Factor analysis of asthma and atopy traits shows 2 major components, one of which is linked to markers on chromosome 5q

Catharine J. Holberg, PhD, Marilyn Halonen, PhD, Susan Solomon, BA, Penelope E. Graves, ScD, Mauro Baldini, MD, Robert P. Erickson, MD, and Fernando D. Martinez, MD *Tucson, Ariz*

Background: A variety of definitions of asthma and atopy traits have been used in genetic studies. The variables used may be correlated, increasing the likelihood of type I error.
Objective: We sought to clarify and quantify phenotypes that may be characterized by related traits. Principal components and factor analysis were applied to the correlation matrix of asthma and atopy traits before linkage analysis.
Methods: Factor analysis was performed on 468 Hispanic and non-Hispanic white children enrolled in the Tucson Children's Respiratory Study, with complete information on 24 items, including skin test response to 7 allergens, total serum IgE levels, presence or absence of asthma attacks, wheezing episodes, hay fever, and cough. Factor score coefficients were then applied to all siblings (n = 877), and quantitative factor scores were derived. Single-point and multipoint nonparametric sib-pair analyses were performed to assess linkage to markers on chromosome 5q31-33. Analyses were also performed for individual items.
Results: Two main factors were identified: Factor I had high loadings on atopic items, including skin test responses, IgE, and hay fever, and Factor II had high loadings that included asthma diagnosis, wheezing, cough, and *Alternaria* species skin test response. Factors I and II were correlated at an *r* value of 0.19. For the quantitative factor scores, significant single-point linkage ($P < .0001$) was demonstrated only for atopic Factor I, and a peak multipoint LOD score of 2.7 was seen for marker D5S479. Multipoint LOD scores for individual items were 1.1 or less.
Conclusion: These analyses suggest evidence for a locus or loci mapping to chromosome 5q31-33 associated with this composite atopic phenotype. (J Allergy Clin Immunol 2001;108:772-80.)

Key words: *Asthma, atopy, genetics, factor analysis, linkage, chromosome 5q*

Abbreviations used
FE: Familial eosinophilia
HE: Haseman-Elston
ML: Maximum likelihood

Successful gene identification relies on basic epidemiologic concepts, such as case identification and accuracy of diagnosis. The mechanisms responsible for the inheritance of asthma, a complex phenotype,[1] are not well understood, despite many reports of linkage of asthma phenotypes to markers mainly on chromosomes 11 and 5.[2-4]

In spite of advances in our knowledge of the genetics and pathogenesis of asthma, a comprehensive definition of asthma has been difficult to achieve. Recent statements have expanded on past definitions, with the addition of certain cellular functions.[5,6] The majority of current definitions and guidelines for asthma diagnosis characterize it as a reversible narrowing of the conducting airways of the lung, occurring either slowly or rapidly,[7] with, as a rule, evidence of airway wall inflammation and eosinophilia.[8] However, in spite of these guidelines, a variety of definitions of asthma have been used in epidemiologic and genetic studies, including subject reports of self or doctor diagnosis, record reviews, and pulmonary function test results.[9] Each method may select different population subsets.[10]

Given the characterization of asthma as a complex phenotype, there has been much emphasis on studying the genetics of phenotypes showing a strong association with asthma, which may also cluster within families. The supposition is that such intermediate phenotypes or, more specifically, the genes regulating their expression may play a crucial role in the pathogenesis of asthma.[11] Thus each intermediate or subphenotype represents a separate specific phenotype or phenotypic marker necessary but not sufficient for the development of asthma.[12] For example, phenotypic markers that have been used or suggested include bronchial hyperresponsiveness, increased total serum IgE levels, specific IgE antibody responsiveness to a single allergen complex, and skin test reactivity.[12-15]

Such intermediate phenotypes, however, are likely to be correlated with one another; for example, skin test response to a specific allergen may be correlated with

From the Arizona Respiratory Center and the Department of Pediatrics, University of Arizona Health Sciences Center, Tucson.
Supported in part by grant No. RO1 HL-56177 from the National Heart, Lung, and Blood Institute. Dr Martinez was also supported by a Research Development Award for Minority Faculty (HL-03154-01). The program package SAGE used in this study is supported by a US Public Health Service Resource Grant (1 P41 RR03655) from the National Center for Research Resources.
Received for publication February 28, 2001; revised July 27, 2001; accepted for publication July 31, 2001.
Reprint requests: Catharine J. Holberg, PhD, Arizona Respiratory Center, PO Box 245030, University of Arizona Health Sciences Center, Tucson, AZ 85724.
Copyright © 2001 by Mosby, Inc.
0091-6749/2001 $35.00 + 0 **1/81/119158**
doi:10.1067/mai.2001.119158

responses to other allergens. Furthermore, testing many correlated variables increases the likelihood of type I error. However, it is possible that sets of such intermediate phenotypic markers may uniquely characterize a complex phenotype more precisely than a diagnostic definition, thereby identifying groups of individuals possessing similar asthmatic or atopic profiles.

The statistical technique of factor analysis provides a method of identifying a relatively small number of factors or factor constructs that can be used to represent the relationships between a large number of variables. These factors, once identified, can then be quantified with a score. Principal components analysis is an initial step in factor analysis. Such methods have been used typically in psychology, educational studies, and the social sciences, where variables are often less precisely defined and the nature of the relations is not clear.[16] The use of factors can also help reduce the type I error by avoiding the use of the Bonferroni correction for multiple statistical tests on related variables.

In this study we apply the statistical techniques of principal components and factor analysis to a set of variable items used in epidemiologic analyses to define asthma or asthma-like symptoms or that have shown associations with these conditions. Factors or composite phenotypes comprising correlated phenotypic markers are identified and quantified by factor scores. Linkage of these composite phenotypes and their component items to genotypic markers on chromosome 5q is assessed by using single-point and multipoint linkage analysis. Chromosome 5q31-33 has long been a focus of interest for assessing linkage of asthma- and atopy-related traits because of the large number of genes in the region that regulate functions known to be associated with atopy.

