Title Heritability and confirmation of genetic association studies for childhood asthma in twins

Short title Genetics of childhood asthma – a twin study

Authors

Vilhelmina Ullemar, BSc[1], Patrik K.E. Magnusson, PhD[1], Cecilia Lundholm, MSc[1], Anna Zettergren, PhD[2,3], Erik Melén, MD, PhD[4,5], Paul Lichtenstein, PhD[1], Catarina Almqvist, MD, PhD[1,6]

1. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.

2. Department of Pharmacology, Institute of Neuroscience and Physiology at the Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden.

3. Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, The Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden.

4. Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

5. Sachs' Children's Hospital, Stockholm, Sweden.

6. Astrid Lindgren Children's Hospital, Lung and Allergy Unit, Karolinska University Hospital, Stockholm, Sweden.

Corresponding author

Vilhelmina Ullemar

Department of Medical Epidemiology and Biostatistics

Karolinska Institutet

Box 281

SE-17177 Stockholm

Telephone +46852482389

E-mail vilhelmina.ullemar@ki.se

Funding sources

Financial support was provided by the Swedish Research Council (grant number 2011-3060) and through the Swedish Initiative for research on Microdata in the Social And Medical Sciences (SIMSAM) framework grant number 340-2013-5867, grants provided by the Stockholm County Council (ALF project), the Strategic Research Program in Epidemiology at Karolinska Institutet, the Swedish Heart-Lung Foundation and the Swedish Asthma and Allergy Association's Research Foundation. The Ministry for Higher Education financially supports the Swedish Twin Registry.

Conflicts of interest The authors declare no conflicts of interest.

Author contributions

All authors listed fulfill the definition of authorship set up by the International Committee of Medical Journal Editors (ICMJE). Vilhelmina Ullemar contributed to conception and design, drafted the article, carried out data analysis and participated in interpretation of results. Patrik K. E. Magnusson and Cecilia Lundholm contributed substantially to conception and design, data analysis and interpretation of data. Anna Zettergren assisted in acquisition of data and interpretation of results. Erik Melén and Paul Lichtenstein made substantial contributions to conception and design of the study and contributed to interpretation of results. Catarina Almqvist initiated the study, provided financial support, made substantial contributions to conception and design, acquisition of data and interpretation of results. All authors revised the article for important intellectual content and approved the final version prior to submission.

Abstract

Background: Although the genetics of asthma has been extensively studied using both quantitative and molecular genetic analysis methods, both approaches lack studies specific to the childhood phenotype and including other allergic diseases. This study aimed to give specific estimates for the heritability of childhood asthma and other allergic diseases, to attempt to replicate findings from genome-wide association studies (GWAS) for childhood asthma and to test the same variants against other allergic diseases.

Methods: In a cohort of 25,306 Swedish twins aged 9 or 12, data on asthma were available from parental interviews and population-based registers. The interviews also inquired about wheeze, hay fever, eczema and food allergy. Through structural equation modeling, the heritability of all phenotypes was calculated. A subset of 10,075 twins was genotyped for 16 single nucleotide polymorphisms (SNPs) selected from previous GWAS; these were first tested for association with asthma and significant findings also against the other allergic diseases.

Results: The heritability of any childhood asthma was 0.82 (95% CI 0.79-0.85). For the other allergic diseases the range was approximately 0.60-0.80. Associations for six SNPs with asthma were replicated, including rs2305480 in the GSDMB gene (OR 0.80, 95% CI 0.74-0.86, p= $1.5*10^{-8}$; other significant associations all below p=$3.5*10^{-4}$). Of these, only rs3771180 in IL1RL1 was associated with any other allergic disease (for hay fever, OR 0.64, 95%CI 0.53-0.77, p=$2.5*10^{-6}$).

Conclusion: Asthma and allergic diseases of childhood are highly heritable, and these high-risk genetic variants associated specifically with childhood asthma, except for one SNP shared with hay fever.

Key words cohort study, epidemiology, heritability, pediatric, population-based twin cohort

Abbreviations

ATC: Anatomical Therapeutic Chemical classification system

CATSS: The Child and Adolescent Twin Study in Sweden

DZ: Dizygotic

GWAS: Genome-wide association study

ICD: International Classification of Disease

ICS: Inhaled corticosteroids

LRT: Likelihood ratio test

LTRA: Leukotriene receptor antagonists

MZ: Monozygotic

NPR: National Patient Register

PCR: Polymerase Chain Reaction

PDR: Prescribed Drug Register

SNP: Single nucleotide polymorphism

Main text total word count: 3479

Introduction

Asthma is a complex, prevalent, and highly heritable disease (1). It affects children as well as adults and is often comorbid with other allergic diseases such as hay fever and atopic eczema (2, 3). Previous studies of the genetic effects behind asthma have applied a variety of different definitions of the disease and study populations often include both children and adults. However, asthma is expressed differently in different age groups (4) and features substantial heterogeneity of phenotypes even within childhood asthma and wheeze (5).

The underlying genetic components of asthma have been demonstrated using quantitative (6-11) as well as molecular genetic methods (12-14). While these principally intend to describe the same phenomenon the perspectives are somewhat different. Quantitative genetic methods estimate the degree to which genes are responsible for the variation of a trait or disease within a population–the heritability – but cannot say which specific genes are involved. Molecular genetics identifies these genetic variants – often specific single nucleotide polymorphisms, SNPs – yet most commonly used methods (including genome-wide association studies, GWAS) do not capture the full extent of human genetic variation.

