

University of Groningen

## Genetics of asthma and atopy

Koppelman, Gerard Henk

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2001

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Koppelman, G. H. (2001). *Genetics of asthma and atopy*. s.n.

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Chapter 5 | Major genes regulating total serum IgE levels in families with asthma

*(Am J Hum Genet 2000; 67:1163-1173)*

Jianfeng Xu<sup>1</sup>, Dirkje S. Postma<sup>2</sup>, Timothy D. Howard<sup>1</sup>, Gerard H. Koppelman<sup>2</sup>, Siqun L. Zheng<sup>1</sup>, O. Colin Stine<sup>3</sup>, Eugene R. Bleecker<sup>1</sup>, and Deborah A. Meyers<sup>1</sup>.

1. Center for Human Genomics, Wake Forest University, Winston-Salem, NC, USA
2. University Hospital, Groningen, the Netherlands
3. Department of Epidemiology, University of Maryland, Baltimore, MD, USA

## ABSTRACT

---

Immunoglobulin E (IgE) has a major role in the pathogenesis of allergic disorders and asthma. Previous data from 92 families, each identified through a proband with asthma, showed evidence for two major genes regulating total serum IgE levels. One of these genes mapped to 5q31-33. In the current study, the segregation analysis was extended by the addition of 108 probands and their families, ascertained in the same manner. A mixed recessive model (i.e. major recessive gene and residual genetic effect) was the best fitting and the most parsimonious one locus model from the segregation analysis. A mixed two-major gene model (i.e. two major genes and residual genetic effect) fit the data significantly better than did the mixed recessive one-major gene model. The second gene modified the effect of the first recessive gene. Individuals with the genotype aaBB (homozygous high-risk allele at the first gene and homozygous low-risk allele at the second locus) had normal IgE levels (mean 23 IU/ml) and only individuals with genotypes aaBb and aabb had high IgE levels (mean 282 IU/ml). A genome-wide screen was performed using variance component analysis. Significant evidence for linkage was found for a novel locus at 7q with a multipoint lod score of 3.36 ( $p = 0.00004$ ). A lod score of 3.65 ( $p = 0.00002$ ) was obtained after genotyping additional markers in this region. Evidence for linkage was also found for two previously reported regions, 5q and 12q, with lod scores of 2.73 ( $p = 0.0002$ ) and 2.46 ( $p = 0.0004$ ), respectively. These results suggest that several major genes, plus residual genetic effects, regulate total serum IgE levels.

## INTRODUCTION

---

Immunoglobulin E (IgE) has important functions in the development of allergic disorders and asthma. High total serum IgE levels have been reported to be correlated with the clinical expression of allergy and asthma.<sup>1-4</sup> Epidemiologic studies have shown that higher IgE levels are associated with bronchial hyperresponsiveness (BHR), a major component of the asthma phenotype.<sup>2, 5, 6</sup> In fact, high total serum levels IgE predicts the development of asthma independent of other allergic factors. Therefore, understanding the genetic mechanisms regulating total serum IgE levels will be important in efforts to dissect the hereditary components of asthma and allergy, complex genetic disorders influenced by the interactions between multiple genes and environmental exposures.<sup>7</sup>

In a study reported elsewhere, our use of two-locus segregation analysis revealed evidence for two major genes and a residual genetic effect regulating total serum IgE levels in the first set of 92 Dutch families ascertained through a parent with asthma.<sup>8</sup> For the first of these loci, evidence of linkage to chromosome 5q was obtained from both one-locus and two-locus analyses based on a candidate gene approach.<sup>8, 9</sup> Now that data are available on the total sample of 200 families ascertained through a parent with asthma (all of which originally had been characterized approximately 25-35 years previously), we have performed a genomewide search for genes regulating total serum IgE levels. The purposes of the current study are: 1) to determine the relationship between total serum IgE levels and other measures of asthma and allergy, 2) to examine the familial aggregation of the total IgE levels by estimating the degree of correlation between relative pairs and to test the fit of various models by performing one- and two-locus segregation analysis, 3) to systematically search for the locations of the major genes across the genome with evenly spaced autosomal markers (~10 cM apart) and variance component analyses. High total serum IgE levels, an important phenotype closely associated with asthma and allergic disorders, are an ideal quantitative trait for use with this analytic approach.

## FAMILIES AND METHODS

---

### Family ascertainment

A total of 200 families (1,171 family members) were ascertained through probands who were initially studied during 1962-75 at Beatrixoord Hospital, Haren, the Netherlands, a regional referral center for patients with asthma and other airway obstruction diseases. Patients who had symptomatic asthma without a current asthma exacerbation were referred to this hospital and were admitted for a standardized, complete evaluation. At the time of initial testing, all probands had asthma symptoms, were hyperresponsive to histamine ( $PC_{20}$  forced expiratory volume in 1s [FEV<sub>1</sub>]  $\leq$  32 mg of histamine/ml, 30-s method), and were < 45 years old. The first 92 probands, together with their spouses, children, children's spouses, and

grandchildren > 6 years old, were recruited and evaluated in the early 1990's.<sup>10</sup> To enlarge the sample, 108 probands and their families were collected during 1994-1999 by use of the same ascertainment scheme that was used for the first 92 families. This study was approved by the Medical Ethics Committee of the University Hospital Groningen, and all participants signed an informed-consent document. Among the 200 families, 166 families consisted of two generations, 33 consisted of three generations; 1 family consisted of four generations.

#### **Clinical evaluation**

All participants answered a modified British Medical Society Respiratory Questionnaire, as well as additional questions pertinent to the diagnosis and assessment of asthma and obstructive pulmonary disease. Pulmonary function was tested using standard methods that included spirometry before and after the administration of inhaled salbutamol (800 mg). Testing of bronchial responsiveness to histamine was performed using the method of De Vries et al.<sup>11</sup>, which had been used to assess the initial participants during the period 1962-75. The reactivity-testing protocol consists of having the subject inhale increasing concentrations of histamine, for 30 seconds of tidal breathing up to a maximum dose of 32 mg of histamine/ml. The test was stopped if FEV<sub>1</sub> decreased  $\geq 20\%$ . Other evaluations included skin tests for responsiveness to 16 common allergens (intracutaneous testing in adults and prick testing in children), a differential blood count (including total eosinophil count), and measures of total serum IgE levels as well as IgE specific to house dust and mixed pollens. A positive skin test was defined as the presence of  $\geq 1$  reaction with a wheal diameter  $\geq 5$ mm. Total serum IgE was measured by solid phase immunoassay (Pharmacia IgE EIA: Pharmacia Diagnostics).

#### **Molecular methods**

Blood samples were shipped from the Netherlands to the molecular genetics laboratory at the University of Maryland at intervals of ~ 2 wk. DNA was isolated by standard protocols using a Puregene kit (Gentra). For the genome wide screen, we used the Weber (version 8) set of markers, which spans the human genome at an average of 10 cM and consists of 366 autosomal markers, 86% of which are tri- and tetra-nucleotide repeats with an average marker heterozygosity of 76%. We performed multiplex PCR using fluorescently labeled primers, separated the resulting amplified fragments on denaturing polyacrylamide gels, detected the fragments with the use of ABI 377 sequencing machines, and scanned and scored the genotypes, using ABI software. A modified version of the program Linkage Designer<sup>12</sup> was used to bin the alleles and check inheritance. The output from Linkage Designer was then analyzed further, for any inconsistencies by operating the LINKAGE software without disease information. As a final check of the data, we used CRIMAP<sup>13</sup> to determine the order and length of the chromosomal map and to detect double recombinants.