TABLE I. Two-factor solution and factor loadings after oblique rotation*

Factor item	Factor I: atopy	Factor II: asthma
Bermuda grass ST	**0.80**	–0.01
Careless weed ST	**0.73**	0.05
Mesquite ST	**0.69**	0.01
Olive ST	**0.67**	0.07
Total serum IgE	**0.64**	0.06
Mulberry ST	**0.61**	0.13
Alternaria ST	**0.54**	**0.19**
House dust ST†	**0.34**	**0.17**
MD rhinitis	**0.23**	**0.31**
Mother past MD asthma	0.13	0.06
Child past MD asthma	0.10	–0.00
Father current MD asthma	0.10	0.03
Male sex	0.09	0.14
Exercise wheeze occasionally	0.08	**0.57**
Eczema	0.07	**0.17**
Exercise wheeze usually	–0.06	**0.27**
Child current MD asthma	0.05	**0.72**
Father non-Hispanic white	0.04	–0.05
Mother non-Hispanic white	0.04	–0.03
Current cough‡	0.02	**0.63**
Mother current MD asthma	0.02	**0.22**
Father past MD asthma	0.02	0.11
Past exposure to ETS	–0.02	–0.03
Current exposure to ETS	0.01	–0.01
Current wheeze	0.01	**0.64**

ST, Skin test; *MD*, physician-diagnosed; *ETS*, environmental tobacco smoke.
*Items with loadings of 0.15 or greater are in boldface; current and past asthma for mother, father, and child were entered into the correlation matrix as sets of dummy variables.
†Positive responses to *Dermatophagoides farinae* and cat dander were combined with positive responses to house dust mix.
‡Responses of 3 or greater on a scale of 1 to 5.

METHODS

The study population was derived from families participating in the Tucson Children's Respiratory Study, a longitudinal study investigating the risk factors for asthma and other acute and chronic illnesses in infancy and childhood.[17] Families were enrolled at random with respect to asthma or other respiratory diseases. Only families in which the biologic parents declared they were either non-Hispanic white or Hispanic were included in this analysis. These are the predominant ethnic groups in the Tucson area. This study was approved by the Human Subjects Committee of the University of Arizona, and parents signed consent forms at all phases of the study.

Blood was drawn for in-depth evaluation, including determination of serum IgE levels, when the mean age of the index children and siblings was 6 and 11 years, respectively. A further in-depth evaluation was performed at 11 years for index children. Log_{10} serum IgE levels were age and sex adjusted by using regression procedures, as previously described.[15] At each in-depth evaluation, individuals were skin tested with extracts of allergens common in the Tucson area (Table I; Hollister-Stier Laboratories, Everett, Wash). Children were considered to have a positive reaction to an allergen if they had a 3-mm or larger wheal reaction to that allergen at either evaluation.

Questionnaires administered at the first in-depth evaluation, completed by parents for themselves and on behalf of their children, provided information on respiratory symptoms and diagnoses, as well as on parental smoking habits. Questionnaires at enrollment provided information on demographic variables.

Blood was drawn for genetic studies from all family members who consented at the time of the second in-depth evaluation. Genomic DNA was extracted from these samples following conventional methods, as previously described by Martinez et al.[18] Genotyping was performed for 11 PCR-based, highly polymorphic, dinucleotide repeat, anonymous markers, spanning approximately 28 cM in chromosome 5q31-33, namely D5S642, D5S393, D5S479, D5S399, D5S414, D5S500, D5S658, D5S436, D5S210, D5S519, and D5S209. Distances between markers (7, 2, 1, 1, 1, 1, 5, 0.5, 7, and 3 cM, respectively) were obtained from published maps or from our own data, as described in former analyses.[18] The genotyping methodology has also been previously described by Martinez et al.[18]

Factor analysis

Briefly, factor analysis is a data-reduction procedure that uses the information from the entire correlation coefficient matrix of a large number of variables to produce a smaller number of hypothetical factor constructs or components that help explain the correlations between the variables. Factors are initially extracted from the correlation matrix until there is no appreciable variance left. Coefficients, weights, or loadings roughly represent the correlation between the variable item and the factor. The factors can be characterized as linear combinations of the original variables, and quantitative factor scores can be estimated.

The Kaiser-Meyer-Olkin (Kaiser, 1974) measure of sampling adequacy and the Bartlett test of sphericity (Norusis, 1994) were calculated as an indication of the appropriateness of the correlation matrix for factor analysis.

Factor analysis was first performed on 468 unrelated Hispanic and non-Hispanic white children for whom we had complete information on a number of variables (Table I). These items were selected as those that have been used in epidemiologic analyses to define asthma, asthma-like symptoms, or atopy or that have shown associations with these conditions.[19-28] The use of unrelated children avoided biases in the correlation matrix because of family size. Resulting factor-score coefficients were then applied to the 468 children and their 409 siblings, and quantitative factor scores were derived for a total of 877 children. Sibling correlations were calculated for the factor scores and for the factor items by using the program FCOR within the Statistical Analysis for Genetic Epidemiology (SAGE) software.[29] Contingency tables were used to characterize the factor scores, grouped in tertiles, versus some of the variable items.

Linkage analysis

Linkage analyses were performed in a group of selected families, with preference given to those with 3 or more children (n = 104 families, with a mean of 3.5 children per family) who had been genotyped for the marker set on chromosome 5q31-33. A total of 288 children from 100 of these families of the original 877 children with available quantitative factor scores were included. Single-point nonparametric sib-pair analyses were performed with the SAGE program SIBPAL.[29] We applied the robust Haseman-Elston (HE) regression method,[30] in which the squared sib-pair difference in the trait of interest is regressed on the estimated proportion of alleles shared identical by descent, and a *t* test for negative slope is performed. For randomly selected pedigrees, this method performs well in providing evidence for linkage when compared with maximum likelihood methods based on knowledge of the correct genetic model.[31]

Single-point linkage analyses were performed with all available sib pairs for each quantitative factor score, respectively, and the 11 markers genotyped on chromosome 5q31-33. Furthermore, single-point analyses were performed by using the extreme tertiles of each factor score, forming a dichotomous variable representing affected and unaffected cases for each factor construct. For comparison, these analyses were also performed for variable items used in the factor analyses.

