Genome-Wide Linkage Analyses of Total Serum IgE Using Variance Components Analysis in Asthmatic Families

R.A. Mathias,1 L.R. Freidhoff,2 M.N. Blumenthal,3 D.A. Meyers,4 L. Lester,5 R. King,3 J.F. Xu,4 J. Solway,6 K.C. Barnes,2 J. Pierce,7 O.C. Stine,4 A. Togias,2 W. Oetting,3 P.L. Marshik,6 J.B. Hetmanski,1 S.K. Huang,2 E. Ehrlich,2 G.M. Dunston,9 F. Malveaux,9 S. Banks-Schlegel,10 N.J. Cox,11 E. Bleecker,4 C. Ober,11 T.H. Beaty,1* S.S. Rich,7 and the CSGA

1Department of Epidemiology, Johns Hopkins University, Baltimore, Maryland
2Clinical Immunology, School of Medicine, Johns Hopkins University, Baltimore, Maryland
3Department of Medicine, University of Minnesota, Minneapolis
4Center for the Genetics of Asthma and Complex Diseases, University of Maryland, Baltimore, Maryland
5Department of Pediatrics, University of Chicago, Chicago, Illinois
6Department of Medicine, University of Chicago, Chicago, Illinois
7Department of Public Health Science, Bowman Gray School of Medicine, Winston-Salem, North Carolina
8College of Pharmacy, University of New Mexico, Albuquerque
9Department of Microbiology, Howard University, Washington, DC
10Division of Lung Diseases, NHLBI, Bethesda, Maryland
11Department of Human Genetics, University of Chicago, Chicago, Illinois

Variance components models were used to analyze total IgE levels in families ascertained though the Collaborative Study of the Genetics of Asthma (CSGA) using a genome-wide array of polymorphic markers. While IgE levels are known to be associated with clinical asthma and recognized to be under strong genetic control (here the heritability was estimated at 44–60% in the three racial groups),


*Correspondence to: Terri H. Beaty, Department of Epidemiology, School of Hygiene and Public Health, Johns Hopkins University, 615 N. Wolfe St., Baltimore, MD 21205. E-mail: tbeaty@jhsph.edu

Received for publication 9 June 2000; revision accepted 2 September 2000

© 2001 Wiley-Liss, Inc.
Variance Components for IgE Levels

Specific genes influencing this trait are still largely unknown. Multipoint analysis of 323 markers yielded little indication of specific regions containing a trait locus controlling total serum IgE levels (adjusted for age and gender). Although a number of regions showed LOD statistics above 1.5 in Caucasian families (chromosome 4) and in African-American families (chromosomes 2 and 4), none yielded consistent evidence in all three racial groups. Analysis of total IgE adjusted for gender, age and Allergy Index (a quantitative score of skin test sensitivity to 14 common aeroallergens) was conducted on these data. In this analysis, a much stronger signal for a trait locus controlling adjusted log[total IgE] was seen on the telomeric end of chromosome 18, but only in Caucasian families. This region accounted for most of the genetic variation in log[total IgE], and may represent a quantitative trait locus for IgE levels independent of atopic response. Oligogenic analysis accounting simultaneously for the contribution of this locus on chromosome 18 and other chromosomal regions showing some evidence of linkage in these Caucasian families (on chromosomes 2, 4 and 20) failed to yield significant evidence for interaction. Genet. Epidemiol. 20:340–355, 2001. © 2001 Wiley-Liss, Inc.

**Key words:** total IgE levels; asthma; variance components; linkage; Allergy Index

**INTRODUCTION**

Twin and family studies have shown that genetic factors are important in the determination of serum IgE levels [Bazaral et al., 1974]. However, studies of specific genetic models have failed to consistently identify a single model of inheritance for IgE levels, and a number of family studies have variously suggested that recessive [Marsh et al., 1974; Gerrard et al., 1978; Meyers et al., 1991], dominant or codominant [Meyers et al., 1982; Sampogna et al., 2000], or polygenic [Hasstedt et al., 1983] models may underlie the distribution of total serum IgE. In these reports the major locus component accounts for between 37 and 84% of the variance in log[total IgE] [Bazaral et al., 1974; Meyers et al., 1987, 1991]. However, it has not yet been possible to map the location of any major gene controlling IgE levels, perhaps because multiple loci may be involved or because environmental factors also play a determining role. Two studies in different populations have suggested that a gene on chromosome 5q may control log[total IgE] [Marsh et al., 1994; Meyers et al., 1994], perhaps in conjunction with a second modifier locus [Xu et al., 1995]. Still, conclusive evidence for the map location of a major locus controlling log[total IgE] levels is lacking.

