Asthma and wheezing are characterised by airway inflammation and reversible airflow obstruction, and seem to have multiple phenotypes that may differ on the basis of age of onset. Boys are more prone to develop wheezing and asthma in the early years, whereas girls are more susceptible later in life. Whether this difference is due to genetic, hormonal, or environmental factors is unclear. We have previously reported an increased risk of wheezing and sensitisation in boys, in agreement with other studies, but could also demonstrate an interaction exceeding additivity between male sex and parental allergic disease, particularly in children with persistent wheezing. This finding raised the hypothesis of a sex specific genetic influence on childhood wheezing.

The interleukin-9 receptor gene (IL9R), located on the pseudoautosomal region of X and Y chromosomes (Xq28 and Yq12), has previously been associated with asthma in two separate family based data sets. Haplotype analyses of polymorphic microsatellite markers in one of the studies revealed significant associations, predominantly in females, which suggest that the influence may be sex specific. However, to our knowledge no study on associations between IL9R single nucleotide polymorphisms (SNPs) and asthma and allergy has yet been reported.

The IL9 receptor belongs to the haematopoietin receptor superfamily and is expressed on T cells, mast cells, macrophages, eosinophils, and neutrophils. The receptor consists of a specific IL9Rα unit and the common γc subunit, which forms a heteromer upon IL9 binding and activates Janus kinases and STAT factors. Signals from the IL9 receptor have been shown to be crucial for immunologic processes such as T cell development and prevention of apoptosis. More specifically, IL9 may stimulate release of chemotactic factors from bronchial epithelial cells and smooth muscle cells. IL9 may also have a key role in the development of allergy, as IL9 can act directly on B lymphocytes (through IL9R) and regulate IgE synthesis.

Thus, IL9/IL9R may be important at several stages of the pathophysiology of asthma and allergy.

Our aim in the present study was to test the hypothesis of sex specific genetic effects on wheezing, by examining the role of IL9R SNPs and haplotype patterns in childhood wheezing and allergic sensitisation. More than 1000 children from the prospective birth cohort, BAMSE, were selected using a case-cohort sampling design. Twenty six potential SNPs were evaluated in order to capture all essential variants of the IL9R gene, and data analyses were performed using the four most informative SNPs in boys and girls separately for evaluation of potential sex specific influence.

**METHODS**

**Study design**

The BAMSE study is based on a well characterised, representative cohort of children followed prospectively from birth, which facilitates generalisation and interpretation of the results. It has previously been presented in detail. In brief, between 1994 and 1996, 4089 newborn infants (2024 girls and 2065 boys) were recruited and data on parental allergic diseases and residential characteristics were requested in questionnaires filled out by the parents when the child was 2 months old (median age). Families planning to move within a year, those without sufficient knowledge of Swedish, and families whose infant suffered from a severe disabling disease were not included. This study base comprised 75% of all eligible children born in predefined areas of both urban and suburban districts of Stockholm. Sixteen per cent of the children have one or two parents born outside Scandinavia. When the children were 1, 2, and 4 years old, new questionnaires were mailed to all parents, this time with the main focus on children’s symptoms related to wheezing and other allergic diseases. Questions related to wheezing followed the ISAAC questionnaire, slightly modified. At approximately 4 years of age (mean 4.3), blood samples were drawn from 2614 children (64%) and analysed for serum IgE antibodies to inhalant and food allergens (using Phadiatop, a mixture of cat, dog, horse, birch, timothy, mugwort, and D. pteronyssinus and C. cladosporiun allergens, and fx5, a mixture of milk, egg white, soy bean, peanut, fish, and wheat allergens; both Pharmacia Diagnostics AB, Uppsala, Sweden).

**Key points**

- The interleukin-9 receptor gene (IL9R), located on the pseudoautosomal region of X and Y chromosomes, has been suggested as a candidate gene for sex specific effects on asthma.
- From a prospective birth cohort study, 542 children with wheezing and 542 controls were selected for analyses of IL9R single nucleotide polymorphisms (SNPs).
- Out of 26 potential SNPs tested, four SNPs were chosen for complete analyses. A specific haplotype (GAGC) had a protective effect against wheezing in boys (odds ratio = 0.58, 95% confidence interval 0.42 to 0.80), but not in girls. We also observed a weak protective effect of the GAGC haplotype against sensitisation to inhalant and/or food allergens.
- These results point to a role of the IL9R gene in the pathophysiology of allergic diseases, which may be sex dependent.

