# Linkage Analysis of Anorexia and Bulimia Nervosa Cohorts Using Selected Behavioral Phenotypes as Quantitative Traits or Covariates

Silviu-Alin Bacanu,  $^1$  Cynthia M. Bulik,  $^2$  Kelly L. Klump,  $^3$  Manfred M. Fichter,  $^4$  Katherine A. Halmi,  $^5$  Pamela Keel,  $^6$  Allan S. Kaplan,  $^{7,8}$  James E. Mitchell,  $^9$  Alessandro Rotondo,  $^{10}$  Michael Strober,  $^{11}$  Janet Treasure,  $^{12}$ D. Blake Woodside,<sup>8</sup> Vibhor A. Sonpar,<sup>1</sup> Weiting Xie,<sup>1</sup> Andrew W. Bergen,<sup>13</sup> Wade H. Berrettini,<sup>14</sup> Walter H. Kaye, 1\* and Bernie Devlin 1\*

To increase the likelihood of finding genetic variation conferring liability to eating disorders, we measured over 100 attributes thought to be related to liability to eating disorders on affected individuals from multiplex families and two cohorts: one recruited through a proband with anorexia nervosa (AN; AN cohort); the other recruited through a proband with bulimia nervosa (BN; BN cohort). By a multilayer decision process based on expert evaluation and statistical analysis, six traits were selected for linkage analysis (1): obsessionality (OBS), age at menarche (MENAR), and anxiety (ANX) for quantitative trait locus (QTL) linkage analysis; and lifetime minimum body mass index (BMI), concern over mistakes (CM), and food-related obsessions (OBF) for covariate-based linkage analysis. The BN cohort produced the largest linkage signals: for QTL linkage analysis, four suggestive signals: (for MENAR, at 10p13; for ANX, at 1q31.1, 4q35.2, and 8q13.1); for covariate-based linkage analyses, both significant and suggestive linkages (for BMI, one significant [4q21.1] and three suggestive [3p23, 10p13, 5p15.3]; for CM, two significant [16p13.3, 14q21.1] and three suggestive [4p15.33, 8q11.23, 10p11.21]; and for OBF, one significant [14q21.1] and five suggestive [4p16.1, 10p13.1, 8q11.23, 16p13.3, 18p11.31]). Results from the AN cohort were far less compelling: for QTL linkage analysis, two suggestive signals (for OBS at 6q21 and for ANX at 9p21.3); for covariate-based linkage analysis, five suggestive signals (for BMI at 4q13.1, for CM at 11p11.2 and 17q25.1, and for OBF at 17q25.1 and 15q26.2). Overlap between the two cohorts was minimal for substantial linkage signals. © 2005 Wiley-Liss, Inc.

**KEY WORDS:** complex disease; endophenotype; liability; mixture model; regres-

#### INTRODUCTION

Eating disorders span a substantial behavioral spectrum. Anorexia nervosa (AN) is typified by rigid maintenance of abnormally low body weight through restriction of food intake, excessive exercise, and/or purging. Bulimia nervosa (BN) is typified by maintenance of normal weight in the presence of binge eating coupled with compensatory behaviors. Individuals with AN tend to be inhibited and over-controlled [Wonderlich et al., 2005]; although some individuals with BN share these traits, others exhibit a more classic pattern of disinhibition and undercontrol [Bulik et al., 1995; Fassino et al., 2004; Steiger et al., 2004]. AN and BN are partially overlapping conditions. Features that unite AN and BN include the common occurrence of diagnostic "crossover" or converting from one disorder to the other, especially from AN to BN [Tozzi et al., 2005]. The fact that eating disorders do not appear to "breed true" in families, with rates of both disorders being elevated in family members of AN and BN probands [Lilenfeld et al., 1998], and the fact that co-twins of individuals with AN are at significantly increased risk for BN [Walters and

<sup>&</sup>lt;sup>1</sup>Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania

<sup>&</sup>lt;sup>2</sup>Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

<sup>&</sup>lt;sup>3</sup>Department of Psychology, Michigan State University, East Lansing, Michigan

 $<sup>^4</sup>$ Klinik Roseneck, Hospital for Behavioral Medicine, Affiliated With the University of Munich, Prien, Germany

 $<sup>^5</sup>$ New York Presbyterian Hospital, Weill Medical College of Cornell University, White Plains, New York

<sup>&</sup>lt;sup>6</sup>Department of Psychology, University of Iowa, Iowa City, Iowa

<sup>&</sup>lt;sup>7</sup>Program for Eating Disorders, Toronto General Hospital, Toronto, Ontario, Canada

<sup>&</sup>lt;sup>8</sup>Department of Psychiatry, Toronto General Hospital, Toronto, Ontario, Canada

<sup>&</sup>lt;sup>9</sup>Neuropsychiatric Research Institute, Fargo, North Dakota

<sup>&</sup>lt;sup>10</sup>Department of Psychiatry, Neurobiology, Pharmacology and Biotechnologies, University of Pisa, Italy

<sup>&</sup>lt;sup>11</sup>Department of Psychiatry and Behavioral Science, University of California at Los Angeles, Los Angeles, California

<sup>&</sup>lt;sup>12</sup>Eating Disorders Unit, Institute of Psychiatry and South London and Maudsley National Health Service Trust, United Kingdom

<sup>&</sup>lt;sup>13</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland

<sup>&</sup>lt;sup>14</sup>Center of Neurobiology and Behavior, University of Pennsylvania, Philadelphia, Pennsylvania

This article contains supplementary material, which may be viewed at the American Journal of Medical Genetics website at http://www.interscience.wiley.com/jpages/1552-4841/suppmat/

Andrew W. Bergen's manuscript does not represent the opinion of the NIH, the DHHS, or the Federal Government.

Silviu-Alin Bacanu's present address is GlaxoSmithKline, 5 Moore Drive, Research Triangle Park, NC 27709.

<sup>\*</sup>Correspondence to: Walter H. Kaye and Bernie Devlin, Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania 15213-2593.

E-mail: whkaye@upmc.edu; devlinbj@upmc.edu

Received 28 January 2005; Accepted 16 June 2005 DOI 10.1002/ajmg.b.30226

Kendler, 1995], suggest that the disorders could share some genetic vulnerabilities.