### Statistical methods

Total serum IgE levels were logarithm transferred ( $\log_{10}$ ) in order to approximate a normal distribution (all analyses were performed using  $\log_{10}$  (IgE) levels). Because log levels were higher in males and in younger individuals, the effects of sex and age was included in the genetic analyses. In the variance component analysis, adjustment of the fixed effects of sex and age was performed simultaneously with fitting of the various models. To estimate the correlation coefficients of IgE for pairs of relatives and for complex segregation analysis, the fixed effects of age and sex were removed by taking the residuals after linear-regression analysis had been performed.

Correlation coefficients of total serum IgE levels among various pairs of relatives were estimated from the sums of squares and cross-products from the pairs, using the computer program FCOR of S.A.G.E. (Statistical Analysis for Genetic Epidemiology). Three distinct weighting methods were used. In the method of equal weight to pairs, every possible pair has equal weight. The pedigrees with large number of pairs will contribute more information than will pedigrees that contain a relatively small number of pairs. In the method of equal weight to pedigrees, each pedigree, regardless of its size contributes equal weight. The data are averaged within pedigrees before they are averaged across pedigrees. In the method of equal weight to nuclear families, only the nuclear families (parents and children) are included.

Complex segregation analyses assuming one- and two-locus models were used to evaluate the transmission of total IgE levels within the 200 families. An ascertainment correction was not used, for the following reasons: 1) the families were ascertained through a parent (not a child) with asthma, not through probands with specific IgE levels; 2) the probands tend to have high IgE levels but with a large range of values; 3) we are interested in major genes in the specific population of families with asthma. In the one-locus segregation analysis, various models were evaluated, including a general model and Mendelian major-gene models, an environmental model, a polygenic model, and mixtures of various polygenic models with either a major-gene model or an environmental model. The likelihoods that each of these models fit the observed family data were computed using the computer software package Pedigree Analysis Package (PAP), revision 4.0. In the cases of mixed models, the likelihoods were approximated by allowing information from previously analyzed family members to represent information about the entire family.<sup>14</sup> The parameters in the general models include (1) one allele frequency  $q_a$  (corresponding to a high value) and three genotypic frequencies ( $F_{AA}$ ,  $F_{Aa}$  and  $F_{aa}$ ) that were assumed to occur in Hardy-Weinberg equilibrium; (2) three arbitrary transmission probabilities  $\tau_{AA}$ ,  $\tau_{Aa}$  and  $\tau_{aa}$  representing the probability that an individual of a given genotype transmitted allele A to the offspring and (3) three arbitrary genotypic means  $\mu_{AA}$ ,  $\mu_{Aa}$  and  $\mu_{aa}$ , a common variance for all the genotypes, and a residual genetic heritability  $h^2$ , which is partitioned from the variance and represents the additive effects of polygenic loci. In the Mendelian models, the three transmission probabilities were fixed to the Mendelian

inheritance, ie.  $\tau_{AA} = 1.0$ ,  $\tau_{Aa} = 0.5$  and  $\tau_{aa} = 0$ . In the environmental models, the three transmission probabilities were set to be the same  $\tau_{AA} = \tau_{Aa} = \tau_{aa}$ , reflecting the independence between parental genotypes and the transmission probabilities. For the Mendelian-only models and environmental only models,  $h^2$  is set to 0. For the mixed models,  $h^2$  is estimated. Within the Mendelian models, a dominant model is derived by setting  $\mu_{Aa} = \mu_{aa}$  and a recessive model is inferred by fixing  $\mu_{AA} = \mu_{Aa}$ .

Two-locus segregation analysis was performed using PAP. Analysis was performed to determine whether two major genes, compared to two environmental factors or one major gene, better modeled the segregation of IgE levels in the families. The parameters in the two-locus segregation analysis included: allele frequencies at each locus ( $q_a$  and  $q_b$ ), the recombination fraction between the two loci ( $\theta$ ), the means ( $\mu$ 's) for the nine possible distributions representing the nine types of individuals (AABB, AABb, Aabb, AaBB, AaBb, Aabb, aaBB, aaBb, and aabb) with a common standard deviation ( $s$ ), and a heritability ( $h^2$ ), which is a measure of residual variance within each type. Mendelian models were inferred by setting the nine transmission probabilities (probability that an individual of a given genotype transmits allele A and B to his or her offspring) to Mendelian expectations. The environmental model was inferred by setting all nine transmission probabilities to equal values, reflecting the independence between parental genotypes and the transmission probabilities.

Two criteria were used to compare these models. For hierarchical models, the likelihood-ratio test was used. Twice the difference in likelihoods ( $-2\ln L$ ) between a restricted and an unrestricted model approximately follows a chi-square statistic, with degrees of freedom equal to the difference in the number of parameters used in the two models. The best fitting model is the one requiring the fewest estimated parameters, while giving a  $\ln$ -likelihood not significantly smaller than the unrestricted model. When comparing non-hierarchical models, the Akaike's<sup>15</sup> Information Criterion (AIC) was used,  $-2\ln L + 2k$ , where  $k$  is the number of parameters estimated in the models. By this criterion, the most parsimonious model is the one with the smallest AIC score.

To estimate the effect of a major gene, we used genotypic probability estimators (GPEs)<sup>16,17</sup>, as implemented in PAP. The genotypic probability ( $p_{ij}$ , the probability that individual  $i$  carries genotype  $j$ ), equals the likelihood conditioned on individual  $i$  carrying genotype  $j$ , divided by the unconditional likelihood and computed with the parameters set at the maximum-likelihood estimates (MLEs). The mean of a trait  $Y$  for genotype  $j$  is calculated as  $\mu_j = \sum P_{ij} / n_j$ , where  $y_i$  is the IgE measured on person  $i$  and  $n_j = \sum P_{ij}$ . A  $t$ -test was used to determine statistical significance.

Variance-component linkage analysis was used to estimate the proportion of the variance attributable to residual genetic effects, random environmental effects, and quantitative trait loci (QTL). By fitting various models, it is possible to make inferences regarding the localization (the chromosomal regions mapped by the genetic markers, i.e. linkage) and the magni-

tude of effect sizes of major genes. Analyses were performed using the computer program package Sequential Oligogenic Linkage Analysis Routines (SOLAR),<sup>18</sup> which uses the computer program FISHER and SEARCH<sup>19</sup> for likelihood optimization in quantitative-trait analysis. For model fitting, the fixed effect of the covariates sex and age were removed by simultaneously including them in the models. In the most basic model, the expected covariance matrix for a pedigree is written as:

$$\Omega = 2\phi\sigma_g^2 + I\sigma_e^2$$

where  $\phi$  is the kinship matrix, and  $I$  is an identity matrix. To test for evidence of major genes (QTLs), the component of QTL is introduced in the model and the expected covariance matrix for a pedigree is written as:

$$\Omega = \Pi\sigma_a^2 + 2\phi\sigma_g^2 + I\sigma_e^2$$

where  $\Pi$  is the matrix whose elements ( $\Pi\sigma_m^2\pi_{jl}$ ) provides the predicted proportion of a gene that individual  $j$  and  $l$  share identity by descent (IBD) at a QTL linked to a genetic marker locus. Marker-specific IBD matrixes ( $\Pi_m$ ) were generated independently for all the 344 markers across the genome. Multipoint IBD matrixes were then generated at 1-cM resolution by incorporating the IBD matrixes at all the neighboring markers and mapping distances between these markers. To test for linkage, the likelihoods of the two models (with the variance due to the QTL estimated or fixed to zero) is compared. Twice the difference in loge likelihood of these two models yields a test statistics that is asymptotically distributed as a  $\frac{1}{2}:\frac{1}{2}$  mixture of a  $\chi_1^2$  variable and a point mass at zero because the estimated variance due to QTL was fixed to a boundary in the nested model.<sup>20</sup> The difference between the two  $\log_{10}$  likelihoods produces a LOD score that is equivalent to the classical LOD score of linkage. Tests for linkage and its effect are repeated throughout the genome.