One approach for modeling variation in a quantitative trait is variance components analysis, which partitions the observed phenotypic variance into components attributable to genetic and environmental causes under a general linear model [Lynch and Walsh, 1998]. Recently, variance components methods have been extended to allow inclusion of a separate component for variation attributed to shared alleles at marker loci, along with additive genetic effects, covariate effects, and random residual effects [Schork, 1993; Amos, 1994; Blangero, 1995]. In the context of gene mapping, several groups have argued that incorporating allele sharing tests for linkage into variance components methods is more powerful than the conventional sib pair method [Haseman and Elston, 1972], and less susceptible to false-positive re-
results when searching for linkage to genes controlling quantitative traits [Goldgar, 1990; Schork, 1993; Blangero, 1995; Williams and Blangero, 1997]. Here we present a variance components analysis designed to identify chromosomal regions containing genes controlling quantitative levels of log[total IgE] in families ascertained through two asthmatic sibs as part of the Collaborative Study of the Genetics of Asthma (CSGA).

There is strong evidence that IgE levels play an important role in the pathogenesis of asthma, with atopic, asthmatic individuals having higher total serum IgE levels than nonallergic, nonasthmatic individuals [Johansson et al., 1972; Burrows et al., 1989]. The CSGA families were ascertained for the purpose of mapping asthma susceptibility genes. Sera were collected to measure IgE levels and skin testing was conducted using 14 common allergens. A genome-wide scan for asthma susceptibility has been presented elsewhere [CSGA, 1997]. In this study, variance components models were used to partition variance in log[total IgE] into components of overall genetic control and components reflecting sharing of marker alleles. We also examined the effect of adjustment for specific immune response using an index of skin test response to an array of 14 common allergens as a covariate. This adjustment should help separate the “cognate” component of IgE from the “non-cognate” component, which might be controlled by separate genetic mechanisms [Marsh et al., 1995]. Comparing chromosomal regions showing evidence for linkage with and without this adjustment should help identify genes controlling total serum IgE (both specific and non-specific IgE) and only non-specific IgE.

MATERIALS AND METHODS

Study Subjects

The CSGA is a study aimed at identifying susceptibility genes for asthma and quantitative phenotypes associated with asthma. Data from both African-American subjects and Caucasian subjects have been collected at four centers: (1) the Johns Hopkins University, in conjunction with Howard University Medical Center in Washington DC, (2) the University of Chicago, (3) the University of Minnesota, and (4) the University of Maryland. Data for all Hispanic subjects were collected through the University of Maryland, in collaboration with the University of New Mexico. All families were ascertained through two asthmatic sibs. There were a total of 107 African-American families, 129 Caucasian families, and 32 Hispanic families (Table I). Clinical characteristics of these subjects have been described elsewhere [CSGA, 1997].

IgE Measurements

Total IgE concentration was measured in duplicate on serum samples using the Sanofi Diagnostic Pasteur’s method [Pierson et al., 1998] by Craig Luehr (Sanofi Diagnostics), and was expressed in International Units/ml (IU/ml). All measurements were repeated, again in duplicate, in a second independent assay. All four repeats were averaged for a final value. Values were log transformed to better approximate a Gaussian distribution. Skewness and kurtosis for total IgE in the combined sample were 7.73 and 107.31, and following transformation, skewness and kurtosis for log[total IgE] were –0.38 and 0.60, respectively.
**Skin prick tests** were performed on subjects using bifurcated smallpox needles (Wyeth Laboratories, Marietta, PA) and standardized glycerinated extracts (ALK Laboratories, Horsholm, Denmark; Greer Laboratories, Lenoir, NC) of common inhaled allergens plus histamine and diluent controls. Reactions were recorded by tracing the perimeter of the wheal 15 minutes after puncturing the skin and transferring this tracing with transparent surgical tape (Transpore; 3M Co., Minneapolis, MN) to paper, creating a permanent record of the skin test. For all wheal reactions, the diameter of the wheal was recorded as the average of the maximum length and the length of the perpendicular bisector of that maximum length. A reaction was considered positive if the final diameter of the wheal was \( \geq 3 \text{ mm} \) after subtracting the area of the wheal for the diluent control from the area resulting from each actual allergen.

The **Allergy Index** (AI), a continuous measure of allergic sensitivity [Freidhoff et al., 1983], was calculated using the following 14 allergens: mites (Dermatophagoides pteronyssinus and Dermatophagoides farinae), molds (Alternaria alternata, Cladosporium herbarum, and Aspergillus fumigatus), cat and dog dander (Felis domesticus and Canis familiaris, respectively), cockroaches (Blattella germanica and Periplaneta americana), grass, weed and tree pollens (Lolium...)