Twin models are frequently applied to estimate the magnitude of genetic effects, using correlations of a disease or trait within monozygotic (MZ) and dizygotic (DZ) twins and an assumption of shared segregating genes (100% within MZ and on average 50% within DZ pairs) to estimate the heritability of the disease (15). In the past, the heritability of asthma was estimated to be between 50 and 90% (2, 6-11, 16-22). However, it is difficult to draw reliable conclusions regarding asthma in children based on these studies, as some concern only adults (8, 9, 17, 18, 20) and the study populations with child subjects often include also adolescents and young adults (6, 17, 21) or address only preschool years (7, 19, 22). The few studies that focus on mid childhood to early adolescence (2, 10, 11) are comparably small. Further, previous research has suggested that patterns of variance components, including heritability,

of asthma may change over time (21). The heritability of allergic and otherwise related phenotypes is less studied, though reported estimates for hay fever (2, 8, 10, 16, 20, 22, 23) and eczema (2, 20, 22) lie in the same approximate range; although as many of these estimates are from the same studies cited previously, similar selection issues exist.

In the recent decade GWAS gained tremendous popularity in molecular genetic studies of complex traits, and typically examine hundreds of thousands of variants in the search for loci associated with various phenotypes. The most notable asthma GWAS have been carried out by the GABRIEL (12) and EVE (13) consortia, respectively. These and other GWAS (14, 24-26) and genetic association studies of candidate genes/ regions (27, 28) found several high risk alleles strongly associated with childhood asthma. In particular variants in genes on chromosomes 2, 5, 6 and 17 replicate between studies. Due to the nature of meta- and pooled analyses, heterogeneity in methodology is inevitable, and usually very general asthma phenotypes have been defined.

Thus, quantitative and molecular genetic methods agree that there are underlying genetic effects behind asthma, but both approaches have suffered from variability in phenotype definitions and study population. Because of this current knowledge is not necessarily applicable or specific to childhood asthma. In addition, the methods aim to explain the genetic effects from different perspectives. By the application of both quantitative and molecular genetic methods within a single study population, where subject characteristics and phenotype definitions align, unique inferences regarding the genetic effects of childhood asthma and allergic diseases may be made.

The dual aim of this study is to provide new heritability estimates for childhood asthma and related phenotypes within a large cohort of Swedish twins, and to test if a selection of SNPs found to be of importance for childhood asthma are also associated with the related phenotypes.

Materials and methods

CATSS-9/12

The Child and Adolescent Twin Study in Sweden (CATSS) is a cohort of twins born in and residents of Sweden on their 9th (and, during the first three years of data collection, 12th) birthday. Parents were interviewed over telephone regarding the twins' health and other characteristics (29). Data collection is still ongoing; so far including over 25,000 twins. The current study included interview data from all participating twins as of October 2014.

The interview contains questions regarding physical similarities between the twins; based on these an algorithm to determine their zygosity with more than 95% accuracy has been constructed (30). Starting in 2008, DNA extracted from saliva samples collected following the CATSS interview was tested using a 46-SNP-array to confirm zygosity (31). When available, DNA-based zygosity assessment took precedence. Overall, 58% (n=14,608) of the study participants had their zygosity assessed by algorithm and 41% (n=10,398) by DNA. The CATSS-9/12 study was approved by the Regional Ethical Review Board of Stockholm, Sweden.

Cohorts

The full CATSS cohort consisted of 25,306 twins from 12,635 pairs. All complete same-sex MZ and DZ pairs of the full cohort were included in main heritability analyses. Twins with available saliva samples and where DNA extraction yielded a sufficient amount of DNA were potential targets for genotyping. In DZ pairs both twins were genotyped whereas in MZ pairs only one twin was genotyped, and SNP-data were imputed for the co-twin. Genotype data was available for 12,388 (49% of the full CATSS cohort) twins. Of these, 8,618 (70%) had zygosity confirmed by DNA, 3,521 (28%) were opposite-sex DZ twins for which confirmation by DNA is unnecessary, and 249 (2%) were assessed by algorithm. Twins of

unknown zygosity (n=35) were treated as DZ in genetic association analysis. One MZ twin of each pair was randomly selected for genetic association analyses. A total of 10,075 twins remained. Out of 2,314 MZ twins randomly selected, 1,356 (59%) had own genotype data. The relationship between subgroups of the cohort is illustrated in Supporting information, Figure 1.

Phenotype definitions

Asthma was defined from interview- and register-based data. The CATSS-9/12 question "Has your child ever had asthma?" was used to define *ever asthma*.

Additional asthma outcomes extracted from national population-based health registers were previously validated against medical records.(32) *Asthma medication* was defined as either 1) ≥2 dispenses of inhaled corticosteroids, ICS (ATC code R03BA), fixed combinations of β2-adrenoreceptor agonists and ICS, β2-ICS (R03AK) and/or leukotriene receptor antagonists, LTRA (R03DC), independent of time between distributions or 2) ≥3 dispenses of ICS, selective β2-adrenoreceptor agonists, β2 (ATC code R03AC), β2-ICS and/or LTRA within a year, registered in the Prescribed Drug Register (PDR) between 4.5 and 18 years of age. *Register-based asthma diagnosis* in the National Patient Register (NPR) was defined by primary ICD-10 diagnosis J45 or J46 or ICD-9 493. The phenotype *any asthma* was defined as presence of any of the aforementioned asthma phenotypes. If *asthma ever* was missing and there were no reports of *asthma medication* or *register-based asthma diagnosis*, *any asthma* was set to missing.

All related phenotypes were defined based on parental reports in the CATSS-9/12 telephone interview. *Wheezing ever* was defined as yes to the question "Has your child ever had whistle breathing?". This question was only presented to parents who had not already reported that their child had asthma. *Wheezing after 3 years of age* was the subgroup of ever wheezers with symptoms after a pre-defined age cutoff. *Hay fever*, *food allergy* and *atopic eczema* ever were

all defined based on the parental telephone interview. I -in the case of eczema, the question was for eczema ever with a follow-up on the specific subtype(s) (33).

