Genome-wide study of early and severe childhood asthma identifies interaction between CDHR3 and GSDMB

Eliasen, Anders U.; Pedersen, Casper Emil T.; Rasmussen, Morten A.; Wang, Ni; Soverini, Matteo; Fritz, Amelie; Stokholm, Jakob; Chawes, Bo L.; Morin, Andréeanne; Bork-Jensen, Jette; Grarup, Niels; Pedersen, Oluf; Hansen, Torben; Linneberg, Allan; Mortensen, Preben B.; Hougaard, David M.; Bybjerg-Grauholm, Jonas; Bækvad-Hansen, Marie; Mors, Ole; Nørdentoft, Merete; Børglum, Anders D.; Werge, Thomas; Agerbo, Esben; Söderhall, Cilla; Altman, Matthew C.; Thysen, Anna H.; McKennan, Chris G.; Brix, Susanne; Gern, James E.; Ober, Carole; Ahluwalia, Tarunveer S.; Bisgaard, Hans; Pedersen, Anders G.; Bønnelykke, Klaus

Published in: Journal of Allergy and Clinical Immunology

DOI: 10.1016/j.jaci.2022.03.019

Publication date: 2022

Document version Peer reviewed version

Document license: CC BY-NC-ND

Genome-wide study of early and severe childhood asthma identifies interaction between \textit{CDHR3} and \textit{GSDMB}.

Genome-wide study of early and severe childhood asthma identifies interaction between
CDHR3 and GSDMB

Population
COPSACserve
Recurrent asthma hospitalizations age 2-6 years
1204 cases & 5328 controls

Search for genetic interactions
Genome-wide search for pairwise interactions using Bayesian logistic regression

Interaction between CDHR3 & GSDMB
Increased risk of early childhood asthma
Viral stimulation in peripheral whole blood
A combination of CDHR3 & GSDMB risk alleles increase IL17-A production during viral infections
## Main tables

### Table 1. Top results from the genome-wide search for genetic interaction effects

<table>
<thead>
<tr>
<th>Interacting loci</th>
<th>RsID</th>
<th>Position</th>
<th>EA (EAF)</th>
<th>Discovery COPSACsevere</th>
<th>Replication</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>OR (95% credible interval)</td>
<td>PP</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IP SYCH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N = 127,914 (5729/122,185)</td>
</tr>
<tr>
<td>CDHR3 (\times) GSDMB*</td>
<td>rs6967330</td>
<td>7q22</td>
<td>A (0.19)</td>
<td>1.28 [1.01;1.63]</td>
<td>90%</td>
</tr>
<tr>
<td></td>
<td>rs2305480</td>
<td>17q21</td>
<td>G (0.57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC22A3 (\times) GSDMB</td>
<td>rs7758229</td>
<td>6q25</td>
<td>T (0.36)</td>
<td>1.25 [1.08;1.46]</td>
<td>96%</td>
</tr>
<tr>
<td></td>
<td>rs2305480</td>
<td>17q21</td>
<td>G (0.57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCNH (\times) ART1</td>
<td>rs2230641</td>
<td>5q14</td>
<td>A (0.76)</td>
<td>1.24 [1.07;1.46]</td>
<td>95%</td>
</tr>
<tr>
<td></td>
<td>rs2280134</td>
<td>11p15</td>
<td>T (0.35)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PP: Posterior probability, representing the probability that the interaction is outside the region of practical equivalence in the discovery analysis. P : represents two sided p-values for the discovery analysis and one-sided p-values for the replication. EA : Effect allele. EAF : Effect allele frequency. The combined replication analysis is based on Bayesian sequential updating, and an inverse-variance weighted meta-analysis to obtain the one sided p-values using a fixed effects model. Numbers in the brackets represent the number of cases and controls respectively. * : discovery based on 379 cases with six or more hospitalizations due to severe asthma exacerbations. ** indicates significant replications after adjusting for multiple testing.
Table 2. **CDHR3 effect and CDHR3 x GSDMB interaction in relation to age of asthma onset.** Analyses are based on the replication cohorts, iPSYCH and UK Biobank. The tables show A) the CDHR3 (rs6967330/A) main effect in strata with different age of asthma onset, and B) the CDHR3 (rs6967330) x GSDMB (rs2305480) interaction effect in strata with different age of asthma onset. Effect estimates are based on additive genetic models. Effects are from additive genetic models with the asthma-associated allele as effect allele.

A) Main effect of **CDHR3** on asthma risk

<table>
<thead>
<tr>
<th>Age of asthma onset</th>
<th>iPSYCH (replication)</th>
<th>UK Biobank (replication)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N_cases</td>
<td>OR   [95% CI]</td>
</tr>
<tr>
<td>0-2</td>
<td>3177</td>
<td>1.30 [1.23; 1.38]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P = 1.59e-18</td>
</tr>
<tr>
<td>2-4</td>
<td>1505</td>
<td>1.28 [1.18; 1.40]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P = 1.16e-08</td>
</tr>
<tr>
<td>4-6</td>
<td>1047</td>
<td>1.06 [1.00; 1.25]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P = 0.04</td>
</tr>
<tr>
<td>&gt; 6</td>
<td>4771</td>
<td>1.04 [0.98; 1.10]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P = 0.16</td>
</tr>
</tbody>
</table>
B) Interaction effect between CDHR3 (rs6967330) and GSDMB (rs2305480) on asthma risk