For our final evaluation of linkage, we used the HE and maximum likelihood (ML) multipoint techniques for quantitative traits found in the software MAPMAKER/Sibs.[32] The ratio of the likelihood of the data under linkage to that under the random Mendelian segregation alternative is calculated for each sib pair. The sum of the logarithms of these likelihood ratios for all sib pairs gives the LOD score (traditionally the log of the odds favoring genetic linkage).[33] The LOD score probability can be calculated by converting it to a *z* score statistic equal to $\sqrt{(\text{LOD} \times 4.6)}$. So doing will confirm that an LOD score of 3.0 corresponds to a *P* value of .0001. There is considerable debate over the appropriate threshold to use to avoid many false-positive claims while not missing a suggestion of linkage.[32]

Multipoint linkage analyses were performed for the composite scores of Factor I (atopy) and Factor II (asthma), respectively, as well as individual items that were significant in the HE regression. It has been shown that only sibling pairs concordant for high or low values or discordant sibling pairs provide substantial power to detect linkage.[34] Therefore we also conducted some multipoint linkage analyses for only those individuals in the upper and lower tertiles of the quantitative scores, representing affected and unaffected cases, respectively. Hence these analyses included sib pairs concordant for high or low values and discordant pairs.

RESULTS

There were 468 unrelated children with data available for factor analysis. Quantitative factor scores were derived for these children, plus 409 siblings with complete data available. Of these 877 children, 288 from 100 families had genotypic data for linkage analyses.

The correlation matrix of items was first subjected to a principal components extraction. A scree test, performed by plotting the Eigen values of each extraction in descending order, suggested a distinct break at 4 factors accounting for 39% of the variance. The Kaiser-Meyer-Olkin measure of sampling adequacy for the correlation matrix was 0.81, and the Bartlett test of sphericity was significant at a *P* value of .00001 or less, indicating that a factor analysis of this correlation matrix would be very appropriate. Several solutions were examined in terms of their adequacy in accounting for the data and approach to simple structure (each factor with high loadings on different variables). After principal axis factoring and orthogonal rotation, a 2-factor solution was found to represent the data, with Factors III and IV being represented only by ethnicity and parental smoking, respectively. The loadings for these 2 factors after principal factor extraction and oblique rotation are given in Table I. This rotation is considered more realistic because the theoretically important underlying dimensions are not assumed to be unrelated to one another.[35] The correlation (ρ) between the two factors was 0.19. Factor I, accounting for 16.4% of the variance, was named "atopy" because of high positive loadings on skin prick test items, IgE, and rhinitis. Factor II, accounting for 6.2% of the variance, was named "asthma" because of high positive loadings on a diagnosis of current asthma, eczema, and a variety of respiratory symptoms, including cough, wheeze, and exercise-induced wheezing, as well as rhinitis. Only rhinitis, house dust, and *Alternaria* species skin tests had loadings greater than 0.15 on both factors. Interestingly, we have previously shown that a positive response to *Alternaria* species is associated with asthma in the Tucson area.[36] Quantitative factor scores were derived for Factor I and Factor II, with high scores representing an increased atopy or asthma profile.

Table II shows sibling correlations for the factor scores and a number of factor items. Twice the sibling correlation provides an estimate of the heritability and represents an upper limit.[37] The heritability estimate is highest for total serum IgE level (75%), with Factor I a close second (74%).

Part of our linkage strategy used only the upper and lower tertiles of the quantitative scores derived from Factors I and II, with the upper tertile representing a more atopic or asthmatic profile for Factors I and II, respectively. Contingency table analyses using the tertiles and the factor items in the 877 children with the data necessary to compute the factor scores helped characterize the factors more clearly. Of the children in the upper tertile of Factor I, 15% had physician-diagnosed current asthma, comprising 55% of the total number of patients with physician-diagnosed current asthma in all 3 tertiles (*P* =

.21, with binomial distribution). In contrast, all of the patients with physician-diagnosed current asthma fell into the upper tertile of Factor II ($P < .0001$, with binomial distribution), making up 26% of the upper tertile. In addition, the mean total serum IgE level (z score) for individuals in the upper tertile of Factor I was 0.90 versus 0.20 and –1.0 for those in the mid and lower tertiles, respectively, with 54% of the variance in IgE accounted for. In contrast, for Factor II, equivalent values were 0.27 versus –0.05 and –0.20, with 3.2% of the variance in IgE accounted for (both $P < .001$).

Of the 877 children with the data necessary to compute quantitative factor scores, 288 belonged to the group of 100 genotyped nuclear families. These 288 children yielded 316 sib pairs, with a total of 144 for Factor I in the upper and lower tertiles (55, 49, and 40 classified as concordant affected, concordant unaffected, and discordant, respectively) and a total of 123 for Factor II (52, 29, and 42 classified as concordant affected, concordant unaffected, and discordant, respectively).

Table III shows the results of the HE linear regression analysis, for Factor I entered both as grouped upper and lower tertiles (ie, affected and unaffected) and as a continuous variable. The negative slope was significant for a number of markers for the upper and lower tertiles, with the highest level of significance (lowest $P = .0001$) for marker D5S479. For the continuous score, once again the lowest P value (.00007) was for marker D5S479. For comparison with the composite score, Table IV shows every significant negative slope from HE linear regression analysis for the individual factor items. None achieve the level of significance found for Factor I, including total serum IgE, which has a heritability level similar to that of Factor I. Furthermore, the lowest P value is seen for mulberry skin test ($P \leq .01$), which is at the same position but over 100 times higher than the lowest P value for Factor I ($P = .00007$). No significant slopes were identified for Factor II, the asthma profile.