To investigate the genetic basis of eating disorders, we have recruited two cohorts of families multiplex for eating disorders: an AN cohort, in which the proband of the family was diagnosed with AN but other affected family members could have any other eating disorder diagnosis (BN or eating disorder not-otherwise-specified); and a BN cohort, which was similarly recruited except that the proband was required to have BN. The structure of these families was usually quite simple, consisting most often of affected sibling pairs, and less often of affected relative pairs.

Psychometric studies have linked AN and BN to a cluster of moderately heritable personality and temperamental traits, such as obsessionality, perfectionism, and harm avoidance [Klump et al., 2000; Halmi et al., 2003; Bulik et al., 2003a; Fassino et al., 2004]. We have measured a wide variety of psychiatric, personality, and temperamental phenotypes on all affected individuals in the multiplex families [Kaye et al., 2000, 2004]. By using a combined approach of genome-wide linkage analysis, teamed with the analysis of phenotypes related to eating disorders, we hope to identify some of the polymorphisms that contribute to eating-disorder liability.

Three linkage analyses were reported on these data. One analysis focused on the first cohort to be collected, namely the AN cohort [Grice et al., 2002]. No substantial linkage signal arose by analysis of the entire cohort, but a suggestive linkage at 1p33-36 emerged by analysis of a small subset of families in which there were multiple individuals with the classic restricting subtype of anorexia nervosa (RAN). RAN individuals maintain abnormally low body weight by restriction of food intake and are well known for extreme behavioral and personality traits, such as drive-for-thinness, perfectionism, and obsession with symmetry and exactness. Analysis of the BN cohort [Bulik et al., 2003b] revealed significant linkage was at 10p13, as well as a suggestive signal at 14q22.2-23.1. By restricting the sample to families in which self-induced vomiting was a salient feature—a behavioral feature that identifies a substantially heritable component of BN [Sullivan et al., 1998]—the linkage signal at 10p13 was amplified substantially.

A third linkage analysis was reported for the AN cohort [Devlin et al., 2002a]. This study evaluated some of the personality and behavioral traits for their potential as covariates in linkage analysis. From a handful of traits chosen for in-depth analysis, two were selected for inclusion in the linkage analysis, namely drive-for-thinness and obsessionality. Together these covariates produced a significant signal for linkage at 1q31.1 [Devlin et al., 2002a; Bacanu, 2005] and two suggestive signals at 2p11.2 and 13q13.3. These analyses employed a relatively novel linkage method in which covariates were assumed to cluster families into linked and unlinked groups, and this feature of the covariates is incorporated into a mixture model [Devlin et al., 2002b].

Until recently, however, the entire set of possible traits had not been explored for their potential for linkage analysis on either the AN or BN cohorts. To select a parsimonious subset of these attributes for linkage analysis, as well as the type of linkage analysis to be applied, we [Bulik et al., 2005] subjected the entire set of variables to a multilayer decision process based on expert evaluation and statistical analysis. Several criteria were critical for trait choice: relevance to eating disorders pathology; demonstrated heritability; and evidence for familiality in our data. Based on statistical diagnostics, six traits were chosen for linkage analysis. Three displayed features of heritable quantitative traits—obsessionality (OBS), age at menarche (MEN), and a composite anxiety measure (ANX)—and seemed best suited for linkage analysis for a quantitative trait locus (QTL). The distributions of the

three other traits in our families—lifetime minimum body mass index (BMI), concern over mistakes (CM), and food-related obsessions (OBF)—differed from that expected under standard quantitative genetic models. Instead, all three clustered families, in that some families showed highly concordant and extreme values for these traits whereas others did not. Thus, data for BMI, CM, and OBF were more compatible with covariate-based linkage analysis.

In this report, we implement the analyses chosen in our previous report [Bulik et al., 2005]. To account for the substantial likelihood of heterogeneity between the AN and BN cohorts, we analyzed the data by cohort and then by aggregating the samples. Therefore,  $6\times 3=18$  linkage analyses will be reported here. We recognize these analyses involve multiple testing. For analyses of data from complex diseases, in which our goal is to find genetic variation affecting the observable phenotypes, data exploration is essential.

#### MATERIALS AND METHODS

# **Samples**

The AN cohort includes psychological assessments and blood samples from 196 probands, all diagnosed with AN, 183 affected full siblings, and 46 other affected second and third degree relatives. Independent from the AN families, the BN cohort include psychological assessments and blood samples from 308 families recruited through a proband diagnosed with BN: 260 ASP; 14 half-siblings; 42 avuncular; 42 cousin; 21 other. Affected relatives could have any eating disorder diagnosis. See Bulik et al. [2005] and Kaye et al. [2000, 2004] for more details. Trait values were also measured in a sample of individuals who were screened to be free from lifetime eating disorders and other major psychopathology [Bulik et al., 2005]. The mean and variance from the control population was used to adjust for ascertainment, by transforming the trait values of the individuals with eating disorders (i.e., by subtracting the mean of the control population and dividing by its standard deviation).

# **Phenotypes**

Our previous research gave detailed description of the evaluated traits, selection of traits for linkage analysis, matching of the traits to linkage method according to the attributes of the traits in families, and citations to reference material [Bulik et al., 2005; Kaye et al., 2000, 2004]. Briefly, phenotypes were derived from the following diagnostic instruments: Structured Interview on Anorexia Nervosa and Bulimic Syndromes; Structured Clinical Interview for DSM IV Axis I Disorders (SCID I) and Axis II Disorders; Yale-Brown Obsessive Compulsive Scale: and Yale-Brown-Cornell Eating Disorder Scale. Phenotypes related to core eating disorder symptoms, mood, temperament, and personality were also derived from the following self-report instruments: Eating Disorders Inventory; State-Trait Anxiety Inventory Form Y; Multidimensional Perfectionism Scale; Temperament and Character Inventory; NEO Personality Inventory; Barratt Impulsivity Scale (BIS-11); and Beck Depression Inventory.

#### **Molecular Analyses**

DNAs from both the AN and BN cohorts were genotyped for the Weber/CHLC Screening Set 9 (http://research.marshfield-clinic.org/genetics/) by the Marshfield Genotyping Service. Error rates were uniformly low (<1%), all markers produced usable data from the AN cohort, but 26 generated incomplete data on the first round of genotyping from the BN cohort, and were excluded from further analysis [Devlin et al., 2002b; Bulik et al., 2003b].