<table>
<thead>
<tr>
<th>Age of asthma onset (years)</th>
<th>iPSYCH (replication)</th>
<th>UK Biobank (replication)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N cases</td>
<td>OR [95% CI]</td>
</tr>
<tr>
<td>0-2</td>
<td>3177</td>
<td>1.15 [1.06; 1.26]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P = 1.15e-03</td>
</tr>
<tr>
<td>2-4</td>
<td>1505</td>
<td>1.01 [0.89; 1.15]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P = 0.83</td>
</tr>
<tr>
<td>4-6</td>
<td>1047</td>
<td>1.13 [0.96; 1.32]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P = 0.13</td>
</tr>
<tr>
<td>&gt; 6</td>
<td>4771</td>
<td>1.08 [0.99; 1.16]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P = 0.06</td>
</tr>
</tbody>
</table>
Genome-wide study of early and severe childhood asthma identifies interaction between CDHR3 and GSDMB

PhD Anders U. Eliasen¹,², PhD Casper Emil T. Pedersen¹, PhD Morten A. Rasmussen¹,³, PhD Ni Wang¹, PhD Matteo Soverini¹, MSc Amelie Fritz¹,², MD, PhD Jakob Stokholm¹, MD, PhD, DMSc Bo L. Chawes¹, PhD Andreanna Morin⁴, PhD Jette Bork-Jensen⁵, MD, PhD, DMSc Niels Grarup⁵, MD, PhD Oluf Pedersen⁵, MD, PhD Torben Hansen⁵, MD, PhD Allan Linneberg⁶,⁷, MD, PhD, DMSc Preben B. Mortensen⁸,⁹,¹⁰, MD, PhD, DMSc David M. Hougaard⁸,¹¹,¹², MSc Jonas Bybjerg-Grauholm⁸,¹², PhD Marie Bækvad-Hansen⁸,¹¹, MD, PhD Ole Mors⁸,¹³, MD, PhD Merete Nordentoft⁸,¹⁴, MD, PhD Anders D. Børglum⁷,⁸,¹⁵, MD, PhD Thomas Werge⁸,¹⁶,¹⁷, MD, PhD, DMSc Esben Agerbo⁸,⁹,¹⁰, PhD Cilla Söderhall¹⁸, MD, PhD Matthew C. Altman¹⁹, PhD Anna H. Thysen¹,²⁰, PhD Chris G. McKennan²¹, PhD Susanne Brix²⁰, PhD James E. Gern²², PhD Carole Ober¹, PhD Tarunveer S. Ahluwalia²³,²⁴, MD, DMSc Hans Bisgaard¹, PhD Anders G. Pedersen²⁵, MD, PhD Klaus Bønnelykke¹

*These authors contributed equally to the study

Affiliation:
¹COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark
²Department of Health Technology, Section for Bioinformatics, Technical University of Denmark, Kgs. Lyngby, Denmark
³Department of Food Science, Faculty of Science, University of Copenhagen, Denmark
⁴Department of Human Genetics, University of Chicago, Chicago, IL, USA
⁵Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Copenhagen, Denmark
⁶Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, The Capital Region, Copenhagen, Denmark
⁷Department of Clinical Medicine, University of Copenhagen, Denmark
⁸iPSYCH, The Lundbeck Foundation Initiative for Integrated Psychiatric Research
⁹NCRR, the National Center for Register-based research, Business and Social Sciences, Aarhus University, Denmark
¹⁰CIRRAU - Center for Integrated Register-Based Research at Aarhus University, Denmark
11 Department for Congenital Disorders, Statens Serum Institut, Copenhagen, Denmark
12 Den neonatale screenings biobank, SSI
13 Psychosis Research Unit, Aarhus University Hospital - Psychiatry, Risskov, Denmark
14 Mental Health Center Copenhagen, Capital Region of Denmark, Copenhagen University Hospital, Denmark
15 Center for Integrative Sequencing, Department of Biomedicine and iSEQ, Aarhus University, Denmark
16 Center for Genomics and Personalized Medicine, Aarhus, Central Region Denmark, Denmark
17 Institute of Biological Psychiatry, Copenhagen Mental Health Services, Copenhagen, Denmark
18 Department of Womens’ and Children’s health, Karolinska Institutet, Stockholm, Sweden
19 Department of Medicine, University of Washington, Seattle WA, USA
20 Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark
21 Departments of Statistics, The University of Chicago, Chicago, Ill
22 Department of Pediatrics, University of Wisconsin, Madison, WI, USA
23 Steno Diabetes Center Copenhagen, Gentofte, Denmark
24 The Bioinformatics Center, Department of Biology, University of Copenhagen, Copenhagen, Denmark

#Corresponding author:
Klaus Bønnelykke, MD, PhD
Address: Ledreborg Allé 34, 2820 Gentofte, Denmark
E-mail: kb@copsac.com
Phone: +45 3867 7360
Fax: (+45) 3977 7129
Website: www.copsac.com