Genotyping

In total 18 single nucleotide polymorphisms (SNPs) previously associated with childhood asthma (12, 13, 27), in addition to two SNPs associated with adult asthma (12) or serum IgE levels (34) were selected for genotyping. The selected SNPs represented the IL6R, IL1RL1, IL18R1, TSLP, SLC22A5, IL13, HLA-DQ, IL33, RORA, SMAD3, Z2BP2, GSMB, GSDMA and IL2RB genes; the location of successfully genotyped SNPs in relation to the genes (in intron, exon, or nearby region) are presented in Table 3. DNA was extracted from saliva samples using the Oragene® DNA self-collection kit (DNA Genotek, Inc. Ottawa, ON, Canada). Genotyping was conducted using KASP(ar) PCR SNP genotyping system (KBioscience, Hoddesdon, Herts, UK). Two of the originally selected polymorphisms (rs9273349 and rs9469220, both located on chromosome 6 in the *HLA-DQ* region) could not be successfully genotyped within the population. Besides these, genotyping success rate was >95%. Individuals with unsuccessful genotyping of more than 3 of the remaining 16 SNPs (n=415) were excluded from genetic association analyses. Following these exclusions, 9,660 individuals remained. Of these, 2,257 (23.4%) were MZ, 7,375 (76.3%) were DZ (of which 4,040 were from same-sex and 3,335 from opposite-sex pairs) and 28 (0.3%) were of unknown zygosity. All SNPs were found to be in Hardy-Weinberg equilibrium. Some of the SNPs were highly correlated in the data (r^2 above 0.85); r^2 between rs2305480 and rs11078927 was 0.99 and 0.91 between rs12936231 and rs726389.

Statistical analysis

Prevalences of all the phenotypes were estimated in the full CATSS cohort, as well as in the groups included in heritability and genetic association analyses, respectively.

Heritability

Twin model analyses were conducted in the R-package OpenMX, an open source structural equation modeling solution (35). Univariate liability threshold models were specified using either an ACE or ADE structural equation. The standardized variance components A (additive genetics), E (nonshared environment) and either C (common environment) or D (genetic dominance, accounting for part of the genetic effect) were estimated with 95% CI. Model selection was based on comparison of phenotype-specific tetrachoric correlations between MZ and DZ twins; an ADE model was selected if the tetrachoric correlation in DZ-same sex (DZ-ss) pairs was less than half of that in MZ pairs. For MZ pairs, the A and D (if applicable) factors are fully correlated; r in DZ pairs there is a 50% correlation for A and 25% for D. The sum of the variance components is standardized to 1, i.e. the combined effects of A, C and D, or E are the total variance of the disease or trait within the population. Likelihood ratio tests (LRTs) were used to compare AE models to full ACE or ADE models. When C or D could be dropped without significant reduction in fit of the model, the AE model was interpreted. Additional sex-limitation models included opposite-sex dizygotic (DZ-os) twins, allowing for either quantitative (the same genes are involved between sexes but effects vary in size; separate estimates of each variance component are needed) or qualitative (different genes are involved for males and females; the variance component A differs by a factor rA) differences in A by sex. LRTs of these models compared to models without and including both parameters were performed. Estimates were presented when relevant.

Genetic association

Genetic association analyses were performed in PLINK (36) v.1.08 using the –logistic command, based on allele frequencies and using the minor allele as the risk allele. The –within command was used to account for clustering within DZ twin pairs through producing robust standard errors using the Huber-White (37) To account for multiple testing, a p-value

below $3.5*10^{-4}$ (calculated based on Bonferroni correction for 9 phenotypes and 16 SNPs - 144 tests - with an alpha level of 0.05) was required for an observed effect to be deemed statistically significant. Genetic association analysis was carried out in two steps. All SNPs were tested for possible association with *any asthma*; then only SNPs significantly associated with *any asthma* were tested for association with the other asthma and related phenotypes. Due to the nature of these data, commonly used methods for calculating a cumulative percentage of the variance explained by the SNPs could not be used. Instead, the joint predictive value of significantly associated SNPs with regard to *any asthma* was examined by fitting a logistic regression model in Stata 13 including the risk alleles as covariates. Then, McFadden's pseudo-R^2 was computed in postestimation, the resulting value constituting a marker of the fit of a model with all the risk alleles included compared to one without. A small estimate indicates low total variance explained (38).

Results

Table 1 shows descriptive characteristics and phenotype prevalences for the full CATSS cohort, as well as for the subgroups included in the heritability (ACE/ADE models) and genetic association analyses, respectively. In the full cohort, 19.6% (n=4,956) had *any asthma,* 14.5% (n=3,663) had *wheezing ever*, and the prevalences of *hay fever*, *food allergy* and *atopic eczema ever* were 7.8 (n=1,969), 8.3 (n=2,105) and 13.1% (n=3,314) respectively.

Tetrachoric correlations within MZ and DZ twin pairs for the different phenotypes are displayed in Table 2. Based on these correlations, ADE models were selected for *asthma ever*, *asthma medication, hay fever ever* and *atopic eczema ever.* For all other phenotypes ACE models were fitted. Based on LRTs there was no significant improvement in the model fit of the ACE/ADE models compared to the simpler AE models for any phenotypes except *register-based asthma diagnosis* and *wheezing after three years of age*.

There were statistically significant additive genetic (A) components for all asthma phenotypes, with estimates ranging from 0.73 (*asthma medication,* 95% CI 0.68-0.78) to 0.83 (*register-based asthma diagnosis,* 95% CI 0.79-0.86), as well as for the related phenotypes, where the highest estimate was 0.78 (95% CI 0.74-0.82) for *food allergy*. Genetic dominance (D) was statistically significant in *atopic eczema ever*, where the estimate for the D component was 0.28, 95% CI 0.03-0.54. Only *register-based asthma diagnosis* (C 0.19, 95% CI 0.07-0.31) and *wheezing after three years of age* (C 0.18, 95% CI 0.02-0.33) had a statistically significant C component. The non-shared environment (E) ranged between 0.13 (*asthma ever,* 95% CI 0.11-0.16) and 0.27 (*asthma medication*, 95% CI 0.22-0.32).

Sex limitation models revealed some minor quantitative and qualitative differences in heritability between males and females for asthma. Sex-specific heritability estimates for the relevant asthma phenotypes are presented in Supporting information, Table S1.