**Table 1** Subject characteristics of Dutch families with asthma

| Characteristic                         | Proband (n=200) | Spouses (n=200) | First-degree offspring (n=530) |
|--|-----------------|-----------------|--------------------------------|
| M:F                                    | 124:76          | 76:124          | 237:293                        |
| Mean age (years) [range]               | 52.1 [37-76]    | 51.1 [33-76]    | 24.0 [6-53]                    |
| PC <sub>20</sub> <32 mg/ml (%) [n]     | 88.2 [170]      | 26.1 [199]      | 46.5 [525]                     |
| FEV <sub>1</sub> predicted (%)         | 69.7            | 98.4            | 93.6                           |
| FEV <sub>1</sub> < 80% predicted (%)   | 61.4            | 9.0             | 12.1                           |
| Reversibility <sup>a</sup> (baseline)  | 77.7            | 21.5            | 31.8                           |
| Reversibility <sup>a</sup> (predicted) | 62.9            | 19.0            | 25.8                           |
| Mean IgE (IU/ml)                       | 92.9            | 26.3            | 64.1                           |
| Positive skin test (%)                 | 81.9            | 30.6            | 54.1                           |

<sup>a</sup> Change > 9%

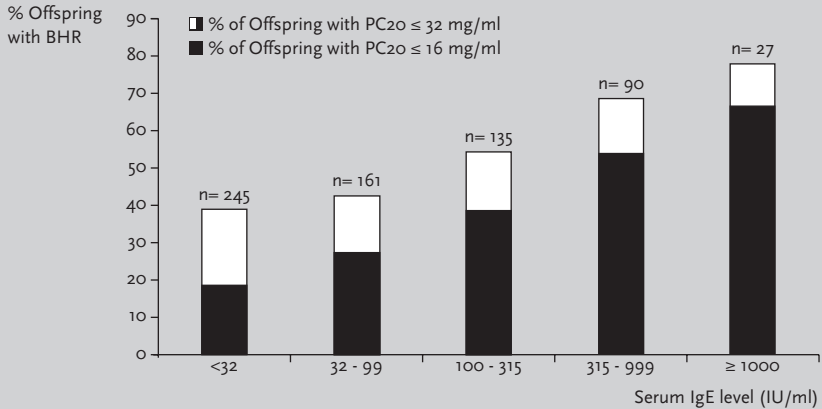
## RESULTS

### Characteristics of patients

Demographic and clinical characteristics for the family members are shown in table 1. There were more male than female probands; the mean age for the probands was 52 years (51 years of age for spouses and 24 years for children). Although all probands were hyperresponsive at the time of original testing, 12% were not hyperresponsive at current testing (30 were not retested due to low lung function, e.g.  $FEV_1 < 40\%$  predicted). A large proportion of probands (82 %) were skin test positive.

Individuals with BHR had significantly higher mean IgE than did individuals without BHR (83 vs. 35 IU/ml,  $t=9.22$ ,  $p < 0.0001$ ). The relationship between total serum IgE levels and the presence and degree of BHR is shown in figure 1. In addition, skin-test-positive individuals had significantly higher IgE levels than individuals with a negative skin test (112 vs. 26 IU/ml,  $t= 17.39$ ,  $p < 0.0001$ ). Results from multiple-regression analysis, where  $\log_{10}(\text{IgE})$  was the dependent variable and BHR, skin test responsiveness, sex, and age (year) were independent variables, also suggested that IgE was independently associated with these clinical and demographic variables. The estimated regression coefficients were  $\beta = 0.20$  ( $p = 0.0001$ ) for BHR,  $\beta = 0.59$  ( $p = 0.0001$ ) for skin test responsiveness,  $\beta = -0.07$  ( $p = 0.04$ ) for gender (male), and  $\beta = -0.006$  ( $p = 0.0001$ ) for age.

**Figure 1** Relationship between total serum IgE levels and bronchial hyperresponsiveness, in the offspring of 200 asthmatic parents





### Correlation coefficients of various relative pairs

Correlation coefficients ( $r$ ) for various pairs of relatives, with equal weights to all pairs, are reported in table 2. The results with equal weights to pedigrees and to nuclear pedigrees were similar (data not shown). There was evidence for familial aggregation of high total serum IgE levels, which was probably the result of a genetic component. The results should be interpreted with caution, since these families were ascertained through an asthmatic parent, rather than being selected from the general population. There was a higher correlation among parent-offspring and sibling pairs, with  $r = 0.24$  and  $r = 0.31$ , respectively, than among more distantly related relative pairs ( $r = 0.12$  for grandparent-grandchildren pairs and  $r$  approached zero for avuncular and first cousin pairs). As expected for unrelated individuals, there was no evidence of correlation between the spouses ( $r = -0.06$ ). The heritability ( $h^2$ ) was estimated to be 0.48 from parent-offspring pairs and 0.62 from sibling pairs.

**Table 2.** Correlation Coefficient of log (IgE) among different relative classes (equal weights to pairs)

| <u>Relationship</u>         | <u>number of Pairs</u> | <u>Correlation Coefficient</u> |
|-----------------------------|------------------------|--------------------------------|
| <i>1st degree relatives</i> |                        |                                |
| Parent-offspring:           |                        |                                |
| Mother-daughter             | 374                    | 0,28                           |
| Mother-son                  | 315                    | 0,20                           |
| Father-daughter             | 363                    | 0,25                           |
| father-son                  | <u>306</u>             | 0,23                           |
| Overall                     | <u>1358</u>            | 0,24                           |
| Sibling:                    |                        |                                |
| Sister-sister               | 195                    | 0,28                           |
| Sister-brother              | 342                    | 0,32                           |
| Brother-brother             | <u>151</u>             | 0,33                           |
| Overall                     | <u>688</u>             | 0,31                           |
| <i>2nd degree relatives</i> |                        |                                |
| Grandparent-grandchild      | 272                    | 0,12                           |
| Avuncular                   | 559                    | 0,02                           |
| <i>3rd degree relatives</i> |                        |                                |
| 1st Cousins                 | 397                    | 0,03                           |
| Unrelated                   |                        |                                |
| Spouses                     | 275                    | - 0,06                         |

