Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 15531–15536, December 1998
Genetics

A genome-wide search for chromosomal loci linked to mental health wellness in relatives at high risk for bipolar affective disorder among the Old Order Amish

EDWARD I. GINNS*†, PAMELA ST. JEAN‡§, ROBERT A. PHILIBERT*§, MARZENA GALDZICKA*, PATRICIA DAMSCHRODER-WILLIAMS*, BONNIE THIEL‡, ROBERT T. LONG*, LORING J. INGRAHAM*, HARNISHA DALWALDI*, MELISSA A. MURRAY*, MELISSA EHLERT*, SHARON PAUL*, BRIAN G. REMORTEL*, ASHIMA P. PATEL*, MARIA C. H. ANDERSON*, CARY SHAIO*, ELAINE LAU*, INNA DYMARSKAIA*, BRIAN M. MARTIN*, BARBARA STUBBLEFIELD*, KATHLEEN M. FALLS¶, JOHN P. CARULLI¶, TIM P. KEITH¶, CATHY S. J. FANN‖, LUCY G. LACY**, CLEONA R. ALLEN**, ABRAM M. HOSTETTER**, ROBERT C. ELSTON‡, NICHOLAS J. SCHORK‡, JANICE A. EGELAND**, AND STEVEN M. PAUL*††

*Clinical Neuroscience Branch, Intramural Research Program, National Institute of Mental Health, Bethesda, MD 20892; ‡Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH 44109; ¶Genome Therapeutics Corporation, Waltham, MA 02154; ‖Department of Epidemiology and Social Medicine, Albert Einstein College of Medicine, New York, NY 10467; **Department of Psychiatry, University of Miami School of Medicine, Miami, FL 33136; and ††Departments of Psychiatry, Pharmacology, and Toxicology, Indiana University School of Medicine and Lilly Research Laboratories, Eli Lilly and Co., Indianapolis, IN 46285

Communicated by Seymour S. Kety, National Institute of Mental Health, Bethesda, MD, October 19, 1998 (received for review May 12, 1998)

ABSTRACT **Bipolar affective disorder (BPAD; manic-depressive illness) is characterized by episodes of mania and/or hypomania interspersed with periods of depression. Compelling evidence supports a significant genetic component in the susceptibility to develop BPAD. To date, however, linkage studies have attempted only to identify chromosomal loci that cause or increase the risk of developing BPAD. To determine whether there could be protective alleles that prevent or reduce the risk of developing BPAD, similar to what is observed in other genetic disorders, we used mental health wellness (absence of any psychiatric disorder) as the phenotype in our genome-wide linkage scan of several large multigeneration Old Order Amish pedigrees exhibiting an extremely high incidence of BPAD. We have found strong evidence for a locus on chromosome 4p at *D4S2949* (maximum GENEHUNTER-PLUS nonparametric linkage score = 4.05, $P = 5.22 \times 10^{-4}$; SIBPAL $P_{empirical}$ value $< 3 \times 10^{-5}$) and suggestive evidence for a locus on chromosome 4q at *D4S397* (maximum GENEHUNTER-PLUS nonparametric linkage score = 3.29, $P = 2.57 \times 10^{-3}$; SIBPAL $P_{empirical}$ value $< 1 \times 10^{-3}$) that are linked to mental health wellness. These findings are consistent with the hypothesis that certain alleles could prevent or modify the clinical manifestations of BPAD and perhaps other related affective disorders.**

Bipolar affective disorder (BPAD) afflicts approximately 1% of the population and is associated with a high risk of suicide (1). Twin, family, and adoption studies have provided strong evidence for an important genetic component in the susceptibility to develop BPAD (2), but unlike for other common medical illnesses, robust biological markers have not been identified for BPAD, and genetic linkage studies have had to rely on categorical diagnoses. Genetic heterogeneity, phenocopies, genotyping errors, and the complexities of performing and interpreting statistical analyses may contribute to some of the inconsistences observed in these linkage studies (3–12). However, because the inheritance of BPAD is probably multifactorial, the possible involvement of multiple genetic components of small effect and/or the occurrence of major allelic effects only in epistasis must be considered. We previously have suggested (13), that in addition to susceptibility alleles, there could be alleles that reduce the risk of developing BPAD in a manner similar to that reported for other complex genetic disorders. If model-based linkage analyses are used, a false-negative linkage finding could result when individuals inherit disease susceptibility alleles but do not manifest the phenotype because of the presence of protective alleles. The inclusion of individuals who inherit susceptibility alleles but do not manifest disease because of protective alleles, or of individuals who inherit protective alleles but nevertheless manifest the disease, also will reduce the power of model-free (allele-sharing) analyses. Thus, regardless of whether model-based or model-free analyses are used, wellness or protective alleles could have a significant impact on linkage analyses.

Ascertainment of psychiatric disorders and health among several large multigenerational Old Order Amish pedigrees covers a period of more than 20 years. Throughout this longitudinal study, procedures for assessing and diagnosing subjects have remained constant and have included a thorough evaluation of all bipolar I (BPI) probands and their relatives (14–16). Morbid risk analyses have demonstrated a high prevalence of affective disorder among first-degree relatives of bipolar probands in these families (17). Because of the long-term, longitudinal nature of the study, the unaffected, mentally healthy individuals in these families also were followed, many for a period of years past the age of risk for BPAD. Consequently, rather than limit our genome-wide search to identifying susceptibility loci for the disease phenotype (BPAD), we tested the hypothesis that protective alleles may contribute to the absence of psychiatric illness (i.e., mental health wellness) in unaffected family members in these high-risk pedigrees. Because the mode of inheritance of any gene(s) modifying the relative risk for affective disorder is unknown (2) we relied exclusively on model-free linkage analyses. We now report strong evidence for linkage of DNA markers on chromosome 4p to mental health wellness in relatives at high risk for, but who did not develop, major affective disorder in several large multigenerational Old Order Amish pedigrees with an extremely high incidence of BPAD.