Of the 16 tested SNPs in the first step of genetic association analysis, six were significantly associated with *any asthma* based on p-values below $3.5*10\text{-}^{4}$ (Table 3). These were rs3771180 in the *IL1RL1* gene on Chromosome 2 ($p= 1.0*10^{-5}$, OR 0.76, 95%CI 0.68-0.86,), and from Chromosome 17 rs12936231 in the *Z2BP2* gene ($p =1.5*10^{-7}$, OR 1.23, 95%CI 1.14-1.33), rs2305480 ($p = 1.5*10^{-8}$, OR 0.80, 95%CI 0.74-0.86), rs11078927 ($p=1.6*10^{-8}$, OR 0.80, 95%CI 0.75-0.86) and rs7216389 ($p=2.0*10^{-7}$, OR 1.23, 95%CI 1.14-1.33) in the *GSDMB* gene, and rs3894194 in the *GSDMA* gene ($p=1.0*10^{-5}$, OR 1.20, 95%CI 1.10-1.29). McFadden's pseudo-R^2 using the significantly associated SNPs was 0.008, indicating that the total predictive value of these genetic variants with regard to *any asthma* was very low.

When the six SNPs associated with *any asthma* were further investigated in relation to other asthma and related phenotypes, all were significantly associated with *asthma ever* (Table 4). Beyond this, the majority of SNPs located on Chromosome 17 were significantly associated with *register-based asthma diagnosis* and *asthma medication*. The only SNP associated with any of the related phenotypes was rs3771180 with *hay fever ever* (Table 4).

Discussion

We studied genetic effects underlying several childhood phenotypes using a combination of quantitative genetic and molecular methods, and saw significant additive genetic effects for asthma and related diseases. Genetic effects were strongest for asthma. Wheezing was among the phenotypes where the effect of genetics was less pronounced. We also confirmed some of the variants previously associated with childhood asthma from GWAS, although, as expected, the predictive value of these SNPs was very low. The SNPs were primarily associated with asthma and did not overlap with any of the related phenotypes save for *hay fever*, as rs3771180 in *IL1RL1* was associated with both.

The magnitude of the additive genetic (A) components for asthma in our study was in line with previous studies including children (2, 6, 7, 10, 11, 17, 19, 22). However, the overall pattern of variance components in asthma differs between studies. We saw no significant variability due to shared environment (C) except for *register-based asthma diagnosis*. This is in agreement with some (2, 10, 17, 19, 22), but not all (6, 7) previous studies involving children. One study of a particularly young population (3 years and below) showed a large contribution of shared environment to phenotypes similar to ours (diagnosis, medication use and hospitalization), indicating that shared environment may be more important at younger ages (7).

The SNPs in our material significantly associated with any asthma included the top two findings for childhood asthma from the GABRIEL consortium (rs2305480 and rs3894194) - one of few GWAS to report estimates separately for childhood asthma. (12) A later study confirmed association of rs2305480 and rs11078927 with asthma diagnosis, bronchial hyper-reactivity and asthma severity (39), providing support that genetic variation in this position is biologically and clinically relevant.

Studies on correlations of asthma, hay fever and eczema within relatives have confirmed that there is large comorbidity and shared additive genetic components in adult (16, 18) and child populations (22, 28), and some GWAS have found loci shared by asthma and atopy (40-42). As we tested a limited number of SNPs our results cannot be used to draw extensive conclusions on genetic variants shared between asthma and related phenotypes, sincecomorbidity between diseases was not used as an outcome in this study.

We could not replicate all the previously genetic associations in our material, including rs744910 in SMAD3. One possible explanation is lack of statistical power. Assuming a phenotype prevalence of 10% and a risk allele frequency of 0.40, we had 76-99% power to detect an OR of 1.20-1.30. These ORs are based on previous literature, and the power presented to detect a statistically significant effect in a material of our size with a desired alpha level of $3.5*10^{-4}$ (44). Considering the actual allele frequencies in this study, the power to confirm previous associations in our material was likely slightly greater. Another potential explanation for non-replication is that these variants, originally identified in GWAS with pooled data from cohorts of diverse origin, may be less important in a more homogenous Swedish population. More studies within populations similar to ours would be required to confirm this.

This is one of the largest studies on heritability of asthma and related allergic diseases in children, providing more robust estimates for many of the phenotypes than were previously available. To our knowledge this is the first study providing an estimate specifically for the heritability of childhood wheeze; previous studies have only included adults (8, 16, 20). The presence of a large, population-based twin cohort assured a representative sample within a highly relevant age range. Information on several phenotypes from questionnaires as well as registers gives us a unique opportunity to draw multifaceted conclusions regarding these

prevalent causes of childhood morbidity. It is sometimes suggested results from twin studies would be less generalisable to the remaining population, but we have recently shown this is not a major concern (43). The register-based asthma outcomes we applied were originally validated based on medical records using clinical criteria, with high predictive values for clinically relevant pediatric asthma (32).

Results from twin models should be interpreted in the light of their assumptions and limitations These include the equal environment assumption (MZ and DZ twins are assumed to have the same degree of shared environment), an absence of gene-environment interaction, and that the specific underlying mechanisms involved in each component (A, C or D, and E) are not identified. In addition, in this study, there could also be a potential effect of rater bias for the parent-reported outcomes. A limitation of our definition of wheezing is that the question was only posed to parents who had not already answered "yes" to the question about asthma ever. This represents therefore a particular subtype of wheeze—one which has not, by the age of 9 or 12, become or at least not been acknowledged as asthma. Such wheezing may be due to different underlying biological mechanisms and thus interesting in its own right; however, this difference between our definition of wheeze as compared to previous studies should be taken into account. It is possible to identify eczema medication use through Swedish registers (45) but we did not have access to a validated medication proxy for eczema, and direct clinical measures such as FeNO or lung function were not available for this population.

