

SPECIAL REPORT

A retrospective collaboration on chromosome 5 by the International Consortium on Asthma Genetics (COAG)

L.J. PALMER^{1,2*}, C. LONJOU^{3*}, K. BARNES⁴, H. CHEN⁵, W.O.C.M. COOKSON⁶, K.A. DEICHMANN⁷, J.W. HOLLOWAY⁸, T. LAITINEN⁹, M. WJST¹⁰ and N.E. MORTON³

¹Channing Laboratory, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA,

²Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio, USA, ³Human Genetics Research Division, University of Southampton, Southampton General Hospital, Southampton, UK, for the Consortium on Asthma Genetics (COAG), ⁴Johns Hopkins University, Baltimore, Maryland, USA, ⁵Department of Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, Beijing, China, ⁶Asthma Genetics Group, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Oxford, UK, ⁷Childrens' Hospital, University of Freiburg, Germany, ⁸Asthma Genetics Group, Human Genetics Research Division, Duthie Building, Southampton General Hospital, UK, ⁹Department of Medical Genetics, University of Helsinki, Finland, ¹⁰Institute of Epidemiology, GSF National Research Center for Environment and Health, Neuherberg, Germany, for the German Asthma Genetics Group. *Equal first authorship

Asthma is the most common chronic childhood disease in developed nations [1], affecting more than 155 million individuals. The cost of treating the disease in the USA alone approximates 6 billion dollars per annum [2]. Asthma is a complex disease that has proven extremely difficult to dissect genetically [3]. Considerable effort is currently being expended in attempts to detect genetic loci contributing to asthma susceptibility [4–7], with the ultimate goal of improving preventive strategies, diagnostic tools and therapies. However, no specific major gene exhibiting functional effects has been mapped to date.

The Consortium on Asthma Genetics (COAG) was established in 1999 as an international effort to define the candidate regions and genes for asthma and atopy with greater precision and reliability than can be achieved by smaller individual studies. Given that one of the most investigated regions of the human genome for linkage to asthma phenotypes has been the 5q region, a pilot retrospective study focusing on combination of evidence for the cytokine region on chromosome 5q was established (COAG-5). The 5q31–33 region is generally regarded as a biologically sensible candidate location for atopy and asthma susceptibility loci [8], and has shown evidence in

many studies of linkage and association to asthma-associated phenotypes [9,10]. However, some studies have shown neither linkage nor association of asthma and associated phenotypes to the 5q region, all studies have had modest sample sizes, and there is likely to have been some level of positive publication bias.

A recent meeting of COAG took place in Southampton, UK (16–17 April 2000) and saw the first presentation of the COAG-5 results. The meeting was attended by 23 delegates representing research groups in Australia, China, Finland, Germany, Italy, Poland, the United Kingdom (UK) and the USA, and presenting data from Australian, Barbados, Chinese, Finnish, German, UK and United States of America (USA) populations. A total of 11 datasets comprising a total of 1070 families and 6277 subjects were analysed. Three of the datasets comprised random population samples and eight were selected for asthma or atopy. A total of 35 markers on chromosome 5q31–33 were typed, although differing numbers of markers were typed in each individual study. There were differences in both marker typing and phenotypic definition that were addressed by the analytical approaches developed for the chromosome 5 analysis.

Two independent analyses were presented. In the first, Lonjou performed an analysis that combined multipoint LOD scores calculated at fixed locations within each dataset, using several definitions for asthma and atopy,

Correspondence: Dr Lyle J. Palmer, Channing Laboratory, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA. E-mail: reljp@channing.harvard.edu.

including affection status and principal components, and different methods to combine lod scores over studies. Study-specific genetic maps were used. This approach is applicable not only to retrospective collaboration, but also to meta-analysis of appropriately summarized data.

After omitting two small samples, Lonjou found that six of the populations gave positive lods for asthma susceptibility in asthmatic sib-pairs over all or most of the cytokine region, while three samples give no positive lods within flanking markers. The combined lod score across all of the study populations was maximal in the proximal area of the cytokine cluster, around reference locus D5S1505 (lod = 2.61, $P = 0.0025$). Although neither the individual samples nor the pooled lod scores in these data reached the conventional lod of 3, they do confirm other claims of linkage to this region. A conservative 90% confidence interval included the IL4 and IL13 candidate genes, but not IL9. The distal part of the cytokine region also gave some evidence of linkage. The maximal lod pooled across all of the study populations was found in the region of the adrenoreceptor locus (ADRB2) (lod = 1.53, $P = 0.03$). This suggested the existence of one or more oligogenes near ADRB2 with modest effects on asthma susceptibility. In contrast, log_e total serum IgE levels did not approach significance by any method in any individual study nor in the overall sample.

There was discussion over the use of principal component analysis to define phenotypes for linkage analysis when phenotypic measurements provided by each group are heterogeneous (some samples having only total IgE, others provided more complete information concerning BHR, wheeze, specific IgE). The advantage of principal components in meta-analysis is to combine into a quantitative score heterogeneous data and to attempt to differentiate between asthma and atopy.