Abbreviations: BPAD, bipolar affective disorder; NPL, nonparametric linkage; BPI, bipolar I; GH-PLUS, GENEHUNTER-PLUS; IBD, identical by descent.

†To whom reprint requests should be addressed at: Clinical Neuroscience Branch, National Institute of Mental Health, Building 49, Room B1EE16, 49 Convent Drive, MSC4405, Bethesda, MD 20892-4405. e-mail: ginnse@irp.nimh.nih.gov.

§P.St.J. and R.A.P. contributed equally to this work.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Diagnostic Assessment. Our genetic-epidemiologic study of BPAD among the Old Order Amish in southeastern Pennsylvania has been described in detail (18), including the methods for ascertainment and diagnostic documentation with informed consent (15, 19). Diagnoses were made, by using strict research diagnostic criteria (16), by a five-member psychiatric review board whose members were blind to pedigree membership, diagnostic opinions, treatment data from abstracted medical records, and genetic marker status. Twenty-five nuclear families were developed based on one or more confirmed cases of BPI, which formed the structure of pedigrees 110, 210, and 310. Our present report uses all of these earlier subjects plus additional expansions, especially in pedigree 410, so that our overall study now contains 346 samples, including those from 50 BPI individuals. The unaffected individuals (mentally well or healthy) are those for whom all SADS-L (schedule for affective disorders and schizophrenia-lifetime version) interview responses were negative, and no contradictory reports were given by family informants. Any individuals for whom some symptomatology was identified, even if it did not meet criteria for a formal diagnosis, were labeled as unknowns in our linkage analyses.

The method used for this longitudinal study is ethnographic and culturally appropriate to the field setting. Several members of each nuclear family with a BPI proband (BPI nuclear family) are seen annually. Those diagnosed with BPI or other major affective disorder undergo a yearly course-of-illness update. Parents of each BPI patient are visited regularly and have proven to be accurate informants. At least one unaffected sibling (control sample) of the married BPI patients has been interviewed yearly since 1990 in connection with a prospective study of children at risk for BPAD. Hence, at least three members and occasionally all members of each BPI nuclear family have been evaluated yearly. Individuals are interviewed anew with the complete SADS-L schedule whenever any abnormal mental or emotional symptoms are identified by the follow-up mechanisms. The long-term, systematic follow-up of the families in our study has demonstrated that onset of illness in the Old Order Amish usually is reported by multiple informants.

Patient Samples. Blood samples were generally collected after each first-degree relative (including parents, siblings, and children older than age 15) of the BPI probands had been interviewed with the complete SADS-L schedule. Samples were obtained with written informed consent and coded to maintain confidentiality. The phlebotomist was kept blind to pedigree relationships and diagnostic status. Lymphoblastoid cell lines, on an average of eight members of each nuclear family, were established at the Coriell Institute for Medical Research, Camden, N.J. and/or the Clinical Neuroscience Branch, Intramural Research Program, National Institute of Mental Health, Bethesda, MD. The NIGMS Human Genetic Mutant Cell Repository catalogue (20) contains updated pedigree and diagnostic information for several of the Amish pedigrees used in our study.

Genotyping. Genomic DNA was obtained from peripheral blood samples and/or immortalized lymphoblastoid cell lines as described (9). The order of typed markers on our mapping panels was obtained from the genetic location database (21).

DNA panels for PCR were in a 96-microtiter plate format, and aliquoted by using a BioMek robot (Beckman Instruments). The DNA samples were PCR-amplified separately with each of the DNA markers and then the PCR products were multiplexed, six markers per lane, for electrophoresis using GENESCAN software on ABI 373 instruments (Applied Biosystems Division, Perkin–Elmer). The genotypes subsequently were analyzed with GENOTYPER. DNA from several individuals was represented multiple times in the genotyping panels as controls to evaluate the consistency of genotypes across different gels.

Genetic Analysis Software (38) was used to identify problematic marker data, and a utility written in SPSS (SPSS, Chicago) generated a list of samples that needed to be rerun because of inheritance discrepancies or unreadable signals. We maximize the useful information by repeating the genotyping/analysis cycles until all possible DNA marker genotypes are obtained. Once genotyping for a marker was finished, the data were reanalyzed with G.A.S., and observed allelic mutations and other noninheritances were zeroed out in the data file and noted. Histograms indicating the marker allele size bins were generated. FASTLINK (22) was used to reanalyze the data before further statistical analyses.

Linkage Analyses. Model-free linkage analyses were conducted by using the two-point affected sib pair analysis program S.A.G.E. SIBPAL (23) and the multipoint analysis program GENEHUNTER-PLUS (GH-PLUS) (24). Because there were a few sibships with incomplete marker information, marker allele frequencies were estimated from the entire Old Order Amish family data set by using a maximum likelihood method implemented in the program MENDEL/USERM13 (25, 26). SIBPAL was used to identify markers showing an excess of alleles shared identical by descent (IBD) among unaffected, mentally healthy sib pairs. Under the null hypothesis of no linkage between a trait and marker, sib pairs would be expected to share on the average 50% of alleles IBD, but when a trait and marker are linked, IBD sharing will be increased in both affected and unaffected sibpairs. Because SIBPAL assumes marker allele frequencies appropriate for random samples, it underestimates the proportion of alleles shared IBD by concordant sib pairs when there is linkage. Multipoint analyses by using the model-free linkage program GH-PLUS produced nonparametric linkage (NPL) scores along points at the chromosomal region of interest. Two scoring functions are available in GH-PLUS: IBD sharing can be assessed among concordant relative pairs (NPL_{pairs}) or it may be assessed among larger groups of concordant relatives (NPL_{all}). Our analyses were conducted by using the NPL_{all} statistics as Kruglyak and colleagues (24) have demonstrated that the NPL_{all} statistic results in a more powerful test than the NPL_{pairs} statistic.