Funding
All funding received by COPSAC is listed on www.copsac.com. The Lundbeck Foundation (Grant no R16-A1694); The Ministry of Health (grant no 903516); Danish Council for Strategic Research (Grant no 0603-00280B) and The Capital Region Research Foundation have provided core support
to the COPSAC research center. The iPSYCH project is funded by the Lundbeck Foundation (grant numbers R102-A9118 and R155-2014-1724) and the universities and university hospitals of Aarhus and Copenhagen. The Danish National Biobank resource was supported by the Novo Nordisk Foundation. National Institutes of Health (NIH) (grant no. HL129735 to H.B, KB, and C.O). The Lundbeck Foundation (grant no. R263-2017-3800 to A.U.E.). The Lundbeck Foundation (grant no. R163-2013-16235 to K.B.). The Novo Nordisk Foundation (grant. no. NNF18OC0031764 to K.B.)

**Competing Interest**

All authors declare no potential, perceived, or real conflict of interest regarding the content of this manuscript. The funding agencies did not have any role in the design and conduct of the study; collection, management, and interpretation of the data; or preparation, review, or approval of the manuscript. No pharmaceutical company was involved in the study.

Hans Bisgaard: Has consultant arrangements with Chiesi Pharmaceuticals and Boehringer Ingelheim.

**Total word count: 4077**
Abstract

Background
Asthma with severe exacerbation is one of the most common causes of hospitalization among young children. Exacerbations are typically triggered by respiratory infections, but the host factors causing recurrent infections and exacerbations in some children are poorly understood. As a result, current treatment options and preventive measures are inadequate.

Objective
To identify genetic interaction associated with the development of childhood asthma.

Methods
We performed an exhaustive search for pairwise interaction between genetic variants, single nucleotide polymorphisms (SNPs), using 1204 cases with a specific phenotype of early childhood asthma with severe exacerbations (age 2-6 years) combined with 5328 non-asthmatic controls. Replication was attempted in 3 independent populations, and potential underlying immune mechanisms were investigated in the COPSAC2010 and COPSAC2000 birth cohorts.

Results
We found evidence of interaction, including replication in independent populations, between the known childhood asthma loci, CDHR3 and GSDMB. The effect of CDHR3 was dependent on the GSDMB genotype and this interaction was more pronounced for severe and early onset of disease. Blood immune analyses suggested a mechanism related to increased IL-17A production after viral stimulation.

Conclusions
We found evidence of interaction between CDHR3 and GSDMB in development of early childhood asthma, possibly related to increased IL-17A response to viral infections. This study demonstrates the importance of focusing on specific disease subtypes for understanding the genetic mechanisms of asthma.

Capsule Summary: We identify genetic interaction between CDHR3 and GSDMB in development of early childhood asthma, possibly related to increased IL-17A response to viral infections. This might lay the foundation for future preventive or therapeutic approaches.
Key Messages:
- We identify genetic interaction between CDHR3 and GSDMB in development of early childhood asthma, possibly related to increased IL-17A response to viral infections.
- This study demonstrates the importance of focusing on specific disease subtypes for understanding the genetic mechanisms of asthma.

Keywords: Childhood asthma, genetic interactions, early-life, immunology

Abbreviations
False discovery rate (FDR)
Genomic-relatedness-based restricted maximum-likelihood (GREML)
Genome-wide association study (GWAS)
Least absolute shrinkage and selection operator (LASSO)
Linkage disequilibrium (LD)
LD-score regression (LDSC)
Nasal epithelial cells (NEC)
Principal component (PC)
Rhinovirus-C (RV-C)
Region of practical equivalence (ROPE)
Single nucleotide polymorphism (SNP)
Introduction

Asthma with severe exacerbation is among the most common causes of hospitalization among young children and has a severe impact on quality of life and health care costs (1,2). Exacerbations are typically triggered by respiratory infections, but the host factors causing recurrent infections and exacerbations in some children are poorly understood. As a result, current treatment options and preventive measures are inadequate.

Genome-wide association studies (GWAS) have identified hundreds of asthma-associated single nucleotide polymorphisms (SNPs) (3). However, these variants only explain a small proportion of the phenotypic variance observed between individuals, a phenomenon known as “missing heritability” (4,5). Uncovering the causes for the missing heritability could result in a better understanding of asthma pathobiology and result in an improvement of personalized preventive measures (5). One factor that is likely to contribute to the missing heritability is genetic interactions, known as epistasis, where one gene modifies the effect of another gene (6). Studies in model organisms have provided biological and statistical evidence of genetic interactions (7), but only few examples have been identified in humans (8-10). The paucity of genetic interaction signals detected in humans has been largely attributed to the statistical power required to detect these interactions, which is difficult to obtain in human studies (7,11). However, another possible explanation is the heterogeneous nature of asthma, which likely represents several functional subtypes (12). It is possible that focusing on specific sub-phenotypes will increase the chance of detecting genetic interaction due to a more uniform underlying disease mechanism and thereby improve the signal-to-noise ratio the statistical power (13). We recently found support for this hypothesis by detecting a biologically plausible genetic interaction between FUT2 and ABO in a GWAS of early childhood asthma with severe exacerbations (14).

