The insulin resistance syndrome (IRS) (DeFronzo and Ferrannini, 1991), also known as Syndrome X (Reaven, 1988), is characterized by a clustering of type 2 diabetes and coronary heart disease (CHD) risk factors, including hypertension, dyslipidemia, glucose intolerance, and central obesity (DeFronzo and Ferrannini, 1991; Ferrannini et al., 1991; Modan et al., 1985; Reaven, 1988). Although genetic influences on individual features of the IRS are well established (Bouchard and Perusse, 1993; Liese et al., 1998), the metabolic and possible genetic basis for the clustering of risk factors remains to be elucidated.

A variety of statistical approaches provide evidence for underlying genetic influences on the interrelated features of the IRS, including body mass index (BMI) and waist-to-hip ratio (WHR) (Cardon et al., 1994) and the clustering of hypertension, diabetes and obesity (Carmelli et al., 1994). Using quantitative genetic analysis, several studies suggest common additive genetic influences (pleiotropy) on pairs of risk factors characterizing the IRS, including high density lipoprotein (HDL) cholesterol and triglyceride (TG), BMI and fat mass (Mahaney et al., 1995), and insulin levels and BMI, WHR, subscapular skinfolds and HDL cholesterol (Mitchell et al., 1996; Rainwater et al., 1997). A recent, trivariate, quantitative genetic analysis in hypertriglyceridemic families also reported strong additive genetic correlations between each pair of lipid and lipoprotein traits, including low-density lipoprotein (LDL) size, TG and HDL (Edwards et al., 1999). Despite this accumulating evidence there have been no reports of specific genes that may underlie these multivariate traits characterizing the syndrome.

Hence, the purpose of this study is to evaluate evidence for genetic linkage, using a sib-pair approach, between previously identified multivariate factors characteristic of the IRS (Edwards et al., 1994), and 9 candidate genes involved in lipid and lipoprotein metabolism, using data from a sample of dizygotic women twins.

Method

Study Subjects

Study subjects were participants in the second examination of the Kaiser Permanente Women Twins Study in Oakland, California. Examination 2 was conducted between 1989 and 1990, and included 704 individuals (206 monozygotic (MZ) and 146 dizygotic (DZ) pairs), representing 81% of the original cohort examined in 1979–1980 (Austin et al., 1987; Selby et al., 1993). Zygosity was determined at
examination 1 based on 20 polymorphic loci, such that the probability of misclassification of a pair who were concordant on all markers as monozygotic was less than 0.001 (Austin et al., 1987). At the time of the second exam, each woman completed a health history questionnaire and a physical examination, including anthropometric and laboratory measurements. The average age of individual women at examination 2 was 51 years, and the majority (90%) of women in the sample were white.

Data Collection
The data for this analysis has previously been described in detail (Edwards et al., 1994). Briefly, weight was measured without shoes and in lightweight clothing. Waist circumference was measured using a standardized protocol. Systolic and diastolic blood pressure was measured after the subject had been seated for five minutes using a mercury sphygmomanometer. Plasma glucose and insulin were both measured following an overnight fast, and again two hours after a 75 g oral glucose load (Glutol, Paddock Laboratories, Minneapolis, MN) as previously described (Edwards et al., 1994). After the overnight fast, 30 ml of whole blood was collected into EDTA-containing tubes for lipid determinations. Plasma was separated by centrifugation within two hours, and stored under refrigeration. Nondenaturing gradient gel electrophoresis was performed on the plasma using 2–16% polyacrylamide gradient gels (Pharmacia) (Nichols et al., 1986). The estimated diameter of the major LDL subclass was calculated based on a calibration curve constructed from high molecular weight standards run on the same gel (Krauss and Burke, 1982). The diameter, denoted LDL peak particle diameter (LDL-PPD), is a continuous variable and is used in the factor analysis as a measure of LDL heterogeneity. Total HDL cholesterol (HDL-C) (Warnick et al., 1984) and triglycerides (Nagele et al., 1984) were determined by standardized methods at the Donner Laboratory (Berkeley, CA).

Candidate Genes
Markers were genotyped for each of the following 9 candidate genes as previously described (Austin et al., 1998; Friedlander et al., 2000; Talmud et al., 2000): apolipoprotein E (apoE), apolipoprotein B (apoB), apolipoprotein CIII (apo CIII), the low density lipoprotein receptor (LDL-R), hormone sensitive lipase (HSL), microsomal triglyceride transfer protein (MTP), hepatic lipase (HL), the insulin receptor (IR), and cholesterol ester transfer protein (CETP). The most highly polymorphic markers available were selected, with the number of alleles ranging from 3 (apoE) to 28 (CETP) (Table 1). Briefly, genotyping of the apoE gene, to identify the 3 alleles ε2, ε3, and ε4, used the method of Bolla et al., (1995). The apoB VNTR in the 3' untranslated region was genotyped according to the method of Boerwinkle et al., (1989). The CETP tetranucleotide (GAAA)n repeat was recently identified and described by Talmud et al., (2000). All other genotypes were carried out using ABI fluorescent labeled forward primers (synthesized by Oswell Ltd, Southampton UK; non fluorescent reverse primers were synthesized by Gibco-BRL, Paisley, UK).