In the second analysis, Palmer used a new variance components implementation of the multipoint Haseman-Elston linkage statistic [11,12] and extended analyses utilizing a Gibbs-sampling-based variance components approach to derive new, ascertainment-adjusted phenotypes [13–15]. Five of the samples gave modest evidence of linkage ($0.01 < \text{empirical-}P < 0.05$) to asthma dichotomy in the 5q31–33 region. Combination of linkage results was investigated using two different approaches: pooling raw data and pooling the location-specific regression coefficients from each of the study populations (analogous to Lonjou's combination of lods). The results regarding asthma affection status and total serum IgE levels were consistent with those of Lonjou. The evidence for linkage to asthma affection was highest in the proximal part of the region, and was maximal at D5S2501 (empirical $P = 0.08$, $t = 1.40$, 2571 full sib-pairs), consistent with previous claims of weak linkage to this region. Log_e total serum IgE

levels did not approach significance by any method in any individual study nor in the overall sample. No evidence of significant linkages to an index of specific IgE immune response or to the dose–response slope for bronchial challenge was found in the overall sample.

This study implemented a number of strategies for dealing with potential issues involved in combining evidence for linkage: the phenotypes examined were chosen to maximize the number of informative sibling pairs; the log_e total serum IgE phenotype was standardized to control for heterogeneity among datasets; advanced phenotypic modelling was applied to extract the component of the binary asthma phenotype that appeared genetic and that minimized the heterogeneity of this phenotype among datasets; a carefully checked genetic map was constructed to facilitate uniform multipoint analysis and interpolation at untyped locations; a state-of-the-art linkage test was used; P -values were generated empirically instead of relying on potentially unrealistic asymptotic estimates of significance; and several data pooling and meta-analytical strategies were applied. Adjusting for possible genotyping errors made little difference to the substantive conclusions.

A primary concern when combining data from diverse populations is the potential interpopulation heterogeneity in study design, phenotype definition, genetic structure, environmental exposures and markers typed. The studies by Lonjou and Palmer implemented a number of different strategies for dealing with these issues, but had similar conclusions. Overall, these complementary analyses of a very large, pooled dataset suggest that loci in the 5q31–33 region have a modest effect on asthma susceptibility and at most a modest effect on atopy. No previous studies of this region have been large enough to detect modest effects with high power at conventional levels of significance (i.e. a lod of 3 or greater).

This study has demonstrated the application of a comprehensive suite of methods to maximize power to detect linkage using multiple datasets. The combination of evidence from multiple studies can be expected to be critical to the successful localization of genes of modest effect for any common, complex disease such as asthma. This study also illustrates the benefits of cooperation between groups studying a complex human disease, that data from diverse studies and populations can be effectively pooled, and that new and important information can result from the analysis of large, pooled datasets. It is hoped that further research and methodological development, together with growing international participation in COAG, will lead to major advances in our understanding of the genetic basis of asthma. COAG now plans to investigate other candidate regions for asthma genes such as the 12q region.

Acknowledgements

COAG was founded under a grant from the National Asthma Campaign, UK. Many people in 14 research groups contributed to this collaboration. The meeting reported here was supported by the Medical Research Council, UK. We are grateful to members of the Collaborative Study on the Genetics of Asthma (CSGA) for their generous contribution of data.

References

- 1 Asher MI, Keil U, Anderson HR et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *Eur Respir J* 1995; 8(3):483–91.
- 2 Smith DH, Malone DC, Lawson KA et al. A national estimate of the economic costs of asthma. *Am J Respir Crit Care Med* 1997; 156(3, Part 1):787–93.
- 3 Sandford A, Weir T, Pare P. The genetics of asthma. *Am J Respir Crit Care Med* 1996; 153:1749–65.
- 4 Daniels S, Bhattacharya S, James A et al. A genome-wide search for quantitative trait loci underlying asthma. *Nature* 1996; 383:247–50.
- 5 CSGA. A genome-wide search for asthma susceptibility loci in ethnically diverse populations. *Nature Genet* 1997; 15:389–92.
- 6 Ober C, Cox NJ, Abney M et al. Genome-wide search for asthma susceptibility loci in a founder population. The Collaborative Study on the Genetics of Asthma. *Hum Mol Genet* 1998; 7(9):1393–8.
- 7 Wjst M, Fischer G, Immervoll T et al. A genome-wide search for linkage to asthma. *Genomics* 1999; 58(1):1–8.
- 8 Corry DB, Kheradmand F. Induction and regulation of the IgE response. *Nature* 1999; 402(6760 Suppl.):B18–23.
- 9 Barnes KC, Marsh DG. The genetics and complexity of allergy and asthma. *Immunol Today* 1998; 19(7):325–32.
- 10 Cookson W. The alliance of genes and environment in asthma and allergy. *Nature* 1999; 402(6760 Suppl.):B5–11.
- 11 Elston RC, Buxbaum S, Jacobs KB et al. Haseman and Elston revisited. *Genet Epidemiol* 2000; 19(1):1–17.
- 12 Xu X, Weiss S, Wei LJ. A unified Haseman-Elston method for testing linkage with quantitative traits. *Am J Hum Genet* 2000; 67(4):1025–8.
- 13 Burton P, Tiller K, Gurrin L et al. Genetic variance components analysis for binary phenotypes using generalized linear mixed models (GLMMs) and Gibbs sampling. *Genet Epidemiol* 1999; 17(2):118–40.
- 14 Palmer L, Tiller K, Burton P. Genome wide linkage analysis using genetic variance components of alcohol-dependency-associated censored and continuous traits. *Genet Epidemiol* 1999; 17(S1):283–8.
- 15 Scurrah KJ, Palmer LJ, Burton PR. Variance components analysis for pedigree-based censored survival data using generalized linear mixed models (GLMMs) and Gibbs sampling in BUGS. *Genet Epidemiol* 2000; 19(2):127–48.