Genetic and epigenetic regulation of YKL-40 in childhood

Stefano Guerra, MD, PhD,[a,b,c,d] Erik Melén, MD, PhD,[e] Jordi Sunyer, MD, PhD,[a,b,c,f] Cheng-Jian Xu, PhD,[g,h,i] Iris Lavi, PhD,[a,b,c] Marta Benet, BStat,[a,b,c] Mariona Bustamante, PhD,[a,b,c,j] Anne-Elie Carsin, MS,[a,b,c,f] Carlota Dobaño, PhD,[k] Mònica Guxens, MD, PhD,[a,b,c,l] Christina Tischer, PhD,[a,b,c] Martine Vrijheid, PhD,[a,b,c] Inger Kull, RN, PhD,[m] Anna Bergström, PhD,[e] Ashish Kumar, MSc,[e,n,o] Cilla Söderhäll, PhD,[p] Ulrike Gehring, PhD,[q] Dorieke J. Dijkstra, MSc,[r] Pieter van der Vlies, BSc,[i] Magnus Wickman, MD, PhD,[e] Jean Bousquet, MD, PhD,[s] Dirkje S. Postma, MD, PhD,[g,h] Josep M. Anto, MD, PhD,[a,b,c,f] and Gerard H. Koppelman, MD, PhD[h,t]

Barcelona, Spain; Tucson, Ariz; Stockholm and Huddinge, Sweden; Groningen, Rotterdam, and Utrecht, The Netherlands; Basel, Switzerland; and Montpellier and Villejuif, France

Background: Circulating levels of the chitinase-like protein YKL-40 are influenced by genetic variation in its encoding gene (chitinase 3–like 1 *[CHI3L1]*) and are increased in patients with several diseases, including asthma. Epigenetic regulation of circulating YKL-40 early in life is unknown.

Objective: We sought to determine (1) whether methylation levels at *CHI3L1* CpG sites mediate the association of *CHI3L1* single nucleotide polymorphisms (SNPs) with YKL-40 levels in the blood and (2) whether these biomarkers (*CHI3L1* SNPs, methylation profiles, and YKL-40 levels) are associated with asthma in early childhood.

Methods: We used data from up to 2405 participants from the Spanish Infancia y Medio Ambiente; the Swedish Barn/Children, Allergy, Milieu, Stockholm, Epidemiological survey; and the Dutch Prevention and Incidence of Asthma and Mite Allergy birth cohorts. Associations between 68 *CHI3L1* SNPs, methylation levels at 14 *CHI3L1* CpG sites in whole-blood DNA, and circulating YKL-40 levels at 4 years of age were tested by using correlation analysis, multivariable regression, and mediation analysis. Each of these biomarkers was also tested for association with asthma at 4 years of age by using multivariable logistic regression.

Results: YKL-40 levels were significantly associated with 7 SNPs and with methylation at 5 CpG sites. Consistent associations between these 7 SNPs (particularly rs10399931 and rs4950928) and 5 CpG sites were observed. Alleles linked to lower YKL-40 levels were associated with higher methylation levels. Participants with high YKL-40 levels (defined as the highest YKL-40 tertile) had increased odds for asthma compared with subjects with low YKL-40 levels (meta-analyzed adjusted odds ratio, 1.90 [95% CI, 1.08-3.36]). In contrast, neither SNPs nor methylation levels at CpG sites in *CHI3L1* were associated with asthma.

Conclusions: The effects of *CHI3L1* genetic variation on circulating YKL-40 levels are partly mediated by methylation profiles. In our study YKL-40 levels, but not *CHI3L1* SNPs

From [a]ISGlobal, Centre for Research in Environmental Epidemiology (CREAL), Barcelona; [b]Universitat Pompeu Fabra, Barcelona; [c]CIBER Epidemiología y Salud Pública (CIBERESP); [d]the Asthma and Airway Disease Research Center, University of Arizona, Tucson; [e]the Institute of Environmental Medicine, Karolinska Institutet, Stockholm; [f]IMIM (Hospital del Mar Medical Research Institute), Barcelona; [g]University of Groningen, University Medical Center Groningen, Department of Pulmonology; [h]the Groningen Research Institute for Asthma and COPD; [i]University of Groningen, University Medical Center Groningen, Department of Genetics; [j]Centre for Genomic Regulation (CRG), the Barcelona Institute of Science and Technology, Barcelona; [k]ISGlobal, Barcelona Center for International Health Research (CRESIB), Hospital Clínic, Universitat de Barcelona; [l]the Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Centre–Sophia Children's Hospital, Rotterdam; [m]Sachs' Children's Hospital, Södersjukhuset, Stockholm, and the Department of Clinical Science and Education, Karolinska Institutet at Södersjukhuset, Stockholm; [n]the Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel; [o]the University of Basel; [p]the Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, and Department of Women's and Children's Health, Karolinska Institutet, Stockholm; [q]the Institute for Risk Assessment Sciences, Utrecht University; [r]the Department of Obstetrics and Gynecology, University of Groningen, University Medical Center Groningen; [s]University Hospital Montpellier, and the Respiratory and Environmental Epidemiology Team, INSERM 1018, CESP Centre, Villejuif; and [t]University of Groningen, University Medical Center Groningen, Beatrix Children's Hospital, Department of Pediatric Pulmonology and Pediatric Allergology.