Here, we performed a genome-wide search for pairwise SNP interactions using a similar subtype-focused approach. We used genotype data from the COPSACsevere cohort, which constitutes children with recurrent hospitalizations due to severe asthma exacerbations during the first six years of life. The search for interactions was based on a Bayesian statistical model. Replication was attempted in the registry-based iPSYCH study (15,16), the COPSAC birth cohorts (17,18), and the UK Biobank (19).
Methods

Discovery study design
The discovery stage was based on 1204 childhood asthma cases with at least two acute hospitalizations due to asthma exacerbations from 2 to 6 years of age identified from national health registries (COPSACsevere) and 5328 non-asthmatic controls from the Danish Inter99 study (20) (Supplementary table 1, supplementary table 2). In addition to the overall case group, we also used a subset of the most pronounced asthma cases characterized by at least six hospitalizations in the discovery analysis.

Replication study design
Top pairwise SNP interactions from the discovery analysis were replicated in independent populations, specifically, the iPSYCH study (15,16), UK Biobank (21), and the COPSAC cohorts (17,18). In the iPSYCH study, childhood asthma cases were defined with at least one hospitalization due to asthma exacerbation in the first six years of life (N = 5729) and individuals without asthma hospitalizations were used as controls (N = 122,185). The UK Biobank analysis was based on retrospective recall using the reported “age of asthma” and doctor-diagnosed asthma based on the “age of doctors diagnosed asthma” information. Individuals with an asthma diagnosis during the first 6 years of life were included as cases (N = 5735). The control group was based on individuals with no doctor-diagnosed asthma (N = 89,924). In the COPSAC cohorts, asthma was diagnosed longitudinally by research physicians following a standardized treatment algorithm (N = 185). Controls were defined as children without an asthma diagnosis (N = 787) (Supplementary table 1, supplementary table 2).

Exhaustive search for pairwise SNP interactions
We used Bayesian logistic regression to estimate both the effect of individual SNPs and the effect of pairwise SNP interactions. The output of a Bayesian analysis is a so-called posterior probability distribution, P(W|D), which quantifies the relative credibility of the different possible parameter values (W) given the data (D), the prior knowledge, and the model assumptions. Specifically, the posterior distribution is computed from the likelihood, P(D|W) and the prior (PW), using Bayes
Theorem: \( P(W|D) = \frac{P(D|W)P(W)}{P(D)} \) (22). We used the software package Stan to estimate the posterior distribution of the parameters in the model (23). To focus on interactions with larger effect sizes, we defined a region of practical equivalence (ROPE) to no effect between 0.91 and 1.10 on the odds ratio scale (24), and estimate the posterior probability that interactions have an effect outside this region (see supplementary methods). Based on simulations (see supplementary methods), we chose to apply the Laplace prior with a scale of 0.05 in the downstream analysis (Supplementary figure 1, Supplementary figure 2, supplementary table 3). Using a Laplace prior in this Bayesian model is equivalent to the use of LASSO (least absolute shrinkage and selection operator) for regularization in frequentist regression. The LASSO has previously been used to search for epistatic effects in a non-Bayesian framework (25), however, the Bayesian LASSO has the advantage of providing straightforward measures of uncertainty for all examined parameters. The final Bayesian model had the following structure:

\[
\text{Phenotype} \sim \text{binomial}(\Theta) \\
\Theta = \logit(\mu + SNP_1 \cdot \beta_1 + SNP_2 \cdot \beta_2 + SNP_1 \cdot SNP_2 \cdot \beta_3 + X_{cov} \cdot \beta_{cov}) \\
\beta_1 \sim \text{Laplace}(0, 0.05) \\
\beta_2 \sim \text{Laplace}(0, 0.05) \\
\beta_3 \sim \text{Laplace}(0, 0.05) \\
\beta_{cov} \sim \text{Normal}(0, 1^2) \\
\mu \sim \text{Normal}(0, 1^2)
\]

\(X_{cov}\) represents a matrix with the 11 covariates as columns (10 principal components (PCs) and gender) and the number of individuals in the analysis as rows. Genotypes in the analysis were represented by the additive allele coding (allele count 0, 1, 2), and using an interaction term representing the product of two SNPs, with the aim of explaining any phenotypic variance not described by the individual additive effects.

To reduce the computational complexity and the number of possible pairwise interactions, we deployed an initial SNP filtering by computing the single SNP associations (GWAS) using the COPSACsevere dataset followed by linkage disequilibrium (LD) pruning, only retaining SNPs with p-values of less than 0.01 and \(r^2\) of less than 0.01 within a 1 MB window (Supplementary figure 3). Each clump was represented by the SNP with the lowest p-value. The LD-pruning was
based on the “clumping”-method implemented in PLINK v. 1.90(26). This reduced the overall number of SNPs to 326, and the number of possible pairwise interactions to 52,975. Based on the simulation results (supplementary methods, Supplementary Figures 4 and 5), the Bayesian model was applied in the downstream analysis to search for pairwise SNP interactions in relation to severe childhood asthma.