RESULTS

First, we analyzed our genome-wide scan dataset looking for evidence of chromosomal loci linked to mental health wellness (the absence of any psychiatric illness). In these analyses, only mental health wellness, in individuals who were over 45 years of age and had a first-degree BPI sibling in their family (pedigrees 110, 210, 310, and 410), was the linkage phenotype of interest (concordantly unaffected pairs) by using SIBPAL. Of more than 980 DNA markers, only five markers representing three chromosome regions had test statistics that were sufficiently outlying and that were likely to represent significant linkage results. Of the markers on chromosome 4p, *D4S2949*, which is located in the vicinity of the BPAD susceptibility locus reported by Blackwood *et al.* (11), had an empirical SIBPAL P value $<3 \times 10^{-5}$ (nominal P value $<1 \times 10^{-7}$). The marker *D4S397* on chromosome 4q had an empirical SIBPAL P value $= 9 \times 10^{-4}$ (nominal P value $= 3 \times 10^{-7}$) On chromosome 11q, two DNA markers (*D11S133* and *D11S29*), located over an approximately 20-cM region, each had a nominal P value $<5 \times 10^{-5}$ (SIBPAL; simulations were not performed). To supplement standard criteria for assessing the significance of our linkage analysis results, we used graphical techniques (Fig. 1) and the empirical assessment of P values (27–29). If each marker assessed in a pairwise linkage analysis is unlinked to the trait, then the P values associated with those markers should be uniformly distributed. In addition, the test statistics used to generate these P values (for instance, t tests in the case of SIBPAL) should follow an appropriate distribution. A plot (generated by using PROC CHART, SAS) of the test statistics obtained from each pairwise linkage analysis is shown in Fig. 1. The plot in the inset depicts a line that should be linear if all markers are unlinked. As seen in Fig. 1, there are outlying test statistic values that likely represent false null hypotheses, that is, evidence for significant linkage results. In addition, in the inset to

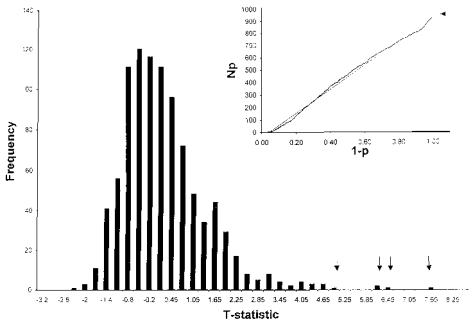

FIG. 1. Plot of test statistics obtained from the pair-wise linkage results. (*Inset*) A cumulative plot of *P*-values whose linearity would reflect uniformity in *P*-values associated with multiple linkage results whose null hypotheses were all true (see text). The outlying test statistics and *P*-values (denoted by arrows) were associated with markers, *D4S107* ($t = 6.24$), *D4S2949* ($t = 7.79$), *D4S2928* ($t = 5.03$), *D11S133* ($t = 6.09$), and *D11S29* ($t = 6.32$).

Fig. 1, the small upturned portion of the *P* value plot near values of $1 - P = 1$ represents departures from uniformity and hence most likely reflects false null hypotheses. Because of the effort required to investigate the significance of these findings and the previous evidence for a BPAD-related locus on chromosome 4 (11), we chose to examine DNA markers on chromosome 4 first for linkage to mental health wellness.

To evaluate the findings on chromosomes 4p and 4q in more detail, we genotyped the subpedigrees and nuclear families containing at least one sibling with BPI (Table 1) by using additional DNA markers in these interesting regions. Compared with our previous report (9) a larger number of individuals was included in these analyses (Table 1). In this report, model-free linkage analyses using SIBPAL and GH-PLUS (24) were performed by using mental health wellness as the linkage phenotype (Tables 2 and 3). In our analyses, individuals having a psychiatric diagnosis other than BPI, as well as those having psychiatric symptoms but no diagnosis, were classified in the unknown category for affected status. In the Amish Study sample of BPI patients ($n = 50$) the mean and median ages of onset (research diagnostic criteria) are 24 and 22 years, respectively. Hence, in all analyses we used a conservative age cutoff of 45 years to define family members with the unaffected wellness phenotype. We also examined the influence of different age cutoffs for defining well individuals and the contribution of different subpedigrees (families from pedigrees 110, 210, 310, and 410 versus only families from pedigree 110) on the test statistics for linkage. Well individuals younger than the specified age cutoff were considered to have an unknown affected status in the analyses.