For multipoint analysis for Factors I and II using the entire range of each quantitative score, the highest LOD score of 1.4, corresponding to a z score of 2.5 ($P = .006$), was seen for Factor I, peaking at marker D5S479. Factor II showed no evidence of linkage, with the highest LOD score of 0.04 at D5S414 equivalent to a z score of 0.43 ($P = .33$). Fig 1 shows the results of the multipoint linkage analyses for items that were significant in the HE single-point test for linkage. The highest LOD score of 0.5, associated with a P value of .06, is seen for mulberry skin testing, peaking at marker D5S479. All other LOD scores are less than 0.4, which is clearly not significant. Restricting the mulberry skin test analysis to only those sibling pairs in the upper and lower tertiles of Factor I increased the peak LOD score at D5S479 to 1.13 ($P = .01$). Similarly, restricting the continuous serum IgE analysis to only those sib pairs in the upper and lower tertile of the IgE distribution increased the peak LOD score for that analysis marginally to 0.43 ($P = .08$) at D5S399.

Fig 2 shows the results of multipoint analysis with the continuous score for Factor I and restricting analysis only to those in the upper and lower tertiles to assess linkage to the 11 markers studied in chromosome 5q. For orientation, the approximate positions of potential candidate gene loci are also shown.[38,39] A clear peak is seen in the area corresponding to marker D5S479, with an LOD score of 2.18 for the HE method and 2.7 for the ML method. The significance levels correspond to normal distribution z scores of 3.2 ($P = .0007$) for HE and 3.5 ($P = .0002$) for ML. These z scores correspond to very significant single-test P values (ie, at a specific locus). In a genome-wide context the threshold for significance would be a z score of 4.1, an LOD score of 3.6 or greater, and a single-test P value of .00002 to avoid the chance encounter of a false-positive result at 5%.[32] Results were entirely comparable when analyses were performed separately for those subjects with both non-Hispanic white or Hispanic parents, respectively. Peak LOD scores of 1.63 and 0.47 were noted at marker D5S467 for families who were non-Hispanic white (69 families) or Hispanic (24 families), respectively.

Finally, there are some closely adjacent markers that have t and P values that run counter to the overall trends (eg, markers D5S399 and D5S414 for the continuous score and markers D5S436 and D5S210 in the tertile analysis; Table III). For consistency, the distances between markers were those used in our former analyses.[18] More recent maps may order the markers and candidate genes slightly differently (www.ensembl.org), with the following caveats: the sequence data are not ideally assembled; not all contaminants have been removed; and gene predictions are not optimally refined by cDNA. Nonetheless, removing D5S414 and D5S436 from the multipoint analysis did not affect the position or magnitude of the peak LOD score.

TABLE II. Sibling correlations for factor scores and items (n = 316 pairs)*

Factor/item	Equal weight to pairs (ρ)	Equal weight to nuclear families (ρ)
Factor I: atopy	0.3115	0.3705
Factor II: asthma	0.1937	0.1744
Bermuda grass ST	0.2082	0.2342
Careless weed ST	0.2366	0.3438
Mesquite ST	0.2104	0.3033
Olive ST	0.1847	0.2497
Total serum IgE	0.3510	0.3745
Mulberry ST	0.1754	0.2412
Alternaria ST	0.1443	0.1711
House dust ST	0.0568	0.0919
MD rhinitis	0.3448	0.2994
Child past MD asthma	0.1508	0.0215
Exercise wheeze	0.1059	0.1014
Eczema	0.2383	0.2415
Exercise wheeze usually	0.1371	0.1126
Child current MD asthma	0.0591	0.0065
Cough	0.1053	0.1801
Current wheeze	0.0570	0.0798

ST, Skin test; *MD*, physician-diagnosed.
*Items measuring parental attributes had high sibling correlations. A sibling correlation for sex was not computed.

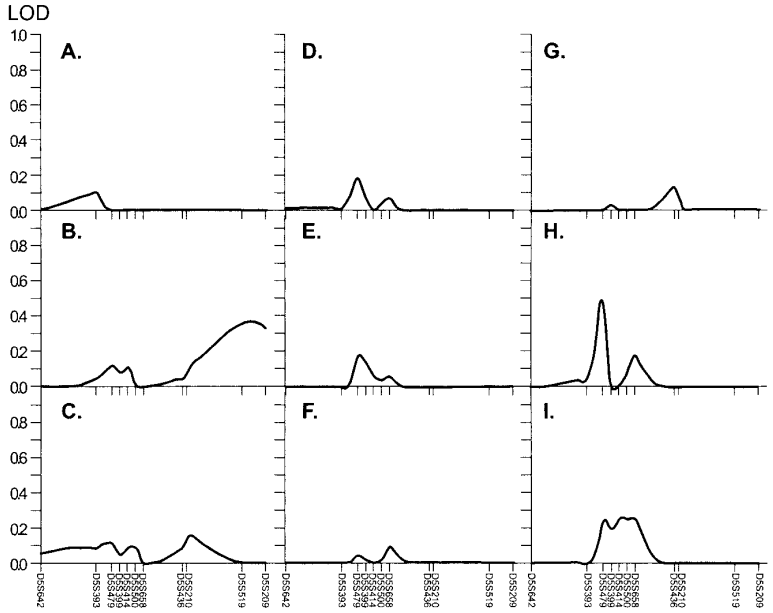


FIG 1. Multipoint HE sib-pair linkage analysis for factor items (phenotypes) significant in single-point HE analysis to markers on chromosome 5q31-33. **A**, Cough; **B**, current wheeze; **C**, exercise-induced wheeze; **D**, Bermuda grass skin test; **E**, careless weed skin test; **F**, mesquite skin test; **G**, house dust skin test; **H**, mulberry skin test; **I**, total serum IgE level.