#### **Statistical Analyses**

We evaluated markers and pedigrees for Mendelian errors using the PedCheck program [O'Connell and Weeks, 1998]. Genotyping errors were set to missing. Nominal and imputed genetic relationships among individuals from the same family were contrasted using the Relpair program [Boehnke and Cox, 1997]. To estimate marker allele frequencies, we counted alleles while ignoring family relationships.

For QTL linkage analysis, we used Merlin software [Abecasis et al., 2002], specifically the regress option, which implements the methods of Sham et al. [2002]. To specify the required population parameters of the trait distribution, we used the distribution from the sample of unaffected individuals described previously [Bulik et al., 2005]. For covariate-based linkage analysis, we chose the pre-clustering method [Devlin et al., 2002b; http://wpicr.wpic.pitt.edu/WPICCompGen/]. For the covariate-based linkage analysis and for families who had more than two affected siblings, we formed all possible pairs of ASP after determining the joint identity-by-descent (IBD) status using GENEHUNTER [Kruglyak et al., 1996] and after determining the probability of membership in the linked group. The latter was calculated as described elsewhere [Devlin et al., 2002b]. We did not correct for these dependent observations because they have little impact on the distribution of the test statistic under the null hypothesis [Greenwood and Bull, 1999].

# **RESULTS**

Consistent with selection of traits by Bulik et al. [2005], we summarize trait distributions by using the data from affected siblings only, which are also used for covariate linkage analyses. QTL linkage analyses include data from all affected individuals. To interpret the linkage findings, we use three descriptors: significant and suggestive linkage (Table I), as customarily defined, and noteworthy for any LOD  $\geq$  1.5. Under the null hypothesis of no linkage, the statistic for covariate linkage analysis is approximately distributed as a mixture of chi-square random variables; to put the QTL and covariate statistics on the same scale, we transform the statistic for covariate linkage to LOD scores. Only in figures (supplementary material, referred to as web-Figure) and tables (Table II, web-Table I) will this transformation be noted. Details about the 10 highest linkage scores, by cohort and trait, can be found in Table II; below we point out linkage scores of LOD  $\geq$  1.5.

# **QTL Linkage Analysis**

QTL linkage analysis seeks to identify regions of the genome in which varying levels of IBD predict variation in the quantitative trait, consistent with the presence of a QTL in the region; i.e., near a QTL, family members showing high IBD will tend to have similar values for the quantitative trait.

**QTL linkage analysis for OBS.** Mean OBS for control, AN and BN cohorts was 4.8 (SD = 3.2), 7.15 (SD = 1.98), and 5.06 (SD = 5.87). QTL linkage analysis (web-Fig. 1, Table II) yielded one noteworthy score for the BN cohort (7p21.3), one suggestive (6q21) and one noteworthy (1q31.1) score for AN

TABLE I. Study-Specific Significant/Suggestive Thresholds for Linkage Calculated From the Realized Linkage Traces\*

	AN cohort	BN cohort
Quantitative trait locus (QTL) linkage Covariate linkage	$3.00/1.71 \\ 2.94/1.65$	2.97/1.68 2.91/1.62

<sup>\*</sup>This approach uses an autoregressive model to estimate the correlation between standard normal statistics at adjacent map points and thereby estimate study-specific critical values [Bacanu, 2005].

cohort, and two suggestive  $(1q31.1,\ 7p21.3)$  scores for the combined cohort.

*QTL* linkage analysis for age at menarche (MENAR). Mean MENAR for control, AN and BN cohorts was 12.55 (SD = 1.5), 13.13 (SD = 1.84), and 13.09 (SD = 1.66). QTL linkage analysis of MENAR (web-Fig. 2, Table II) yielded one suggestive (10p13) score for the BN cohort, nothing noteworthy for the AN cohort, and three suggestive (10p14, 4a25.14a21.1) scores for the combined cohort.

**QTL linkage analysis for ANX.** ANX is a composite variable of trait anxiety and harm avoidance. Mean trait anxiety for control, AN and BN cohorts was  $40.4~(\mathrm{SD}=10.2)$ ,  $51.8~(\mathrm{SD}=13.8)$ , and  $48.4~(\mathrm{SD}=13.4)$ . Mean harm avoidance showed a similar pattern,  $2.5~(\mathrm{SD}=1.9)$ ,  $7.0~(\mathrm{SD}=3.0)$ , and  $5.8~(\mathrm{SD}=2.9)$ . QTL linkage analysis of trait anxiety (web-Fig. 3, Table II) yielded three suggestive (1q31.1, 4q35.2, 8q13.1) and one noteworthy (15q24.1) scores for the BN cohort, one suggestive (9p21.3) and one noteworthy (9q21.33) score for the AN cohort, and three suggestive (1q25.1, 9p21.3, 8q13.1) scores for the combined cohort.

# Covariate-Linkage Analysis

For covariate-linkage analysis, we use the "pre-clustering" method from Devlin et al. [2002b], which assumes covariate values can be used to probabilistically cluster families into "linked" and unlinked groups. Families are assumed to belong to the linked group if they cluster by a trait typically extreme in people with eating disorders. Pre-clustering assigns to each family an a priori probability they belong in the linked group.

Covariate linkage analysis for BMI. Mean BMI for control, AN and BN cohorts was  $20.14~(\mathrm{SD}=3.3),\ 15.38~(\mathrm{SD}=2.68),\ \mathrm{and}\ 17.23~(\mathrm{SD}=2.75).$  Covariate linkage analysis of BMI (web-Fig. 4, Table II) yielded one significant (14q21.1), three suggestive (3p23, 10p13, 5p15.3), and two noteworthy (3p26, 3q12.3) scores for the BN cohort, one suggestive (4q13.1) and one noteworthy (9q21.33) score for the AN cohort, and one suggestive (5p15.33) and one noteworthy (10p14) score for the combined cohort.