Table 1. Old Order Amish subjects included in linkage analysis

Subjects	Analysis categories	
	Mentally healthy	Unknowns
Pedigrees 110, 210, 310, 410		
≥25 years old	138	85
≥35 years old	109	114
≥45 years old	74	149
≥55 years old	52	171
Pedigree 110 only		
≥25 years old	45	32
≥35 years old	37	40
≥45 years old	31	46
≥55 years old	23	54

The category of unknowns includes individuals of unknown phenotype, individuals with psychiatric diagnoses other than BPI, and individuals who are mentally healthy but are younger than the particular age cut-off used in analyses. BPI individuals are not included in the unknown phenotype category. In pedigrees 110, 210, 310, and 410, 39 people were diagnosed with BPI, eight with BPII, 21 with recurrent depressive disorder, two with unipolar depressive disorder, and 15 with other psychiatric illness. In pedigree 110 alone, 18 people were diagnosed with BPI, two with BPII, 10 with major depressive disorder, and five with other psychiaric illness. Note: the individuals used in these linkage analyses represent only a subset of the entire Amish bipolar pedigrees because only nuclear families and subpedigrees containing a sibling with BPI were included.

On chromosome 4p, the maximum multipoint NPL values (GH-PLUS; including only individuals > age 45 years) were 4.05 ($P = 5.22 \times 10^{-4}$) for pedigree 110 and 4.05 ($P = 1.84 \times 10^{-4}$) for all pedigrees, respectively. The maximum multipoint NPL values (GH-PLUS; including individuals > age 45 years) for markers on chromosome 4q were 3.29 ($P = 2.57 \times 10^{-3}$) for pedigree 110 and 2.82 ($P = 4.43 \times 10^{-3}$) for all pedigrees, respectively. The GH-PLUS $-\log_{10} P$ value as a function of the map position at these locations on chromosome 4 is shown in Fig. 2. SIBPAL test statistics for markers on chromosomes 4p and 4q are shown in Tables 2 and 3. On chromosome 4 the lowest (nominal) *P* values obtained from the SIBPAL test statistics were for markers *D4S2949* (4p; $P < 1 \times 10^{-7}$) and *D4S397* (4q; $P = 3 \times 10^{-7}$). The maximum multipoint NPL values (GH-PLUS; including individuals > age 45 years) for markers on chromosome 11q were 2.43 for pedigree 110 and 2.49 for all pedigrees, respectively.

To obtain empirical *P*-values, we simulated genotype data by randomly assigning marker alleles to the founders and then assigning alleles to their descendants following Mendelian inheritance. Allowing for consanguineous matings, the entire family structure was used in marker assignment, thus taking into account all relationships between individuals in the dataset. For each simulation, after marker assignment, the pedigrees were trimmed down to that of the nuclear families used in the linkage analysis. SIBPAL then was run on the trimmed dataset, and test statistics for concordant and discordant sib pairs were obtained. The true *P* value is simply estimated as the proportion of replicates in which the simulated statistic is greater than or equal to the observed statistic, i.e., the probability that the observed result or something more extreme would be obtained by chance alone. Simulations were conducted for markers on chromosomes 4p and 4q. For each marker, 100,000 replicates were obtained. The empirical *P* value on chromosome 4p clearly meets the proposed criteria of significance for linkage (30).

DISCUSSION

If alleles exist that are associated with mental health wellness, then the identification of chromosome regions containing these alleles would be enhanced by studying the genetically at risk, mentally healthy members of large, multigenerational pedigrees like our Old Order Amish families. However, in trying to identify protective or wellness alleles, one must recognize that there are phenocopies that need to be considered in our analyses. Despite the extremely high risk for developing disease, some individuals are undoubtedly well because they do not inherit any (or all) of the requisite susceptibility alleles for BPAD. Because the age of greatest liability for onset of BPAD in the Old Order Amish is from early teens through 24 years of age, the misspecification of the well phenotype for individuals who eventually will develop BPAD would be greatest through this age period. In these Old Order Amish families susceptibility alleles for BPAD probably occur in very high frequency. An important step in demonstrating that there are protective alleles will be to show that there are mentally healthy individuals who share marker alleles that should increase the risk of developing BPAD, and yet, in the presence of protective alleles, these individuals do not manifest BPAD. For many of the markers, $\hat{\Pi}$, an underestimate of the proportion of alleles shared IBD in well sibpairs, increases with increasing age, i.e., a more stringent definition of the well phenotype. For example, with respect to marker *D4S2949* on 4p, $\hat{\Pi}$ is 0.60, 0.65 and 0.71 for age cutoff points of 25, 35, and 45 years, respectively. This finding suggests that increasing the age for inclusion eliminates some age-related well phenocopies.