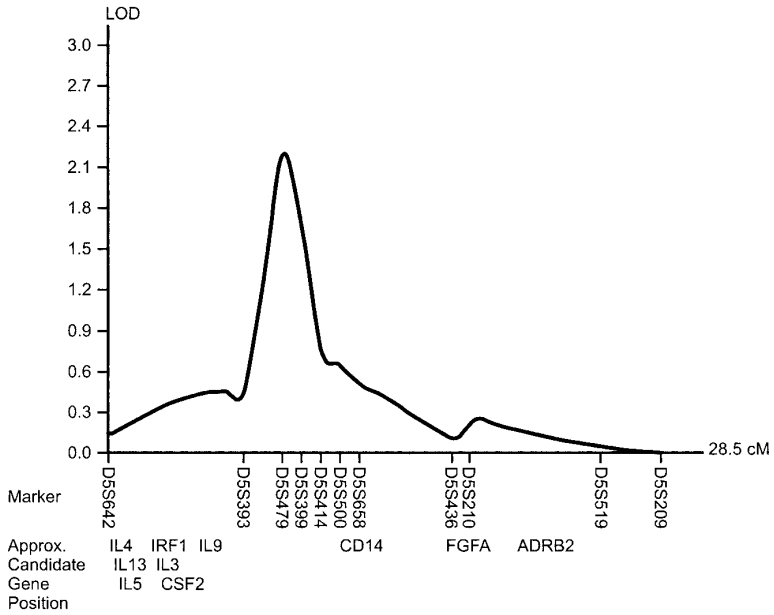


FIG 2. Multipoint HE sib-pair linkage analysis for atopy profile (Factor I) to markers on chromosome 5q31-33 using only those sib pairs in the upper and lower tertile of the quantitative score distribution. The approximate location of candidate genes is shown: *IRF1*, IFN regulatory factor 1; *CSF2*, colony-stimulating factor 2; *FGFA*, fibroblast growth factor acidic; *ADRB2*, β-adrenoreceptor.

DISCUSSION

The composite atopy-related phenotype obtained by means of factor analysis shows significant evidence for linkage with markers in chromosome 5q31-33, but the asthma phenotype derived as a second factor after removing the variance associated with the atopy factor showed no evidence for linkage with these same markers. These results suggest that variants in genes in chromosome 5q may determine atopy and, through this mechanism, increase susceptibility to asthma. However, there appear to be no variants that determine asthma independent of the

TABLE III. HE linear regression analysis for Factor I, discrete and continuous variables: Squared sib-pair trait difference versus estimated proportion of alleles shared identical by descent

	Factor I (upper and lower tertiles grouped)				Factor I (continuous score)			
Marker	**Effective *df***	**Slope**	***t* Value**	***P* value**	**Effective *df***	**Slope**	***t* Value**	***P* value**
D5S642	68	–0.0333	–0.2261	.4109	135	–0.4407	–1.1446	.1272
D5S393	71	–0.1455	–1.0997	.1376	136	–0.6267	–1.7317	.0428*
D5S479	72	–0.5450	–3.9032	.0001†	147	–1.4749	–3.9056	.00007†
D5S399	71	–0.4015	–2.9257	.0023†	144	–1.1515	–3.2115	.0008†
D5S414	67	–0.4116	–2.8023	.0033†	133	–0.2700	–0.7181	.2370
D5S500	73	–0.2922	–2.2250	.0146*	146	–0.4877	–1.3936	.0828
D5S658	79	–0.2841	–2.1409	.0177*	151	–0.8784	–2.3195	.0109*
D5S436	83	0.0722	0.5051	.6926	158	0.2038	0.5652	.7136
D5S210	86	–0.3437	–2.6597	.0047†	169	–0.4257	–1.3057	.0967
D5S519	79	–0.2538	–1.9679	.0263*	161	0.0712	0.2043	.5808
D5S209	79	–0.1894	–1.1791	.1209	155	–0.5854	–1.4726	.0714

*$P \leq 5\%$.
†$P \leq 1\%$.

TABLE IV. HE linear regression analysis: Squared sib-pair trait difference versus estimated proportion of alleles shared identical by descent—Factor items (phenotypes) showing significant slopes for markers*

Marker	**Slope**	***P* value**	**Phenotype**	**Slope**	***P* value**	**Phenotype**
D5S642						
D5S393	–0.1521	.0180†	Cough			
D5S479	–0.1999	.0202†	Careless weed	–0.2099	.0094‡	Mulberry
D5S399	–0.2240	.0118†	Bermuda			
D5S414	–0.5165	.0373†	IgE			
D5S500	–0.0711	.0313†	Exercise-induced wheezing			
D5S658	–0.1914	.0163†	Mesquite			
D5S436						
D5S210	–0.1393	.0252†	House dust			
D5S519	–0.1303	.0469†	Current wheezing	–0.0633	.0441†	Exercise-induced wheezing
D5S209						

*There were no significant findings for physician-diagnosed rhinitis, IgE upper and lower tertile categories, olive or *Alternaria* species skin prick tests, child current or past asthma, eczema, or occasional exercise-induced wheezing. There is no variability for items measuring parental attributes.
†$P \leq 5\%$.
‡$P \leq 1\%$.

atopic status in this region. This finding is in accordance with recent studies showing that although sensitization to allergens is associated with asthma, there is no relation between allergen exposure and the prevalence of asthma.[40] This suggests that the development of childhood asthma and the induction of specific IgE responses are determined by independent factors, implying that the genetic determinants of asthma and atopy differ. Thus although asthma is an atopic disease, the exact nature of the link between asthma and atopy remains to be determined.

The use of composite variables may help in analyses of complex phenotypes in which a number of different but related symptoms and signs can be combined through use of the correlations in the empirical data to produce fewer, less correlated dimensions that may prove to be more useful in genetic studies. Affected sibling pairs for Factor I would have high serum IgE levels, with or without asthma symptoms, whereas for Factor II, they would have asthma diagnoses or symptoms, with or without a high serum IgE level. In the case of the composite atopic profile, the evidence for linkage is considerably clearer and more statistically significant when compared with linkage analyses for its component items, which themselves are correlated with other variable items. The testing of many correlated variables increases the likelihood of type I error, which is avoided with the use of the quantitative composite score.

There are relatively few studies that have used the quantitative score approach in the study of the genetics of asthma and atopy. Lawrence et al[41] initially pioneered the use of principal components analysis and logistic regression to study familial correlations and segregation analyses of asthma- and atopy-related traits. Morton[42] has enumerated several advantages of the use of quantitative scores, including the applicability of quantitative trait locus theory, allowing for the adjustment for covariates and the reduction of multiple traits to a smaller number of scores, optimizing some relevant property like heritability. With the use of all the family data, Morton has suggested an approach maximizing heritability. In this context Wilkinson et al[43] have used principal components analysis to derive quantitative scores for asthma and atopy, athough IgE was selected as the best indicator of atopy on the basis of its high value for heritability.

