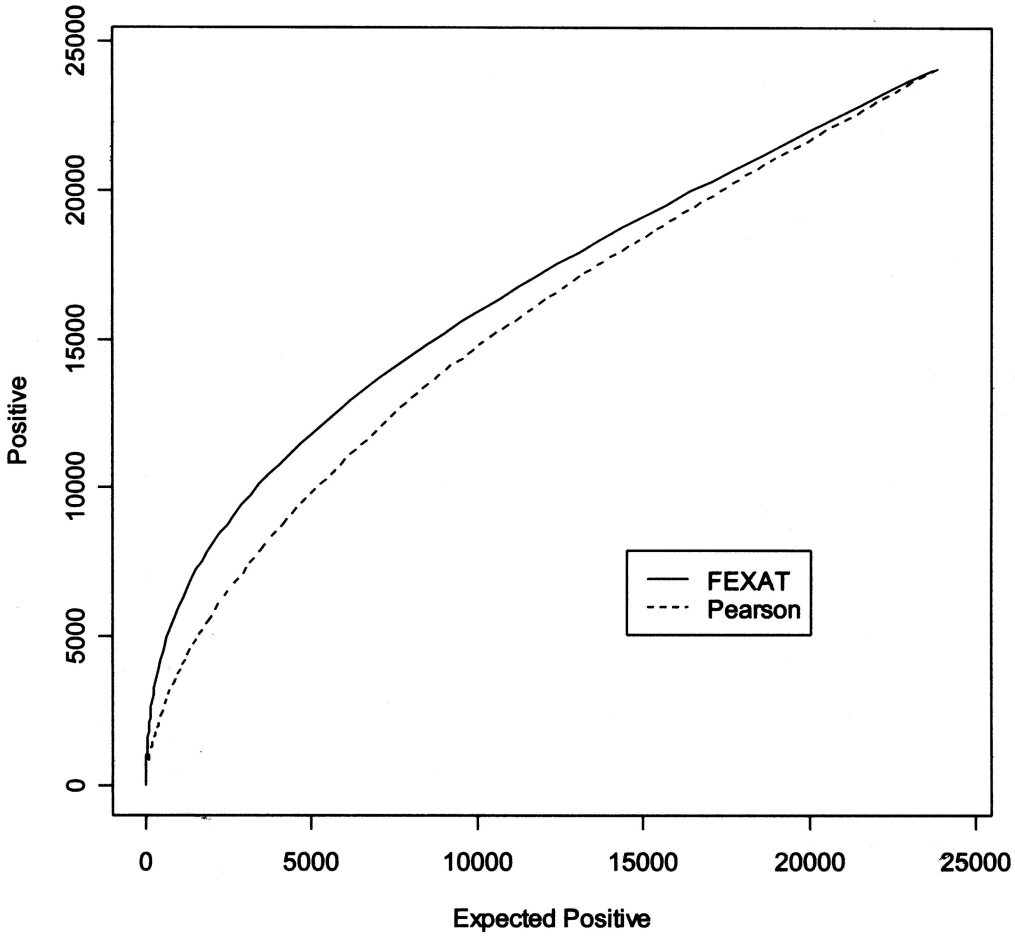

Figure 3 Number of positive FEXAT and Pearson tests versus the expected number of positive results under the null of no within-family association (calculated via the permutation procedure described in the text) for critical values, range 0–16.

APC can interact with beta-catenin independently of the other genes involved in the deconstruction complex (Seidensticker et al. 2000). On the other hand, when testing nine *TCF*/beta-catenin complexes that bind to DNA and activate *WNT* targets, we found significant results for *CTNNB1* interacting with *LEF1* (GenBank accession number NM_016269), *TCF4* (GenBank accession number NM_003199), *CTBP2* (GenBank accession number NM_001329), *WNT11* (GenBank accession number NM_004626), *WISP2* (GenBank accession number NM_003881), and several of the mitogen-activated protein 3-kinases, all known to be involved in the *WNT* activation process (see table 3). Further, we note that, for most of these genes, the association was significant only with respect to the FEXAT statistic and not the Pearson correlation test statistic, so the associations would have been missed had family structure not been taken into account. For instance, the FEXAT statistic for the association between *TCF4* and beta-catenin has a *P* value of .04, whereas the *P* value for the corresponding Pearson correlation test statistic was .34. *TCF4* and beta-catenin complexes are among those most central to the activation of the *WNT* pathway. (The FEXAT does not achieve the Bonferroni-corrected level of .05/16 = .003 for any of the genes listed in table 3 other than *WNT11*. However, more tests [8 = 4 + 4 MAP3K genes] are significant at the .05 level than expected by chance [.8].)