**Abbreviations:** CI, confidence interval; HWE, Hardy-Weinberg equilibrium; LD, linkage disequilibrium; MALDI-TOF mass spectrometry; matrix-assisted laser desorption/ionisation-time of flight mass spectrometry; OR, odds ratio; SNP, single nucleotide polymorphism; SR, success rates.
A positive test result was defined as IgE antibody levels >0.35 kU/L. A total of 2298 of these blood samples were available for genetic analyses after exclusion of 69 samples because of too little blood, 81 samples due to lack of questionnaire data, and 166 samples because parental consent to genetic analyses was not obtained.

Definitions of cases and controls

In a case-cohort sampling design, a sequential random sample of 709 children (357 girls, 352 boys) from the “genetics cohort” (n = 2298) was selected as a subcohort (fig 1) until 542 children with no defined wheezing were included. These were used as random controls (282 girls, 260 boys), whereas children who fulfilled any wheezing criterion up to the age of 4 were identified as wheezing cases (n = 167). An additional 375 children were identified as wheezing cases from the overall “genetics cohort”, resulting in 542 wheezing cases and 542 controls.

Figure 1. Selection of cases and controls using a case-cohort sampling design. A sequential random sample of 709 children from the “genetics cohort” (n = 2298) was selected as a subcohort. From this random subcohort, children with no defined wheezing were used as controls (n = 542), whereas children who fulfilled any wheezing criterion up to the age of 4 were identified as wheezing cases (n = 167). An additional 375 children were identified as wheezing cases from the overall “genetics cohort”, resulting in 542 wheezing cases and 542 controls.

Figure 2. Allelic association between the four chosen SNPs (numbered 1–4 according to location in the gene), measured by D’ for the IL9R gene area (in both cases and controls). D’ values for each shade are indicated in the column to the right.

A total of 195 (27.5%) in the subcohort were sensitised to either inhalant or food allergens, whereas 514 children were not sensitised. The randomly selected subcohort appeared representative of the original cohort in that no significant differences were seen concerning, for example, ethnic background, sex, parental allergic diseases, smoking mothers, and wheezing prevalence. Children in the subcohort were somewhat more often sensitised compared to the full cohort (27.5% v 24.0%, p = 0.06) and a minimal difference was seen regarding age at investigation (4.31 v 4.29 years, p = 0.01). Due to occasional failure in DNA extraction and genotyping.
the number of individuals presented in the tables is somewhat lower than stated above.

SNPs and genotyping
DNA was extracted from peripheral blood leukocytes using a standard non-enzymatic method or the PUREGENE Kit (Gentra Systems, Minneapolis, MN, USA). Initially, 26 potential SNPs were selected from the NCBI and Sequenom Real SNP databases. Primers for multiplexing PCR and extension reactions were designed by the SpectroDesigner software (Sequenom, San Diego, CA, USA). PCR and extension reactions were performed according to standard protocols and detailed information is available on request (erik.melen@imm.ki.se).

The SNP analysis was performed by MALDI-TOF (matrix-assisted laser desorption/ionisation-time of flight) mass spectrometry (Sequenom) and the genotyping results were also manually double checked. All samples were analysed with a sex specific assay (Amelogenin) in order to guarantee accuracy in blood and DNA sample management. The assay indicated that sampling and genotyping errors were clearly below 5%.