Table 1
Candidate Gene Polymorphisms Typed for Women Twins

<table>
<thead>
<tr>
<th>Candidate Gene</th>
<th>Chromosome</th>
<th>Marker</th>
<th>No. of Alleles</th>
<th>Heterozygosity Index*</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApoE</td>
<td>19</td>
<td>Triallelic: Isoforms: ε2, ε3, and ε4</td>
<td>3</td>
<td>0.43</td>
</tr>
<tr>
<td>ApoB</td>
<td>2</td>
<td>VNTR at 3' end</td>
<td>14</td>
<td>0.74</td>
</tr>
<tr>
<td>Apo CIII</td>
<td>11</td>
<td>Tetranucleotide repeat in intron 20</td>
<td>6</td>
<td>0.60</td>
</tr>
<tr>
<td>LDL-R</td>
<td>19</td>
<td>D19S394, 250 kb 5' to LDL-R gene</td>
<td>21</td>
<td>0.91</td>
</tr>
<tr>
<td>HSL</td>
<td>19</td>
<td>GT repeat in intron 7</td>
<td>14</td>
<td>0.67</td>
</tr>
<tr>
<td>MTP</td>
<td>4</td>
<td>CA repeat in intron 10</td>
<td>13</td>
<td>0.74</td>
</tr>
<tr>
<td>HL</td>
<td>15</td>
<td>CA repeat in intron 8</td>
<td>7</td>
<td>0.63</td>
</tr>
<tr>
<td>IR</td>
<td>19</td>
<td>CT repeat in intron 2</td>
<td>12</td>
<td>0.51</td>
</tr>
<tr>
<td>CETP</td>
<td>16</td>
<td>GAAA repeat in promoter region</td>
<td>28</td>
<td>0.73</td>
</tr>
</tbody>
</table>

* Based on all individual women typed in the study. (Adapted from Austin et al., 1998)
each factor of interest, that were used as continuous variables in the sib-pair linkage analysis.

Sib Pair Analysis. Quantitative sib-pair linkage analysis was used to test for genetic linkage between the factors characterizing the IRS and the candidate genes. The sib-pair method has the advantage of being a nonparametric method of linkage analysis that requires no a priori knowledge of the mode of inheritance of the phenotype under investigation. The rationale underlying this method is that, in the presence of linkage between a marker locus and the quantitative trait, sib pairs having more similar phenotypes will, on average, share a greater proportion of genes (alleles) identical by descent (IBD) at the marker locus. The sib-pair method used here is implemented in the SIBPAL program of the S.A.G.E. package (Release 2.2, 1994). In this regression approach, quantitative variation in the trait (factor scores) is examined in sibships as a function of shared alleles for the polymorphic marker of each candidate gene. A regression line is then calculated with the squared difference in the factor score between sibs (squaring is used to obtain only positive values) as the dependent variable and the estimated number of shared alleles IBD as the independent variable (Haseman & Elston, 1972). If the slope of this line is negative, and significantly different from zero with the appropriate degrees of freedom (Blackwelder & Elston, 1982), the result is interpreted as evidence for linkage between the phenotype and the candidate gene. In addition, the slope of the regression line reflects the magnitude of evidence for linkage.

In this sib-pair analysis, because parental genotypes were not available observed allele sharing represents identity by state (IBS). However, by using allele frequency estimates from each marker IBD can be estimated from IBS. This estimation procedure is implemented in the SIBPAL subroutine of S.A.G.E., and is based on an algorithm described by Amos and Elston (1989). Using this procedure the resulting IBD values differ from 0, 1, and 2.

Finally, because population-based allele frequencies were not available for several of the markers used in this study, all allele frequency estimates for IBD estimation were based on the sample of DZ women twins.

Due to the requirement of intact twin pairs for genetic analysis, pairs were excluded if one or both co-twins were missing data for any variable included in the factor analysis, were diabetic, or if either twin had triglyceride values greater than 400 mg/dl. After exclusions there were a total of 126 pairs of DZ women twins available for analysis. However, the sample size varies slightly for each candidate gene due to missing genotype data. No adjustment for age or gender is needed since all pairs are perfectly matched on these characteristics. Because the analysis includes multiple comparisons, a p-value of < 0.01 was considered statistically significant evidence for linkage.