Through the use of a combination of twin methodology and molecular genetics, this study confirms underlying additive genetic effects behind asthma and related phenotypes in mid childhood. We were able to replicate previously strong associations between variants at the *ORMDL3* and *IL1RL1* regions and asthma across several different definitions. These genetic variants appeared to be specific to asthma with the exception of one SNP that was also

associated with hay fever. Our findings support these highly correlated SNPs remain as suitable targets for the study of genetic variation in relation to epidemiological and clinical studies of asthma.

References

1. Melen E, Pershagen G. Pathophysiology of asthma: lessons from genetic research with particular focus on severe asthma. J Intern Med. 2012;272(2):108-20.

2. Lichtenstein P, Svartengren M. Genes, environments, and sex: factors of importance in atopic diseases in 7-9-year-old Swedish twins. Allergy. 1997;52(11):1079-86.

3. Pinart M, Benet M, Annesi-Maesano I, von Berg A, Berdel D, Carlsen KC, et al. Comorbidity of eczema, rhinitis, and asthma in IgE-sensitised and non-IgE-sensitised children in MeDALL: a population-based cohort study. The Lancet Respiratory medicine. 2014;2(2):131-40.

4. Henderson J, Granell R, Sterne J. The search for new asthma phenotypes. Arch Dis Child. 2009;94(5):333-6.

5. Spycher BD, Henderson J, Granell R, Evans DM, Smith GD, Timpson NJ, et al. Genome-wide prediction of childhood asthma and related phenotypes in a longitudinal birth cohort. J Allergy Clin Immunol. 2012;130(2):503-9 e7.

6. Thomsen SF, van der Sluis S, Kyvik KO, Skytthe A, Backer V. Estimates of asthma heritability in a large twin sample. Clin Exp Allergy. 2010;40(7):1054-61.

7. Bunyavanich S, Silberg JL, Lasky-Su J, Gillespie NA, Lange NE, Canino G, et al. A twin study of early-childhood asthma in Puerto Ricans. PloS one. 2013;8(7):e68473.

8. Duffy DL, Martin NG, Battistutta D, Hopper JL, Mathews JD. Genetics of asthma and hay fever in Australian twins. Am Rev Respir Dis. 1990;142(6 Pt 1):1351-8.

9. Harris JR, Magnus P, Samuelsen SO, Tambs K. No evidence for effects of family environment on asthma. A retrospective study of Norwegian twins. Am J Respir Crit Care Med. 1997;156(1):43-9.

10. Fagnani C, Annesi-Maesano I, Brescianini S, D'Ippolito C, Medda E, Nistico L, et al. Heritability and shared genetic effects of asthma and hay fever: an Italian study of young twins. Twin Res Hum Gen. 2008;11(2):121-31.

11. Laitinen T, Rasanen M, Kaprio J, Koskenvuo M, Laitinen LA. Importance of genetic factors in adolescent asthma: a population-based twin-family study. Am J Respir Crit Care Med. 1998;157(4 Pt 1):1073-8.

12. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. N Engl J Med. 2010;363(13):1211-21.

13. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. Nat Genet. 2011;43(9):887-92.

14. Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souef P, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. Lancet. 2011;378(9795):1006-14.

15. Posthuma D, Beem AL, de Geus EJ, van Baal GC, von Hjelmborg JB, Iachine I, et al. Theory and practice in quantitative genetics. Twin Res. 2003;6(5):361-76.

16. Thomsen SF, Ulrik CS, Kyvik KO, Ferreira MA, Backer V. Multivariate genetic analysis of atopy phenotypes in a selected sample of twins. Clin Exp Allergy. 2006;36(11):1382-90.

17. Skadhauge LR, Christensen K, Kyvik KO, Sigsgaard T. Genetic and environmental influence on asthma: a population-based study of 11,688 Danish twin pairs. Eur Resp J. 1999;13(1):8-14.

18. Willemsen G, van Beijsterveldt TC, van Baal CG, Postma D, Boomsma DI. Heritability of self-reported asthma and allergy: a study in adult Dutch twins, siblings and parents. Twin Res Hum Genet. 2008;11(2):132-42.

19. Koeppen-Schomerus G, Stevenson J, Plomin R. Genes and environment in asthma: a study of 4 year old twins. Arch Dis Child. 2001;85(5):398-400.

20. Nystad W, Roysamb E, Magnus P, Tambs K, Harris JR. A comparison of genetic and environmental variance structures for asthma, hay fever and eczema with symptoms of the same diseases: a study of Norwegian twins. Int J Epidemiol. 2005;34(6):1302-9.

21. Thomsen SF, van der Sluis S, Kyvik KO, Skytthe A, Skadhauge LR, Backer V. Increase in the heritability of asthma from 1994 to 2003 among adolescent twins. Respir. 2011;105(8):1147-52.

22. van Beijsterveldt CE, Boomsma DI. Genetics of parentally reported asthma, eczema and rhinitis in 5-yr-old twins. Eur Resp J 2007;29(3):516-21.

23. Rasanen M, Laitinen T, Kaprio J, Koskenvuo M, Laitinen LA. Hay fever--a Finnish nationwide study of adolescent twins and their parents. Allergy. 1998;53(9):885-90.

24. Galanter JM, Gignoux CR, Torgerson DG, Roth LA, Eng C, Oh SS, et al. Genome-wide association study and admixture mapping identify different asthma-associated loci in Latinos: the Genes-environments & Admixture in Latino Americans study. J Allergy Clin Immunol. 2014;134(2):295-305.

25. Wan YI, Shrine NR, Soler Artigas M, Wain LV, Blakey JD, Moffatt MF, et al. Genome-wide association study to identify genetic determinants of severe asthma. Thorax. 2012;67(9):762-8.

26. Mathias RA, Grant AV, Rafaels N, Hand T, Gao L, Vergara C, et al. A genome-wide association study on African-ancestry populations for asthma. J Allergy Clin Immunol. 2010;125(2):336-46.e4.

27. Verlaan DJ, Berlivet S, Hunninghake GM, Madore AM, Lariviere M, Moussette S, et al. Allele-specific chromatin remodeling in the ZPBP2/GSDMB/ORMDL3 locus associated with the risk of asthma and autoimmune disease. Am J Hum Genet. 2009;85(3):377-93.