Table 2. Results of SIBPAL analysis of 4p markers

	Pedigree 110			Pedigrees 110, 210, 310, 410		
		P-value			*P*-value	
Marker	Î (s.e.)	Nominal	Simulated	Î (s.e.)	Nominal	Simulated
D4S412	.4749 (.0621)	.6555	np	.5116 (.0539)	.4154	np
D4S431	.5734 (.0441)	.0523	np	.5921 (.0388)	.0110	np
D4S2366	.6781 (.0452)	.0002	.0005	.6024 (.0356)	.0027	.0094
D4S2935	.5066 (.0218)	.3825	np	.4998 (.0198)	.5043	np
D4S3007	.6233 (.0386)	.0014	.0023	.5632 (.0337)	.0330	.0496
D4S394	.6782 (.0513)	.0007	.0012	.5955 (.0421)	.0135	.0249
D4S2983	.7219 (.0484)	$<1 \times 10^{-4}$	np	.6090 (.0377)	.0025	np
D4S2923	.6661 (.0446)	.0003	np	.5902 (.0307)	.0022	np
D4S615	.7161 (.0393)	$<1 \times 10^{-4}$	np	.6223 (.0324)	.0002	np
Afma184za9	.7396 (.0446)	$<1 \times 10^{-4}$	np	.6220 (.0370)	.0008	np
D4S2928	.7333 (.0257)	$<5 \times 10^{-5}$	np	.6369 (.0272)	$<5 \times 10^{-5}$	np
D4S1605	.5453 (.0258)	.0440	.0472	.5795 (.0244)	.0011	0.0058
D4S1582	.6787 (.0616)	.0032	.0112	.6269 (.0557)	.0139	.0510
D4S107	.6557 (.0246)	$<5 \times 10^{-5}$	.0029	.6514 (.0243)	$<5 \times 10^{-5}$	.0088
D4S3009	.7325 (.0552)	.0001	np	.6237 (.0379)	.0008	np
D4S2906	.6460 (.0396)	.0004	np	.5853 (.0327)	.0055	np
D4S2949	.7077 (.0202)	$<1 \times 10^{-7}$	$<3 \times 10^{-5}$	.6888 (.0243)	$<1 \times 10^{-7}$	$<3 \times 10^{-5}$
Afm087zg5	.5229 (.0368)	.2686	np	.5114 (.0246)	.3218	np
D4S2944	.5647 (.0263)	.0093	np	.5428 (.0255)	.0483	np
D4S403	.6032 (.0492)	.0217	.0233	.5989 (.0443)	.0232	.0350
D4S2942	.7196 (.0308)	$<1 \times 10^{-4}$	np	.6627 (.0243)	$<1 \times 10^{-4}$	np
D4S2984	.5510 (.0396)	.1032	np	.5493 (.0297)	.0505	np
D4S1602	.6001 (.0561)	.0412	np	.5703 (.0383)	.0356	np
D4S1511	.6242 (.0489)	.0077	np	.5779 (.0315)	.0079	np
D4S2311	.7429 (.0279)	$<5 \times 10^{-5}$	np	.6327 (.0336)	.0001	np
D4S3048	.6628 (.0573)	.0036	np	.5998 (.0403)	.0078	np
D4S419	.5981 (.0270)	.0004	.0010	.5772 (.0319)	.0100	.0201
D4S404	.6785 (.0489)	.0004	.0010	.6428 (.0470)	.0020	.0072
D4S391	.7008 (.0487)	.0001	.0003	.6585 (.0470)	.0008	.0035

Î is the estimated proportion of alleles shared identical by descent. np, simulations not performed.

It is conceivable that virtually all cases of affective disorder in these families are caused by a common set of susceptibility alleles. The wellness or protective loci that we tentatively have identified could harbor alleles that prevent the manifestation of a BPAD spectrum phenotype, which also could include major depressive disorder. In our analyses the strongest evidence for protective alleles comes from pedigree 110, suggesting that such alleles may be more likely in this branch of the family. However, highly significant test statistics and multipoint logarithm of odds scores (using GH-PLUS) also are observed when pedigrees 110, 210, 310, and 410 are used for analyses (Fig. 2 *A* and *B*). The decreased sharing in proportion of alleles IBD for discordant pairs provides further support for the existence of alleles associated with the absence of affective disorder (mental health wellness) in these families (data not shown). In addition, epistatic interactions between alleles also could prevent or delay onset of an illness such as major depressive disorder from developing into BPAD. Indeed, as we increase the age-of-risk cutoff for defining the well phenotype from 25 to 45 years in our linkage analyses, the number of mentally healthy members decreases as expected, yet the evidence for linkage increases (data not shown).

There is some debate on the analysis of sibling pairs as to whether the use of inbred sibling pairs results in an increased number of false-positives if allele-sharing-based statistical methods are used (31). However, the arguments that (*i*) inbred sibling pairs are likely to share more genes than noninbred sibling pairs (i.e., have a kinship factor greater than 0.5) and (*ii*) that greater regions of the genome would show significant deviations from the expected noninbred sibling sharing value of 0.5 are incorrect when one is merely considering an analysis of sibling pairs involving only the transmission of alleles from parents to offspring. The transmission of alleles from parents to offspring will follow Mendelian ratios, and thus the null values for 0, 1, or 2 IBD sibling allele sharing in any population will be 0.25, 0.50, and 0.25, whenever only parental and sibling genotype information is used. However, if the origin of the parental alleles is taken into consideration, then there will be greater information about alleles shared by sibling pairs from inbred populations. For example, this increased information has the potential to resolve ambiguities in the sharing of alleles transmitted from homozygous parents, because the two copies of the allele in an inbred homozygous parent could be IBD. This information also could help resolve alleles shared by siblings identical in state into alleles shared IBD, showing that alleles transmitted to two offspring from different parents may be copies of the same allele because of the relatedness of the parents. If genealogy is taken into account, then the increased ability to resolve ambiguities in allele sharing would result in greater power in the analysis of inbred sibling pairs (31).

Ultimately, if inbreeding exists in a population from which sibling pairs have been gathered, but one ignores genealogical information by merely studying the transmission of alleles from parents to offspring, then no increase in false-positive linkage results will occur, because Mendel's law applies to inbred as well as outbred parent-offspring allele transmission studies. On the contrary, a decrease in power may result from inbred sibling pair analyses because spouses may manifest greater homozygosity and therefore provide less informative genotypes for parent-offspring-based linkage studies.