### TABLE I. Descriptive Statistics of Families and Phenotypes in the Three Ethnic Groups

<table>
<thead>
<tr>
<th></th>
<th>African Americans</th>
<th>Caucasians</th>
<th>Hispanics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Families (subjects)</td>
<td>Families (subjects)</td>
<td>Families (subjects)</td>
</tr>
<tr>
<td>All families</td>
<td>107 (708)</td>
<td>129 (1,104)</td>
<td>32 (205)</td>
</tr>
<tr>
<td>Hopkins</td>
<td>51 (335)</td>
<td>39 (244)</td>
<td>—</td>
</tr>
<tr>
<td>Maryland</td>
<td>24 (149)</td>
<td>38 (210)</td>
<td>32 (205)</td>
</tr>
<tr>
<td>Chicago</td>
<td>32 (224)</td>
<td>35 (291)</td>
<td>—</td>
</tr>
<tr>
<td>Minnesota</td>
<td>—</td>
<td>17 (359)</td>
<td>—</td>
</tr>
<tr>
<td><strong>All subjects</strong></td>
<td><strong>24.1 ± 15.0</strong></td>
<td><strong>32.1 ± 18.8</strong></td>
<td><strong>24.5 ± 15.5</strong></td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>females</td>
<td>285 (51.2)</td>
<td>472 (51.1)</td>
<td>85 (49.7)</td>
</tr>
<tr>
<td>males</td>
<td>273 (48.8)</td>
<td>452 (48.9)</td>
<td>86 (50.3)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>females</td>
<td>227 (79.4)</td>
<td>342 (72.5)</td>
<td>70 (82.4)</td>
</tr>
<tr>
<td>males</td>
<td>218 (79.8)</td>
<td>343 (75.9)</td>
<td>80 (93.0)</td>
</tr>
<tr>
<td><strong>Skin-test positive</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>females</td>
<td>2.14 ± 0.70</td>
<td>1.85 ± 0.75</td>
<td>2.04 ± 0.69</td>
</tr>
<tr>
<td>males</td>
<td>2.28 ± 0.70</td>
<td>1.98 ± 0.74</td>
<td>2.34 ± 0.62</td>
</tr>
<tr>
<td><strong>Log(1 + Allergy Index)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>females</td>
<td>0.17 ± 0.14</td>
<td>0.169 ± 0.15</td>
<td>0.214 ± 0.16</td>
</tr>
<tr>
<td>males</td>
<td>0.18 ± 0.15</td>
<td>0.185 ± 0.15</td>
<td>0.293 ± 0.16</td>
</tr>
<tr>
<td><strong>Skin-test positive subjects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>females</td>
<td>2.45 ± 0.59</td>
<td>2.08 ± 0.66</td>
<td>2.15 ± 0.67</td>
</tr>
<tr>
<td>males</td>
<td>2.29 ± 0.64</td>
<td>2.19 ± 0.67</td>
<td>2.42 ± 0.54</td>
</tr>
<tr>
<td><strong>Log(1 + Allergy Index)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>females</td>
<td>0.21 ± 0.12</td>
<td>0.23 ± 0.13</td>
<td>0.26 ± 0.14</td>
</tr>
<tr>
<td>males</td>
<td>0.23 ± 0.14</td>
<td>0.24 ± 0.12</td>
<td>0.32 ± 0.15</td>
</tr>
</tbody>
</table>

*Values represent number of subjects (percentage) or mean ± SD.
perenne, Ambrosia artemisiifolia, Artemisia vulgaris, Quercus alba, Betula verrucosa). In the Hispanic subjects, an additional three allergens [Bermuda grass (Cynodon dactylon), Russian thistle (Salsola kali), and Mountain Cedar (Juniperus ashei)] were included in determination of the AI since the highest response rates occurred for these three local allergens in the Hispanic subjects. Briefly, after skin testing with the above allergens, each patient’s allergic sensitivity to each particular allergen (i.e., wheal diameter at a specific concentration) was expressed as the allergen concentration giving rise to a wheal of a particular diameter in an intradermal test [Squire, 1950; Marsh et al., 1973]. Distributions of log[antigen concentration]s for all allergens were adjusted to the mean of the distribution of the most prevalent allergen in order to consider quantitative differences in the subjects’ responses toward each extract (dependent on allergen quality, etc.) [Freidhoff et al., 1983]. The resulting set of distributions was rescaled and the final AI was generated by averaging separate indices over all allergens for each individual. Finally, the distribution of AI was then log transformed (after adding 1 to each value) to reduce skewness. Skewness and kurtosis for AI in the combined sample were 1.17 and 1.33, and following transformation, skewness, and kurtosis for log[1 + AI] were 0.48 and –0.68, respectively.

Genotyping

DNA was extracted from whole blood, cell pellets, or lymphoblastoid lines by standard methods. Genotypes were determined at the Mammalian Genotyping Service in Marshfield, WI, using a fluorescent detection system. The screening set (Weber v8.0) included a mixture of dinucleotide, trinucleotide, and tetranucleotide polymorphisms, with average distance between markers of 10 cM (for details see Marshfield web site; http://www.marshmed.org/genetics/). Map distances used in this analysis were obtained from the analysis of the combined CSGA families (for details see CSGA web site; http://www.csga.org/).