**Replication analysis**

Top interactions from the discovery analysis were replicated in three data sets, the COPSAC birth cohort, the iPSYCH study, and the UK Biobank. Analyses in the COPSAC cohorts were adjusted for gender and cohort status e.g. whether children were included in COPSAC2000 or COPSAC2010 and 10 PCs. Results from the iPSYCH population were adjusted for 10 PCs, gender, and whether the individuals had a psychiatric diagnosis. UK Biobank results were adjusted for gender and 10 PCs. We report the estimated interaction effect on the odds ratio scale, the 95% confidence interval, and one-sided P-values in relation to the direction of the discovery interaction estimates.

**Analysis of RNA-sequence data**

Analysis of the effect of genetic interaction on the expression of *CDHR3* was done based on RNA-sequencing data from respiratory (nasal) epithelial cells (NEC) obtained from 352 children in the COPSAC2010 cohort at 6 years of age, and from 187 children from the URECA study (27) at 11 years of age. We also investigated *CDHR3* expression in blood leukocytes from 116 children from the URECA study (27) obtained at 2 years of age, and 67 children of similar age from a Swedish study (28).

**Analysis of functional immune types**

We used immunology data from the COPSAC2010 cohort, where extensive immune profiling has been performed at age 18 months in 541 children as previously described in (29). We examined the fold increase in cytokine levels, defined as the ratio between the response in whole blood stimulated with viral-mimicking ligands (R848 (single-stranded RNA) or Poly(I:C) (double-stranded RNA)) and unstimulated whole blood. These ratios were subsequently log-transformed. The *CDHR3* x *GSDMB* interaction effect was estimated by linear regression adjusting for gender.
Results were replicated in the COPSAC2000 cohort containing similar immunological profiling of 286 children based on blood samples obtained at 6 months of age (30). The replication analysis was adjusted for gender, and we report one-sided p-values for replication analysis in relation to the direction for the discovery estimates.

**Results**

**Heritability of asthma was larger among the most severe cases**

We have previously shown how the magnitude and statistical significance of estimated main effects for genome-wide significant asthma-SNPs increase with increasing disease severity, quantified by the number of hospitalizations due to severe asthma exacerbations (31). We investigated whether this phenomenon would manifest itself on a genome-wide scale as well. The variance explained by the SNPs was highest in the most severe group with six or more hospitalizations with a heritability estimate of 0.30 (95%CI = 0.22 ; 0.37) (**Supplementary table 4**). The LD score regression (LDSC) intercept was similar across the severity strata indicating that the observed increase in heritability was not due to population stratification or an increase in cryptic relatedness (32) (**Supplementary table 4**). Based on this, we hypothesized that genetic interactions might be more likely identified using the most severe asthmatics.(14) and therefore ran the Bayesian model in both the overall case group as well as within the most severe disease stratum in the search for pairwise SNP interactions.

**Interaction between CDHR3 and GSDMB increases the risk of severe childhood asthma**

**Discovery**

We applied the LASSO model to search for pairwise interactions among the 326 SNPs using the overall case group (N\textsubscript{cases} = 1204) and the most severe disease stratum (N\textsubscript{cases} = 379). Among the 52,975 investigated pairwise interactions, three were found to have more than 90% posterior probability of association (**Table 1, Figure 1**). One of these was in the severe disease stratum and was between rs6769330 in CDHR3 and rs2305480 in GSDMB (OR = 1.28, 95% credible interval = 1.01 – 1.63, posterior probability = 90%). The remaining two were discovered in the overall case group and were between rs7758229 in SLC22A3 and rs2305480 in GSDMB (OR = 1.25, 95% credible interval = 1.08 – 1.46, posterior probability = 96%), and between rs2230641 in CCNH5 and rs2280134 in ART1 (OR = 1.24, 95% credible interval = 1.07 – 1.46, posterior probability =
95%; Table 1). Case-only analyses supported the interactions, while there was no evidence of association between genotypes in controls (Supplementary table 5). The estimated effect of the CDHR3 x GSDMB interaction increased with disease severity and was strongest in the most severe children (Supplementary table 6). Stratified analysis showed that the CDHR3 effect was strongest and only statistically significant (OR = 2.08 (95%CI = 1.63 ; 2.66), p-value = 4.77e-09) in children who were homozygous for the GSDMB risk allele (rs2305480, G) (Figure 2A), and children with a GSDMB GG and CDHR3 AA/AG genotype combination had a substantially increased risk of severe asthma compared to children who were homozygous for both CDHR3 and GSDMB non-risk alleles (OR = 11.0 (95%CI = 6.75 ; 18.8), p-value = 3.53e-20) (Supplementary table 7A). We observed a general increase in the number of interactions surpassing various posterior probability thresholds as the cases became more severe (Supplementary table 8), supporting that focusing on a severe phenotype can increase statistical power for detecting genetic interactions.