Genetic mapping of complex disorders with multifactorial inheritance could be especially difficult if, in addition to susceptibility alleles, individuals inherit protective alleles that prevent or reduce the risk of manifesting the disease phenotype. Even though model-based linkage analyses that do not allow for a multifactorial component are of only limited usefulness in these circumstances, they are still used frequently. In these instances, a false-negative linkage finding (type 2 error) could result when

Table 3. Results of SIBPAL analysis of 4q markers

	Pedigree 110			Pedigrees 110, 210, 310, 410		
		P-value			*P*-value	
Marker	Î (s.e.)	*Nominal*	Simulated	Î (s.e.)	Nominal	Simulated
D4S2303	.4616 (.0423)	.8145	np	.4670 (.0305)	.8585	np
D4S2985	.5754 (.0255)	.0027	np	.5403 (.0139)	.0025	np
D4S2423	.5445 (.0415)	.1453	.1373	.5445 (.0304)	.0743	.0846
D4S2286	.5533 (.0522)	.1570	np	.5225 (.0381)	.2780	np
D4S2959	.5035 (.0359)	.4619	np	.4906 (.0268)	.6370	np
D4S175	.5960 (.0558)	.0471	.0636	.5995 (.0484)	.0231	.0348
D4S422	.6198 (.0500)	.0108	np	.5685 (.0386)	.0403	np
D4S1576	.5290 (.0509)	.2861	np	.5377 (.0367)	.1545	np
D4S2294	.4960 (.0446)	.5351	np	.4867 (.0381)	.6358	np
D4S1579	.6206 (.0381)	.0015	np	.5740 (.0298)	.0077	np
D4S397	.7511 (.0449)	3×10^{-7}	.0009	.6586 (.0376)	5×10^{-6}	.0002
D4S3089	.4544 (.0348)	.9013	np	.4768 (.0261)	.8120	np
D4S2965	.5296 (.0581)	.3068	np	.5267 (.0366)	.2340	np
D4S192	.5135 (.0408)	.3715	np	.5040 (.0337)	.4525	np
D4S420	.5595 (.0539)	.1384	np	.5462 (.0389)	.1200	np
D4S1644	.5224 (.0521)	.3351	.2870	.5503 (.0362)	.0845	.0925
D4S3334	.5491 (.0254)	.0304	.0497	.5258 (.0287)	.1858	.1769
D4S1565	.5091 (.0373)	.4042	np	.5040 (.0271)	.4420	np
D4S1625	.5433 (.0454)	.1730	np	.5533 (.0339)	.0603	np
D4S424	.5901 (.0527)	.0481	np	.5950 (.0461)	.0226	np
D4S1604	.5501 (.0473)	.1480	np	.5095 (.0345)	.3919	np
D4S1548	.5597 (.0356)	.0511	np	.5814 (.0267)	.0016	np

Î is the estimated proportion of alleles shared identical by descent. np, simulations not performed.

individuals inherit disease susceptibility alleles but do not manifest the phenotype because of the simultaneous presence of protective alleles. If model-based methods are used, it is important to provide a reasonably low estimate of penetrance and include a multifactorial component in the model.

In the initial stages of analyzing a disorder like BPAD, which most likely displays multifactorial inheritance, robust model-free (allele sharing) methods are usually more useful than model-based linkage analysis (32). Concordant individuals should demonstrate excess allele sharing, even with the occurrence of phenocopies, genetic heterogeneity, high frequency of susceptibility alleles, and incomplete penetrance. Individuals who inherit susceptibility alleles but do not manifest disease because of protective alleles and individuals who inherit protective alleles but nevertheless manifest the disease will reduce the power of these analyses. Thus, regardless of the type of linkage analysis performed, the presence of protective alleles could have a major impact on identifying susceptibility loci.

Although the idea that protective alleles could modify (or even prevent) a behavioral phenotype like BPAD is relatively novel, there are examples where such protective alleles can affect the expression or inheritance of other Mendelian and multifactorial disorders. The severity of sickle cell anemia is influenced by genes that increase the amount of circulating fetal hemoglobin (33). Similarly, the genotype of the chemokine receptor CCR5 dramatically influences the kinetics of HIV-1 infection, where most individuals who are homozygous for a 32-bp deletion in the CCR5 gene encoding the coreceptor for macrophage-tropic HIV-1 are protected from virus infection (34). In Alzheimer's disease, apolipoprotein (Apo) E2, in contrast to ApoE4, appears to reduce the risk of developing the disease and may protect individuals who inherit a disease-associated ApoE4 allele (35). In an extended Italian family, Apo $A\text{-}I_{MILANO}$ protects against the development of both clinical and pathologic signs of atherosclerosis, despite significantly elevated plasma triglycerides and a markedly decreased level of high density lipoprotein-cholesterol (36). In the nonobese diabetic mouse model of human autoimmune insulin-dependent diabetes mellitus, partial protection from disease is provided by resistance alleles occurring singly at either the Idd3 or Idd10 non-major histocompatibility complex loci, whereas epistatic interaction between resistance alleles at these two loci produces nearly complete protection from diabetes (37).

There are several mechanisms by which wellness or protective alleles could affect the clinical manifestations of BPAD in the Old Order Amish. One possibility is that dominant acting protective alleles, either singly or acting together in epistasis, could prevent or modify the BPAD phenotype. The variable penetrance of illness or its heterogeneous clinical manifestations could result from resistance or protective alleles that alone provide only partial protection, while together with other genes produce epistatic interactions, resulting in a greater degree of modification of the phenotype. Alternatively, there also could be cellular target molecules, e.g., mood effectors, having forms that are either resistant or susceptible to the genetic and/or environmental susceptibility factors for BPAD. Individuals having resistant mood effectors would be protected from the effects of susceptibility alleles and/or environmental factors that result in the BPAD phenotype. In contrast, individuals with sensitive forms of these mood effectors would be vulnerable to developing the BPI phenotype when requisite BPAD susceptibility alleles and/or environmental factors are present. If epistatic interactions are required for manifestation of the effects of either susceptibility or protective alleles, the existence of resistant and sensitive forms of cellular effectors or protective alleles would be most apparent in families (or populations) where there is a high density of affected individuals such as the Old Order Amish in the present study. Regardless of the mechanism, the presence of wellness or protective alleles would have a significant impact on linkage analyses as evidenced by preventing the appearance of the BPAD phenotype (or its presentation as a forme fruste) in individuals who are otherwise genetically predisposed to developing illness.

