Evidence for a Susceptibility Gene for Anorexia Nervosa on Chromosome 1


Eating disorders, such as anorexia nervosa (AN), have a significant genetic component. In the current study, a genomewide linkage analysis of 192 families with at least one affected relative pair with AN and related eating disorders, including bulimia nervosa, was performed, resulting in only modest evidence for linkage, with the highest nonparametric linkage (NPL) score, 1.80, at marker D4S2367 on chromosome 4. Since the reduction of sample heterogeneity would increase power to detect linkage, we performed linkage analysis in a subset of families in which at least two affected relatives had diagnoses of restricting AN, a clinically defined subtype of AN characterized by severe limitation of food intake without the presence of binge-eating or purging behavior. When we limited the linkage analysis to this clinically more homogeneous subgroup, the highest multipoint NPL score observed was 3.03, at marker D1S3721 on chromosome 1p. The genotyping of additional markers in this region led to a peak multipoint NPL score of 3.45, thereby providing suggestive evidence for the presence of an AN-susceptibility locus on chromosome 1p.

Eating disorders—psychiatric disorders characterized by severe disturbances in eating behavior that typically have onset during late adolescence and young adulthood—are divided into three major types: anorexia nervosa (AN), bulimia nervosa (BN), and eating disorder not otherwise specified (EDNOS). AN is characterized by obsessive fear of weight gain, severely restricted eating, and low body weight, and, in women, AN has the highest mortality among the psychiatric disorders (Sullivan 1995); AN is divided into two clinical subtypes: restricting anorexia nervosa (RAN) and binge-eating/purging anorexia nervosa (BPAN). A related disorder (Hsu et al. 1979; Casper et al. 1980; Strober et al. 1982; Bulik et al. 1997), BN, can occur at any body weight and is characterized by binge-eating and compensatory weight-loss behaviors. Family studies have shown significantly increased prevalence of eating disorders in relatives of probands with AN, compared to relatives of control probands (Gershon et al. 1983; Strober et al. 1985; Hudson et al. 1987; Lilenfeld et al. 1998), thereby providing evidence for the familiality of eating disorders. Twin studies of AN have included both clinically derived samples (Garfinkel and Garner 1982; Holland et al. 1984, 1988) and population-based investigations (Wade et al. 2000). In these studies, concordance rates for MZ twins with AN have been estimated to be 52%–56%, whereas concordance rates for DZ twins with AN have been estimated to be 5%–11% (Garfinkel and Garner 1982; Holland et al. 1984, 1988; Wade et al. 2000). These studies strongly suggest that the familiality observed in family studies is primarily due to genetic causes.

We undertook a multicenter study with an allele-
sharing linkage strategy to identify loci that may contribute to AN. Clinical and diagnostic information on the sample is summarized in Table 1. This sample was composed of 192 families with at least two affected relatives, all of which were ascertained through a proband who met DSM-IV criteria (American Psychiatric Association 1994, pp. 544–545) for >3 years when criterion D, amenorrhea, for AN was waived (because some subjects were menstruating owing to treatment with exogenous hormone replacement and because some subjects were male) and who reported a family history of eating disorders in a first-degree relative (for a detailed description of this sample, see Kaye et al. [2000]). Affected relatives met DSM-IV (American Psychiatric Association 1994, pp. 544–545) criteria for AN, BN, or EDNOS (Kaye et al. 2000). Full assessments were completed on the proband and affected relative(s) (for details, see Kaye et al. 2000). Eating-disorder diagnoses were confirmed by trained raters with the Structured Interview for Anorexia Nervosa and Bulimic Syndromes (Fichter et al. 1998). Of the 230 relative pairs, 88.8% were female-female, 10.8% were female-male, and 0.4% was male-male, consistent with the increased prevalence of eating disorders in females. Mean ± SD age of the probands was 27.8 ± 9.4 years. More than 98% of the families were European ancestry, with 138 families from North America, 37 families from Germany, and 21 families from England. Informed consent was obtained from all study participants, and all sites received approval from their local institutional review board.

In the entire sample of families, the only multipoint NPL score >1.5 was observed at marker D4S2367 on chromosome 4 (see boldface italic values under the heading “Entire Sample” in Table 2). Values from multipoint LOD analysis that allows for heterogeneity (HLOD), under both dominant and recessive inheritance are shown. The estimated α values (i.e., estimates of the fraction of families showing linkage at that locus) under dominant inheritance was 22% at this position. Multipoint NPL scores >1.0 were observed at marker D4S403 on chromosome 4 (NPL score 1.42; P = .07), at marker D11S1392 on chromosome 11 (NPL score 1.20; P = .11), at marker D13S796 on chromosome 13 (NPL score 1.03; P = .15), and at marker D15S657 on chromosome 15 (NPL score 1.23; P = .11). This modest evidence for linkage could reflect either multiple interacting genes, each with weak to moderate effect, and/or sample heterogeneity.