Results
The three factors of interest are shown in Table 2 and have been previously described (Edwards et al., 1994). Briefly, factor 1 is characterized by large positive factor loadings for body weight, waist circumference and fasting insulin, and is interpreted as a weight/fat factor. Factor 2 is characterized by large positive factor loadings for fasting insulin, fasting glucose, postload insulin and postload glucose, and is interpreted as an insulin/glucose factor. Factor 3 is characterized primarily by TG, HDL and LDL-PPD and is interpreted as a lipid factor. For the present analysis, the resulting factors are interpreted as composite variables representing multivariate phenotypes of the IRS.

Table 3 presents the sib-pair results for each of the three factors and each of the 9 candidate genes. The strongest evidence for linkage was observed between the lipid factor and the CETP gene (Figure 1). The significantly negative slope indicates that those pairs who are more alike at the CETP marker locus (slope = –1.94, p = 0.002) have more similar scores for the lipid factor. The apoE gene showed modest

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results of Factor Analysis: Factors and Factor Loadings*</td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Weight</td>
</tr>
<tr>
<td>Waist</td>
</tr>
<tr>
<td>F. Ins</td>
</tr>
<tr>
<td>F. Glu</td>
</tr>
<tr>
<td>PL. Ins</td>
</tr>
<tr>
<td>PL. Glu</td>
</tr>
<tr>
<td>TG</td>
</tr>
<tr>
<td>HDL</td>
</tr>
<tr>
<td>LDL-PPD</td>
</tr>
<tr>
<td>DBP</td>
</tr>
<tr>
<td>SBP</td>
</tr>
</tbody>
</table>

*Factor loadings represent the correlation between the individual variable and each factor. Components are rotated using Varimax rotation to obtain factors; F., fasting; PL., postload; TG, triglyceride; HDL, high density lipoprotein cholesterol; and LDL-PPD, low-density lipoprotein peak particle diameter; DBP, diastolic blood pressure; SBP, systolic blood pressure. (Adapted from Edwards et al., 1994)
evidence for linkage to the weight/fat factor. The negative slope (slope = –4.3, \( p = 0.01 \)) indicates that those pairs who are more alike at the apoE locus have more similar scores for the weight/fat factor, consistent with linkage.

**Discussion**

These results represent the first report of linkage between multivariate phenotypes of the IRS, defined by factor analysis, and candidate genes for lipid and lipoprotein metabolism. Specifically, these results are consistent with linkage between the lipid factor and the CETP gene, and possible genetic linkage between the body weight/fat factor and the apoE gene. Importantly, these results extend previous work from quantitative genetic analyses (Mahaney et al., 1995; Mitchell et al., 1996; Edwards et al., 1999) and suggest that single major genes, such as the CETP gene, contribute to the well-established covariation between LDL size, TG and HDL.

The primary finding in this report is the apparent linkage between the lipid factor and the highly polymorphic marker within the CETP gene promoter. This result is consistent with several previous reports based on the individual lipid and lipoproteins characterizing the lipid factor. Rotter et al. reported evidence for linkage with LDL size and an anonymous marker (D16S313) within 6 cM of the CETP gene on chromosome 16 (Rotter et al., 1996). Allayee et al. (1998) recently reported linkage to microsatellite markers flanking the CETP and lecithin:cholesterol acyltransferase (LCAT) genes to small dense LDL in familial combined hyperlipidemic families. Studies have also demonstrated an association between variation in the CETP gene and HDL levels (Gudnason et al., 1999; Kuivenhoven et al., 1998).

Importantly, these results extend the previous report by Talmud et al. showing evidence for linkage between the CETP gene and several individual phenotypes, including LDL size, TG and Apo B levels in the same sample of women twins. Taken together, these results suggest that variation in the CETP gene influences not only LDL size, TG, and possibly HDL and ApoB levels, but also a composite phenotype characterized by these lipids and lipoproteins. Together, these results suggest that the CETP gene could have pleiotropic effects on these risk factors.

Adding support to this hypothesis is the work by Comuzzie et al. (1997), who indicates that a synthetic phenotype extracted using principal component analysis is of greater utility in linkage analysis when the majority of pleiotropy is attributable to major loci, rather than polygenes. Thus, the CETP gene may account for a portion of the well-established statistical and metabolic association observed between these lipid and lipoprotein risk factors.