Covariate linkage analysis for CM. Mean CM for control, AN and BN cohorts was  $26.9~(\mathrm{SD}=8.6),~31.37~(\mathrm{SD}=9.10),$  and  $30.01~(\mathrm{SD}=9.65).$  Covariate linkage analysis of CM (web-Fig. 5, Table II) yielded two significant (14q21.1, 16p13.3) and three suggestive (4p15.33, 8q11.23, 10p11.21) scores for the BN cohort, two suggestive (11p11.2, 17q25.1) and one noteworthy (10q22.3) scores for the AN cohort, and three suggestive (10q21.3, 16p13.3, 17q24.2) and one noteworthy (6q16.3) scores for the combined cohort.

Covariate linkage analysis for food obsessions OBS. Mean OBF for control, AN and BN cohorts was 1.58 (SD=3.2), 23.97 (SD=5.89), and 22.63 (SD=6.67). Covariate linkage analysis of OBF (web-Fig. 6, Table II) yielded one significant (14q21.1) and five suggestive (4p16.1, 10p13.1, 8q11.23, 16p13.3, 18p11.31) scores from the BN cohort, two suggestive (17q25.1, 15q26.2) and two noteworthy (5p14.1, 1q32.1) scores from the AN cohort, and five suggestive (3q13.32, 4p16.1, 5p15.31, 10p14, 10q21.3) and three noteworthy (5q31.1, 6p24.1, 18p11.31) scores for the combined cohort.

# DISCUSSION

Both QTL and covariate-based linkage analyses produced substantial linkage signals from the BN cohort. QTL linkage analysis produced four suggestive signals, one when MENAR was the outcome and the remainder when ANX was the outcome. Covariate linkage analyses revealed a number of substantial linkage scores from the BN cohort: for minimum BMI, one significant and three suggestive; for CM, two

TABLE II. Ten Highest Linkage Scores, by Cohort and Trait

					QTL	$\Gamma$					
	Obsessionality (OBS)	ity (OBS)			Age at menarche (MENAR)	he (MENAR)			Anxiety (ANX)	(ANX)	
Chr	Position	ГОО	pval	Chr	Position	ГОД	pval	Chr	Position	LOD	pval
7 8	21(7p21.3)	1.558	0.004	10	36(10p13)	2.899	0.00013	1 4	198(1q31.1) 209(4q35.2)	1.865	0.002
16	101(16a23.1)	1.256	0.008	14	56(14a22.2)	1.112	0.012	+ oo	81(8a13.1)	1.729	0.002
11	100(11q22.3)	1.039	0.014	50	126(5q23.1)	1.108	0.012	$1\overline{5}$	68(15q24.1)	1.62	0.003
9	179(6q26)	0.944	0.05	12	56(12q12)	1.096	0.012	œ j	1(8p23.3)	1.121	0.012
ကေ	40(5p14.1)	0.922	0.02	4 5	105(4q23)	1.05	0.014	13	33(13q13.3)	0.973	0.02
16	150000303	0.88	0.0z 0.0z	2 81	17(10p15.3) $13(18p11.31)$	$0.816 \\ 0.715$	0.03	ი <del>-</del>	76(1p33)	0.932	0.02
12	150(12q24.32) 208(1q31.3)	0.849 $0.773$	0.02	3 10	$191(3q26.31) \\ 104(10q22.3)$	$0.714 \\ 0.593$	0.03	8 2	262(2q37.3) 137(6q23.3)	0.884 0.873	0.02
					Covariate	iate					
	Body mass index (BMI)	idex (BMI)			Concern over mistakes (CM)	nistakes (CM)			Food-related obsessions (OBF)	essions (OBF)	
Chr	Position	Eq LOD	pval	Chr	Position	Eq LOD	pval	Chr	Position	Eq LOD	pval
14	54(14q22.2)	2.983	0.00021	16	11(16p13.3)	2.965	0.0002	14	54(14q22.2)	2.965	0.00022
٠ د	63(3p23)	1.953	0.00271	14	56(14q22.2)	2.965	0.00022	4.5	13(4p16.1)	2.721	0.00033
0 73	34(10p13) 12(5n15 32)	1.869	0.00335 0.0063	4 oc	25(4p15.33) 66(8a11-23)	1.737	0.00468	2 ∝	56(10p13) 67(8a11 23)	2.168	0.00158
က	23(3p26.1)	1.578	0.00703	10	64(10p11.21)	1.678	0.00544	16	11(16p13.3)	1.745	0.00423
က	119(3q12.3)	1.57	0.00717	10	38(10p13)	1.445	0.00988	18	17(18p11.31)	1.667	0.00559
∞ ≺	77(8q12.1)	1.177	0.01988	17 6	117(17q25.3)	1.369	0.01203	ကဏ	152(5q31.3)	1.457	0.00959
# <del>4</del>	19(4p10.1)	1.114	0.01388	<b>O</b> 67	37(3n25.2)	1.203	0.01621	ט גנ	22(5n15.20)	1.450	0.01011
т —	260(1943)	1	0.03186	υro	150(5q31.3)	1.078	0.02587	o —	102(1p31.1)	1.327	0.01342
AN-ARP											
					QTL	T					
	OBS	s			MENAR	AR			ANX	X	
Chr	Position	TOD	pval	Chr	Position	LOD	pval	Chr	Position	TOD	pval
9	129(6q23.1)	1.774	0.002	14	63(14q23.1)	1.437	0.005	60	52(9p21.3)	1.821	0.002
- 4	200(1451.1) 153(6a25 2)	1.049 1.052	0.004	1.4	30(21422.3) 111(4025)	1.326	0.009	n 4	00(3421.55) 127(6023.1)	0.922	0.003
12	25(12p13.2)	0.78	0.03	10	175(10q26.3)	1.037	0.014	0 01	4(2p25.3)	0.748	0.03
2	134(2q14.3)	0.73	0.03	20	45(20p12.1)	1.032	0.015	14	115(14q32.13)	0.715	0.03
2 5	250(2q37.3)	0.715	0.03	$\frac{12}{2}$	117(12q24.21)	0.891	0.02	10	109(10q22.3)	0.592	0.05
8 2	124(18q23) 65(12a13.13)	0.625	0.04	9 9	$53(6p21.2)$ 87(16 $\alpha$ 22.2)	$0.81 \\ 0.737$	0.03	4 -	63(4p13) $148(1n12)$	0.487 $0.418$	0.07
13	15(13q12.12)	0.525	0.06	9	83(6q12)	0.721	0.03	9	49(6p21.2)	0.37	0.1
11	136(11q24.3)	0.406	60.0	4	155(4q32.1)	0.654	0.04	19	$35(1\overline{9}q13.2)$	0.335	0.11