Disclosure of potential conflict of interest: U. Gehring receives grant support from EU FP7 grant no. 261357. J. Bousquet serves on the board for Almirall, Meda, Merck, MSD, Novartis, Sanofi-Aventis, Takeda, Teva, and Uriach and receives payment for lectures from Almirall, AstraZeneca, Chiesi, GlaxoSmithKline, Meda, Menarini, Merck, MSD, Novartis, Sanofi-Aventis, Takeda, Teva, and Uriach. D. S. Postma serves as a consultant for Takeda, TEVA, Chiesi, GlaxoSmithKline, AstraZeneca, and Boehringer Ingelheim and receives grant support from Astra Zeneca, Chiesi, Genentech, GlaxoSmithKline, and Roche. G. H. Koppelman receives grant support from the Lung Foundation of the Netherlands, TEVA, Ubbo Emmius Foundation, and TETRI Foundation and receives travel support from GlaxoSmithKline. The rest of the authors declare that they have no relevant conflicts of interest.

Supported by grant agreement no. 261357 (Mechanisms of the Development of Allergy [MeDALL]) from the EU Seventh Framework Programme. Additional funding for individual cohorts was provided as follows: *INMA* (Instituto de Salud Carlos III [Red INMA G03/176; CB06/02/0041; PI041436; PI081151 incl. FEDER funds; MS13/00054], Generalitat de Catalunya-CIRIT 1999SGR 00241, Fundació La marató de TV3 [090430]); *BAMSE* (the Swedish Heart-Lung Foundation, Swedish Research Council, Freemason Child House Foundation in Stockholm, Stockholm County Council [ALF], Swedish Foundation for Strategic Research [SSF, RBc08-0027, EpiGene project] and the Strategic Research Programme [SFO] in Epidemiology at Karolinska Institute); and *PIAMA* (The Netherlands Organization for Health Research and Development; The Netherlands Organization for Scientific Research; The Netherlands Lung Foundation (AF 4.1.14.001); The Netherlands Ministry of Spatial Planning, Housing, and the Environment; and The Netherlands Ministry of Health, Welfare, and Sport).

Received for publication March 31, 2016; revised May 30, 2017; accepted for publication June 14, 2017.

Available online July 21, 2017.

Corresponding author: Stefano Guerra, MD, PhD, Asthma and Airway Disease Research Center, 1501 N Campbell Ave, Tucson, AZ 85724. E-mail: stefano.guerra@isglobal.org; stefano@email.arizona.edu.

The CrossMark symbol notifies online readers when updates have been made to the article such as errata or minor corrections

0091-6749/$36.00

http://dx.doi.org/10.1016/j.jaci.2017.06.030

or methylation levels, were associated with childhood asthma. (J Allergy Clin Immunol 2018;141:1105-14.)

Key words: *YKL-40,* CHI3L1*, asthma, epigenetics, DNA methylation, genetics*

YKL-40, a chitinase-like protein, is upregulated in patients with asthma, cancer, and other diseases characterized by inflammation and tissue remodeling.[1] In adults YKL-40 levels are increased in the blood and lungs of patients with asthma[2-4] and correlate with lung function deficits, disease severity and persistence.[2-10] However, to date, findings on YKL-40 and asthma in childhood have been conflicting.[11-13]