### One-locus segregation analyses

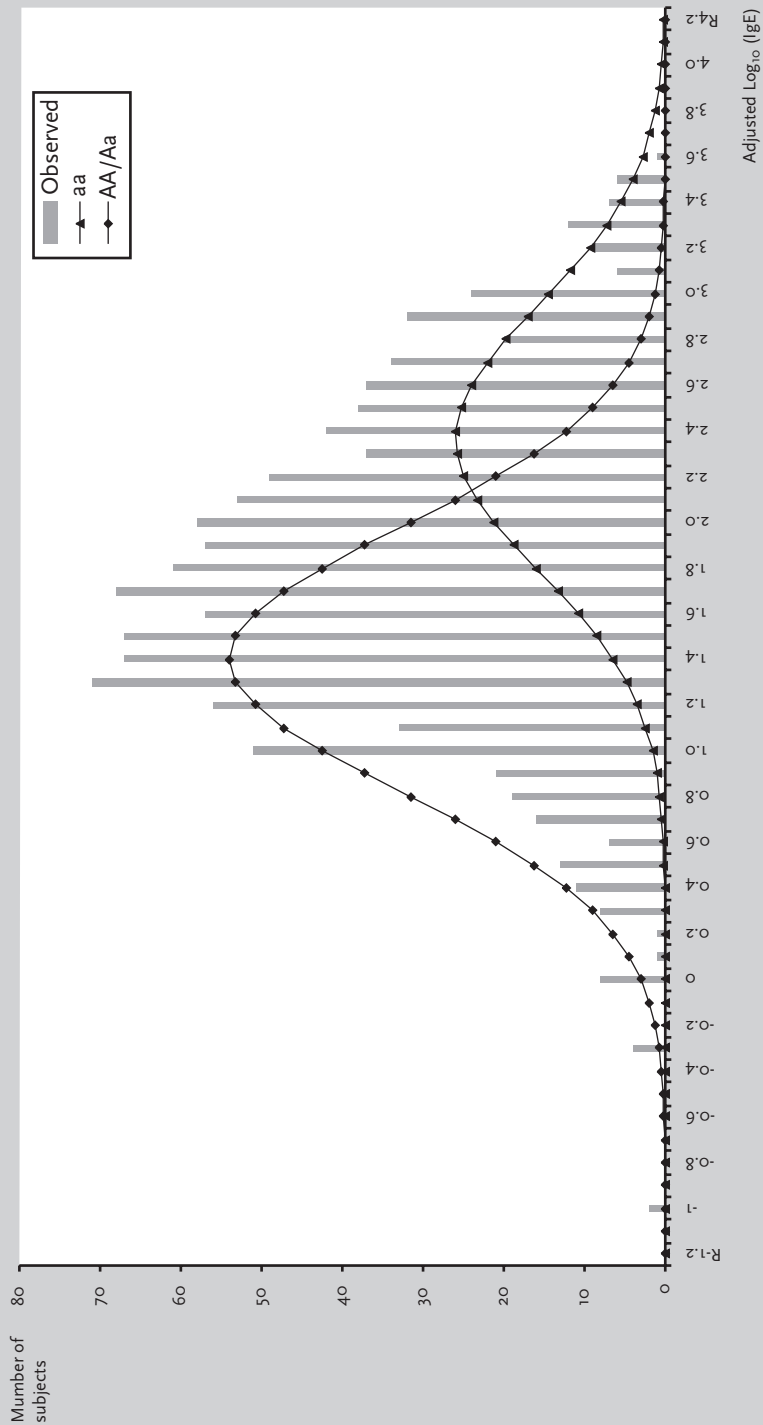
To test whether the familial aggregation of high total IgE levels is due to shared genes or shared environments, complex segregation analyses were performed. In the one-locus segregation analysis, both the sporadic model and mixed environmental model were rejected (table 3). Although Mendelian-only models were rejected, several mixed Mendelian models were not. These models fit the observed distribution of adjusted  $\log_{10}(\text{IgE})$  levels and were not significantly different from the general model ( $p > 0.05$ ). Of the mixed Mendelian models, the mixed recessive model (a

**Table 3.** Segregation Analysis of Adjusted  $\log_{10}(\text{lgE})$  in 200 Dutch Families

| Model                    | $q_A$ | Transmission Probability |      |      |      | Mean |      |     | SD   | $h^2$   | -2lnL   | AIC    | chisq | df       | P-value |
|--------------------------|-------|--------------------------|------|------|------|------|------|-----|------|---------|---------|--------|-------|----------|---------|
|                          |       | AA                       | Aa   | aa   | AA   | Aa   | aa   |     |      |         |         |        |       |          |         |
| General                  | 0,54  | [1]                      | 0,47 | [0]  | 1,63 | 1,41 | 2,34 | .58 | 0,57 | 2356,52 | 2374,52 | 9,43   | 2     | 0,009    |         |
| Mix. Env. (Env. + Poly.) | 0,55  | 0,47                     | = AA | = AA | 1,37 | 1,65 | 2,15 | .65 | 0,65 | 2365,95 | 2379,95 | 0,17   | 3     | 0,98     |         |
| Mix. Cod. (Cod. + Poly.) | 0,55  | [1]                      | [5]  | [0]  | 1,63 | 1,41 | 2,34 | .58 | 0,57 | 2356,69 | 2368,69 | 9,38   | 4     | 0,052    |         |
| Mix. Dom. (Dom. + Poly.) | 0,21  | [1]                      | [5]  | [0]  | 1,49 | 2,14 | = Aa | .63 | 0,47 | 2365,90 | 2375,90 | 1,90   | 4     | 0,754    |         |
| Mix. Rec. (Rec. + Poly.) | 0,57  | [1]                      | [5]  | [0]  | 1,45 | = AA | 2,32 | .58 | 0,49 | 2358,42 | 2368,42 | 25,89  | 4     | 0,00003  |         |
| Cod. Only                | 0,54  | [1]                      | [5]  | [0]  | 1,02 | 1,64 | 2,42 | .5  | [0]  | 2372,41 | 2382,41 | 50,63  | 5     | 1,02E-10 |         |
| Rec. only                | 0,66  | [1]                      | [5]  | [0]  | 1,35 | = AA | 2,22 | .55 | [0]  | 2399,15 | 2407,15 | 10,03  | 6     | 0,04     |         |
| Polygenic                | [1]   | n/a                      | n/a  | n/a  | 1,73 | = AA | = AA | .71 | 0,55 | 2366,55 | 2372,55 | 130,52 | 7     | 0        |         |
| Sporadic                 | [1]   | n/a                      | n/a  | n/a  | 1,74 | = AA | = AA | .7  | [0]  | 2487,04 | 2491,04 |        |       |          |         |

Note: Env = Environment, Poly = Polygene, Cod = Codominant, Dom = Dominant, Rec = Recessive.

**Figure 2.** Distributions of adjusted  $\log_{10}(\text{IgE})$  in 200 Dutch families.



Vertical bar represents the distribution of the observed adjusted  $\log_{10}(\text{IgE})$ .

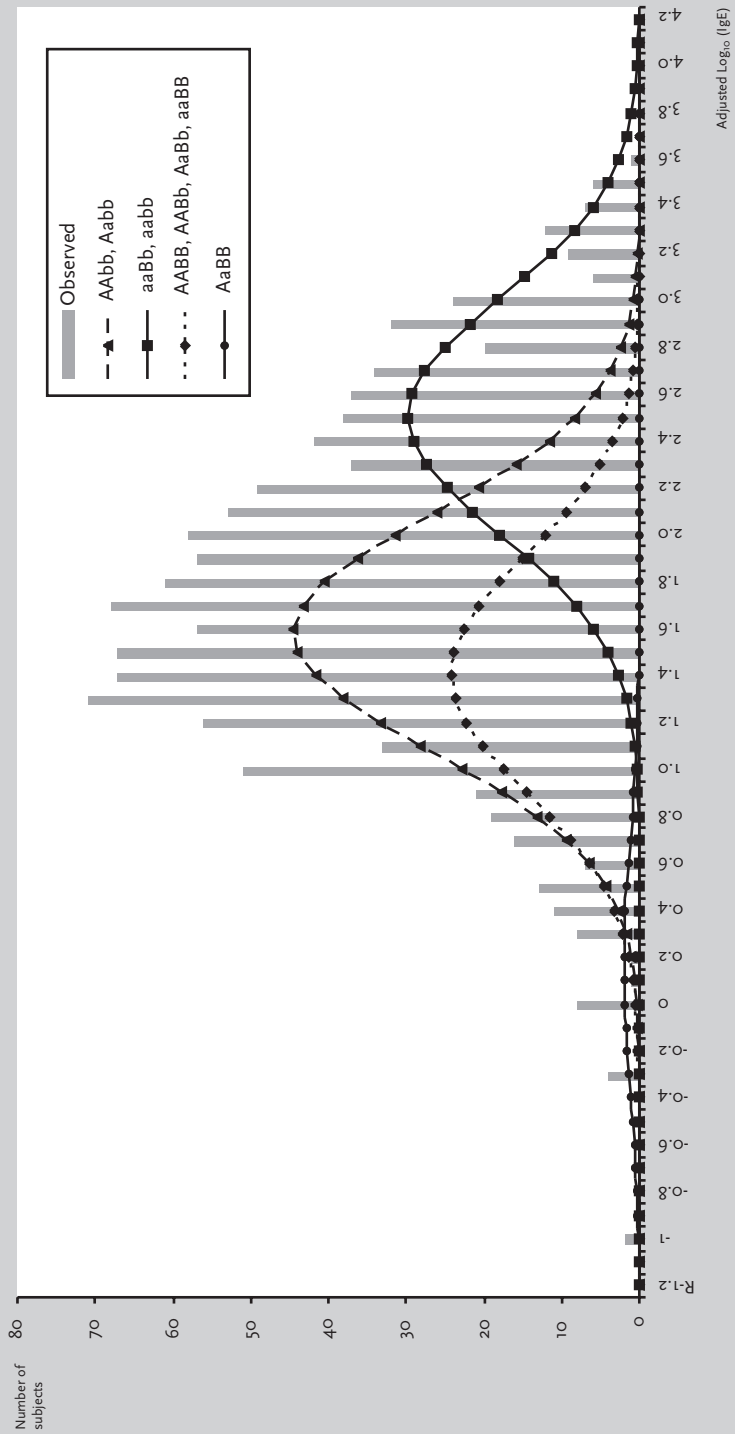
The two line curves represent the distributions of the three genotypes under the best fitting mixed recessive model.