Covariate

	BMI				CM				OBF		
Chr	$\mathbf{Position}$	$E_{\rm q}$ LOD	pval	$\operatorname{Chr}$	Position	$E_{\rm q}$ LOD	pval	$\operatorname{Chr}$	Position	$E_{\rm q}  { m LOD}$	pval
4	75(4q13.1)	1.972	0.00258	111	48(11p13)	2.004	0.00238	17	94(17q25.1)	1.746	0.00457
6	90(9q21.33)	1.583	0.00693	17	94(17q25.1)	1.788	0.00411	15	101(15q26.2)	1.7	0.00514
17	38(17p12)	1.38	0.01169	10	113(10q23.31)	1.603	0.00659	ರ	51(5p13.3)	1.563	0.0073
15	101(15q26.2)	1.349	0.01268	9	173(6q26)	1.269	0.01561		214(1q31.3)	1.55	0.00754
7	184(2q32.1)	1.349	0.01269	9	129(6q23.1)	1.165	0.02053	15	21(15q13.3)	1.389	0.01143
19	59(19q13.11)	1.337	0.01309	$\frac{13}{1}$	99(13q33.3)	1.034	0.02907	12	19(12p13.31)	1.331	0.01331
6	46(9q21.3)	1.216	0.01795	7	136(7q32.3)	0.921	0.03942	10	119(10q23.33)	1.312	0.01395
4	23(4p15.33)	1.214	0.01807	15	117(15q26.3)	0.83	0.05059	œ	142(8q24.21)	1.261	0.01595
18	18(18p11.31)	1.212	0.01816	12	35(12p12.3)	0.772	0.05928	2	190(2q32.1)	1.053	0.02764
11	88(11q14.1)	1.122	0.02304	12	59(12q12)	0.766	0.06031	11	48(11p13)	1.033	0.02921
AN- and	AN- and BN-ARP										
					QTL	ت					
	OBS	70			MENAR	AR			ANX		
Chr	Position	LOD	pval	Chr	Position	TOD	pval	Chr	Position	TOD	pval
-	900(1~91.9)	1 004	0.00981	10	95/10-19)	9 00 4	260000	-	106/1~95 1)	1 006	0.00049
1 1	200(1451.5) 91(7n91.3)	1.304	0.00231	10	99(10p19) 105(4093)	2.304 9.008	0.00020	- О	100(1425.1) 59(0p91-3)	1.990	0.00245
- 1	00(1693 1)	1 364	0.0041	۲ -	55(170999)	2000	0.00205	o 0	80(8g13 1)	1.021	0.00310
2 2	78(18c21.31)	1.503	0.0122	ן על	125(5a23.1)	1 234	0.00348	. <del>.</del> .	65(15024 1)	1 494	0.0040
1	$136(11\alpha24.3)$	1.151	0.02132	12	55(12a12)	1.133	0.02236	9 9	131(6023.1)	1.367	0.01211
; <sub></sub>	98(1p31.1)	1.092	0.02493	4	13(4p16.1)	1.112	0.02364	000	0(8p23.3)	1.121	0.02308
9	177(6q26)	0.985	0.03319	10	175(10q26.3)	1.037	0.02887	9	49(6p21.2)	1.043	0.02841
12	55(12q12)	0.813	0.053	20	45(20p11.22)	0.99	0.03274	13	33(13q13.3)	1.039	0.02871
10	159(10q26.3)	0.798	0.05524	12	109(12q23.2)	96.0	0.03549	14	117(14q32.13)	0.941	0.03737
$\infty$	32(8p22)	0.783	0.05758	9	81(6q12)	0.92	0.03956	-	76(1p33)	0.927	0.03881
					Covariate	iate					
	BMI				CM				OBF		
Chr	Position	Eq LOD	pval	$\operatorname{Chr}$	Position	Eq LOD	pval	Chr	Position	$\operatorname{Eq}\operatorname{LOD}$	pval
5	2(5p15.33)	1.709	0.00502	10	96(10q22.1)	2.258	0.00126	4	2(4p16.1)	2.631	0.0005
10	38(10p13)	1.614	0.00641	16	2(16p13.3)	2.02	0.00229	10	96(10q22.1)	2.138	0.0017
14	$38(14q\bar{1}3.1)$	1.434	0.01017	17	80(17q23.2)	1.858	0.00344	က	$138(3q\overline{1}3.33)$	1.843	0.00358
10	96(10q22.1)	1.362	0.01227	9	142(6q23.3)	1.525	0.00804	10	36(10p13)	1.756	0.00446
-	252(1q42.3)	1.256	0.01616	4	8(4p16.1)	1.491	0.01109	ចេ	26(5p15.2)	1.666	0.00561
00 (	92(8q21.11)	1.22	0.01779	$\frac{12}{1}$	46(12p11.23)	1.401	0.01346	က	136(5q31.1)	1.621	0.00629
က I	134(3q13.32)	1.214	0.01807	∞ <u>(</u>	66(8q11.23)	1.326	0.0193	9 ;	16(6p25.1)	1.619	0.00632
٠ ٢	10(5p15.33)	1.176	0.01993	13	94(13q33.3)	1.189	0.02907	20 1	14(18p11.31)	1.602	0.00661
17	48(12p11.23)	1.164	0.02059		140(5q31.1)	1.034	0.02926	17	86(17q24.2)	1.352	0.01258
D)	74(9q21.2)	1.111	0.02369	14	44(14q21.1)	1.032	0.0298	cI	8(15q12)	1.263	0.01589
Positions	Positions are given in centiMorgans according to the Marshfield m	ane according to	the Marchfold		m Cmmm/:uttq) b tos bo	areh fioldelinie or	A/BESEABCH/C	FNETIOS/204	on commoning cat 0 (http://umm. monch fieldelinie and REGE ARCH/CENETICS/cate/Sat OS mean Erromes htm)		The arbitraries of less is a superior

Positions are given in centiMorgans according to the Marshfield map, screening set 9 (http://www2.marshfieldclinic.org/RESEARCH/GENETICS/sets/Set9ScreenFrames.htm). The physical locations, in parentheses, are given according to the UCSD Human genome browser build 35 (http://genome.ucsc.edu/).

significant and three suggestive; and for OBF, one significant and five suggestive. Analysis of the AN cohort was less fruitful. QTL linkage analysis provided two suggestive signals (for OBS and ANX) and five suggestive signals for covariate-based linkage analysis (1 for BMI, 2 for CM, and 2 for OBF). The AN cohort is roughly one-half the size of the BN cohort, and this difference might account for the limited linkage signals.