Replication
We tried to replicate the top 3 putative pairwise interactions (with posterior probability of > 90%) in independent cohorts. The interaction between CDHR3 and GSDMB showed evidence of interaction in relation to a hospital diagnosis with asthma during the first 6 years of life in the iPSYCH cohort (OR = 1.12 (95% CI = 1.05 ; 1.20), p-value = 3.4e-04) (Table 1, supplementary table 1, supplementary table 2), and to an asthma diagnosis with onset in the first 6 years of life in the UK Biobank (OR = 1.07 (95% CI = 1.00 ; 1.14), p-value = 0.03) (Table 1, supplementary table 1, supplementary table 2). The COPSAC birth cohorts showed similar interaction in terms of directionality, but the effect was not statistically significant (OR = 1.14 (95% CI = 0.76 ; 1.76), p-value = 0.28) (Table 1, supplementary table 1, supplementary table 2). The interaction was statistically significant in a combined analysis (meta-analysis) of all the replication cohorts (OR = 1.10 (95% CI = 1.05-1.15), p-value = 6.69e-05) (Table 1) and remained significant after adjusting for the number of candidate interactions tested for replication.

Despite the known higher prevalence of asthma hospitalization among individuals with a psychiatric diagnosis (33), this did not seem to influence the interaction effect, which was significant in both individuals with a psychiatric diagnosis as well as in healthy controls (N = 85,828, OR = 1.11 (95% CI = 1.04 ; 1.19), p-value = 0.002) and healthy controls (N = 42,086, OR = 1.19 (95% CI = 1.03 ; 1.44), p-value = 0.01), respectively. We further examined the interaction
between rs6967330 in CDHR3 and other variants in the 17q21 region, using SNPs from the full non-pruned genotype data set. We also found evidence of interaction with rs12936231 that was comparable in terms of effect size and significance to the interaction with rs2305480 in GSDMB (Supplementary table 9, supplementary figure 6).

The CDHR3 x GSDMB interaction pattern was consistent across the discovery stage and replication cohorts, showing that the CDHR3 effect was strongest in children carrying GSDMB risk alleles (rs2305480, G) whereas the effect of CDHR3 was nearly absent and not statistically significant in individuals with 0 GSDMB risk alleles (Figure 2B, supplementary table 7B, supplementary table 10). The GSDMB effect did not seem to depend equally on the CDHR3 genotype and increased the risk of childhood asthma in all CDHR3 genotype strata (Supplementary table 11).

The two large replication cohorts, iPSYCH and UK Biobank, allowed further exploration of the effect of age of asthma onset defined as the age of first hospital diagnosis with asthma in iPSYCH and the reported age of the first asthma diagnosis in the UK Biobank. The main effect of CDHR3 showed a strong age-dependency with the largest effect seen in children with asthma onset during the two first years of life, and declining hereafter to no effect on asthma with onset after 4 years of age (Table 2A). A similar pattern was seen for the interaction with consistent evidence of interaction for asthma onset in the first 2 years of life in both replication cohorts but not thereafter (Table 2B). Stratified analyses showed a strong effect of CDHR3 risk variants on asthma with onset before age 2 years in children who were homozygous for GSDMB risk alleles (iPSYCH: OR=1.48 (95%CI = 1.34 ; 1.64), p-value = 4.42e-15 and UK Biobank: OR=1.51 (95%CI = 1.26 ; 1.80, p-value 4.63e-06))(Supplementary table 12).

Interaction between CDHR3 and GSDMB and respiratory tract infections

Given the known involvement of CDHR3 in risk of respiratory infections (34), specifically rhinovirus-C (RV-C) infections, we investigated if the interaction could be explained by increased risk of acute respiratory illnesses triggered by specific infections in early life in the COPSAC birth cohorts. However, the interaction was not associated with the rate of either bacterial infections (IRR = 0.99 (95%CI = 0.79 ; 1.23), p-value = 0.90), viral infections (IRR = 0.96 (95%CI = 0.79 ; 1.20), p-value = 0.72), or RV-C infections (IRR = 0.74 (95%CI = 0.34 ; 1.46), p-value = 0.37).
Interaction between *CDHR3* and *GSDMB* and gene expression

Based on RNA-sequence data from blood leukocytes, we found evidence of interaction in relation to *CDHR3* expression (Estimate = 0.92 (95%CI = 0.34 ; 1.5), p-value = 0.002), where the *CDHR3* variant (rs6967330) was identified as an eQTL for *CDHR3* only in children homozygous for the *GSDMB* risk allele (rs2305480 = GG) (*Supplementary figure 7*). We found no association between the interaction and *CDRH3* expression in respiratory epithelial cells (*Supplementary table 13*).

Interaction between *CDHR3* and *GSDMB* was associated with increased viral-induced IL-17A response

Given the evidence of interaction in blood leukocytes and the known associations of *CDHR3* and *GSDMB* with the risk of viral infections (34,35), we investigated whether the interaction influenced cytokine levels during stimulation with viral stimuli in terms of single-stranded RNA (R848, targeting TLR7/8) and double-stranded RNA (Poly(I:C), targeting TLR3) at 18 months of age in COPSAC2010. We found that the *CDHR3* x *GSDMB* interaction was significantly associated with increased IL-17A levels during TLR3 stimulation (Estimate = 1.56 (95% CI = 1.24 ; 1.95), p-value = 1.26e-04, FDR = 0.005) (*Figure 3*). The *CDHR3* risk variant was only associated with increased IL-17A levels upon TLR3 stimulation in the children homozygous for the *GSDMB* risk allele (rs2305480 = GG) (*Supplementary table 14A*). Complementary, the effect of *GSDMB* was only found in children with at least one *CDHR3* risk allele (rs6967330 = GA, rs6967330 = AA) (*Supplementary table 14A*). We replicated this interaction between *CDHR3* x *GSDMB* on Poly(I:C) stimulated IL-17A cytokine levels in the COPSAC2000 cohort at age 6 months (Estimate = 1.89 (95% CI = 1.00 ; 3.59), p-value = 0.03). The stratified results in COPSAC2000 were similar to what we observed for the discovery analysis (*Supplementary table 14B*).