Statistical analyses
Allele frequencies and success rates (SR) were calculated according to standard procedures, and deviations of observed genotype frequencies from the Hardy-Weinberg equilibrium (HWE) were identified by the $\chi^2$ test. Block-wise inheritance of the four markers was checked by estimating the relative linkage disequilibrium ($D'$) (fig 2). Since pedigree information was unavailable, haplotype frequencies were estimated by an EM algorithm as previously described. The haplotypes were inferred for both cases and controls separately and together and there was no significant difference in the observed haplotype frequencies between the two approaches. Each individual contributed two haplotypes (one from each chromosome) to the analyses. $p$ Values for differences in allele/haplotype frequencies between cases and controls were obtained using the $\chi^2$ test. Odds ratios (OR) and 95% confidence intervals (CI) for comparison of haplotype frequency in wheezing cases versus controls and sensitised versus non-sensitised children were calculated using standard procedures. The effect of a particular haplotype, GAGC, on wheezing and sensitisation was also assessed by a multiple (and multinomial) regression model and controlled for potential confounders (socio-economic status, maternal smoking during pregnancy or at enrolment, mother’s age at enrolment, and pet ownership) previously identified as confounders for wheezing. The regression analysis was also adjusted for ethnic background. However, none of these variables had any confounding effect alone or together (the $\beta$ coefficient for GAGC changed less than 10% when entering a variable or when comparing the crude and adjusted models) and the results are therefore presented unadjusted. No adjustment was made for sex, since the data were analysed separately for boys and girls, or maternal/paternal heredity, since the haplotypes can be viewed as a form of heredity and adjustment may in such case lead to over adjustment. Since haplotype association tests were performed multiple times, a permutation test was used to estimate the overall significance (empirical $p$ value) for the whole haplotype set, that is, to test whether the observed distribution of haplotypes could be expected by chance. Thus, the haplotypes were treated as fixed, while the phenotypes (case or control) of the haplotypes (chromosomes) were randomised. The proportion of 50 000 such iterations (randomised $\chi^2$ tests) where a stronger association was found than in the actual data, provided the empirical $p$ value of the observation.

Figure 3 Location and allele frequencies of 26 potential SNPs in 276 subjects (184 cases and 92 controls). The estimated block region is indicated by a box with broken lines. *Included in the haplotype analyses. †Excluded due to Hardy Weinberg $p$ values <0.01 (rs2077051: 0.003, rs2037999: 0.006). ‡Excluded due to low information content in the haplotype analyses. Minor allele frequencies of the excluded SNPs: rs1973881: 0.01 (C), rs1883079: 0.04 (A), rs2077051: 0.34 (A), and rs2037999: 0.35 (C).
Haplotype analyses

Haplotype analyses showed that the four major observed haplotypes constituted >95% of all haplotypes and therefore, by virtue of the high LD in this region, can be expected to account for virtually all essential variations in this gene at a population level (table 2). Analyses of boys and girls separately revealed that the GAGC haplotype was less prevalent in male cases (all types of wheezing) than in male controls (OR = 0.58, 95% CI 0.42 to 0.80), but no such difference could be found between female cases and controls. The empirical p value for difference in haplotype distribution (the whole haplotype set) between cases and controls was significant in boys (p = 0.01) but not in girls (p = 0.33).

Further analyses of the GAGC haplotype showed that the genetic effect in boys was compatible with a dominant mode of inheritance, as a protective effect could be observed even in the presence of one such haplotype, OR = 0.59, 95% CI 0.40 to 0.86 (OR = 0.46, 95% CI 0.17 to 1.24 for homozygous GAGC

### Table 1

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Minor allele</th>
<th>Controls</th>
<th>All cases</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs731477 (C/G)</td>
<td>G</td>
<td>0.20</td>
<td>0.18</td>
<td>0.40</td>
</tr>
<tr>
<td>rs731476 (A/G)</td>
<td>A</td>
<td>0.38</td>
<td>0.31</td>
<td>0.003</td>
</tr>
<tr>
<td>rs731478 (C/G)</td>
<td>G</td>
<td>0.19</td>
<td>0.14</td>
<td>0.004</td>
</tr>
<tr>
<td>rs1973880 (C/T)</td>
<td>T</td>
<td>0.17</td>
<td>0.16</td>
<td>0.60</td>
</tr>
</tbody>
</table>

*p Values for differences in allele frequencies between cases and controls using the x^2 test.