Under the null hypothesis of no linkage and a single scan, we expect only one linkage score to exceed the suggestive threshold. The distribution of the number of regions in which the scores exceed the threshold can be modeled as a Poisson with mean one, and it is simple under this assumed distribution to calculate the probability of regional scores exceeding the threshold at least X times for a single scan. The probability of exceeding the threshold at least X=3 times is  $\sim\!0.06$ , so the results for BMI, CM, and OBF are striking for the BN cohort (Tables I and II), even after accounting for multiple tests. Likewise the number of suggestive signals for both the AN and BN cohorts for CM and OBF are highly significant (P<0.001 for either).

Disappointingly, however, all significant signals from the BN cohort were dampened slightly or substantially when the BN and AN cohorts were combined. Diminished linkage signals could result from either statistical or biological causes, or both. In terms of statistical causes for the disparity, the simplest explanation is that the substantial linkage signals in the BN sample are false positives. A more subtle explanation appeals to the observation that large signals from linkage scans are expected to be highly biased, greatly exaggerating the true effect of the locus (or loci) generating the signal [Goring et al., 2001]. This phenomenon might account for some of what we observe. Nonetheless, based on statistical theory, for any liability locus in common in the two cohorts, one expects linkage signals in both datasets to produce positive evidence for linkage, even if the magnitudes of the signals are quite different. This expectation is not realized; instead the complementary sample tends to produce weak evidence against linkage when the other is strongly positive.

We suspect a better explanation for the disparity lies in biology: AN and BN cohorts differ at a fundamental biological and genetic level, and that these differences are not easily resolved by the quantitative traits/covariates we used in the analyses. Biologically, the trait might not have the same salience for AN and BN cohorts. BMI is an example. Individuals with consistent AN, almost by definition, display lower lifetime BMI than individuals with consistent BN. The ability to attain and maintain extremely low body weights is apparently a fundamental biologically- and genetically-based difference between individuals with AN and BN. Our genetic explanation hypothesizes that a locus that generates substantial liability to BN will often—but not always—fail to generate liability to AN. For example, by using QTL linkage analysis, significant linkage is obtained on 10p for MENAR in the BN cohort, but LOD = 0.01 for the AN cohort. As described in Bulik et al. [2003b], the same region of 10p, which produced significant linkage based on diagnosis, overlaps substantially with regions showing linkage for obesity [Hager et al., 1998; Hinney et al., 2000]. Rates of obesity higher than that expected by chance are observed in families accessed through BN probands, but not in families accessed through AN probands. Thus, while speculative, the putative QTL on 10p could be genetically correlated with the disinhibition characteristic of BN and obesity, but have no impact on the inhibition characteristic of AN. It is interesting to note that early ageof-menarche has been associated on a population level with disinhibition [Johansson and Ritzen, 2005] and with the development of binge-eating in the absence of compensatory behaviors independent of BMI [Reichborn-Kjennerud et al., 2004]. It might also be important to note that OBF produces a suggestive signal in the region from the BN cohort, but OBS does not.

The other two traits producing significant linkage in the BN sample were CM and OBF, both detected by covariate-based linkage analysis. Again the putative loci underlying these traits could pleiotropically impact liability to BN. It is also possible that different values of CM and OBF are associated with liability to AN versus BN (e.g., low OBF for AN and high OBF for BN). If so, this might be detectable by QTL linkage analysis. To investigate the latter possibility, we tested chromosomes 14 and 16 for QTL linkage to CM and OBF. We find supporting evidence for QTL at 14q22.2 for the BN cohort (CM, LOD = 2.70; OBF, LOD = 1.20), but not the AN cohort (CM, LOD = -0.46; OBF, LOD = -0.14). We also find some weak support for a QTL at 16p13.3 region for the BN cohort (CM, LOD = 1.39; OBF, LOD = 0.50), but not the AN cohort (CM, LOD = -0.43; OBF, LOD = -0.23). Our data, therefore, support the idea that loci near 14q21.1 and 16p13.3 affect liability to BN, but have little impact on liability to AN.

On the other hand, consistent with our genetic hypothesis that some loci do confer liability to both AN and BN, some linkage signals were amplified by combining the datasets (Table II). For OBS, the LOD score at 1q31.1 climbs from 1.55 for the AN cohort to 1.98 for the combined sample and, at 7p21.2, from 1.56 for the BN cohort to 1.79. Another notable increase occurs at 4q23 for MENAR, which achieves LOD=2.01 for the combined sample from scores for the individual BN and AN cohorts of roughly 1.0. Other regions/ traits showing similar changes include 5p15.33/BMI (LOD=1.71), 10q21.3/OBF (LOD=2.14) and 3q13.32/OBF (LOD=1.84). See Table II for more details.

Certain regions of the genome repeatedly show positive linkage signals for multiple traits with different samples (web-Table I). An obvious example is the 14q21 region, which showed significant linkages in the BN cohort for BMI, CM, and OBF (Tables I and II). Because of the pattern of linkage signals, we were curious if these covariates up-weighted/down-weighted the same families for linkage analysis (the probability of membership of the families in the linked group), even though the covariates were largely uncorrelated (see Figs. 3 and 4 in Bulik et al. [2005]). Thus, we evaluated the correlation of the weights, and found they were not highly correlated across these three traits. Maximum pairwise correlation was 0.21 and minimum was 0.10. This is the region that showed suggestive linkage in a previous analysis of diagnosis and IBD-sharing [Bulik et al., 2003b]. Clustering families by these covariates gives greater weight to families with higher IBD-sharing, without information about IBD-sharing per se.