The reduction of sample heterogeneity should increase the likelihood of identification of susceptibility loci for eating disorders. In this study, all families were ascertained through a proband who met criteria for AN; however, the criteria for affection status were broader for affected relatives. With the goal of examining the effect that reduction of clinical heterogeneity has on the genetic linkage analysis of this sample, we used additional clinical data to define a subgroup (n = 37) from the original

### Table 1

<table>
<thead>
<tr>
<th>Characteristics of Probands Partitioned by AN Subtype</th>
<th>Sample with AN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ENTIRE SAMPLE</td>
</tr>
<tr>
<td>Mean ± SD age (years)</td>
<td>27.76 ± 9.38</td>
</tr>
<tr>
<td>Gender (% males)</td>
<td>5.6</td>
</tr>
<tr>
<td>Mean ± SD body-mass index (kg/m²):</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>18.62 ± 2.55</td>
</tr>
<tr>
<td>Minimum (past)</td>
<td>14.27 ± 2.28</td>
</tr>
<tr>
<td>Maximum (past)</td>
<td>21.27 ± 2.88</td>
</tr>
<tr>
<td>No. of families:</td>
<td></td>
</tr>
<tr>
<td>With one affected sibling pair∗</td>
<td>147</td>
</tr>
<tr>
<td>With three or four affected siblings∗</td>
<td>16</td>
</tr>
<tr>
<td>With other affected relative pairs∗</td>
<td>29</td>
</tr>
<tr>
<td>Total</td>
<td>192</td>
</tr>
</tbody>
</table>

**NOTE.**—The original sample was composed of 196 families (4 of which were not used in the genetic analysis, as detailed in the note to Table 2), each including both a proband with a diagnosis of either the RAN or BPAN subtype and one or more family members affected with AN, BN, or EDNOS. For the families in the subgroup with RAN, two or more individuals had a diagnosis of RAN. For the families in the subgroup with AN, two or more individuals had a diagnosis of AN of either subtype. For both the RAN and AN subgroups, any additional family members with BPAN, BN, or EDNOS were identified as affected, in this table and subsequent analyses.

∗ Families may have had additional affected relatives.

† Families included, for example, cousin-cousin relative pairs, aunt-niece relative pairs, or half sibs.
Table 2
Summary of Highest Multipoint NPL Scores

<table>
<thead>
<tr>
<th>CHROMOSOME (MARKER)</th>
<th>DISTANCE (CM)</th>
<th>MULTIPLET NPL z (P)</th>
<th>ENTIRE SAMPLE</th>
<th>HLOD</th>
<th>MULTIPLET NPL z (P)</th>
<th>RECESSIVE</th>
<th>DOMINANT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (D1S3721)</td>
<td>72.37</td>
<td>0.84 (.20)</td>
<td>0.19</td>
<td>0.17</td>
<td>3.03 (.001)</td>
<td>2.0</td>
<td>1.5</td>
</tr>
<tr>
<td>2 (D2S3911)</td>
<td>186.45</td>
<td>0.12 (.45)</td>
<td>0.00</td>
<td>0.00</td>
<td>1.70 (.05)</td>
<td>.43</td>
<td>.70</td>
</tr>
<tr>
<td>4 (D4S403)</td>
<td>25.97</td>
<td>1.42 (.08)</td>
<td>0.26</td>
<td>0.79</td>
<td>1.82 (.03)</td>
<td>.59</td>
<td>1.54</td>
</tr>
<tr>
<td>4 (D4S2367)</td>
<td>78.50</td>
<td>1.80 (.03)</td>
<td>0.23</td>
<td>1.05</td>
<td>2.44 (.008)</td>
<td>.65</td>
<td>2.23</td>
</tr>
<tr>
<td>5 (D5S1501)</td>
<td>85.25</td>
<td>-24 (.59)</td>
<td>0.00</td>
<td>0.00</td>
<td>1.63 (.05)</td>
<td>.02</td>
<td>1.06</td>
</tr>
<tr>
<td>8 (D8S264)</td>
<td>4.04</td>
<td>0.11 (.45)</td>
<td>0.00</td>
<td>0.03</td>
<td>1.90 (.03)</td>
<td>1.36</td>
<td>.70</td>
</tr>
<tr>
<td>9 (D9S922)</td>
<td>80.08</td>
<td>.78 (.22)</td>
<td>0.09</td>
<td>0.21</td>
<td>1.51 (.07)</td>
<td>.28</td>
<td>.78</td>
</tr>
<tr>
<td>16 (D16S748)</td>
<td>22.19</td>
<td>-.78 (.78)</td>
<td>0.00</td>
<td>0.00</td>
<td>1.70 (.05)</td>
<td>.55</td>
<td>.34</td>
</tr>
<tr>
<td>22 (D22S6383)</td>
<td>36.16</td>
<td>-.63 (.73)</td>
<td>-.01</td>
<td>0.00</td>
<td>1.50 (.07)</td>
<td>.44</td>
<td>.13</td>
</tr>
</tbody>
</table>