**Table 4.** Segregation analysis of adjusted  $\log_{10}(\text{IgE})$  assuming one- and two-locus models in 200 Dutch Families

|   | Mean for each genotype |       |         |      |      |         |      |      |         |         |           |     |
|---|------------------------|-------|---------|------|------|---------|------|------|---------|---------|-----------|-----|
|   | $q_a$                  | $q_b$ | Locus B |      |      | Locus A |      |      | SD      | $h^2$   | $-2\ln L$ | AIC |
|   |                        |       | BB      | Bb   | bb   | AA      | Aa   | aa   |         |         |           |     |
| One-locus                                 |                        |       |         |      |      |         |      |      |         |         |           |     |
| Mixed codominant                          | 0,55                   |       |         | 1,63 | 1,41 | 2,34    | 0,58 | 0,57 | 2356,69 | 2368,69 |           |     |
| Mixed recessive                           | 0,57                   |       |         | 1,45 | = AA | 2,32    | 0,58 | 0,49 | 2358,42 | 2368,42 |           |     |
| Two-locus                                 |                        |       |         |      |      |         |      |      |         |         |           |     |
| Two major gene model 1                    | 0,57                   | 0,83  |         |      |      |         | 0,51 | 0,68 | 2329,19 | 2355,19 |           |     |
| locus A codominant                        |                        |       | BB      | 1,51 | 0,03 | 1,43    |      |      |         |         |           |     |
| locus B codominant                        |                        |       | Bb      | 1,58 | 1,44 | 2,59    |      |      |         |         |           |     |
|   |                        |       | bb      | 1,53 | 1,37 | 2,33    |      |      |         |         |           |     |
| Two major gene model 2                    | 0,55                   | 0,8   |         |      |      |         | 0,5  | 0,52 | 2331,25 | 2347,25 |           |     |
| locus A recessive                         |                        |       | BB      | 1,36 | 0,14 | 1,36    |      |      |         |         |           |     |
| locus B dominant                          |                        |       | Bb      | 1,36 | 1,36 | 2,45    |      |      |         |         |           |     |
|   |                        |       | bb      | 1,53 | 1,53 | 2,45    |      |      |         |         |           |     |
| Two major environmental risk factor model | 0,58                   | 0,83  |         |      |      |         | 0,53 | 0,9  | 2349,98 | 2371,98 |           |     |
|   |                        |       | BB      | 1,37 | 0,98 | 1,37    |      |      |         |         |           |     |
|   |                        |       | Bb      | 1,37 | 1,37 | 2,57    |      |      |         |         |           |     |
|   |                        |       | bb      | 1,97 | 1,37 | 2,22    |      |      |         |         |           |     |

major recessive gene and residual genetic effects) was the best fitting ( $p = 0.75$ ) and the most parsimonious (smallest AIC) model, although it was only slightly better than the mixed additive model. The MLE of the gene frequency  $q_a$  (corresponding to a high IgE level) under the mixed recessive model was 0.57, which results in genotype frequencies of 0.68 for AA/Aa, and 0.32 for genotype aa. When we used GPEs, this recessive gene had a large effect on IgE levels and was responsible for 32.4% of the adjusted  $\log_{10}(\text{IgE})$  levels in these families. The mean IgE level for genotype aa ( $\mu_{aa}$ ) was estimated to be 209 IU/ml, significantly higher than the mean (29

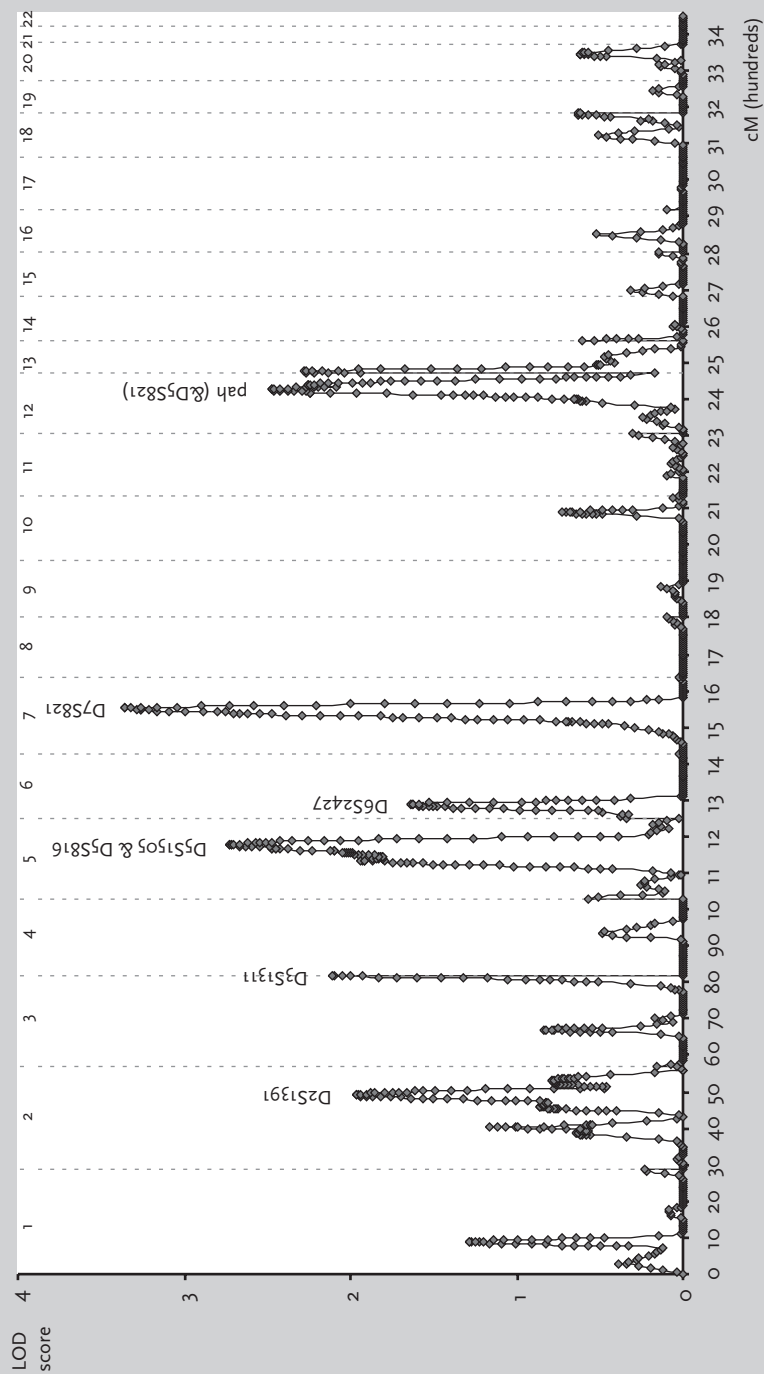
**Figure 3.** Distributions of adjusted  $\log_{10}(\text{IgE})$  in 200 Dutch families.