28. Fuertes E, Soderhall C, Acevedo N, Becker A, Brauer M, Chan-Yeung M, et al. Associations between the 17q21 region and allergic rhinitis in 5 birth cohorts. J Allergy Clin Immunol. 2014.

29. Anckarsater H, Lundstrom S, Kollberg L, Kerekes N, Palm C, Carlstrom E, et al. The Child and Adolescent Twin Study in Sweden (CATSS). Twin Res Hum Genet 2011;14(6):495-508.

30. Magnusson PK, Almqvist C, Rahman I, Ganna A, Viktorin A, Walum H, et al. The Swedish twin registry: establishment of a biobank and other recent developments. Twin Res Hum Genet. 2013;16(1):317-29.

31. Hannelius U, Gherman L, Makela VV, Lindstedt A, Zucchelli M, Lagerberg C, et al. Large-scale zygosity testing using single nucleotide polymorphisms. Twin Res Hum Genet 2007;10(4):604-25.

32. Ortqvist AK, Lundholm C, Wettermark B, Ludvigsson JF, Ye W, Almqvist C. Validation of asthma and eczema in population-based Swedish drug and patient registers. Pharmacoepidemiol Drug Saf. 2013;22(8):850-60.

33. Lundholm C, Ortqvist AK, Lichtenstein P, Cnattingius S, Almqvist C. Impaired fetal growth decreases the risk of childhood atopic eczema: a Swedish twin study. Clin Exp Allergy. 40.2010. p. 1044-53.

34. Levin AM, Mathias RA, Huang L, Roth LA, Daley D, Myers RA, et al. A meta-analysis of genome-wide association studies for serum total IgE in diverse study populations. J Allergy Clin Immunol. 2013;131(4):1176-84.

35. Boker S, Neale M, Maes H, Wilde M, Spiegel M, Brick T, et al. OpenMx: An Open Source Extended Structural Equation Modeling Framework. Psychometrika. 2011;76(2):306-17.

36. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559-75.

37. Livshits G, Peter I, Vainder M, Hauspie R. Genetic analysis of growth curve parameters of body weight, height and head circumference. Ann Hum Biol 2000;27(3):299-312.

38. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, editor. Frontiers in Econometrics: Academic Press; 1974. p. Pp. 105-42.

39. Tulah AS, Holloway JW, Sayers I. Defining the contribution of SNPs identified in asthma GWAS to clinical variables in asthmatic children. BMC Med Genet. 2013;14:100.

40. Ferreira MA, Matheson MC, Tang CS, Granell R, Ang W, Hui J, et al. Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. J Allergy Clin Immunol. 2014;133(6):1564-71.

41. Weidinger S, Willis-Owen SA, Kamatani Y, Baurecht H, Morar N, Liang L, et al. A genome-wide association study of atopic dermatitis identifies loci with overlapping effects on asthma and psoriasis. Hum Mol Genet. 2013;22(23):4841-56.

42. Tang HY, Tang XF, Zuo XB, Gao JP, Sheng YJ, Li Y, et al. Association analysis of single nucleotide polymorphisms at five loci: comparison between atopic dermatitis and asthma in the Chinese Han population. PloS one. 2012;7(4):e35334.

43. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics. 2003;19(1):149-50

44. Ullemar V, Lundholm C, Almqvist C. Twins' Risk of Childhood Asthma Mediated By Gestational Age and Birth Weight. Clin Exp Allergy 2015.

45. Henriksen L, Simonsen J, Haerskjold A, Linder M, Kieler H, Thomsen SF, et al. Incidence rates of atopic dermatitis, asthma, and allergic rhinoconjunctivitis in Danish and Swedish children. J Allergy Clin Immunol 2015;136(2):360-366 e362.

Table 1 Background characteristics of the study population and phenotype prevalences as distributed in the subgroups within the study; the full cohort, those selected for ACE/ADE models, those with genotype data available and finally those randomly selected from this group for genetic association analysis (removing one MZ twin from each MZ pair).

	Full CATSS cohort N (%)	Selected for ACE / ADE models N (%)	Genotype data available N (%)	Randomly selected for genetic association analysis N (%)
All	25,306 (100)	15,896 (100)	12,388 (100)	10,075 (100)
Characteristics				
Gender				
Male	12,867 (50.9)	8,168 (51.4)	6,196 (50.0)	5,104 (50.7)
Female	12,439 (49.1)	7,728 (48.6)	6,192 (50.0)	4,971 (49.3)
Age at telephone interview, years				
9	18,786 (74.2)	11,546 (72.6)	9,674 (70.0)	7,121 (70.7)
12	6,520 (25.8)	4,350 (27.4)	3,714 (30.0)	2,954 (29.3)
Zygosity				
MZ	7,042 (27.8)	7,042 (44.3)	4,627 (37.4)	2,314 (23.0)
DZ-same sex	8,854 (35.0)	8,854 (55.7)	4,205 (34.0)	4,205 (41.7)
DZ-opposite sex	8,834 (34.9)	-	3,521 (28.4)	3,521 (35.0)
Unknown	576 (2.3)	-	35 (0.3)	35 (0.3)
Asthma				
Interview-based				
Asthma ever				
Yes	3,519 (13.9)	2,181 (13.7)	1,725 (13.9)	1,416 (14.1)
No	21,582 (85.3)	13,584 (85.5)	10,579 (85.4)	8,592 (85.3)
Missing	205 (0.8)	131 (0.8)	84 (0.7)	65 (0.7)
Register-based				
Asthma medication				
Yes	2,479 (9.8)	1,507 (9.5)	1,304 (10.5)	1,102 (10.9)
No	22,827 (90.2)	14,389 (90.5)	11,084 (89.5)	8,973 (89.1)
Asthma diagnosis				
Yes	2,743 (10.8)	1,710 (10.8)	1,402 (11.3)	1,169 (11.6)