### Table 2

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Controls, boys (n = 246)</th>
<th>Wheezing cases, boys (n = 314)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGCC</td>
<td>0.481</td>
<td>0.506</td>
<td>1.10 (0.87 to 1.40)</td>
</tr>
<tr>
<td>CACC</td>
<td>0.158</td>
<td>0.170</td>
<td>1.09 (0.79 to 1.50)</td>
</tr>
<tr>
<td>GAGC</td>
<td>0.204</td>
<td>0.129</td>
<td>0.58 (0.42 to 0.80)</td>
</tr>
<tr>
<td>CGCT</td>
<td>0.127</td>
<td>0.149</td>
<td>1.20 (0.85 to 1.69)</td>
</tr>
<tr>
<td>Other</td>
<td>0.030</td>
<td>0.046</td>
<td>1.54 (0.79 to 3.14)</td>
</tr>
</tbody>
</table>

*p Values for differences in allele frequencies between cases and controls using the x^2 test.

### Table 3

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>No sensatisation</th>
<th>Sensitisation</th>
<th>OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys</td>
<td>Frequency</td>
<td>OR (95% CI)*</td>
<td></td>
</tr>
<tr>
<td>O GAGC</td>
<td>(n = 174)</td>
<td>(n = 72)</td>
<td>1.00</td>
</tr>
<tr>
<td>1 or 2 GAGC</td>
<td>(n = 211)</td>
<td>(n = 59)</td>
<td>0.60 (0.33 to 1.09)</td>
</tr>
<tr>
<td>Girls</td>
<td>Frequency</td>
<td>OR (95% CI)*</td>
<td></td>
</tr>
<tr>
<td>O GAGC</td>
<td>(n = 211)</td>
<td>(n = 157)</td>
<td>72.0</td>
</tr>
<tr>
<td>1 or 2 GAGC</td>
<td>(n = 211)</td>
<td>(n = 53)</td>
<td>0.91 (0.48 to 1.75)</td>
</tr>
</tbody>
</table>

*p Values for differences in allele frequencies between cases and controls using the x^2 test.

**RESULTS**

**SNP validation and allele frequencies**

It was possible to genotype 25 of 26 potential SNPs, and eight of those were found to be polymorphic (minor allele >1%, fig 3) based on analyses in a test sample of 276 individuals (184 cases and 92 controls). SNPs that were found to be monomorphic, had low minor allele frequencies (<5%), or had low information content in preliminary haplotype analyses were excluded, leaving six intronic SNPs for further analyses in all cases and controls. The overall success rate for each SNP was >90%. Notably, none of the coding SNPs had a high enough minor allele frequency to be able to contribute substantially to a genetic risk in our material. HWE tests suggested that allele readout problems occurred in two assays (rs2077051 and rs2037999; p < 0.01 for controls) and these two SNPs were therefore excluded. The four remaining SNPs were genotyped and were found to be in strong LD and within one haplotype block, spanning an approximately 1400 bp region on the gene (fig 3). Even when the two SNPs were excluded, the remaining four tagging SNPs (rs731477, rs731476, rs731478, and rs1973880) were sufficient to describe all relevant variations in this DNA segment, since the two markers did not add further subdivision of haplotypes (not shown). Allele frequencies in cases and controls for each SNP are presented in table 1.

**Haplotype analyses**

Haplotype analyses showed that the four major observed haplotypes constituted >95% of all haplotypes and therefore, by virtue of the high LD in this region, can be expected to account for virtually all essential variations in this gene at a population level (table 2). Analyses of boys and girls separately revealed that the GAGC haplotype was less prevalent in male cases (all types of wheezing) than in male controls (OR = 0.58, 95% CI 0.42 to 0.80), but no such difference could be found between female cases and controls. The empirical p value for difference in haplotype distribution (the whole haplotype set) between cases and controls was significant in boys (p = 0.01) but not in girls (p = 0.33).