The FEXAT was able to detect biologically plausible correlations among expression levels for genes involved in the *WNT* pathway that Pearson's correlation did not. This suggests that gene-expression variations are associated with one another by way of within-family differences for those genes involved in this pathway, not across-family differences. For example,figure 2 plots the joint distribution of *CTNNB1* and *WNT11* expression. The FEXAT found these two genes to be statistically significantly associated ($P = .0002$), whereas the standard correlation test did not ($P = .44$). There appears to be a positive correlation in at least three of the four sibships—with the largest sibship having the strongest correlation—although there is no clear overall correlation.

Conversely, the FEXAT did not find sorbitol dehydrogenase (*SORD*) and *CTNNB1* to be significantly associated ($P = .76$), although the standard test did ($P = .02$). Visual inspection suggests an overall negative correlation in *SORD* and *CTNNB1,* although they are uncorrelated in each of the sibships (fig. 2). Sorbitol dehydrogenase is an enzyme involved in fructose and mannose metabolism (the polyol pathway). We have no a priori reason to believe it should be associated with *WNT* signaling. Given current knowledge, it is difficult to definitively exclude this apparent association as a false positive, but the significant Pearson test should be interpreted cautiously, since it may be a statistical artifact owing to the chance alignment of sibship means.

We also calculated the expected number of false-positive tests under the null hypotheses of no within-family association between *CTNNB1* and the other measured mRNAs. We did this by averaging >500 random permutations of the trait (*CTNNB1*) values, where trait values were permuted within sibships but not between sibships. This procedure assumes no within-family correlation in trait and expression while conditioning on possible between-family differences. We emphasize that the terminology "false positive" is context dependent. Here, the focus is on within-family correlation, so tests called significant owing to between-family differences are "false positives." If the focus were instead on correlation across families, such significant tests would be interesting and not "false positives."

In particular, for a statistic $T(X_i, Y)$ testing association between the expression for gene i and the trait and a given critical value t^*, the estimated expected number of significant results under the null is

$$\frac{\sum_{ij} 1_{\{T[\mathbf{X}_i,\pi_j(\mathbf{Y})]>t^*\}}}{J},$$

where j indexes the random permutations $\pi_1,\ldots,\pi_J$, and $1_{[*]}$ is the indicator function. This procedure is similar to that proposed by Tusher et al. (2001). Permutation *P* values can be calculated by comparing the observed statistic $T(X,Y)$ with the permuted values $T(X_i, Y)$.

Figure 3 plots the number of positive FEXAT and Pearson tests for association between *CTNNB1* and 24,479 mRNAs versus the number of false positives over the range of t^*. The FEXAT generally has a lower ratio of expected false positive results to observed positive results. This is because many of the significant Pearson statistics reflect between-family associations between the trait and expression levels. Thus, if within-family correlation is the focus, the FEXAT can be more accurate (the ratio of "false positive" to positive results is smaller).

These results suggest that statistical methodologies that take the stratified nature of family-based studies into account can usefully complement standard correlation measures. When the number of independent families sampled is small—as will often be the case, because of the expense of measuring gene-expression levels—and the proportion of trait and expression variance owing to shared genetic and/or environmental factors is high, standard methods may detect an association between trait and expression when there is none. Standard methods may also miss associations primarily owing to within-family correlation in a trait and gene expression. We have shown via simulation and a data example that the proposed FEXAT performs better than standard methods in these situations.

Further methodological developments are required, however. The assumption that all trait and expression combinations within a sibship are equally likely conditional on the order statistics may not hold. For example, the gene under study may be linked to a causal gene, even though it does not play a role in determining the trait itself. If gene-expression levels depend on a particular polymorphism in the gene under study, then all trait and expression combinations will not be equally likely for sibships of size >2 (Kraft 2001). Replacing the variance estimates in the denominator of the FEXAT with the empirical estimates $[\sum_j X_{ij}(Y_{ij} - \bar{Y}_{i\cdot})]^2$ does not require that the combinations within a family be equally likely, but the resulting statistic performed poorly when the number of families sampled was small. It had very low power (results not shown).