Vertical bar represents the distribution of the observed adjusted  $\log_{10}(\text{IgE})$ .

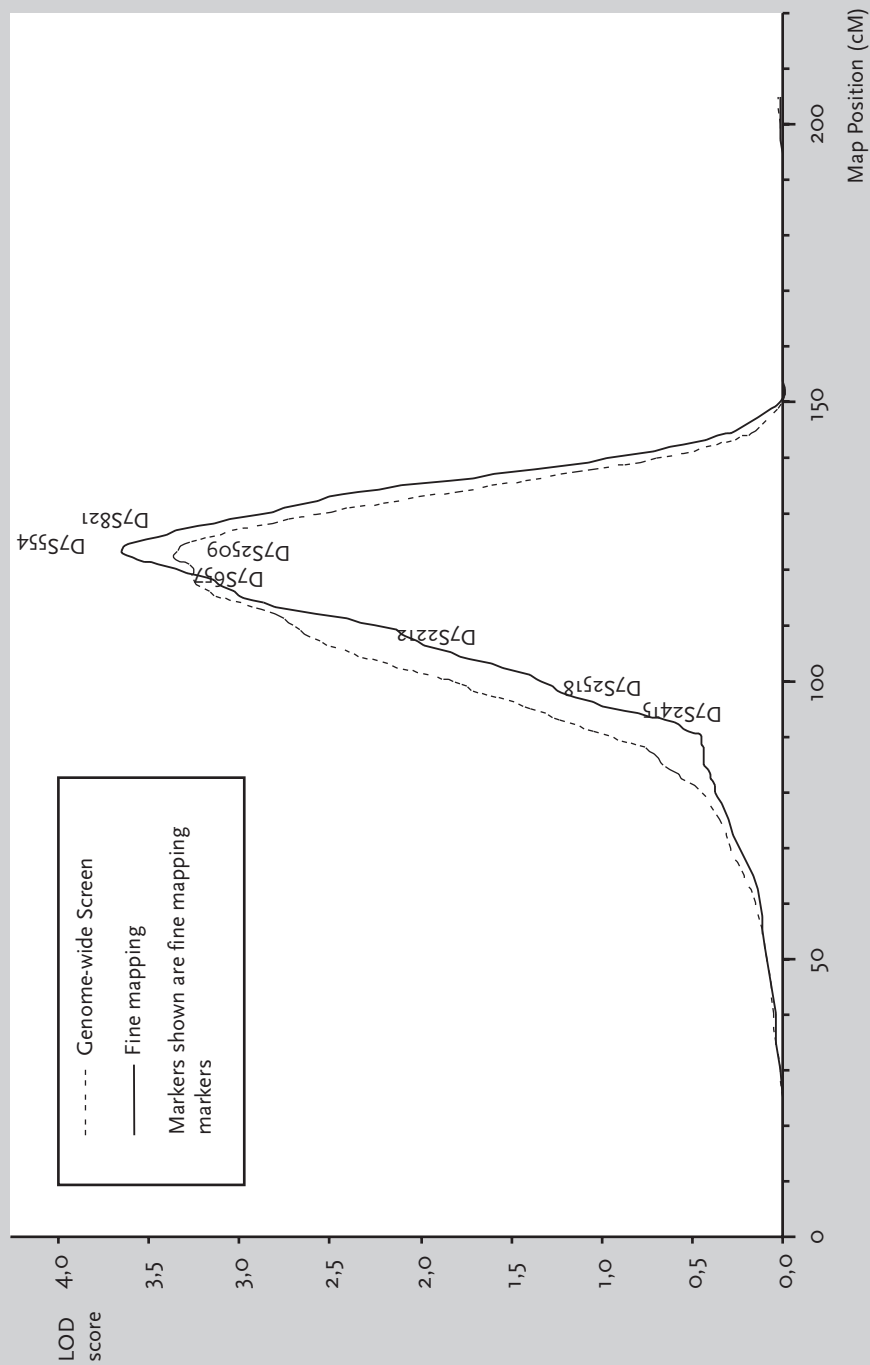
The four line curves represent the distributions of the nine genotypes under the best fitting mixed two major gene model.

**Figure 4.** Genome-wide screen for major genes regulating adjusted  $\log_{10}(\text{IgE})$  levels in 200 Dutch families using variance component analysis.



Three hundred and forty-four evenly spaced autosomal markers were genotyped. Vertical dotted lines divided the genome into 22 chromosomes.

**Figure 5.** Multipoint linkage analysis of  $\text{Log}_{10}(\text{lgE})$  on chromosome 7 using variance component approach



IU/ml) for the other genotypes ( $\mu_{AA}$ ,  $\mu_{Aa}$ ) (t-statistic = 23.89,  $p = 6.2E-10^9$ ). This model is plotted in figure 2, illustrating the fit of the model. The residual polygenic component was large, which was evidenced by a high heritability estimate (MLE of  $h^2 = 0.49$ , 95% CI: 0.35-0.63) and a significant improvement in model fitting compared with a recessive only model ( $\chi^2 = 40.73$ ,  $df = 1$ ,  $p = 1.7 \times 10^{-6}$ ).

#### Two locus segregation analysis

When two-locus segregation analysis was performed, the mixed two-major-gene models (two major genes and residual genetic effects) fit the data significantly better than did the one-major-gene models (table 4). Of the several mixed two-major-gene models that were tested, a model with a major recessive gene and a dominant modifier gene was the most parsimonious model. This model (two-major gene model 2) fit the data significantly better than did the one-major-recessive-gene model ( $\chi^2 = 27.17$ ,  $df = 3$ ,  $p = 5.42 \times 10^{-6}$ ). It also had a much lower AIC value (2347.25) than did the two-major-environmental-risk-factor model (AIC = 2371.98). According to this model, the MLE, gene frequency for the first recessive gene ( $q_a$ ) was 0.55, similar to the gene frequency of the mixed recessive model under one-locus segregation analysis. The MLE of the gene frequency for the second dominant gene ( $q_b$ ) was 0.8. The second gene modifies the effect of the first recessive gene so that some of the individuals who were homozygous for the high-risk allele at the first gene did not have high IgE levels. Individuals with genotype aaBB (homozygous high-risk allele at the first gene and homozygous low-risk allele at the second locus) had normal IgE levels (mean = 23 IU/ml) and only individuals with genotypes aaBb and aabb (29.6% of the total sample) had high IgE levels (mean = 282 IU/ml) (figure 3). The two major genes were responsible for 51.3% the adjusted  $\log_{10}(\text{IgE})$  variance. The first gene was responsible for 40.6% and the second gene was responsible for 9.0% of the adjusted  $\log_{10}(\text{IgE})$  variance.

#### Linkage analysis using variance component analysis

Linkage analysis, using the variance components approach, was performed to systematically scan the genome for the locations of the major genes regulating total serum IgE levels. Results of the multipoint linkage analysis for the 344 evenly spaced autosomal markers are presented in figure 4. One chromosomal region with a lod score of 3.36 ( $p = 0.00004$ ) was observed at 7q and flanked by markers D7S820 and D7S821. This locus explained 39% of adjusted  $\log_{10}(\text{IgE})$  variance and the residual genetic effect explained 17% of the variance. The evidence for linkage at this region was further supported by two additional analyses. First, the lod score was strengthened by genotyping more markers in the regions. After 7 markers were added to the 30 cM regions, the lod score increased to 3.65 ( $p = 0.00002$ ) at the same region (figure 5). Second, both the first 92 families and the second 108 families provided a positive lod score at the same region. The lod score at this region was 2.48 ( $p = 0.0003$ ) for the first 92 families and 1.88 ( $p = 0.002$ ) for the second 108 families.