Discussion

Primary Findings

To our knowledge, this is the first successful genome-wide search for gene-gene interactions in asthma. We identified and replicated interaction between variants in *CDHR3* and the 17q12-21 locus near *GSDMB* and found that this was most pronounced for asthma with severe exacerbations.
and onset in the first years of life. Finally, we propose a potential mechanism related to increased IL-17A production during viral infections.

**Strength and limitations**

There is still no unified agreement on how to characterize the early childhood asthma phenotype. One of the main strengths of the present study is the strict inclusion criteria for the asthma cases within the discovery cohort, resulting in a more homogeneous case group with a specific asthma phenotype.

One limitation of the study is that we only investigated interactions between SNPs with significant main effects. This pre-filtering facilitated an exhaustive search for pairwise interactions, but it is likely that it removed true interactions. Another limitation is that the replication studies used different phenotype definitions. The asthma phenotype in the UK Biobank replication study was based on questionnaires and retrospective recall, making it a less well-defined phenotype compared to the other large replication resource in the iPSYCH study where the phenotype was based on asthma hospitalizations. Similarly, the asthma cases in the COPSAC cohorts did not all have severe asthma exacerbations. Nevertheless, the interaction between *CDHR3* and *GSDMB* also replicated with similar age at onset pattern in UK biobank. Another limitation is that IPSYCH controls will have included some children with milder asthma and asthma exacerbations that haven’t required hospitalization, leading to an increased risk of type II errors in the replication.

**Interpretation**

We found evidence that gene-gene interactions play a role in childhood asthma and that it can be detected by focusing on a specific sub-phenotype. The interaction between *CDHR3* and *GSDMB* was specifically detected in children with many (≥ 6) severe asthmatic episodes requiring hospitalization in the first 6 years of life. We previously observed the same severity pattern for interaction between *FUT2* and *ABO* functional gene variants gene-gene interactions (14). The interaction replicated in independent (larger) populations using milder disease criteria and even using retrospective recall of the asthma onset as in the UK biobank. This demonstrates that genetic signals detected in very severe phenotypes can also be transferred to milder disease once detected. Age at onset was another important phenotypic parameter for this interaction. The interaction was discovered and replicated specifically when restricted to onset in the first 6 years of life and was
strongest for onset in the first 2 years of life. A similar age-dependency was seen for the CDHR3 main signal. It is plausible that gene-gene interactions take place in relation to specific ages and/or mechanisms for which both genes have strong effects. The finding of a strong age-dependency of the main CDHR3 effect also emphasizes the importance of phenotype specificity for asthma genetics in general. The CDHR3 risk variant was first discovered in the COPSACsevere study but was not observed in a later large GWAS from the UK biobank focusing on asthma without age restriction or childhood-onset asthma defined as onset before 12 years of age (3,36). However, when focusing on asthma with early-onset (the first 2 years of life), the CDHR3 effect was evident in both iPSYCH and UK biobank with an effect size in the same range (OR 1.27 - 1.29, Table 2A) as the strongest asthma genes normally found in GWAS(3,36) and even higher (OR 1.49-1.51) in children who were also homozygous for the 17q21, GSDMB variant (Table 2, Supplementary table 10). This demonstrates the existence of genetic effects on specific asthma subtypes and mechanisms (endotypes), and how this can be missed in studies with broader phenotypic definitions.

The interaction between the CDHR3 and GSDMB variants is biologically plausible, because both are associated with a clinical phenotype of early-onset childhood asthma, and specifically with rhinovirus-induced asthmatic episodes (34,35), making it likely that they share pathogenetic pathways with potential of effect modification. Our results suggest that CDHR3 risk variants only increase the risk of childhood asthma in children who also carry risk variants at GSDMB. The rs2305480 variant in the GSDMB gene is located in the 17q21 region, which is the most well-established childhood asthma locus (3,36,37), particularly associated with childhood-onset asthma (3,38). Despite the strong involvement in childhood asthma, determination of the causal variant at this locus has been difficult due to extensive LD in the region in individuals of European ancestry (37). The earliest studies pointed to ORMDL3 as the most likely asthma candidate in the region (39), while a recent study suggested that GSDMB is instead likely to be the causal gene (40). In the present study, the rs2305480 variant was the SNP with the strongest association to the phenotype in this region. However, we cannot rule out that the mechanism behind the interaction is related to ORMDL3 since a similar interaction effect was seen for the SNP rs12936231, with a functional role in relation to ORMDL3 expression in immune cells (Supplementary table 9, Supplementary figure 6) (39).
The role of CDHR3 in relation to asthma and infections has primarily been thought to involve the airway epithelium (41). However, CDHR3 has also been found to be expressed in immune cells, particularly in T-cells (42). Interestingly, this is also the case for GSDMB and ORMDL3, the putative functional genes in the 17q21 region (42), and the 17q21 (GSDMB) locus also seems to play a functional role in immune cells (28,39).