Using this approach, they have found evidence of linkage for an asthma score to markers on chromosome 12.

We found that the heritability of the factor score was equivalent with that of some of the factor items, although for assessing linkage to markers on chromosome 5q, the use of the factor score gave clearer evidence of linkage. Interestingly, total serum IgE levels, which have a high loading on Factor I and a high heritability (75%) and have shown evidence of linkage in the 5q31-33 region in other studies,[44,45] showed little evidence of linkage in this region in our population. Our estimate of heritability is higher than many estimates, which range from 40% to 50%.[46-48] It is, however, more compatible with heritability estimates from twin studies of 61% for total serum IgE[49] and 80% for cord-blood IgE.[50] Nonetheless, only when total serum IgE level was included with other items as part of a complete atopic profile was there strong evidence for linkage in this region. The atopic profile represented by Factor I appears to provide a parsimonious summary of many related traits, with a high heritability of 74%.

Chromosome 5q31-33 has been a focus of interest for assessing linkage of asthma- and atopy-related traits because of the large number of genes in this region regulating functions known to be associated with atopy. Genes in this area include *IL4*, *IL5*, *IL9*, *IL13*, *GMCSF*, *IL12B*, and *ITK*, and the genes encoding both α- and β_2-adrenoceptors and the glucocorticosteroid receptor. The cytokines formed by these genes participate in proliferation, differentiation, and activation of hematopoetic cells and cells of the immune system. Activities associated with atopy and inflammatory processes include production of IgE (IL-4 and IL-13); eosinophil growth, maturation, differentiation, activation, and survival (IL-3, IL-5, IL-9, and GM-CSF); and inhibition of T-cell differentiation (IL-12).[51,52] A number of recent studies have indicated linkage to markers on chromosome 5q31-33 of asthma- and atopy-related phenotypes. Thus evidence indicates the presence of genes in the chromosome 5q cytokine cluster, encoding for IL-3, IL-4, IL-13, and IL-5, associated with total serum IgE,[44,45] bronchial hyperreactivity,[39] asthma and atopy traits,[53] and specific IgE response to American cockroach and dog.[54] Genes in the 5q31-33 region have also been shown to control the intensity of infection by *Schistosoma mansoni*,[55] suggesting they may play a general role in regulating IgE–mediated responses. A recent genome–wide search plus subsequent flanking markers to localize a gene for familial eosinophilia (FE) demonstrated significant evidence for an FE locus between the chromosome 5q31-33 markers D5S642 and D5S816.[56] In contrast, a study in a Finnish population found no evidence of linkage in this area to either total serum IgE levels or asthma.[57]

We have recently shown significant evidence for linkage among sib pairs concordant for low levels of circulating eosinophils by using the same set of markers as the present study, with a peak for marker D5S500.[18] This peak is very close to the peak found for Factor I at D5S479, although eosinophil levels were not included in the present analysis. Including eosinophil levels did not change our results; loadings for eosinophil levels were 0.32 for the atopy Factor I and 0.10 for the asthma Factor II.

Association studies in this area have been inconsistent and have not always provided evidence of strong allelic association. For example, ensuing studies by Rioux et al[56] to identify polymorphisms responsible for the FE linkage signal were unable to identify any functional association with sequence polymorphisms within 3 genes whose products are important players in the development and proliferation of eosinophils, namely *IL3*, IL5, and *GMCSF*. Association studies have shown that polymorphisms in the promoter region of *IL4*, located on chromosome 5q31-33 near marker D5S642, may be associated with the development of asthma in Japanese children, although no significant differences were noted in either total or specific IgE levels.[58] Another study found no clear evidence of allelic association between 11 polymorphic markers spanning the 5q31-33 region and the traits of total serum IgE and bronchial hyperreactivity.[59] In the Children's Respiratory Study population we have reported association between a polymorphism[60] in the β_2-adrenoreceptor and a positive response to bronchodilators[61] and association of a polymorphism in the 5′ flanking region of the *CD14* gene, between markers D5S658 and D5S436, with circulating soluble CD14 levels and with total serum IgE levels.[62] Others have reported a significant association between an allele at the *IL9* locus, centromeric to marker D5S393, and total serum IgE levels.[63] We have recently reported 7 polymorphisms (6 novel) in *IL13*, telomeric to *IL4*.[64] Four of these are tightly linked to a variant in the terminal portion of the coding region of the gene, which is strongly associated with high total serum IgE levels in 3 populations of children. In the present study there were no significant relations between the atopic or asthmatic profiles considered in tertiles of the factor score distributions, respectively, and the polymorphisms in the β_2-adrenoceptor, *CD14* or *IL13* genes (data not shown), which we have reported on previously,[61,62,64] indicating that these variants are not responsible for the linkage signal seen with the atopy phenotype in chromosome 5q31-33.

Clearly, identifying the gene or genes responsible for the linkage signal with the compound atopy phenotype may help in elucidating the complex interactions between atopic status and other airway-related alterations in determining the asthma phenotype. In addition, studies in other areas of the genome could provide valuable information on linkage sites for Factor II, the asthmatic profile, characterized by symptoms and diagnoses that intriguingly showed little indication of linkage in this area.

We thank M. A. Smith, RN, and L. L. De la Ossa, RN, for their work as study nurses and B. W. Saul, MS, for database management.

REFERENCES

1. Elston RC. Genetic analysis of complex phenotypes: family studies of total IgE levels and bronchial hyperreactivity. In: Marsh DG, Lockhart A, Holgate ST, editors. The genetics of asthma. Oxford (UK): Blackwell Scientific Publications; 1993. p. 143-51.