We have focused on continuous traits. The FEXAT can be easily modified to accommodate categorical outcomes Y_{ij}. For example, in the dichotomous disease, Y_{ij} could be set to 1 for subjects with disease and to 0 for those without. Further extensions might include tests for effect modification owing to measured environmental or genetic covariates.

The primary application of the FEXAT described in this report is to test for correlations between gene expression levels and continuous traits, possibly including other gene-expression levels, as a method to identify candidate genes associated with a clinical trait of interest. The FEXAT approach could also be used to define a proximity measure between different genes on the basis of the within-family correlation in their expression levels; this proximity measure could then be used to cluster genes, just as Pearson's correlation coefficient is used now for tissue samples from unrelated individuals (Quackenbush 2001). The authors' FEXAT function currently estimates the within-family correlation as:

$$\frac{\sum_i \sum_j X_{ij}(Y_{ij} - \bar{Y}_{i\cdot})}{\sqrt{\sum_i \sum_j (X_{ij} - \bar{X}_{i\cdot})^2 \sum_j (Y_{ij} - \bar{Y}_{i\cdot})^2}} .$$

This is simply the average within-family covariance divided by the square root of the average product of the within-family trait and expression variances. The FEXAT correlation may uncover patterns in gene expression in families missed by Pearson's correlation coefficient.

Acknowledgments

Part of this work was funded by National Institute of Environmental Health Sciences grant R21 ES 011667. J.A. was supported by UCLA-NSF/IGERT bioinformatics training award DGE-99987641.

Electronic-Database Information

Accession numbers and URLs for data presented herein are as follows:

GenBank, http://www.ncbi.nlm.nih.gov/GenBank/ (for *CTNNB1* [accession number NM_001904], *GSK3B* [NM_002093], *AXIN2* [NM_004655]), *APC* [NM_000038], *DVL1–3* [NM_004421–3], *LEF1* [NM_016269], *TCF4* [NM_003199], *CTBP2* [NM_001329], *WNT11* [NM_004626], *WISP2* [NM_003881], and (*AXIN1* [AF009674])

R Archive Network, http://cran.r-project.org/

R Code, http://www.bol.ucla.edu/~pkraft/soft.htm (for the FEXAT software)

References

Brem R, Yvert G, Clinton R, Kruglyak L (2002) Genetic dissection of transcriptional regulation in budding yeast. Science 296:752–755

Clerget-Darpoux F, Selinger-Leneman H, Babron M-C (2001) Why do complex traits resist DNA analysis? IJHG 1:55–63

Cox D, Hinkley D (1974) Theoretical statistics. Chapman and Hall, London

Dalma-Weiszhaus D, Chicurel M, Gingeras T (2002) Microarrays and genetic epidemiology: a multipurpose tool for a multifacted field. Genet Epidemiol 23:4–21

Hedge P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes J, Snesrud E, Lee N, Quackenbush J (2000) A concise guide to cDNA microarray analysis. Biotechniques 29:548–556

Horvath S, Baur M (2000) Future directions of research in statistical genetics. Stat Med 19:3337–3343

Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. J Comput Graph Stat 5:299–314

Kraft P (2001) A robust score test for linkage disequilibrium in general pedigrees. Genet Epidemiol 21:S403–S408

Mantel N (1963) Chi-square tests with one degree of freedom: extensions of the Mantel-Haenszel procedure. JASA 58:690–700

Quackenbush J (2001) Computational analysis of microarray data. Nat Rev Genet 2:418–427

Reya T, O'Riordan M, Okamura R, Devaney E, Willert K, Nusse R, Grosschedl R (2000) Wnt signalling regulates B lymphocyte proliferation through a LEF dependent mechanism. Immunity 13:15–24

Schadt EE, Monks S, Drake T, Lusis A, Che N, Colinayo V, Ruff T, Milligan S, Lamb J, Cavet G, Linsley P, Mao M, Stoughton R, Friend S (2003) The genetics of gene expression surveyed in maize, mouse and man. Nature 422:297–302

Seidensticker M, Behrens J (2000) Biochemical interactions in the Wnt pathway. Biochim Biophys Acta 1495:168–182

Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 98:5116–5121

Wayne ML, McIntyre LM (2002) Combining mapping and arraying: an approach to candidate gene identification. Proc Natl Acad Sci USA 99:14903–14906