Further analyses of the GAGC haplotype showed that the genetic effect in boys was compatible with a dominant mode of inheritance, as a protective effect could be observed even in the presence of one such haplotype, OR = 0.59, 95% CI 0.40 to 0.86 (OR = 0.46, 95% CI 0.17 to 1.24 for homozygous GAGC.
carriers). We also calculated the estimated prevented fraction (of the GAGC haplotype) in the full cohort (boys only) according to previously described formulae. Due to the exposure of having one or two copies of the GAGC haplotype in the genome, approximately 11% of the wheezing cases were prevented (that is, these cases would have been present in the population without such exposure).

Joint analyses of both boys and girls together showed that the frequencies of the four most common haplotypes did not significantly differ in wheezing subjects compared to controls when disregarding sex differences (empirical p value = 0.37), although the difference in frequency of the GAGC haplotype was of borderline significance (0.15 v 0.17, respectively; OR = 0.82, 95% CI 0.65 to 1.03). The cases were further classified as having recurrent transient wheezing, early persistent wheezing, or late onset wheezing, according to the predefined criteria. Due to the limited number of children in each subgroup of wheezing, no stratification by sex could be made and haplotype analyses revealed no significant effect in any of the subgroups.

Taking both wheezing and sensitisation status into account, associations of borderline significance were observed between the presence of one or two GAGC haplotypes and sensitisation in boys without wheezing symptoms as compared to boys without both wheezing and sensitisation (OR = 0.60, 95% CI 0.33 to 1.09; table 3). The haplotype effect was even more pronounced in boys who wheezed up to the age of 4 but were not sensitised, whereas a less pronounced effect was observed in those with concomitant wheezing and sensitisation. Again, no particular effect was seen in girls. Haplotype analyses were also performed in the randomly selected subcohort of 709 children with sensitisation as an independent outcome. Sensitised children had significantly lower frequency of the GAGC haplotype than non-sensitised children (0.13 v 0.17, OR = 0.71, 95% CI 0.50 to 0.99, boys and girls together).

DISCUSSION
Childhood wheezing and atopy are to a large extent heritable, but these conditions are genetically complex and may be influenced by several genetic and environmental components. Linkage analyses for asthma susceptibility and IgE regulation loci performed some years ago were initially promising in the search for a major gene, but strong candidate genes have been ruled out. We also calculated the estimated prevented fraction of the full cohort (11%) due to exposure to one or two copies of the GAGC haplotype. For comparison, the population attributable fraction of exposure to environmental risk factors for wheezing up to the age of 2 has been estimated as 14.5% (the fraction of the disease in the population that would be avoided if the studied exposures were eliminated: maternal tobacco smoke, being raised in a damp home, and breastfeeding less than 4 months). The GAGC haplotype was found to be independently associated with both wheezing (in boys) and sensitisation, which strengthens the associations. The results are biologically plausible, as the IL9 signal pathway is involved in both airway and immunological processes. For both outcomes, the GAGC haplotype was protective. None of the other haplotypes were identified as being significantly predisposing.

The observation of conserved haplotype block structures has greatly facilitated genetic association studies. To study the risk effects of the IL9R gene, we considered 26 potential SNPs altogether, including both coding and non-coding variants. Of these, four were found to be sufficiently common to be informative and these tags defined four major haplotypes representing >95% of all chromosomes. It is possible that none of the four SNPs used to tag the haplotypes are functionally involved in regulating the genetic effect, but any functionally important polymorphisms should be embedded in one of the four major haplotypes to have a reasonable effect on the population level.

The observed associations are thus indirect, but constitute evidence that IL9R variants may influence disease susceptibility and that the analysed SNPs are in LD with potential causal SNP(s). A few known changes in IL9R function have been reported elsewhere; a splice variant containing an intron deletion of a single residue at codon 173 (DeltaQ) has been reported to alter structure and function of the receptor, resulting in inability to bind IL9.Functional changes in IL9R may also derive from processes affecting key amino acids essential for STAT activation and signalling, such as tyrosine 116 and 336. Such changes are likely to have an impact on allergic diseases, as increased expression of IL9 mRNA and enhanced IL9/IL9R immunoreactivity have been demonstrated, for example, in the Airways of asthmatic subjects. However, polymorphisms affecting tyrosine 116 and 336 have to our knowledge not yet been identified.