To examine the impact that outliers had on the results, the linkage data were reanalyzed for all the chromosome 7 markers after deletion of the six outliers with the highest  $\log_{10}(\text{IgE})$  values. The lod score changed minimally from 3.65 to 3.74, suggesting that the results were stable and not driven by the few outliers.

Evidence for linkage at two other regions was also observed. A peak lod score of 2.73 ( $p = 0.0002$ ) was found at 5q31, flanked by markers D5S666 and D5S402, a result that was consistent with previous findings.<sup>9</sup> This locus explained 37% of adjusted  $\log_{10}(\text{IgE})$  variance. A peak lod score of 2.46 ( $p = 0.0004$ ) was found at 12q, flanked by markers PAH and D12S2070. There were several other regions with lod score of  $\sim 2$ . Some of these regions (3 qter and 13pter) were at the tip of the chromosomes and were supported mainly by one marker. The significance and the interpretation of these regions were uncertain and genotyping more markers in these regions is necessary. Other regions (2q, and 6p) were supported by multiple markers.

## Discussion

---

The family-ascertainment scheme of the current study was appropriate for performing both segregation analysis, and variance component analysis for total serum IgE levels. In this study all the probands were recruited from a well-defined population sample in the northeastern Netherlands, where little immigration occurs. The probands were recruited because of prior diagnosis of asthma 25-35 years ago and were evaluated using a standardized protocol.<sup>10</sup> All the children and grandchildren ( $\geq 8$  years old) of the probands were studied, regardless of their phenotypic status.

Although many studies have demonstrated heritability of serum IgE levels, the precise mode of genetic control has remained elusive. There have been multiple complex segregation analyses investigating the mode of inheritance of total serum IgE within families. Several previous studies have found evidence for a recessive gene regulating high IgE levels, with different estimates of gene frequencies and mean IgE levels.<sup>9,21,22,23</sup> Evidence for a codominant mode of inheritance was reported by Martinez et al.<sup>24</sup> suggesting that homozygotes and heterozygotes have different mean levels although there is clearly overlap between the distributions of total serum IgE levels.

The segregation analysis of the current study represented an extension of the segregation analyses in the first 92 families.<sup>8,9</sup> In the initial study, evidence was found for a mixed recessive gene (a recessive gene and residual genetic effect) and a mixed model with two major genes (two major genes and residual genetic effect) regulating total IgE levels, for a one-locus model and two-locus model, respectively. After including an additional 108 families, we obtained increased evidence supporting these models. Under a one-locus segregation analysis in the total sample of 200 families, the mixed recessive model fit the data as well as the general model ( $\chi_3^2 = 1.90$ ,  $p = 0.59$ ). This could be compared with the relatively poor fit of the mixed

recessive model for the first 92 families ( $\chi_3^2 = 6.90$ ,  $p = 0.07$ ).<sup>9</sup> In the total sample, the mixed two-major-gene model fit the data significantly better than did the mixed one-major gene model ( $\chi_3^2 = 27.17$ ,  $p = 5.4 \times 10^{-6}$ ), whereas, in the first 92 families, there was a marginal improvement of the best two-major gene model over one-major gene model, ( $\chi_5^2 = 11.9$ ,  $p = 0.04$ ).

For our previous linkage analyses, we had used a candidate gene approach and had found evidence for linkage to chromosome 5q in the first 92 families.<sup>8,9</sup> We have now completed a genomewide search approach in 200 Dutch families with asthma, using a new analytic method appropriate for investigation of quantitative traits. A region on chromosome 7q reached the criteria for a genomewide significant linkage and provided the strongest evidence for linkage, with a peak lod score of 3.65,  $p = 0.00002$ .<sup>25</sup> Linkage-analysis results also identified two regions with evidence for linkage at 5q and 12q, both regions have been reported elsewhere and are rich with appropriate candidate genes for immunologic and allergic responses. There were two linkage reports that systematically searched across the genome for the loci regulating total serum IgE. In a linkage study of 364 subjects in 80 nuclear families subselected from a population sample of 230 families in Busselton in Western Australia, Daniels et al.<sup>26</sup> used 20cM intervals for their genome screen and, after genotyping 274 autosomal markers, identified four regions that are likely to contain the genes regulating total serum IgE levels; the four regions were 6p, 7p, 11q13, and 16q2. In another genomewide screen linkage study, carried out in Germany in a smaller sample (97 families), Wjst et al.<sup>27</sup> reported four regions (2p, 6p, 9q2, and 12q) linked to loci regulating total serum IgE.

The possible explanations for the varying results from the various studies include the differences in study populations, ascertainment schemes, sample sizes, and analytical methodology. The assumptions required for some of the analytical methods oversimplify the complexity of the trait and are probably inappropriate for modeling the genetic regulation of total serum IgE levels. This is especially true for segregation analysis where most of the approaches assume only one major factor (a gene). Obviously, this is not realistic for a common trait such as high total serum IgE levels. Alternatively, segregation analysis assuming two major genes, although probably still too simple a model, may significantly improve our ability to model the trait. The effects of two genes and their interactions can be explored under the two-locus model.

There have been difficulties in linkage analysis of quantitative traits in humans. Parametric linkage analysis of quantitative traits has been rarely used, because it requires specifying a genetic model to describe the mode of inheritance, which is usually uncertain. Furthermore, when the SD within the genotype is too large and the distribution of the trait overlaps among the genotypes, the parametric approach is usually uninformative. In some of the earlier studies, non-parametric analysis was performed using the phenotypic information of sib-pairs only, either because other relatives

were not characterized or the all relative pair approach was not yet available for quantitative traits. Recently, an alternative quantitative trait linkage analysis, a variance-component method, was developed.<sup>28-31</sup> The variance-component method allows for marker-specific effects, residual additive genetic effects, and random environmental effects. The variance component approach for a linkage study of a quantitative trait has a number of compelling features.<sup>32,33</sup> It uses all of the available inheritance information in a pedigree of any size or structure. This not only avoids the violation of true bivariate structure of a sib-pair, which can occur in analyses of the phenotypic sib-pair difference, but also uses the available information more efficiently and therefore can achieve greater power to detect linkage. Second, the variance-component approach resolves the problem of independence of sib-pairs within a family by maximizing the likelihood of a pedigree that is jointly conditional on all members of the pedigree.<sup>34</sup> The violation of independence of sib-pairs has been a significant problem, because different weighting schemes have tended to produce strikingly different results.<sup>35</sup>

The results of the present study can be summarized as follows. First, the results confirm the strong association between total serum IgE levels and bronchial hyperresponsiveness and allergy. Second, as observed in previous studies, there is strong aggregation of total serum IgE levels within families. Third, the segregation analysis provides evidence that major genes with residual genetic effects are responsible for the aggregation of total serum IgE levels. The presence of at least two major genes, one behaving as a recessive gene and another behaving as a dominant modifier gene, is consistent with the observed distribution of IgE levels in these families. Finally, a genomewide search using variance-component approaches identified several regions that are likely to contain the major genes regulating total serum IgE levels. Regions on chromosomes 5q, 12q, and 6p have been reported elsewhere.<sup>8,9,21,26,27,36,37</sup> The novel region on chromosome 7q was confirmed by typing additional markers and needs to be replicated in other populations.

#### **Acknowledgements**

---

We thank all of the families whose participation made this project possible. The research was supported in part by Netherlands Asthma Foundation and by National Institutes of Health (NIH) grant R01HL/48341. Some of the results of this paper were obtained by using the program package S.A.G.E., which is supported by a U.S. Public Health Service Resource Grant (1 P41 RR03655) from the National Center for Research Resources.