STATISTICAL METHODS

The variance component models and multipoint IBD estimation methods implemented in the Sequential Oligogenic Linkage Analysis Routines or SOLAR package [Almasy and Blangero, 1998] were used. SOLAR, which incorporates a modified version of the FISHER program [Lange et al., 1987], provides maximum likelihood estimates of each variance component. This approach is based on specifying the expected genetic covariances among “n” relatives in an extended family as a function of background polygenic factors, fixed effects of covariates and residual error, along with a component due to an unobserved quantitative trait locus (QTL) linked to an observed marker locus. Under this model, the (n × n) covariance matrix for a pedigree can be partitioned as

\[ \Omega = \prod \sigma^2_m + 2\Phi \sigma^2_e + \sigma^2_g \]

where \( \sigma^2_m \) is the genetic variance due to an unobserved QTL tightly linked to the observed marker m; \( \Pi \) is the (n × n) matrix whose elements represent the probability that any two individuals i and j are identical by descent (IBD) for this observed genetic marker m; \( \sigma^2_e \) is the background polygenic variance influencing the trait; \( \Phi \) is
Variance Components for IgE Levels

The variance components were estimated simultaneously along with fixed effects of observed covariates (e.g., age, sex, and conditionally the log[1 + AI]), using maximum-likelihood methods assuming the residual error term followed a multivariate normal distribution within families. A series of three models for each racial subgroup were considered: (1) a sporadic model with an overall mean, fixed effects of covariates, residual variance component, but neither the polygenic nor the marker locus component; (2) a polygenic model with trait mean, covariates, and an additive polygenic component along with the residual variance but no marker information considered, and hence no QTL; and (3) a linked model with the trait mean, covariates, residual polygenic variance, and error variance plus the variance due to the observed marker locus. This third model had 10 parameters: (1) \( \mu \) = the overall trait mean; (2) \( \beta(\text{sex}) \) = the regression coefficient for sex; (3) \( \beta(\text{age}_m) \) = the regression coefficient for age in males; (4) \( \beta(\text{age}_f) \) = the regression coefficient for age in females; (5) \( \beta(\text{age}_m^2) \) = the regression coefficient for age squared in males; (6) \( \beta(\text{age}_f^2) \) = the regression coefficient for age squared in females; (7) \( \sigma \) = the trait standard deviation; (8) \( h^2 \) = the proportion of variance due to additive polygenic factors; (9) \( h^2_m \) = the proportion of the total variance due to the observed marker locus (acting through its linkage to an unobserved QTL); and (10) \( \beta(\text{AI}) \) = the regression coefficient for log[1 + AI] for those models incorporating AI.

Each of the three models was used to estimate genetic parameters with and without AI as a covariate to determine the effects of adjusting for specific allergic response on a genome-wide search for unobserved QTLs controlling serum IgE. The null hypothesis of no linkage, i.e., where the genetic variance due to the marker locus equals zero, was tested by comparing each linked model to the polygenic model where only a background genetic component was estimated. Twice the difference in log, likelihoods of these two models provides a likelihood ratio test statistic that is asymptotically distributed as a mixture of a \( \chi^2 \) with one degree of freedom and a degenerate \( \chi^2 \) with zero degrees of freedom [Self and Liang, 1987]. Alternatively, converting this difference to log_{10} scale provides a LOD score statistic for evidence of linkage to an individual marker for two-point analysis or to an imputed chromosomal position in multipoint analysis. This variance components strategy can be applied to each marker individually or in a multipoint analysis by using an array of markers to estimate the probability of IBD sharing at arbitrary points along a chromosomal region. In our own multipoint analysis, a total of 323 markers over the entire genome were used in a multipoint variance components analysis.

As these CSGA families were ascertained through sibling pairs with asthma and our phenotype of interest, log[total IgE], is associated with asthma, we also performed analyses in which an ascertainment correction was applied. Comuzzie and Williams [1999], in the analysis of an alcoholism data set (COGA), showed how ascertainment correction could alter the magnitude of evidence for linkage, although the location of their linkage peak varied only slightly. In that context, no ascertainment correction produced results intermediate to several different ascertainment correction strategies. In contrast to their analysis of a dichotomous affection status, our analysis of log[total IgE] as a quantitative trait provided little difference in evidence for linkage before and after ascertainment correction. This may be because there was
little difference in the distribution of log[total IgE] between asthmatic probands and other family members in our own data. For this reason, our results are presented without ascertainment correction.