No	22,563 (89.2)	14,186 (89.2)	10,986 (88.7)	8,906 (88.4)
Any asthma (interview or register)				
Yes	4,956 (19.6)	3,056 (19.2)	2,493 (20.1)	2,061 (20.5)
No	20,217 (79.9)	12,757 (80.3)	9,845 (79.5)	7,975 (79.2)
Missing[1]	133 (0.5)	83 (0.5)	50 (0.4)	39 (0.4)
Related phenotypes (interview)				
Wheezing ever				
Yes	3,663 (14.5)	2,336 (14.7)	1,838 (14.8)	1,479 (14.7)
No	17,978 (71.0)	11,290 (71.0)	8,770 (70.8)	7,129 (70.8)
Missing	3,665 (14.5)	2,270 (14.3)	1,780 (14.4)	1,467 (14.6)
Wheezing after 3 years of age				
Yes	1,962 (7.8)	1,223 (7.7)	967 (7.8)	790 (7.8)
No	19,173 (75.8)	12,065 (75.9)	9,352 (75.5)	7,593 (75.4)
Missing	4,171 (16.5)	2,608 (16.4)	2,069 (16.7)	1,692 (16.8)
Hay fever ever				
Yes	1,969 (7.8)	1,234 (7.8)	990 (8.0)	813 (8.1)
No	23,050 (91.1)	14,473 (91.1)	11,278 (91.0)	9,157 (90.1)
Missing	287 (1.1)	189 (1.2)	120 (1.0)	105 (1.0)
Food allergy ever				
Yes	2,105 (8.3)	1,315 (8.3)	1,023 (8.3)	828 (8.2)
No	23,010 (90.9)	14,473 (91.1)	11,276 (91.0)	9,177 (91.1)
Missing	191 (0.8)	108 (0.7)	89 (0.7)	70 (0.7)
Atopic eczema ever				
Yes	3,314 (13.1)	2,093 (13.2)	1,769 (14.3)	1,428 (14.2)
No	21,848 (86.3)	13,712 (86.3)	10,554 (85.2)	8,602 (85.4)
Missing	144 (0.6)	91 (0.6)	65 (0.5)	45 (0.5)

Table 2 Tetrachoric correlations and twin models for asthma and related phenotypes using full MZ and same-sex DZ twin pairs from CATSS-9/12 (n=15,896)

					Variance component estimate (95% CI)			Fit of model	
	N (%)	rMZ	rDZ-ss	Model	A	C or D	E	-2lnL	P-value*
Total	15,896 (100)								
Asthma									
Interview-based									
Asthma ever	2,181 (13.7)	0.87	0.42	ADE	0.79 (0.55-0.89)	0.08 (0.00-0.33)	0.13 (0.10-0.16)	11543.67	0.50
				AE	0.87 (0.84-0.89)	-	0.13 (0.11-0.16)	11544.12	
Register-based									
Asthma medication	1,507 (9.5)	0.74	0.32	ADE	0.51 (0.22-0.76)	0.23 (0.00-0.54)	0.25 (0.21-0.31)	9505.79	0.12
				AE	0.73 (0.68-0.78)	-	0.27 (0.22-0.32)	9508.17	
Asthma diagnosis	1,710 (10.8)	0.81	0.51	ACE	0.62 (0.48-0.76)	0.19 (0.07-0.31)	0.19 (0.15-0.23)	10033.95	<0.01
				AE	0.83 (0.79-0.86)	-	0.17 (0.14-0.21)	10043.39	
Any asthma (interview or register)	3,056 (19.2)	0.81	0.45	ACE	0.73 (0.61-0.84)	0.09 (0.00-0.19)	0.19 (0.16-0.22)	14351.11	0.10
				AE	0.82 (0.79-0.85)	-	0.18 (0.15-0.21)	14353.82	
Related phenotypes						-			
Wheezing ever	2,336 (14.7)	0.75	0.44	ACE	0.63 (0.49-0.78)	0.12 (0.00-0.24)	0.25 (0.21-0.29)	11791.01	0.07
				AE	0.76 (0.72-0.80)	-	0.24 (0.20-0.28)	11794.18	
Wheezing after 3 years of age	1,223 (7.7)	0.77	0.49	ACE	0.60 (0.42-0.78)	0.18 (0.02-0.33)	0.22 (0.17-0.28)	7698.00	0.03
				AE	0.80 (0.75-0.84)	-	0.20 (0.16-0.25)	7702.98	
Hay fever ever	1,234 (7.8)	0.78	0.39	ADE	0.78 (0.48-0.82)	0.00 (0.00-0.31)	0.22 (0.18-0.27)	8108.884	0.99
				AE	0.78 (0.73-0.82)	-	0.22 (0.18-0.27)	8108.884	
Food allergy ever	1,315 (8.3)	0.78	0.40	ACE	0.75 (0.57-0.82)	0.03 (0.00-0.19)	0.22 (0.18-0.27)	8492.27	0.67
				AE	0.78 (0.74-0.82)	-	0.22 (0.18-0.26)	8492.45	
Atopic eczema ever	2,093 (13.2)	0.83	0.35	ADE	0.56 (0.30-0.81)	0.28 (0.03-0.54)	0.16 (0.13-0.19)	11436.76	0.03
				AE	0.83 (0.80-0.86)	-	0.17 (0.14-0.20)	11441.51	

Notes:
* P-value for likelihood ratio test comparing the null model (AE) to the alternate model (ACE or ADE). A significant result in the likelihood ratio test means the alternate model has a better fit to the data.
Abbreviations:
-2lnL: minus twice the log likelihood
A: additive genetic component

C: shared environment component
D: genetic dominance component
E: unique environment component

Table 3 Association of the successfully genotyped SNPs with *any asthma* in genotyped twins and (one randomly selected MZ twin per pair), Eexcluding individuals with missing data on more than 3 SNPs, n=415, after which n=9,660 individuals remained. Of these, 2,257 (23.4%) were MZ, 7,375 (76.3%) were DZ and 28 (0.3%) were of unknown zygosity. Controls are unaffected by all other phenotypes (including within same "group" of phenotypes)