Thus, we suggest that the identification of chromosomal loci harboring genes that contribute to the clinical manifestations of BPAD will likely require a multilocus approach that considers both additive and subtractive influences of alleles on the BPAD phenotype. The involvement of protective or wellness alleles in determining the manifestation of the BPAD phenotype would provide an attractive, testable explanation for at least some of the difficulty encountered in searches for BPAD susceptibility alleles.

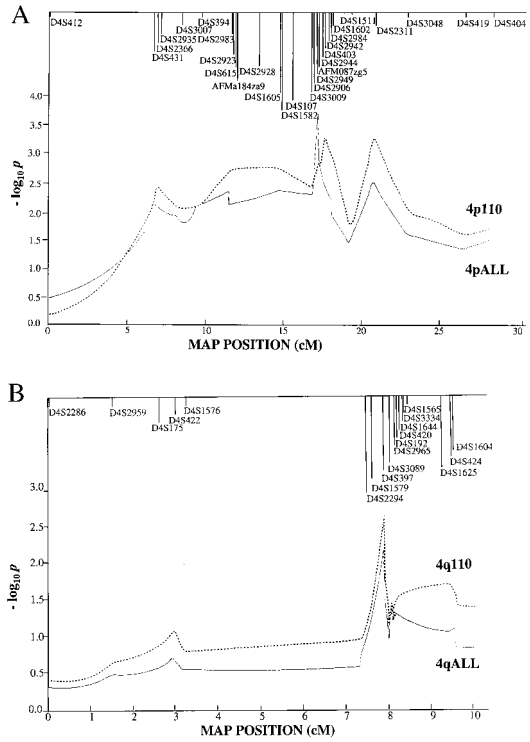

FIG. 2. Model-free linkage analysis of wellness by using GH-PLUS. Map position is in Kosambi centimorgans. The $-\log_{10} P$ was calculated by using P values generated by GH-PLUS on the assumption that the NPL score is standard normally distributed. A $-\log_{10} P$ of 4.0 corresponds asymptotically to a logarithm of odds score of 3.0. Only mentally healthy individuals 45 years of age or older were classified as being well (see text). (*A*) $-\log_{10} P$ for markers on chromosome 4p: dotted line, pedigree 110 only; and solid line, pedigrees 110, 210, 310, and 410. (*B*) $-\log_{10} P$ for markers on chromosome 4q: dotted line, pedigree 110 only, and solid line, pedigrees 110, 210, 310, and 410.

The test statistics from our analyses for alleles linked to the absence of psychiatric illness in the Old Order Amish are at least as significant as those reported for any susceptibility locus, and further investigations into the significance of these findings in the inheritance of BPAD are warranted. The identification and characterization of protective alleles and their gene products could lead to the development of a more rational and direct approach to effective therapy for affective disorders.

We thank the Amish Study Psychiatric Board (Drs. J.N. Sussex, J.J. Schwab, J. Endicott, D.R. Offord, and A.M. Hostetter) and the Amish families for more than 20 years of participation; J.R. Engle for administrative help; Alma Becker for phlebotomy; the Coriell Institute for Medical Research for establishing and maintaining Amish Study collection of cell lines; Azita Kashani, Misha Schmilovish, and Yu-Hsien Lee for technical assistance; B. Kuhns for help with manuscript preparation; and Dr. Seymour Kety and many other colleagues for critical and helpful discussions. The Amish Study (1976–1994) was supported by National Institute of Mental Health Grant MH28287 to J.A.E., with additional support from the Stanley Foundation, National Alliance for the Mentally Ill (1996–1998). Support to R.C.E., N.J.S., and P.St.J. was, in part, through National Center for Research Resources Grant RR03655, and to R.C.E. through National Institute of General Medical Sciences Grant 28356. A portion of this work was supported by a Cooperative Research and Development Agreement (CRADA) with Eli Lilly and Company.