While there was no evidence of interaction in relation to CDHR3 expression in respiratory (nasal) epithelial cells, we did see evidence of interaction in blood leukocytes (Supplementary table 13), suggesting that the interactions might involve immune cells. In support of this hypothesis, we found that the interaction increased IL-17A cytokine levels in blood after stimulation with viral ligands. IL-17A is a hallmark cytokine of Th17 cells, involved in a neutrophil-mediated immune response (43), which is a characteristic of early life wheezing (44). Increased levels of IL-17A correlate with severe forms of asthma (45,46), and an IL-17A dominated immune response is detrimental during viral infection (47), and has previously been linked to early transient asthma in the COPSAC2010 cohort (29). Interestingly, 17q21 gene variation has been associated with increased IL-17 production in early life (48), and ORMDL3 expression is able to activate T-cells and increase IL-17 production in experimental settings(49). It is, therefore, possible that the CDHR3 x GSDMB interaction is specifically involved in a neutrophil-dominated immune response to viral infections, which increases the risk of asthma in early life (29). Based on these results, it can be speculated that IL-17A, or related immune mediators, could be a future preventive or therapeutic target in childhood asthma.

**Conclusion**

Using a well-defined sample of severe childhood asthma cases and careful statistical modeling, we were able to detect an example of gene-gene interaction between SNPs at the CDHR3 and GSDMB loci, with a potential underlying mechanism related to excessive IL-17A cytokine production following viral infection. The ability to detect these interactions largely depended on the phenotype specificity, both in terms of severity and disease onset. These results suggest that a key ingredient for understanding the genetic mechanisms of asthma is detailed longitudinal clinical phenotyping, hereby focusing on more homogenous case populations.
Acknowledgements
We express our deepest gratitude to the children and families of the COPSAC2000, COPSAC2010, and COPSAC Exacerbation cohort studies for all their support and commitment. We acknowledge and appreciate the unique efforts of the COPSAC research team. This research has been conducted using data from UK Biobank database (www.ukbiobank.ac.uk), project ID numbers 32683 and 54273.

Authors Contributions
The guarantors of the study are A.U.E., A.G.P, and K.B., from conception and design to conduct of the study and acquisition of data, data analysis, and interpretation of data. A.U.E. has written the first draft of the manuscript.

All authors guarantee that the accuracy and integrity of any part of the work have been appropriately investigated and resolved and all have approved the final version of the manuscript. The corresponding author had full access to the data and had final responsibility for the decision to submit for publication. No honorarium, grant, or other forms of payment was given to any of the authors to produce this manuscript.

Author Information
Data deposition statement
Summary statistics are available upon request to the corresponding author K.B. (kb@copsac.com).

Corresponding author line
Correspondence and requests for materials should be addressed to K.B. (kb@copsac.com).

Governance
We are aware of and comply with recognized codes of good research practice, including the Danish Code of Conduct for Research Integrity. We comply with national and international rules on the safety and rights of patients and healthy subjects, including Good Clinical Practice (GCP) as defined in the EU's Directive on Good Clinical Practice, the International Conference on Harmonisation's (ICH) good clinical practice guidelines and the Helsinki Declaration. Privacy is important to us which is why we follow national and international legislation on General Data Protection Regulation (GDPR), the Danish Act on Processing of Personal Data and the practice of the Danish Data Inspectorate.
References


24. Kruschke JK. Rejecting or accepting parameter values in Bayesian estimation [Internet]. Available from: http://dx.doi.org/10.31219/osf.io/s5vdy


Figure legends

**Figure 1. Results of the genome-wide search for pairwise SNP interactions.** Results are shown for A) analysis using the complete group of asthma cases (Ncases = 1204, Ncontrols = 5328) and B) analysis using the most severe group of asthmatics (Ncases = 379, Ncontrols = 5328). Bold red lines represent interactions with a posterior probability > 90%. Thin red lines represent interactions with posterior probability between 90% and 75%. Grey lines represent interactions with posterior probability < 75%.

**Figure 2. CDHR3 effect on asthma risk stratified by GSDMB genotype.** Results are shown for A) the discovery stage (Ncases = 379, Ncontrols = 5328) and B) the replication stage (meta-analyses of iPSYCH, UK biobank and COPSAC birth cohorts) (Ncases = 11,649, Ncontrols = 212,869). CDHR3 (rs6967330) effects are based on an additive genetic model with the asthma-associated allele (A) as effect allele. GSDMB = 0: rs2305480 AA genotype. GSDMB = 1: rs2305480 AG genotype. GSDMB = 2: rs2305480 GG genotype. Error bars represent 95% confidence intervals.

**Figure 3. CDHR3 and GSDMB interaction effects on cytokine response during viral stimuli.** Analyses are based on immunological profiling from the COPSAC2010 cohort (N = 566). Circles represent interaction effects on log fold cytokine levels (stimulated/unstimulated). Size of circles is scaled according to the interaction p-value, with larger circles illustrating smaller p-values.