Tests for possible epistatic interactions between genes on different chromosomes were performed for those loci with LOD scores >2 on the genome wide screen, with any loci showing a LOD score >1. This test for interaction uses an extension of the variance components model where interactions between two markers or two chromosomal regions are considered simultaneously. In the two-locus tests for interaction, IBD matrices for both the selected chromosomal regions (\(\Pi_1\) and \(\Pi_2\)), are used to estimate the additive-by-additive component of the epistatic interaction. The element by element matrix multiplication of \(\Pi_1\) and \(\Pi_2\) serves as the coefficient for this epistatic component \(\sigma_{m1m2}^2\) [Blangero and Almasy, 1997]. A series of three models were considered in tests for epistatic interaction: (1) two separate one-locus models with marker effects due to each marker locus individually; (2) a two-locus additive model with marker effects for two loci simultaneously, but no interaction term; and (3) a two-locus epistatic model with both marker effects (\(\sigma_{m1}^2\) and \(\sigma_{m2}^2\)), in addition to the epistatic term (\(\sigma_{m1m2}^2\)). A sequence of likelihood ratio tests, provide assessment of one locus “adjusted” for the other (i.e., comparing the two-locus model with no interaction to a one-locus model). The significance of the final epistatic term is assessed by comparing the epistatic model to the additive model. To account for the constraint of the variance components to be positive in the above models, all hypotheses tested were assessed at a type I error rate of 0.05 following generation of \(P\) values using a one-tailed test.

RESULTS

Log[total IgE] appears to be normally distributed in all three racial groups in the CSGA, and has a strong genetic component in families from all three racial groups. The residual heritability (after the removal of age and gender effects) was 60.3 ± 7.1% for African Americans, 55.9 ± 5.7% for the Caucasians, and 43.5 ± 15.6% for Hispanics. When log[1 + AI] was added as a covariate in the model, the corresponding residual heritabilities were 52.4 ± 7.3%, 42.2 ± 6.1%, and 44.1 ± 15.1% for the African Americans, Caucasians, and Hispanics, respectively. Thus, both total IgE levels and adjusted IgE levels appear to be under strong genetic control. These modest differences in the heritability of log[total IgE] with and without adjustment for AI provide no indication of the potential evidence for linkage to any specific genes/regions.

Genome-wide LOD scores for log[total IgE] are presented in Figure 1 for each racial group separately, without (Fig. 1A) and with (Fig. 1B) covariate adjustment for log transformed [1 + AI]. Heterogeneity in evidence for linkage to particular loci is apparent by comparing these linkage results across the three racial groups. Without adjustment for AI (Fig. 1A), the peak LOD scores in the African-American families (row 1) were 1.80 at 156 cM on chromosome 4 (closest marker D4S2431) and 1.52 at 102 cM on chromosome 2 (closest marker D2S1270). In the Caucasian families (row 2), the highest LOD score was 1.70 at 50 cM on chromosome 4 (closest marker D4S1627). The Hispanic families (row 3) showed no LOD scores greater than 1.5, and the peak LOD of 1.40 was seen both at 76 cM on chromosome 1 (closest marker D1S2134) and at 102 cM on chromosome 14 (closest marker D14S611).

The results of the analysis of log[total IgE] with covariate adjustment for spe-
Variance Components for IgE Levels

Fig. 1. Genome-wide multipoint LOD scores for African Americans, Caucasians, and Hispanics, (A) without adjustment for specific immune response and (B) with adjustment for specific immune response. Note: Vertical lines distinguish chromosomes.

cific immune response using log(1 + AI) are shown in Figure 1B. Comparing the results of these analyses with and without this adjustment among African-American families, the peak LOD on chromosome 4 seen without this covariate adjustment decreased from 1.80 to 1.11 at the same location (158 cM) and from 1.52 to 1.13 at 102 cM on chromosome 2. In the Caucasian families, the peak on chromosome 4 decreased from 1.70 to 1.23, while the peak LOD score on chromosome 18 increased.
from 1.32 to 2.75 (with the closest marker being D18S844). In the Hispanic families, the maximum LOD of 1.51 occurred at 94 cM on chromosome 9 (closest marker D9S938) and 1.48 at 22 cM on chromosome 12 (closest marker D12S391), regions not implicated in the previous analyses without adjustment for AI.

Figure 2A presents LOD scores over chromosome 18 for all three racial groups (using 12 markers, with D18S844 at 109.4 cM as the terminal marker). There are obvious ethnic differences, with significant LOD scores evident only in the Cauca-
Variance Components for IgE Levels

sian group of families. Multipoint analyses were then performed for chromosome 18 for the Caucasian families stratifying by the four recruiting centers. For this analysis, a new terminal marker was typed (D18S1905 at 125.5 cM) in the CSGA Caucasian subjects to verify that the peak at D18S844 was not an artifact of its being the terminal marker. In the combined sample from all centers, the LOD at 124 cM is 1.18, while the LOD at D18S844 remains almost the same (LOD at 108 cM was 2.75 before adding D18S1905 and 2.61 after adding D18S1905).