#### **Electronic-Database Information**

---

Accession number and URLs for data in this article are as follow:  
Online Mendelian Inheritance in Man (OMIM),

<http://www.ncbi.nlm.nih.gov/Omim> (for IgE [MIM 147061] and asthma [MIM 600807])

LINKAGE: <ftp://linkage.rockefeller.edu/software/linkage>

PAP: Pedigree Analysis Package,

<ftp://ftp.genetics.utah.edu/pub/software/pap>

S.A.G.E. Statistical Analysis for Genetic Epidemiology,

<http://darwin.cwru.edu/pub/sage.html>

SOLAR: Sequential Oligogenic Linkage Analysis Routines,

<http://www.sfbr.org/sfbr/public/software/solar/index.html>

## References

- 1 Johansson SG, Bennich HH, Berg T. The clinical significance of IgE. *Prog Clin Immunol* 1972; 1:157-181.
- 2 Burrows B, Martinez FD, Halonen M, Barbee R, Cline MG. Association of asthma with serum IgE levels and skin-test reactivity to allergens. *N Engl J Med* 1989; 320(5):271-277.
- 3 Sears MR, Burrows B, Flannery EM, Herbison GP, Hewitt CJ, Holdaway MD. Relation between airway responsiveness and serum IgE in children with asthma and in apparently normal children. *N Engl J Med* 1991; 325(15):1067-1071.
- 4 Halonen M, Stern D, Taussig LM, Wright A, Ray CG, Martinez FD. The predictive relationship between serum IgE levels at birth and subsequent incidences of lower respiratory illnesses and eczema in infants. *Am Rev Respir Dis* 1992; 146(4):866-870.
- 5 Hopp RJ, Bewtra AK, Nair NM, Townley RG. Specificity and sensitivity of methacholine inhalation challenge in normal and asthmatic children. *J Allergy Clin Immunol* 1984; 74(2):154-158.
- 6 Burrows B, Lebowitz MD, Barbee RA, Cline MG. Findings before diagnoses of asthma among the elderly in a longitudinal study of a general population sample. *J Allergy Clin Immunol* 1991; 88(6):870-877.
- 7 Wiesch DG, Meyers DA, Bleecker ER. Genetics of asthma. *J Allergy Clin Immunol* 1999; 104(5):895-901.
- 8 Xu J, Levitt RC, Panhuysen CI, Postma DS, Taylor EW, Amelung PJ et al. Evidence for two unlinked loci regulating total serum IgE levels. *Am J Hum Genet* 1995; 57(2):425-430.
- 9 Meyers DA, Postma DS, Panhuysen CI, Xu J, Amelung PJ, Levitt RC et al. Evidence for a locus regulating total serum IgE levels mapping to chromosome 5. *Genomics* 1994; 23(2):464-470.
- 10 Panhuysen CIM, Bleecker ER, Koëter GH, Meyers DA, Postma DS. Characterization of obstructive airways disease in family members of probands with asthma; an algorithm for the diagnosis of asthma. *Am J Respir Crit Care Med* 1998; 157:1734-1742.
- 11 de Vries K, Goei JT, Booy-Noord H, Orie NGM. Changes during 24 hours in the lung function and histamine hyperreactivity of the bronchial tree in asthmatic and bronchitic patients. *Int Arch All* 1962; 20:93-101.
- 12 Van Camp G, Balemans W, Wellems PJ. Linkage designer and linkage reporter software for automated gene localization studies. *Trends Genet* 1997; 13:82.
- 13 Lander ES, Green P. Construction of multilocus genetic linkage maps in humans. *Proc Natl Acad Sci U S A* 1987; 84(8):2363-2367.
- 14 Hasstedt SJ. A mixed-model likelihood approximation on large pedigrees. *Comput Biomed Res* 1982; 15(3):295-307.
- 15 Akaike H. A new look at the statistical model identification. *IEEE Trans Automatic Control* 1974; 19:719-723.

- 16 Elston RC, Stewart J. A general model for the genetic analysis of pedigree data. *Hum Hered* 1971; 21(6):523-542.
- 17 Hasstedt SJ, Moll PP. Estimation of genetic model parameters: variables correlated with a quantitative phenotype exhibiting major locus inheritance. *Genet Epidemiol* 1989; 6(2):319-332.
- 18 Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 1998; 62(5):1198-1211.
- 19 Lange K, Weeks D, Boehnke M. Programs for Pedigree Analysis: MENDEL, FISHER, and dGENE. *Genet Epidemiol* 1988; 5(6):471-472.
- 20 Self SG, Kiang KY. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under non-standard conditions. *J Am Stat Assoc* 1987; 82:605-610.
- 21 Marsh DG, Bias WB, Ishizaka K. Genetic control of basal serum immunoglobulin E level and its effect on specific reagent sensitivity. *Proc Natl Acad Sci U S A* 1974; 71(9):3588-3592.
- 22 Gerrard JW, Rao DC, Morton NE. A genetic study of immunoglobulin E. *Am J Hum Genet* 1978; 30:46-58.
- 23 Meyers DA, Beaty TH, Freidhoff LR, Marsh DG. Inheritance of total serum IgE (basal levels) in man. *Am J Hum Genet* 1987; 41(1):51-62.
- 24 Martinez FD, Holberg CJ, Halonen M, Morgan WJ, Wright AL, Taussig LM. Evidence for Mendelian inheritance of serum IgE levels in Hispanic and non-Hispanic white families. *Am J Hum Genet* 1994; 55(3):555-565.
- 25 Lander ES, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; 11:241-247.
- 26 Daniels SE, Bhattacharya S, James A, Leaves NI, Young A, Hill MR et al. A genome-wide search for quantitative trait loci underlying asthma. *Nature* 1996; 383(6597):247-250.
- 27 Wjst M, Fischer G, Immervoll T, Jung M, Saar K, Rueschendorf F et al. A genome-wide search for linkage to asthma. *Genomics* 1999; 58:1-8.
- 28 Goldgar DE. Multipoint analysis of human quantitative genetic variation. *Am J Hum Genet* 1990; 47(6):957-967.
- 29 Schork NJ. Extended multipoint identity-by-descent analysis of human quantitative traits: efficiency, power, and modeling considerations. *Am J Hum Genet* 1993; 53(6):1306-1319.
- 30 Amos CI. Robust variance-components approach for assessing genetic linkage in pedigrees. *Am J Hum Genet* 1994; 54(3):535-543.
- 31 Blangero J, Almasy L. Multipoint oligogenic linkage analysis of quantitative traits. *Genet Epidemiol* 1997; 14(6):959-964.
- 32 Williams JT, Blangero J. Comparison of variance components and sibpair-based approaches to quantitative trait linkage analysis in unselected samples. *Genet Epidemiol* 1999; 16(2):113-134.
- 33 Williams JT, Duggirala R, Blangero J. Statistical properties of a variance components method for quantitative trait linkage analysis in nuclear families and extended pedigrees. *Genet Epidemiol* 1997; 14(6):1065-1070.
- 34 Amos CI, Zhu DK, Boerwinkle E. Assessing genetic linkage and association with robust components of variance approaches. *Ann Hum Genet* 1996; 60 (Pt 2):143-160.

- 35 Sham PC, Zhao JH, Curtis D. Optimal weighting scheme for affected sib-pair analysis of sibship data. *Ann Hum Genet* 1997; 61 ( Pt 1):61-69.
- 36 Marsh DG, Neely JD, Breazeale DR, Ghosh B, Freidhoff LR, Ehrlich-Kautzky E et al. Linkage analysis of IL4 and other chromosome 5q31.1 markers and total serum immunoglobulin E concentrations. *Science* 1994; 264(5162):1152-1156.
- 37 Barnes KC, Neely JD, Duffy DL, Freidhoff LR, Breazeale DR, Schou C et al. Linkage of asthma and total serum IgE concentration to markers on chromosome 12q: evidence from Afro-Caribbean and Caucasian populations. *Genomics* 1996; 37(1):41-50.