The wheezing cases in the present study were identified on the basis of the number of wheezing episodes up to the age of 4. A certain number of wheezing episodes is considered to be the major criterion for childhood asthma. We have previously reported that already at the ages of 2 and 4 years, different risk factors and disease characteristics, such as sensitisation and lung function, are associated with the three phenotypes transient, persistent, and late onset wheezing. Haplotype associations in these separate groups (boys and girls together), however, showed no significant effects and the joint analyses including all cases showed that the protective effect seemed to be first and foremost sex specific. For the two SNPs excluded due to deviation from the HWE, the observed numbers of heterozygous individuals were lower than expected. This observation indicates that allele readout problems occurred in these assays, as we do not...
believe that conditions related to the studied population, for example, non-random mating, inbreeding etc., influenced the HWE to any large extent. Possible reasons for the problems include non-specific assays (although primers were controlled), genotyping errors, DNA quality, unknown issues, or simply deviation from the HWE equilibrium by chance."

In the present study on allergic diseases in children, 26 potential SNPs in the complex IL9 gene region located on the X and Y chromosomes were screened and evaluated. Using four polymorphic SNPs with adequate information content, we show a significant protective effect of a specific haplotype against wheezing in boys and also demonstrate protective effects against the development of sensitisation. This points to a role of the IL9 gene in the pathophysiology of allergic diseases, but additional studies on the IL9 gene are needed to further investigate the functional consequences of the identified haplotypes.

ACKNOWLEDGEMENTS

We wish to thank all children and parents in the BAMSE cohort, all nurses at the Health Care Centres and the BAMSE staff, especially Stina Gustavsson, at the Department of Occupational and Environmental Health, Stockholm County Council. We also thank Dr Myriam Peyrard-Janvid and Hannele Anthoni for help with DNA extraction, Päivi Kiviluoma for database management, Dr Päivi Onkamo for study design, and Dr Cecilia Lindgren for genotyping advice.

Authors’ affiliations

E Melén, F Nyberg, M Wickman, G Pershagen, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
E Melén, Centre for Allergy Research, Karolinska Institutet, Sweden
H Gullénst, A Lindstedt, J Kere, Clinical Research Centre, Karolinska University Hospital, Huddinge, Sweden

H Gullénst, J Kere, Department of Medical Genetics, University of Helsinki, Finland
M Zucchini, J Kere, Department of Biosciences, Karolinska Institutet, Sweden
F Nyberg, AstraZeneca R&D Malmö, Malmö, Sweden
M Wickman, Department of Occupational and Environmental Health, Stockholm County Council, Stockholm, Sweden

Fredrik Nyberg is employed by AstraZeneca and AstraZeneca also supports his part time adjunct position as Lecturer in Molecular Epidemiology at Karolinska Institutet. AstraZeneca has not contributed any direct financing to this study. Sources of research support: Stockholm County Council, The Swedish Asthma and Allergy Association, The Swedish Foundation for Health Care Science and Allergy Research, The Swedish Medical Research Council, The Swedish Heart-Lung Foundation, The Swedish Committee for Medical Research.

Conflict of interest: none declared.

Swedish Society of Medicine, Konsul Th.C. Bergh’s Foundation and The Swedish Foundation for Health Care Science and Allergy Research, The County Council, The Swedish Asthma and Allergy Association, The Swedish Centre for Epidemiology at Karolinska Institutet. AstraZeneca has not contributed supports his part time adjunct position as Lecturer in Molecular Epidemiology at Karolinska Institutet.


Sex specific protective effects of interleukin-9 receptor haplotypes on childhood wheezing and sensitisation

E Melén, H Gullstén, M Zucchelli, A Lindstedt, F Nyberg, M Wickman, G Pershagen and J Kere

J Med Genet 2004 41: e123
doi: 10.1136/jmg.2004.023135

Updated information and services can be found at:
http://jmg.bmj.com/content/41/12/e123

These include:

References
This article cites 38 articles, 16 of which you can access for free at:
http://jmg.bmj.com/content/41/12/e123#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

- Immunology (including allergy) (603)
- Asthma (21)
- Genetic screening / counselling (886)
- Molecular genetics (1254)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/