A likelihood ratio test for heterogeneity at locus 108 cM on chromosome 18 was performed by summing the log-likelihoods at locus 108 cM from each center and comparing this sum to the log-likelihood computed on the combined sample. Twice this difference yielded a $\chi^2$ test of 50.45 with 30 degrees of freedom ($P = 0.01$). Hence, statistically significant evidence of heterogeneity among the four centers at position 108 cM on chromosome 18 was observed. This evidence of statistical heterogeneity is consistent with the observed LOD scores from each center for Caucasian families, where only two centers (Maryland and Chicago) appear to have families contributing evidence for linkage (Fig. 2B).

Multi-locus analyses to test for epistasis were then performed among the Caucasian families using chromosome 18 (the position showing the strongest evidence for linkage) paired with all other loci yielding a LOD score above 1.0 in the genome-wide multipoint analysis of the Caucasian families (i.e., chromosome 2, position 106 cM; chromosome 4, position 48 cM; and chromosome 20, position 32 cM). Three levels of modeling were performed: (1) single-locus models where each locus was analyzed separately, (2) two-locus models with only additive effects for each pair of loci; and (2) two-locus models with additive effects as well as an epistatic term for interaction between the two loci. One-tailed $P$ values were generated for all hypotheses tested to account for the constraint of the variances components to be positive in the above models, and significance was assessed at a type I error rate of 0.05.

Table II shows the results of this interaction analysis. Each single-locus model was compared to the baseline polygenic model to test for significance of the individual marker loci. All four of these selected loci were statistically significant when considered alone, explaining between 27 and 40% of the total variance. Then, each two-locus additive model was tested against the 1-locus model with chromosome 18 alone, providing a test for significant additive effects of the second locus (either chromosome 2, 4, or 20). Chromosome 4 still showed a statistically significant additive effect when chromosome 18 was considered, although the estimated heritability due to sharing of marker alleles at this position on chromosome 4 explained considerably less variance than did the region on chromosome 18 (18 vs. 30%, respectively). The epistatic term allowing interaction between chromosome 18 and 4 was not significant ($P > 0.05$ comparing models 8 and 7 in Table II). Chromosomes 2 and 20 have borderline significant additive effects ($P = 0.045$ and $P = 0.050$) comparing the respective two-locus additive models to the two locus model for chromosome 18. None of the models with a separate term for interaction showed a significant improvement, suggesting epistasis between these regions may not be a major factor in the genetic control of IgE.

DISCUSSION

Given the strength of the association between clinical asthma and elevated IgE levels, we have chosen to use families ascertained by the Collaborative Study of the
<table>
<thead>
<tr>
<th>No.</th>
<th>Model</th>
<th>Chromosome</th>
<th>$\mu$</th>
<th>$h^2$</th>
<th>$h^2_{m1}$</th>
<th>$h^2_{m2}$</th>
<th>$h^2_{m1m2}$</th>
<th>$e$</th>
<th>Loglikelihood</th>
<th>$\chi^2$, 1 df (P value)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Polygenic</td>
<td>—</td>
<td>1.838</td>
<td>0.422</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.578</td>
<td>69.635</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>1-locus</td>
<td>18</td>
<td>1.833</td>
<td>0.041</td>
<td>0.399</td>
<td>—</td>
<td>—</td>
<td>0.561</td>
<td>75.641</td>
<td>12.012 (2 v 1)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>3</td>
<td>1-locus</td>
<td>2</td>
<td>1.845</td>
<td>0.104</td>
<td>0.315</td>
<td>—</td>
<td>—</td>
<td>0.580</td>
<td>72.665</td>
<td>6.060 (3 v 1)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>4</td>
<td>2-locus, additive</td>
<td>18,2</td>
<td>1.839</td>
<td>0.000</td>
<td>0.297</td>
<td>0.159</td>
<td>—</td>
<td>0.544</td>
<td>77.077</td>
<td>2.872 (4 v 2)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>5</td>
<td>2-locus, epistatic</td>
<td>18,2</td>
<td>1.838</td>
<td>0.000</td>
<td>0.258</td>
<td>0.119</td>
<td>0.146</td>
<td>0.476</td>
<td>77.239</td>
<td>0.324 (5 v 4)</td>
<td>(&gt;0.05)</td>
</tr>
<tr>
<td>6</td>
<td>1-locus</td>
<td>4</td>
<td>1.833</td>
<td>0.128</td>
<td>0.279</td>
<td>—</td>
<td>—</td>
<td>0.593</td>
<td>72.488</td>
<td>5.706 (6 v 1)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>7</td>
<td>2-locus, additive</td>
<td>18,4</td>
<td>1.831</td>
<td>0.000</td>
<td>0.291</td>
<td>0.180</td>
<td>—</td>
<td>0.529</td>
<td>78.585</td>
<td>5.888 (7 v 2)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>8</td>
<td>2-locus, epistatic</td>
<td>18,4</td>
<td>1.827</td>
<td>0.000</td>
<td>0.220</td>
<td>0.102</td>
<td>0.270</td>
<td>0.408</td>
<td>79.243</td>
<td>1.316 (8 v 7)</td>
<td>(&gt;0.05)</td>
</tr>
<tr>
<td>9</td>
<td>1-locus</td>
<td>20</td>
<td>1.842</td>
<td>0.161</td>
<td>0.272</td>
<td>—</td>
<td>—</td>
<td>0.567</td>
<td>72.312</td>
<td>5.354 (9 v 1)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>10</td>
<td>2-locus, additive</td>
<td>18,20</td>
<td>1.836</td>
<td>0.000</td>
<td>0.299</td>
<td>0.164</td>
<td>—</td>
<td>0.537</td>
<td>77.072</td>
<td>2.862 (10 v 2)</td>
<td>(0.050)</td>
</tr>
<tr>
<td>11</td>
<td>2-locus, epistatic</td>
<td>18,20</td>
<td>1.835</td>
<td>0.000</td>
<td>0.219</td>
<td>0.082</td>
<td>0.302</td>
<td>0.397</td>
<td>77.833</td>
<td>1.522 (11 v 10)</td>
<td>(&gt;0.05)</td>
</tr>
</tbody>
</table>