2. Sandford A, Weir T, Paré P. The genetics of asthma. Am J Respir Crit Care Med 1996;153:1749-65.
3. The Collaborative Study on the Genetics of Asthma (CGSA). A genome-wide search for asthma susceptibility loci in ethnically diverse populations. The Collaborative Study on the Genetics of Asthma (CGSA). Nat Genet 1997;15:389-92.
4. Boushey HA, Fahy JV. Basic mechanisms of asthma. Environ Health Perspect 1995;106(Suppl 6):229-33.
5. National Asthma Education and Prevention Program. Expert panel report 2: guidelines for the diagnosis and management of asthma. Bethesda (MD): National Institutes of Health; April 1997. Publication No. 97-4051.
6. National Institutes of Health. NHLBI/WHO Workshop Report. Bethesda (MD): National Heart Lung and Blood Institute; March 1993.
7. American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. Am Rev Respir Dis 1987;136:225-44.
8. Bousquet J, Chanez P, Lacoste YJ. Eosinophilic inflammation in asthma. N Engl J Med 1990;323:1033-9.
9. Gregg I. Epidemiologic aspects. Philadelphia: WB Saunders; 1983.
10. Samet JM. Epidemiologic approaches to the identification of asthma. Chest 1987;91(Suppl 6):74S-8S.
11. Martinez FD. Complexities of the genetics of asthma. Am Rev Respir Crit Care Med 1997;156(Suppl):S117-22.
12. Levitt RC, Mitzner W, Kleeberger SR. A genetic approach to the study of lung physiology: understanding biological variability in airways repsonsiveness. Am J Physiol 1990;258:L157-64.
13. Cookson WOCM, Sharp PA, Faux JA, Hopkin JM. Linkage between immunoglobulin E responses underlying asthma and rhinitis and chromosome 11q. Lancet 1989;1:1292-5.
14. Marsh DG. General discussion and summing-up. In: Marsh DG, Lockhart A, Holgate ST, editors. The genetics of asthma. Oxford (UK): Blackwell Scientific Publications; 1993. p. 321.
15. 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;55:555-65.
16. Bowden SC, Dodds B, Whelan G, et al. Confirmatory factor analysis of the Wechsler Memory Scale-revised in a sample of clients with alcohol dependency. J Clin Exp Neuropsychol 1997;19:755-62.
17. Taussig LM, Wright AL, Morgan WJ, Harrison HR, Ray CG. The Tucson Children's Respiratory Study. I. Design and implementation of a prospective study of acute and chronic respiratory illness in children. Am J Epidemiol 1989;129:1219-31.
18. Martinez FD, Solomon S, Holberg CJ, Graves PE, Baldini M, Erickson RP. Linkage of circulating eosinophils to markers in chromosome 5q. Am J Respir Crit Care Med 1998;158:1739-44.
19. Wright A, Holberg C, Martinez F, Halonen M, Morgan W, Taussig L. Epidemiology of physician-diagnosed allergic rhinitis in childhood. Pediatrics 1994;94:895-901.
20. Wright A, Holberg C, Morgan W, Taussig L, Halonen M, Martinez F. Recurrent cough in childhood and its relation to asthma. Am J Respir Crit Care Med 1996;153:1259-65.
21. Martinez F, Wright A, Taussig L, Holberg C, Halonen M, Morgan W. Asthma and wheezing in the first six years of life. N Engl J Med 1995;332:133-8.
22. Martinez F, Cline M, Burrows B. Increased incidence of asthma in children of smoking mothers. Pediatrics 1992;89:21-6.
23. Halonen M, Stern D, Taussig L, Wright A, Ray C, Martinez F. 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;146:866-70.
24. Burrows B, Martinez FD, Halonen M, Barbee RA, Cline MG. Association of asthma with serum IgE levels and skin test reactivity to allergens. N Engl J Med 1989;320:271-7.
25. Samet J. Respiratory Disease in a New Mexico population sample of Hispanic and non-Hispanic whites. Am Rev Respir Dis 1982;125:152-7.
26. Sherrill D, Stein R, Halonen M, Holberg C, Wright A, Martinez F. Total serum IgE and its association with asthma symptoms and allergic sensitization among children. J Allergy Clin Immunol 1999;104:28-36.
27. Stein R, Holberg C, Sherrill D, et al. Influence of parental smoking on respiratory symptoms during the first decade of life. Am J Epidemiol 1999;149:1030-7.
28. Dodge R. A comparison of the respiratory health of Mexican-American and non-Mexican-American white children. Chest 1983;84:587-92.
29. S.A.G.E. Statistical Analysis for Genetic Epidemiology. Computer Program. Cleveland (OH): Department of Epidemiology and Biostatistics, Rammelkamp Center for Education and Research, MetroHealth Campus, Case Western Reserve University; 1997.
30. Haseman JK, Elston RC. The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 1972;2:3-19.
31. Demenais F, Amos C. Power of the sib-pair and lod score methods for linkage analysis of quantitative traits. Prog Clin Biol Res 1989;329:201-6.
32. Kruglyak L, Lander ES. Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 1995;57:439-54.
33. Kruglyak L, Daly M, Reeve-Daly M, Lander E. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 1996;58:1347-63.
34. Risch N, Zhang H. Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science 1995;268:1584-9.
35. Kim J-O, Nie N, Hadlai Hull C, Jenkins J, Steinbrenner K, Bent D. Factor analysis. Statistical Package for the Social Sciences. 2nd ed. New York: McGraw Hill; 1975.
36. Halonen M, Stern DA, Wright AL, Taussig LM, Martinez FD. Alternaria as a major allergen for asthma in children raised in a desert environment. Am J Respir Crit Care Med 1997;155:1356-61.
37. Falconer DS, Mackay TFC. Introduction to quantitative genetics. 4th ed. Harlow, United Kingdom: Longman Group Ltd; 1996.
38. Garcia A, Marquet S, Bucheton B, et al. Linkage analysis of blood Plasmodium falciparum levels: interest of the 5q31-33 chromosome region. Am J Trop Med Hyg 1998;58:705-9.
39. Postma DS, Bleeker ER, Amelung PJ, et al. Genetic susceptibility to asthma—bronchial hyperresponsiveness coinherited with a major gene for atopy. N Engl J Med 1995;333:894-900.
40. Lau S, Illi S, Sommerfeld C, et al. Early exposure to house-dust mite and cat allergens and development of childhood asthma: a cohort study. Multicentre Allergy Study group. Lancet 2000;356:1392-7.
41. Lawrence S, Beasley R, Doull I, et al. Genetic analysis of atopy and asthma as quantitative traits and ordered polychotomies. Ann Hum Genet 1994;58(Suppl):359-68.
42. Morton NE. Quantitative scores for asthma and atopy. Clin Exp Allergy 1998;28(Suppl 1):95-7.
43. Wilkinson J, Grimley S, Collins A, Thomas NS, Holgate ST, Morton N. Linkage of asthma to markers on chromosome 12 in a sample of 240 families using quantitative phenotype scores. Genomics 1998;53:251-9.
44. Marsh DG, Neely JD, Breazeale DR, et al. Linkage analysis of IL4 and other chromosome 5q31.1 markers and total serum immunoglobulin E concentrations. Science 1994;264:1152-6.
45. Meyers DA, Postma DS, Panhuysen CIM, et al. Evidence for a locus regulating total serum IgE levels mapping to chromosome 5. Genomics 1994;23:464-70.
46. Blumenthal M, Namboordiri K, Mendell N, Gleich G, Elston R, Yunis E. Genetic transmission of serum IgE levels. Am J Med Genet 1981;10:219-28.
47. Gerrard J, Rao D, Morton N. A genetic study of IgE. Am J Hum Genet 1978;30:46-78.
48. Palmer L, Burton P, Faux J, James A, Musk A, Cookson W. Independent inheritance of serum immunolglobulin E concentrations and airway responsiveness. Am J Respir Crit Care Med 2000;161:1836-43.
49. Hopp R, Bewtra A, Watt G, Nair N, Townley R. Genetic analysis of allergic disease in twins. J Allergy Clin Immunol 1984;73:265-70.
50. Husby S, Holm N, Christenson K, Skov R, Morling N, Peterson P. Cord blood immunoglobulin E in like-sexed momzygotic and dizygotic twins. Clin Genet 1996;50:332-8.
51. Coyle AJ, Ackerman S, Burch R, Proud D, Irwin CG. Human eosinophil granule major basic protein and synthetic polycations induce airway hyperresponsiveness in vivo dependent on bradykinin generation. J Clin Invest 1995;95:1735-40.
52. Punnonen J, Yssel H, de Vries JE. The relative contribution of IL4 and IL13 to human IgE synthesis induced by activated CD4+ or CD8+ T cells. J Allergy Clin Immunol 1997;100:792-801.
53. Noguchi E, Shibasaki M, Arinami T, et al. Evidence for linkage between asthma/atopy in childhood and chromosome 5q31-33 in a Japanese population. Am J Respir Crit Care Med 1997;156:1390-3.
54. Hizawa N, Freidoff LR, Ehrlich E, et al. Genetic influences of chromosomes 5q31-33 and 11q13 on specific IgE responsiveness to common inhaled allergens among African American families. Collaborative Study on the Genetics of Asthma. J Allergy Clin Immunol 1998;102:449-53.