Note.—A peripheral-blood sample was collected from each proband, affected relative, and participating biological parent, and DNA was extracted according to standard methods (Lahiri and Nurnberger 1991). The Weber screening set, version 9 (Center for Medical Genetics, Marshfield Medical Research Foundation), which is a screening panel of 386 fluorescently tagged markers dispersed at a density slightly less than 10 cM, was used for the genomewide scan. Mean ± SD marker heterozygosity was 0.77 ± 0.06. Genotype information and reported family structures were compared for inconsistencies (including paternity relationships, twin status, etc.). DNA from four families yielded genotypes that were inconsistent with the stated family structure and were excluded from subsequent genetic analysis. Multipoint NPL scores were determined using GENEHUNTER (Kruglyak et al. 1996). The z score and associated P value for markers with peak NPL scores ≥1.5 for each sample are indicated in boldface italic. In addition, peak scores and from multipoint HLOD analysis are shown (Smith 1963), under dominant and recessive models and with penetrance set at 50% (Vieland et al. 1992; Greenberg et al. 1998; Dürner et al. 1999). For these parametric analyses, disease-allele frequency was arbitrarily set at 0.1 for recessive inheritance and .006 for dominant inheritance.

In the current study, we looked for susceptibility loci for eating disorders. This study represents the first genomewide linkage analysis of eating disorders. To date, population-based association studies have been the primary approach for identifying genetic factors contributing to these disorders (Kaye et al. 1997). However, genetic heterogeneity can make these approaches challenging, as different subtypes of eating disorders may have distinct genetic etiologies.

In the current study, we focused on a sample of families in which two or more affected individuals had the RAN subtype (table 1). We defined the RAN subtype as individuals who met the diagnostic criteria for RAN and had at least one affected relative. This definition was chosen because RAN is a relatively homogenous subtype compared to other eating disorders, such as BN or EDNOS, which have more heterogeneous diagnostic criteria. The RAN subtype includes individuals with a diagnosis of AN and one or more relatives with AN of either the RAN or BPAN subtype, with a diagnosis of RAN (fig. 1). Two of these peaks (at markers D4S403 and D4S2367 on chromosome 4) were observed in the sample with RAN (table 1). In the entire sample, the NPL score was 0.83 (table 1). In the subgroup with AN, the NPL score was 1.42, whereas, in the entire sample, the NPL score was 0.83 (fig. 1). Thus, as the affection model was narrowed, the peak NPL scores at this location increased. This increase in evidence for linkage is consistent with there being decreasing genetic heterogeneity as the phenotype is refined. At the second-highest peak (on chromosome 4 at marker D4S2367) in the original sample, there was a more modest increase in NPL score as the affection model was narrowed (table 2).

To increase the genetic information concerning the observed peak on chromosome 1, we genotyped three additional markers (D1S255, D1S5198, and D1S447) in the families with RAN. Additional markers in this region resulted in a peak multipoint NPL score of 3.45 (P = .00038) at marker D1S3721 in the subgroup with RAN (fig. 2). Multipoint HLOD scores were 1.68 under the dominant model and 2.43 under the recessive model (fig. 2). Estimated a values were 0.45 under the dominant model and 0.47 under the recessive model.

In the current study, we looked for susceptibility loci for eating disorders. This study represents the first genomewide linkage analysis of eating disorders. To date, population-based association studies have been the pri-

Sample of families in which two or more affected individuals had the RAN subtype (table 1). We defined the RAN subtype as individuals who met the diagnostic criteria for RAN and had at least one affected relative. This definition was chosen because RAN is a relatively homogenous subtype compared to other eating disorders, such as BN or EDNOS, which have more heterogeneous diagnostic criteria. The RAN subtype includes individuals with a diagnosis of AN and one or more relatives with AN of either the RAN or BPAN subtype, with a diagnosis of RAN (fig. 1). Two of these peaks (at markers D4S403 and D4S2367 on chromosome 4) were observed in the sample with RAN (table 1). In the entire sample, the NPL score was 0.83 (table 1). In the subgroup with AN, the NPL score was 1.42, whereas, in the entire sample, the NPL score was 0.83 (fig. 1). Thus, as the affection model was narrowed, the peak NPL scores at this location increased. This increase in evidence for linkage is consistent with there being decreasing genetic heterogeneity as the phenotype is refined. At the second-highest peak (on chromosome 4 at marker D4S2367) in the original sample, there was a more modest increase in NPL score as the affection model was narrowed (table 2).