A complicated region is 4p16.1-4p15.33. Within 4p16.1, suggestive and close to significant linkage occurs for OBF in the BN cohort; two other traits/analyses yield LOD scores greater than 1.0, for BMI (1.14; BN) and CM (1.49; BN/AN). At the adjacent chromosome band, 4p15.33, suggestive linkage occurs for CM in the BN/AN cohorts and LOD = 1.21 occurs for BMI in the AN cohort. Thus, ignoring the possibility of no liability locus in 4p16.1-4p15.33, the same locus could be generating all the signals. An alternative interpretation puts a locus in 4p16.1 affecting liability to BN and/or a locus in 4p15.33 affecting liability more generally. See web-Table I for more regions of overlap.

A few regions of overlapping linkage signals coincide with our previous research. In addition to 10p13 and 14q21.1, the 1q31 region again comes to the fore (Table II, web-Table I). In our first exploratory analysis using covariates, 1q31 showed significant linkage [Devlin et al., 2002a; Bacanu, 2005] when both OBS and drive-for-thinness were used as covariates. (Unfortunately, drive-for-thinness was not measured in the BN cohort, and no combination of traits measured in both samples accurately predicts drive-for-thinness in the AN

cohort.) Our current analyses, using OBS in a QTL linkage analysis, produces a strong suggestive signal for linkage at 1q31.1 in the BN/AN cohorts. Complementing this linkage is another suggestive linkage at 1q31.1 for ANX in the BN cohort, which shifts to 1q25.1 (a difference of only 12 cM) and increases to suggestive (LOD = 2.0) when this sample is combined with the AN cohort.

Intriguingly, 11q22 shows some overlap of linkage signals for the BN and AN cohorts for OBS and BMI (Table II, web-Table I). This region contains DRD2, polymorphisms which demonstrate linkage and association in our analyses [Bergen et al., in press]. The –141 C/– insertion/deletion (–141 Indel) polymorphism, which affects DRD2 transcription efficiency, shows significant association with diagnosis at the level of alleles, genotypes and haplotypes; the insertion C allele, which appears to increase expression of DRD2, is transmitted from parents to their affected offspring at rates significantly greater than that expected by chance; and haplotypes containing the insertion C allele and other SNP variants show even greater transmission distortion. Therefore, the linkage results (Table II, web-Table I) could be attributable to the impact of DRD2 polymorphisms.

After measuring 100 features relevant for eating disorders in multiplex families for eating disorders, we used a multilayer decision process to select six traits for linkage analysis and team the traits with an appropriate analytic method [Bulik et al., 2005]. Insofar as we are aware, this is the first study to explore the phenotypic space in this way, and it could prove a useful blueprint for other studies of its kind, such as ongoing studies of the genetic basis of Type II diabetes and hypertension. When the results of the phenotypic analyses were applied to the genetic data from two cohorts of multiplex families, a number of linkage signals worthy of follow-up study arise. It is tempting to conclude that our approach to these complex data has been successful because we have identified a greater number of significant and suggestive linkages than that expected by chance. Nonetheless we are cognizant that proof of success will only come when alleles generating liability to eating disorders—or affecting the traits underlying liability to eating disorders—are convincingly identified under our linkage peaks. We have pursued two approaches to achieve this goal: bolstering our linkage results by linkage studies on new samples of multiplex families; and by direct identification of genetic variation generating liability to disease in these linkage regions. Looking at the results from both cohorts, two promising features stand out: regions of the genome which repeatedly show positive linkage signals for different traits (14g21.1, 4p16.1-4p15.33, 1g31, 11g22, 10p13) and different samples; and suggestive linkage signals that emerge by combination of the two samples (1q31.1, 7p21.2, 4q25, 5p15.33, 10q21.3, 3q13.32).

# ACKNOWLEDGMENTS

The authors thank the Price Foundation for the support of the clinical collection of subjects and genotyping, and contribution to the support of data analysis. Data analysis was supported by grants from the National Institute of Health MH057881 and MH066117 (to BD), the latter part of a collaborative R01 grant to study the genetic basis of eating disorders. S-AB was also supported by a NARSAD Young Investigator Award. 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 contribution of time and effort in support of this study, and to

an anonymous reviewer who significantly improved the manuscript.