55. Marquet S, Abel L, Hillaire D, et al. Genetic localization of a locus controlling the intensity of infection by Schistosoma mansoni on chromosome 5q31-q33. Nat Genet 1996;14:181-4.
56. Rioux JD, Stone VA, Daly MJ, et al. Familial eosinophilia maps to the cytokine gene cluster on human chromosome region 5q31-33. Am J Hum Genet 1998;63:1086-94.
57. Laitinen T, Kauppi P, Ignatius J, et al. Genetic control of serum IgE levels and asthma: linkage and linkage disequilibrium studies in an isolated population. Hum Mol Genet 1997;6:2069-76.
58. Noguchi E, Shibasaki M, Arinami T, et al. Association of asthma and the interleukin-4 promoter gene in Japanese. Clin Exp Allergy 1998;28:449-53.
59. Mansur AH, Bishop DT, Markham AF, Britton J, Morrison JF. Association study of asthma and atopy traits and chromosome 5q cytokine cluster markers. Clin Exp Allergy 1998;28:141-50.
60. Liggett SB. Polymorphisms of the beta2-adrenergic receptor and asthma. Am J Respir Crit Care Med 1997;156(Suppl):S156-62.
61. Martinez FD, Graves PE, Baldini M, Erickson R. Association between genetic polymorphisms of the b2-Adrenoceptor and response to albuterol in children with and without a history of wheezing. J Clin Invest 1997;100:3184-8.
62. Baldini M, Lohman IC, Halonen M, Erickson RP, Holt PG, Martinez FD. A polymorphism in the 5' flanking region of the CD14 gene is associated with circulating soluble CD14 levels and with total serum immunoglobulin E. Am J Respir Cell Mol Biol 1999;20:976-83.
63. Doull IJ, Lawrence S, Watson M, et al. Allelic association of gene markers on chromosomes 5q and 11q with atopy and bronchial hyperresponsiveness. Am J Respir Crit Care Med 1996;153:1280-4.
64. Graves PE, Kabesch M, Halonen M, et al. A cluster of seven tightly linked polymorphisms in the IL-13 gene is associated with total serum IgE levels in three populations of Caucasian children. J Allergy Clin Immunol 2000;105:506-13.

ON THE MOVE?

Send us your new address at least six weeks ahead

Don't miss a single issue of the journal! To ensure prompt service when you change your address, please photocopy and complete the form below.

Please send your change of address notification at least six weeks before your move to ensure continued service. We regret we cannot guarantee replacement of issues missed due to late notification.

JOURNAL TITLE:
Fill in the title of the journal here. _____

OLD ADDRESS:
Affix the address label from a recent issue of the journal here.

NEW ADDRESS:
Clearly print your new address here.

Name _____

Address _____

City/State/ZIP _____

COPY AND MAIL THIS FORM TO:
Mosby
Subscription Customer Service
6277 Sea Harbor Dr.
Orlando, FL 32887

OR FAX TO:
407-363-9661

Mosby

OR PHONE:
800-654-2452
Outside the U.S., call
407-345-4000