To increase the genetic information concerning the observed peak on chromosome 1, we genotyped three additional markers (D1S255, D1S5198, and D1S447) in the families with RAN. Additional markers in this region resulted in a peak multipoint NPL score of 3.45 (P = .00038) at marker D1S3721 in the subgroup with RAN (fig. 2). Multipoint HLOD scores were 1.68 under the dominant model and 2.43 under the recessive model (fig. 2). Estimated a values were 0.45 under the dominant model and 0.47 under the recessive model.

In the current study, we looked for susceptibility loci for eating disorders. This study represents the first genomewide linkage analysis of eating disorders. To date, population-based association studies have been the pri-
mary approach to the study of the genetics of eating disorders. Thus far, these candidate-gene analyses have demonstrated only modest success in the identification of genes that may be related to AN susceptibility.

The peak scores that we identified in the entire sample of affected families may indicate the presence of susceptibility loci at these positions but may also represent false positives. The absence of higher peaks in this large sample could be due to a number of potential factors:

1. Although eating disorders are substantially heritable, the genetic susceptibility may be due to a number of interacting loci of small-to-modest effect. If eating disorders arise from a number of interacting genes, then susceptibility loci may reflect variants of small effect that are fairly common in the general population. These variants therefore are difficult to identify by linkage analysis.

2. As discussed above, there may be substantial genetic heterogeneity in eating disorders—a possibility that is supported by the low \( \alpha \) values observed parametric linkage analysis in the entire sample.

The reduction of sample heterogeneity by narrowing the affection-status model resulted in a peak NPL score of 3.03, at marker D1S3721 on chromosome 1, which increased to 3.45 when additional markers were considered. Evidence for linkage in the subgroup with RAN was also seen on chromosomes 2, 4, 5, 8, 9, 16, and 22. The evidence for linkage in the subgroup with RAN must be tempered by our study design, which included multiple tests. Therefore, until replicated in an independent sample, these results can be taken as only suggestive linkage.

That a more restrictive phenotype increases NPL scores and HLODs at marker D1S3721 on chromosome 1 sug-

![Figure 1](image1)

**Figure 1** Evidence for linkage on chromosome 1 in the subgroup with AN. Multipoint NPL \( z \) scores (Kruglyak et al. 1996) are plotted, as a function of chromosomal position, for the three categories of families defined in table 1: the subgroup with RAN (\( n = 37 \) families; **solid line**), the subgroup with AN (\( n = 127 \) families; **dashed line**), and the entire eating-disorder sample (\( n = 192 \) families; **dotted line**).

![Figure 2](image2)

**Figure 2** Evidence for linkage on chromosome 1p, with additional markers in the subgroup with RAN. Multipoint NPL scores (**dotted line**) were calculated across the indicated interval, as were multipoint HLOD analyses performed under dominant (**solid line**) and recessive (**dashed line**) models, as detailed in the note to table 2. Across this interval, markers were D1S552, D1S1622, D1S255, D1S598, D1S3721, D1S447, and D1S2134. Positions of markers are indicated. The highest score was over the position of marker D1S3721.
gests that there is decreased heterogeneity in the subgroup with RAN, which is consistent with the clinical presentation of RAN. Decreased heterogeneity is also supported by the higher \( \alpha \) values observed in the sample with RAN, compared to both the sample with AN and the entire sample. To evaluate whether the evidence for linkage at marker D1S3721 was derived primarily from the subgroup with RAN, we performed linkage analysis in the subset of all families except the 37 families with RAN \((n = 155)\). By Morton’s predivided sample test, these comparisons provided significant evidence for heterogeneity, for both the peak NPL score and HLOD (with \( \chi^2 = 10.13, df = 1, and P = .0015 \) and \( \chi^2 = 8.43, df = 1, and P = .0037 \), respectively), and support genetic heterogeneity in eating disorders (Morton 1956). Interestingly, <25% of relative-pair families with RAN had additional family members in whom other eating disorders had been diagnosed. This suggests that the genetic vulnerability in these families may lead to manifestation of eating disorders beyond RAN. In summary, our data support the presence, on chromosome 1, of a gene for AN and further suggest that the use of restrictive diagnoses can decrease heterogeneity and thereby facilitate the identification of such genes.

Acknowledgments

The authors wish to thank the Price Foundation for the support of the clinical collection of subjects, genotyping, and data analysis. Genotypic markers, allele frequencies, and genetic maps were generated at the Center for Medical Genetics, Marshfield Medical Research Foundation, with support from the National Heart, Lung and Blood Institute. The authors are indebted to the participating families for their contributions of time and effort in support of this study.

Electronic-Database Information

The URL for data in this report is as follows:

Center for Medical Genetics, Marshfield Medical Research Foundation, http://research.marshfieldclinic.org/genetics/ (for Weber screening set, version 9)

References


Smith CAB (1963) Testing for heterogeneity of the recombi-