#### REFERENCES

- Abecasis GR, Cherny SS, Cookson WO, Cardon LR. 2002. Merlin—Rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97–101.
- Bacanu S-A. 2005. Robust estimation of critical values for genome scans to detect linkage. Genet Epidemiol 28:24-32.
- Bergen AW, Yeager M, Welch RA, Haque K, Ganjei JK, van den Bree MBM, Mazzanti C, Nardi I, Fichter MM, Halmi KA, Kaplan AS, Strober M, Treasure J, Woodside DB, Bulik CM, Bacanu S-A, Devlin B, Berrettini WH, Goldman D, Kaye WH. 2005. Association of multiple DRD2 polymorphisms with anorexia nervosa. Neuropsychopharmacology 30: 1703–1710.
- Boehnke M, Cox NJ. 1997. Accurate inference of relationships in sib-pair linkage studies. Am J Hum Genet 61:423–429.
- Bulik C, Sullivan P, Carter F, Joyce P. 1995. Temperament, character, and personality disorder in bulimia nervosa. J Nerv Ment Dis 183:593– 598
- Bulik CM, Sullivan PF, Kendler KS. 2003a. Genetic and environmental contributions to obesity and binge eating. Int J Eat Disord 33:293–298
- Bulik CM, Devlin B, Bacanu SA, Thornton L, Klump KL, Fichter MM, Halmi KA, Kaplan AS, Strober M, Woodside DB, Bergen AW, Ganjei JK, Crow S, Mitchell J, Rotondo A, Mauri M, Cassano G, Keel P, Berrettini WH, Kaye WH. 2003b. Significant linkage on chromosome 10p in families with bulimia nervosa. Am J Hum Genet 72:200–207.
- Bulik CM, Bacanu S-A, Klump KL, Fichter MM, Halmi KA, Keel P, Kaplan AS, Mitchell JE, Rotondo A, Strober M, Treasure J, Woodside DB, Sonpar VA, Xie W, Bergen A, Berrettini AH, Kaye WH, Devlin B. 2005. Selection of eating-disorder phenotypes for linkage analysis. Am J Med Genet B 139:81–87.
- Devlin B, Bacanu S-A, Klump KL, Bulik CM, Fichter MM, Halmi KA, Kaplan AS, Strober M, Treasure J, Woodside DB, Berrettini WH, Kaye WH. 2002a. Linkage analysis of anorexia nervosa incorporating behavioral covariates. Hum Mol Genet 11:689–696.
- Devlin B, Jones B, Bacanu S-A, Roeder K. 2002b. Mixture models for linkage analysis of affected sibling pairs and covariates. Genet Epidemiol 22: 52–65.
- Fassino S, Amianto F, Gramaglia C, Facchini F, Abbate Daga G. 2004. Temperament and character in eating disorders: Ten years of studies. Eat Weight Disord 9:81–90.
- Goring HH, Terwilliger JD, Blangero J. 2001. Large upward bias in estimation of locus-specific effects from genome wide scans. Am J Hum Genet 69:1357–1369.
- Greenwood CM, Bull SB. 1999. Down-weighting of multiple affected sib pairs leads to biased likelihood-ratio tests, under the assumption of no linkage. Am J Hum Genet 64:1248–1252.
- Grice DE, Halmi KA, Fichter MM, Strober M, Woodside DB, Treasure JT, Kaplan AS, Magistretti PJ, Goldman D, Bulik CM, Klump K, Fichter M, Halmi K, Kaplan A, Strober M, Treasure J, Woodside B, Kaye WH. 2002. Evidence for a susceptibility gene for anorexia nervosa on chromosome 1. Am J Hum Genet 70:787–792.
- Hager J, Dina C, Francke S, Dubois S, Houari M, Vatin V, Vaillant E, Lorentz N, Basdevant A, Clement K, Guy-Grand B, Froguel P. 1998. A genome-wide scan for human obesity genes reveals a major susceptibility locus on chromosome 10. Nat Genet 20:304–308.
- Halmi KA, Sunday SR, Klump KL, Strober M, Leckman JF, Fichter M, Kaplan A, Woodside B, Treasure J, Berrettini WH, Al Shabboat M, Bulik CM, Kaye WH. 2003. Obsessions and compulsions in anorexia nervosa subtypes. Int J Eat Disord 3:308–319.
- Hinney A, Ziegler A, Oeffner F, Wedewardt C, Vogel M, Wulftange H, Geller F, Stubing K, Siegfried W, Goldschmidt HP, Remschmidt H, Hebebrand J. 2000. Independent confirmation of a major locus for obesity on chromosome 10. J Clin Endocrinol Metab 85:2962–2965.
- Johansson T, Ritzen EM. 2005. Very long-term follow-up of girls with early and late menarche. Endocr Dev 8:126–136.
- Kaye WH, Lilenfeld LR, Berrettini WH, Strober M, Devlin B, Klump KL, Goldman D, Bulik CM, Halmi KA, Fichter MM, Kaplan A, Woodside DB, Treasure J, Plotnicov KH, Pollice C, Rao R, McConaha CW. 2000. A search for susceptibility loci for anorexia nervosa: Methods and sample description. Biol Psychiatry 47:794–803.

- Kaye WH, Devlin B, Barbarich N, Bulik CM, Thornton L, Bacanu S-A, Fichter MM, Halmi KA, Kaplan AS, Strober M, Woodside DB, Bergen AW, Crow S, Mitchell J, Rotondo A, Mauri M, Cassano G, Keel P, Plotnicov K, Pollice C, Klump KL, Lilenfeld LR, Ganjei JK, Quadflieg N, Berrettini WH. 2004. Genetic analysis of bulimia nervosa: Methods and sample description. Int J Eat Disord 35:556–570.
- Klump KL, Bulik CM, Pollice C, Halmi KA, Fichter MM, Berrettini WH, Devlin B, Strober M, Kaplan A, Woodside DB, Treasure J, Shabbout M, Lilenfeld LR, Plotnicov KH, Kaye WH. 2000. Temperament and character in women with anorexia nervosa. J Nerv Ment Dis 188:559– 567
- Kruglyak L, Daly MJ, Reeve-Daly MJ, Lander ES. 1996. Parametric and nonparametric linkage analysis: A unified multipoint approach. Am J Hum Genet 58:1347–1363.
- Lilenfeld L, Kaye W, Greeno C, Merikangas K, Plotnikov K, Pollice C, Rao R, Strober M, Bulik CM, Nagy L. 1998. A controlled family study of restricting anorexia and bulimia nervosa: Comorbidity in probands and disorders in first-degree relatives. Arch Gen Psychiatry 55:603–610
- O'Connell JR, Weeks DE. 1998. PedCheck: A program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63: 259–266.

- Reichborn-Kjennerud T, Bulik CM, Sullivan PF, Tambs K, Harris JR. 2004. Psychiatric and medical symptoms in binge eating in the absence of compensatory behaviors. Obes Res 12:1445–1454.
- Sham PC, Purcell S, Cherny SS, Abecasis GR. 2002. Powerful regression-based quantitative-trait linkage analysis of general pedigrees. Am J Hum Genet 71:238–253.
- Steiger H, Gauvin L, Israel M, Kin N, Young S, Roussin J. 2004. Serotonin function, personality-trait variations, and childhood abuse in women with bulimia-spectrum eating disorders. J Clin Psychiatry 65:830–837.
- Sullivan PF, Bulik CM, Kendler KS. 1998. The genetic epidemiology of binging and vomiting. Br J Psychiatry 173:75–79.
- Tozzi F, Klump KL, Thornton LM, Bulik CM, Devlin B, Fichter MM, Halmi KA, Kaplan AS, Strober M, Woodside DB, Crow S, Mitchell J, Rotondo A, Mauri M, Cassano G, Keel P, Plotnicov KH, Pollice C, Lilenfeld LR, Berrettini WH, Kaye WH. 2005. Symptom fluctuation in eating disorders: Correlates of diagnostic crossover. Am J Psychiatry 162:732–740.
- Walters EE, Kendler KS. 1995. Anorexia nervosa and anorexic-like syndromes in a population-based female twin sample. Am J Psychiatry 152:64–71.
- Wonderlich S, Lilenfeld L, Riso L, Engel S, Mitchell J. 2005. Personality and anorexia nervosa. Int J Eat Disord 37(Suppl 1):S68–S71.