* $\mu$ = trait mean; $h^2$ = proportion of variance due to residual additive polygenic effects; $h^2_{m1}$ = proportion of variance due to the first marker locus; $h^2_{m2}$ = proportion of variance due to the second marker locus; $h^2_{m1m2}$ = proportion of variance due to the epistatic effect of locus 1*locus 2; $e$ = residual variance. In each 2-locus model, $h^2_{m1}$ is the proportion of variance due to chromosome 18, while $h^2_{m2}$ is the proportion due to the second locus (i.e., chromosome 2, 4, or 20).
Genetics of Asthma [CSGA, 1997] in a genome-wide screen to identify genes controlling the quantitative trait of serum IgE levels using variance component models. Since the CSGA families were recruited from different racial/ethnic groups and four different geographic sites, there is substantial potential for heterogeneity. These families were all ascertained through a pair of asthmatic sibs, and thus are not representative of the general population or even all families with a single asthmatic proband. These highly selected multiplex asthmatic families from the CSGA limit our inferences about possible genes controlling IgE levels due to the ascertainment scheme, but they do offer an excellent opportunity to identify genes controlling IgE levels that may be part of the complex pathogenetic process for asthma.

Several published reports have suggested that either selected genes or chromosomal regions may be involved in direct genetic control of serum IgE levels. To date, some evidence of linkage to total IgE levels has been reported in the following chromosomal regions: 2pter, 4q35, 5q23-q31, 6p21.2-p23, 7p15.2, 9q31.1, 11q13, 12q14-q24.33, and 16q22.1-q24.2 [Marsh et al., 1994; Hizawa et al., 1995; Barnes et al., 1996; Daniels et al., 1996; Nickel et al., 1997; Wjst et al., 1999]. Evidence for linkage to 2pter and 9q31.1 was provided from a genome-wide screen among German families [Wjst et al., 1999] and the evidence for chromosomes 4q35, 7p15.2, and 16q22.1-q24.2 came from a genome-wide screen among families from the United Kingdom and Australia [Daniels et al., 1996]. In our own study, the best evidence for linkage of genes controlling total IgE levels (i.e., without covariate adjustment for AI) occurred at two markers in distinct regions of chromosome 4 (D4S1627 and D4S2431 in the Caucasian and African-American families, respectively). Interestingly, D4S2431 (at 176 cM pter) is approximately 24 cM centromeric from the locus where the Oxford group reported evidence for linkage to asthma and IgE (200 cM pter) [Daniels et al., 1996]), according to the Marshfield map (http://www.marshmed.org/genetics/).

In this study, we did not observe linkage to some regions previously reported to contain genes controlling total IgE in other populations, including 5q23-q31, 6p21-p23, 11q13, and 12q14-q24.33. The evidence for linkage to total IgE as a quantitative trait has been well established on chromosomes 5q, 6p, and 11q among multiple well-characterized populations, several of which were genetically restricted and/or homogeneous (e.g., Amish, Dutch). A number of different possibilities may account for these inconsistencies, including differences in sampling designs and statistical methods. The primary methods of analysis for total IgE as a quantitative trait have been sib-pair regression analysis and its multipoint extensions as implemented in MAPMAKER/SIBS. This is the first report of linkage analysis for total IgE using variance components analysis. Evidence for linkage to total IgE and markers in chromosome 12q has been shown primarily in studies treating total IgE as a qualitative trait, with the exception of the German study [Wjst et al., 1999].