CHR	Gene	Location	SNP	Minor allele	MAF in cases	MAF in controls	P	OR	(95% CI)
1	IL6R	Intron	rs4129267	T	0.39	0.39	0.55	1.02	(0.95 - 1.11)
2	IL1RL1	Intron	rs3771180	T	0.11	0.14	**$1.0 * 10^{-5}$**	0.76	(0.68 - 0.86)
2	IL18R1	Intron	rs3771166	A	0.36	0.37	0.64	0.98	(0.91 - 1.06)
5	TSLP	Nearby region	rs1837253	T	0.27	0.27	0.58	0.98	(0.89 - 1.07)
5	SLC22A5	Intron	rs2073643	T	0.50	0.49	0.10	1.07	(0.99 - 1.15)
5	IL13	Intron	rs1295686	T	0.22	0.21	0.15	1.07	(0.98 - 1.18)
9	IL33	Nearby region	rs1342326	C	0.16	0.14	0.02	1.14	(1.02 - 1.26)
9	IL33	Nearby region	rs2381416	C	0.26	0.24	0.04	1.10	(1.00 - 1.20)
15	RORA	Intron	rs11071559	T	0.12	0.13	0.09	0.90	(0.80 - 1.02)
15	SMAD3	Intron	rs744910	A	0.50	0.49	0.86	1.01	(0.93 - 1.09)
17	Z2BP2	Intron	rs12936231	C	0.52	0.47	**$1.5 * 10^{-7}$**	1.23	(1.14 - 1.33)
17	GSDMB	Exon	rs2305480	A	0.42	0.47	**$1.5 * 10^{-8}$**	0.80	(0.74 - 0.86)
17	GSDMB	Intron	rs11078927	T	0.42	0.47	**$1.6 * 10^{-8}$**	0.80	(0.74 - 0.86)
17	GSDMB	Intron	rs7216389	T	0.53	0.47	**$2.0 * 10^{-7}$**	1.23	(1.14 - 1.33)
17	GSDMA	Exon	rs3894194	A	0.50	0.45	**$1.0 * 10^{-5}$**	1.20	(1.10 - 1.29)
22	IL2RB	Intron	rs2284033	A	0.44	0.45	0.10	0.94	(0.87 - 1.01)

CHR: Chromosome
SNP: Single nucleotide polymorphism
MAF: Minor allele frequency
L95 and U95: Lower and upper bound of the 95% confidence interval of the odds ratio

Table 4 Association of the six SNPs most strongly associated with *any asthma* with all other phenotypes in the first half of the twin data, excluding individuals with missing data on more than 3 SNPs, n=415, after which n=9,660 individuals remained. Of these, 2,257 (23.4%) were MZ, 7,375 (76.3%) were DZ and 28 (0.3%) were of unknown zygosity. Controls are unaffected by all other phenotypes (including within same "group" of phenotypes)-.

	rs3771180		rs12936231		rs2305480		rs11078927		rs7216389		rs3894194	
Phenotype	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
Asthma ever	0.73 (0.63-0.84)	**$1.5*10^{-5}$**	1.27 (1.17-1.40)	**$1.3*10^{-7}$**	0.76 (0.70-0.83)	**$2.8*10^{-9}$**	0.76 (0.70-0.83)	**$3.9*10^{-9}$**	1.29 (1.18-1.41)	**$5.1*10^{-8}$**	1.31 (1.19-1.43)	**$9.8*10^{-9}$**
Asthma medication	0.81 (0.70-0.94)	$6.6*10^{-3}$	1.24 (1.12-1.38)	**$2.5*10^{-5}$**	0.80 (0.73-0.89)	**$1.3*10^{-5}$**	0.80 (0.73-0.88)	**$1.8*10^{-5}$**	1.22 (1.10-1.35)	**$9.9*10^{-5}$**	1.17 (1.06-1.28)	$3.3*10^{-3}$
Register-based asthma diagnosis	0.78 (0.67-0.90)	$1.1*10^{-3}$	1.24 (1.13-1.37)	**$1.1*10^{-5}$**	0.78 (0.71-0.86)	**$8.8*10^{-7}$**	0.78 (0.71-0.86)	**$1.0*10^{-6}$**	1.25 (1.14-1.38)	**$6.0*10^{-6}$**	1.19 (1.08-1.31)	$5.7*10^{-4}$
Wheezing ever	0.97 (0.86-1.10)	0.61	1.02 (0.94-1.11)	0.65	0.95 (0.87-1.04)	0.25	0.95 (0.87-1.43)	0.24	1.01 (0.93-1.10)	0.79	1.00 (0.91-1.09)	0.95
Wheezing after 3 years of age	0.96 (0.82-1.12)	0.60	1.04 (0.93-1.16)	0.54	0.92 (0.83-1.03)	0.16	0.92 (0.83-1.04)	0.17	1.00 (0.90-1.12)	0.94	0.99 (0.89-1.11)	0.90
Hay fever ever	0.64 (0.53-0.77)	**$2.5*10^{-6}$**	1.10 (0.99-1.23)	0.10	0.89 (0.80-0.99)	0.04	0.89 (0.80-0.99)	0.04	1.10 (0.98-1.23)	0.11	1.14 (1.23-1.28)	0.02
Food allergy ever	0.89 (0.76-1.05)	0.17	1.12 (1.00-1.25)	0.06	0.89 (0.80-0.99)	0.03	0.88 (0.79-0.99)	0.03	1.12 (1.00-1.25)	0.04	1.12 (1.00-1.25)	0.05
Atopic eczema ever	0.92 (0.81-1.05)	0.21	1.01 (0.93-1.10)	0.82	0.96 (0.88-1.05)	0.42	0.96 (0.88-1.05)	0.38	1.04 (0.95-1.13)	0.44	1.07 (0.98-1.17)	0.14