A biologic basis for the linkage of the CETP gene with the lipid factor is supported by numerous studies. CETP activity depends on the size of the pool of TG-rich lipopro-
teins, and may explain the association between hypertriglyceridemia and reduced HDL levels (Tall, 1986). Further, Ambrosch et al. (1998) showed that increasing TG and CETP activity levels are associated with decreasing LDL particle diameters and a preponderance of a monodispersed LDL pattern, particularly in insulin-resistant subjects. In diet studies Lagrost et al. (1999) demonstrated that CETP activity influences the size distribution of LDL and HDL particles in normolipidemic individuals, and is correlated negatively with HDL cholesterol, but positively with triglyceride concentrations. Guerin et al. (1994) showed that the capacity of LDL particles to accept the cholesterol ester from HDL was highly correlated with the LDL-TG content. Finally, an in vitro study suggested that the plasma TG-rich lipoprotein level and the extent of intraplasma LCAT, CETP, lipoprotein lipase and likely HL reactions in vivo may play a role in determining the LDL phenotype (Chung et al., 1998). Thus, involvement of the CETP gene in the covariation of LDL-PPD, TG and HDL is plausible.

In this study borderline evidence for linkage was also observed between the apoE gene and the body weight/fat factor. This result is consistent with previous association studies. For example, in a study based on the Bogalusa Heart Study apoE phenotypes were significantly associated with BMI, percent body fat and plasma insulin levels, with the apo ε group having significantly lower values of BMI, body fat, and plasma insulin levels than the ε3 and ε4 groups (Srinivasan et al., 1996). Another study suggested that the ε4 phenotype significantly modifies the central obesity-induced changes in metabolic and hemodynamic variables characteristic of insulin resistance (Uusitupa et al., 1996). In the current study individuals with an apo e2 allele had lower weight/fat factor scores than individuals with either ε3 or ε4, however, the result was not statistically significant. If confirmed, these results in healthy women could indicate that genes involved in lipid metabolism are also involved in determining differences in body fat/obesity and insulin resistance, suggesting multilocus determination of factors contributing to the IRS.

Suggestive evidence for linkage between the lipid factor and both the apoB and HSL genes was found. Although these findings are consistent with previous reports of genetic linkage with the individual lipids and lipoproteins characterizing this multivariate factor (Austin et al., 1998; Houlston et al., 1991; Rotter et al., 1996), as well as changes in levels (Friedlander et al., 2000), the evidence for linkage is weak and should be interpreted with caution. However, it is possible that the power to detect linkage to genes with more modest effects was increased when using the multivariate phenotypes. Several other groups have shown that multivariate approaches, including using a linear combination of multiple phenotypic values, can increase power to detect linkage (Allison et al., 1998; Blangero et al., 1993; Boomsma, 1996; Jiang and Zeng, 1995; Lange and Boehnke, 1983).

Although the results from this study support the presence of genetic linkage between multivariate factors and candidate genes, several important limitations should be noted. First, an alpha value of 0.01 was used to indicate statistical significance. However, multiple comparisons and the finite nature of the genome could lead to false-positive results in this setting. Secondly, it is important to note that this analysis was based entirely on women, whereas previous reports included both men and women. In addition, the limited number of dizygotic women twin pairs available for analysis may have reduced the ability to detect true genetic linkage. Third, because parental information was not available, IBD was estimated from both IBS and the allele frequency for each marker, using a subroutine in the SIBPAL module of S.A.G.E. Further, because population allele frequencies were not available for many of the markers included in this study, estimates for all candidate genes were obtained from the sample of women twins. ApoE allele frequencies in this sample of women were similar to previous reports (Cauley et al., 1993; Jarvik et al., 1997; Sing and Davigon, 1985), thus allele misspecification is not likely to account for the modest linkage between the apoE gene and the weight/fat factor. Finally, a disadvantage of sib-pair linkage analysis is that environmental and behavioral covariates cannot be simultaneously incorporated into the analysis. Combined multivariate segregation and linkage analysis using full pedigree data should be used to confirm these results and to further characterize the genetic basis of these multivariate traits.

In conclusion, evidence for genetic linkage of multivariate factors of the IRS to candidate genes has been presented, contributing to the growing body of evidence suggesting multilocus determination of the IRS. This study adds additional information by providing evidence for underlying susceptibility genes, particularly the CETP gene, influencing the covariation of these interrelated risk factors. Because these multivariate factors are characterized by risk factors for type 2 diabetes and CHD, confirming these results may have important implications for understanding the genetic basis of both conditions.

Acknowledgments
The authors would like to thank Andy Louie for his important contributions to the analysis, Le Ahn Luong, Lema Haddad and Clair Turner for genotype determinations. This research was supported by National Institutes of Health grants HL-41830 and HL-46880, and was performed during Dr. Austin’s tenure as an Established Investigator of the American Heart Association. Sib-pair linkage results were obtained by use of the S.A.G.E. program package, which is supported by a U.S. Public Health Service Resource Grant (1 P41 RR03655) from the National Center for Research Resources. PT is supported by the British Heart Foundation, and KE receives support from an American Diabetes Association Research Award.

References

https://doi.org/10.1375/twin.4.1.41 Published online by Cambridge University Press


S.A.G.E. Statistical Analysis for Genetic Epidemiology, release 2.2 (1994). Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland.