Log[total IgE] clearly has a strong genetic component with heritability estimates between 44 and 60% in the three racial groups studied here. These estimates agree with published estimates from other populations. However, this estimated heritability includes both major locus effects and true polygenic components, as these two are perfectly confounded in the general linear model underlying this variance components analysis [Amos, 1994]. Some previous segregation analyses on IgE have suggested a major gene effect [Meyers et al., 1991; Xu et al., 1995; Sampogna et al.,...
Mathias et al. 2000], while others found only polygenic components [Hasstedt et al., 1983]. The premise of the variance components approach used here is that major gene effects can be separated from background polygenic factors through linkage and allele sharing at observed marker loci.

As seen in Figure 1A, a genome-wide screen of 323 markers provided only modest evidence of linkage between these markers and an unobserved QTL controlling total serum IgE. Furthermore, these findings varied across the three racial/ethnic groups available in the CSGA. In the analysis of total IgE without covariate adjustment for AI, for African-American families, the maximum LOD was 1.8 (near marker D4S2431); for Caucasian families, the maximum LOD was 1.7 (near D4S1627); and for Hispanic families, the LOD scores of 1.4 occurred on two distinct regions (near D1S2134 and D14S611). We chose not to combine the three racial/ethnic groups due to the potential for spurious results when genetically distinct sub-populations are combined because IBD sharing is somewhat dependent on estimated marker allele frequencies, which do differ across the three racial groups.

To further refine our analysis of log[total IgE], we incorporated the AI as an observed measure of specific immune response to a number of common antigens. It is important to realize that adjusting for log[1 + AI] as a covariate in this linear model alters the interpretation of these variance components models somewhat. The phenotype being analyzed becomes the residual variance in log[total IgE] after fixed effects of a combined skin test response to a panel of 14 specific allergens have been removed. To the extent that both log[total IgE] and AI are controlled by the same gene(s), this adjustment will reduce the overall heritability and potentially remove evidence for linkage. In fact, a bivariate model that included both log[total IgE] and log[1 + AI] showed a substantial genetic correlation (0.58) between these two observed phenotypes, as well as a strong correlation (0.35) in their residual error terms. Thus, we would expect heritability for adjusted log[total IgE] to be smaller in the models with AI as a covariate. This was the case in two of the three racial/ethnic groups when a simple polygenic model was fit: the heritability dropped from 60 to 52% in African Americans and from 56 to 42% in Caucasians, while the Hispanic families showed no real change (44% both with and without this covariate).

The most striking evidence for linkage to a nonspecific IgE locus (after covariate adjustment for AI) was an increase in the LOD score from 1.33 to 2.74 at D18S844 among the Caucasian families. No prior linkage to this region has been reported, although it is of interest that D18S844 is within 1 cM of the gene encoding the cytosolic component of nuclear factor of activated T cells (NFATC1). Suppressed activity of NFATC1 was associated with highly increased serum IgE levels in patients with severe atopic dermatitis [Wierenga et al., 1999], thus making it an intriguing candidate for control of IgE levels.

Clinical aspects of this sample should be considered when addressing issues related to genetic contribution to a complex phenotype such as log[total IgE]. Since only the Caucasian families yield evidence of linkage to a region on chromosome 18, the question arises of whether these subjects are clinically distinct from the African-American and Hispanic subjects. As previously shown (Table I), the Caucasian family members are slightly older (32 ± 19 years) than the African-American (24 ± 15 years) and Hispanic (25 ± 16 years) subjects. Further, the Caucasian group appears to be less
Variance Components for IgE Levels

allergic (having lower rates of skin test positivity and lower log[total IgE]). In those subjects who are positive for at least one skin test, the Caucasians appear less “reactive” with slightly lower log[total IgE] levels, although their AI values were intermediate. Thus, one interpretation of these results is that a gene in this region of chromosome 18 could represent an “allergen-insensitive” IgE response locus.

In summary, dissection of the genetic and environmental factors contributing to variation of IgE is complex and will require multiple studies in different ethnic groups and populations. Characterization of the environmental factors that affect IgE levels, including allergen exposures, will be necessary to fully explore both clinical and genetic heterogeneity in this trait. Once specific IgE genes are identified for control of specific or non-specific IgE levels, the role of IgE and atopy in the pathogenesis of asthma may become clearer.

ACKNOWLEDGMENTS

We extend our gratitude to the CSGA study coordinators, Maria Stockton-Porter, Rebecca Brown, Shannon Gierczale, and Christina Hueber, for their work in recruiting the study participants and in performing all phenotype tests. We also thank Dr. Elizabeth Pugh for her invaluable advice on analytical methods and the implementation of SOLAR, and Fang-Chi Hsu and Xin-Liu for their help in analysis. This work was supported by United States Public Health Service grants HL-49612 to T.H.B., HL-49596 to C.O., HL-49609 to M.N.B., HL-49602 to E.R.B., DK595522 to N.J.C, HL-58977 to the S.S.R, and HV-48141 to the Mammalian Genotyping Service.

REFERENCES