1. Goodwin, F. K. & Jamison, K. R. (1990) *Manic-Depressive Illness* (Oxford Univ. Press, New York).
2. Craddock, N. & McGuffin, P. (1993) *Ann. Med.* **25,** 317–322.
3. Egeland, J. A., Gerhard, D. S., Pauls, D. L., Sussex, J. N., Kidd, K. K., Allen, C. R., Hostetter, A. M. & Housman, D. E. (1987) *Nature (London)* **325,** 783–787.
4. Kelsoe, J. R., Ginns, E. I., Egeland, J. A., Gerhard, D. S., Goldstein, A. M., Bale, S. J., Pauls, D. L., Long, R. T., Kidd, K. K., Conte, G., *et al.* (1989) *Nature (London)* **342,** 238–243.
5. Berrettini, W. H., Ferraro, T. N., Goldin, L. R., Weeks, D. E., Detera-Wadleigh, S., Nurnberger, J. I., Jr. & Gershon, E. S. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5918–5921.
6. Straub, R. E., Lehner, T., Luo, Y., Loth, J. E., Shao, W., Sharpe, L., Alexander, J. R., Das, K., Simon, R., Fieve, R. R., *et al.* (1994) *Nat. Genet.* **8,** 291–296.
7. Pekkarinen, P., Terwilliger, J., Bredbacka, P. E., Lonnqvist, J. & Peltonen, L. (1995) *Genome Res.* **5,** 105–115.
8. Pauls, D. L., Ott, J., Paul, S. M., Allen, C. R., Fann, C. S., Carulli, J. P., Falls, K. M., Bouthillier, C. A., Gravius, T. C., Keith, T. P., *et al.* (1995) *Am. J. Hum. Genet.* **57,** 636–643.
9. Ginns, E. I., Ott, J., Egeland, J. A., Allen, C R., Fann, C. S. J., Pauls, D. L., Weissenbach, J., Carulli, J. P., Falls, K. M., Keith, T. P. & Paul, S. M. (1996) *Nat. Genet.* **12,** 431–435.
10. National Institute of Mental Health Genetics Initiative Bipolar Group (1997) *Am. J. Med. Genet.* **74,** 227–269.
11. Blackwood, B. D. R., He, L., Morris, S. W., McLean, A., Whitton, C., Thomson, M., Walker, M. T., Woodburn, K., Sharp, C. M., Wright, A. F., *et al.* (1996) *Nat. Genet.* **12,** 427–430.
12. Freimer, N. B., Reus, V. I., Escamilla, M. A., McInnes, L. A., Spesny, M., Leon, P., Service, S. K., Smith, L. B., Silva, S., Rojas E., *et al.* (1996) *Nat. Genet.* **12,** 436–441.
13. Philibert, R. A., Egeland, J. A., Paul, S. M. & Ginns, E. I. (1997) *J. Affective Disorders* **43,** 1–3.
14. Egeland, J. A., Sussex, J., Endicott, J., Hostetter, A., Offord, D., Schwab, J. & Pauls, D. (1990) *Psychiatr. Genet.* **1,** 5–18.
15. Hostetter, A. M., Egeland, J. A. & Endicott, J. (1983) *Am. J. Psychiatry* **140,** 62–66.
16. Spitzer, R., Endicott, J. & Robins, E. (1978) *Arch. Gen. Psychiatry* **35,** 773–782.
17. Pauls, D. L., Morton, L. & Egeland, J. A. (1992) *Arch. Gen. Psychiatry* **49,** 703–708.
18. Egeland, J. A. (1994) in *Genetic Studies in Affective Disorders*, eds. Papolos, D. F. & Lachman, H. M. (Wiley, New York), pp. 70–90.
19. Endicott, J. & Spitzer, R. (1978) *Arch. Gen. Psychiatry* **35,** 837–844.
20. Egeland, J. A. (1994) in *1994–1995 Catalog of Cell Lines* (NIGMS Human Genetic Mutant Cell Repository, Camden, N.J.), pp. 408–428 and 992–999.
21. Collins, A., Frezal, J., Teague, J. & Morton, N. E. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 14771–14775.
22. Schaffer, A. A. (1996) *Hum. Hered.* **46,** 226–235.
23. Elston, R., Bailey-Wilson, J., Bonney, G., Tran, L., Keats, B. & Wilson, A. (1997) S.A.G.E., Statistical Analysis for Genetic Epidemiology, version 3.1 (Case Western Reserve University, Cleveland, OH).
24. Kruglyak, L., Daly, M. J., Reeve-Daly, M. P. & Lander, E. S. (1996) *Am. J. Hum Genet.* **58,** 1347–1363.
25. Lange, K., Weeks, D. & Boehnke, M. (1988) *Genet. Epidemiol.* **5,** 471–472.
26. Boehnke, M. (1991) *Am. J. Hum. Genet.* **48,** 22–25.
27. Schweder, T. & Spjotvoll, E. (1982) *Biometrika* **69,** 493–502.
28. Witte, J. S., Elston, R. C. & Schork, N. J. (1996) *Nat. Genet.* **12,** 355–358.
29. Drigalenko, E. L. & Elston, R. C. (1997) *Genet. Epidemiol.* **14,** 779–784.
30. Lander, E. S. & Kruglyak, L. (1995) *Nat. Genet.* **11,** 241–247.
31. Genin, E. & Clerget-Darpoux, F. (1996) *Am. J. Hum. Genet.* **59,** 1149–1162.
32. Elston, R. C. (1995) *Exp. Clin. Immunogenet.* **12,** 129–140.
33. Perrine, R. P., Brown, M. J., Clegg, J. B., Weatherall, D. J. & May, A. (1972) *Lancet* **2,** 1163–1167.
34. Picchio, G. R., Gulizia, R. J. & Mosier, D. E. (1997) *J. Virol.* **71,** 7124–7127.
35. Corder, E. H., Saunders, A. M., Risch, N. J., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Jr., Rimmler, J. B., Locke, P. A., Conneally, P. M., Schmader, K. E., *et al.* (1994) *Nat. Genet.* **7,** 180–183.
36. Franceschini, G., Sirtori, C. R., Capurso, A. 2nd, Weisgraber, K. H. & Mahley, R. W. (1980) *J. Clin. Invest.* **66,** 892–900.
37. Wicker, L. S., Todd, J. A., Prins, J. B., Podolin, P. L., Renjilian, R. J. & Peterson, L. B. (1994) *J. Exp. Med.* **180,** 1705–1713.
38. Young, A. (1993–1995) Genetic Analysis Software (G.A.S.), version 2.0 (Oxford University, Oxford).