The mechanisms linking YKL-40 with asthma remain to be determined, although *in vitro* and animal studies support its role in T_H2 adaptive immune responses.[14] The association between genetic variation in chitinase 3–like 1 (*CHI3L1*, the gene encoding YKL-40) and asthma susceptibility/severity suggests a possible causal link.[11,15-18] A group of 4 single nucleotide polymorphisms (SNPs) tagged by rs4950928 were first associated with asthma in 3 of 4 tested populations, with the major alleles conferring increased risk.[11] In the same study serum YKL-40 levels were found to be a highly heritable quantitative trait in the general population and to be directly associated with the same C allele at rs4950928 that was associated with asthma risk. Subsequent studies showed that another SNP, rs10399931, which is in strong linkage disequilibrium (LD) with rs4950928, had similar, if not stronger, effects on gene expression,[16] plasma YKL-40 levels,[19] and asthma.[16] However, although the association between *CHI3L1* genetic variation and YKL-40 levels has been conclusively established, the relation of *CHI3L1* variation to asthma remains controversial because other reports,[20-22] including a large study of more than 6500 Danish adults,[20] failed to replicate the aforementioned genetic associations with asthma.

Identification of the mechanisms by which genetic variation in *CHI3L1* regulates YKL-40 levels might have important implications for understanding the potential effect of this gene on human disease. Epigenetic regulation of gene expression is one of the possible mechanisms by which genetic variation can affect protein levels and disease susceptibility,[23] including childhood respiratory diseases.[24] DNA sequence variants across the genome have been shown to have *cis*- (and to less extent *trans*-) effects on methylation levels at specific CpG sites.[23,25-30] Yet, to date, no study has addressed DNA methylation as a possible intermediary mechanism of the relation between *CHI3L1* genetic variation and upregulated YKL-40 protein levels.

The primary goal of the present study was to determine whether methylation levels at *CHI3L1* CpG sites mediate the association of *CHI3L1* genetic variation to YKL-40 levels in blood. In secondary analyses we also sought to assess whether the *CHI3L1* genotype and methylation levels that regulate YKL-40 levels are associated with asthma in early childhood.

Abbreviations used

adjOR:	Adjusted odds ratio
BAMSE:	Swedish Barn/Children, Allergy, Milieu, Stockholm, Epidemiological survey
CHI3L1:	Chitinase 3–like 1
GWAS:	Genome-wide association study
INMA:	Spanish Infancia y Medio Ambiente
LD:	Linkage disequilibrium
MeDALL:	Mechanisms of the Development of Allergy
PCA:	Principal component analysis
PIAMA:	Prevention and Incidence of Asthma and Mite Allergy
SNP:	Single nucleotide polymorphism

METHODS

Study populations and design

This study was part of the Mechanisms of the Development of Allergy (MeDALL) project,[31] which included analyses on YKL-40 as an *a priori* biomarker candidate. The design and available data for the present study are summarized in Fig 1.

Primary analyses (Fig 1, *A*) included molecular and phenotypic data from 433 participants who were 4 years of age from the Spanish Infancia y Medio Ambiente (INMA; n = 203)[32] and the Swedish Barn/Children, Allergy, Milieu, Stockholm, Epidemiological survey (BAMSE; n = 230)[33] birth cohorts. These children were selected from their original cohorts for epigenetic and YKL-40 studies in MeDALL based on the following sampling strategy. Among participants who provided paired DNA samples at 2 time points (birth and 4 years for INMA and 4 and 8 years for BAMSE), we selected children who had 1 or more of 3 diagnoses (asthma, eczema, and allergic rhinitis) and a similar number of randomly selected control subjects (none of the 3 diagnoses, nested case-control design). Of these children, 172 INMA and 78 BAMSE participants also had available genome-wide association study (GWAS) data.

In secondary analyses of association with asthma (Fig 1, *B*), in addition to INMA and BAMSE, we also included 4-year-old participants from the Prevention and Incidence of Asthma and Mite Allergy (PIAMA) study to increase statistical power.[34,35] No serum YKL-40 levels were available for PIAMA, but epigenetic studies were completed on 193 PIAMA participants who were selected based on the same sampling strategy described previously. For analyses of genetic association, we used data from all participants from the 3 cohorts who had both GWAS data and asthma information (n = 336 for INMA, n = 467 for BAMSE, and n = 1602 for PIAMA; total n = 2405).

A detailed description of the 3 birth cohorts and additional information on molecular assays are provided in the Methods section in this article's Online Repository at www.jacionline.org. For all cohorts, parents provided written informed consent, and the local ethics review boards approved the performed studies and procedures.

Definition of asthma

In MeDALL an asthma definition[36] was used that included positive answers to at least 2 of the 3 following questions: (1) "Has your child ever been diagnosed by a doctor as having asthma?"; (2) "Has your child taken any medicines for asthma (including inhalers, nebulizers, tablets, or liquid medicines) or for breathing difficulties (chest tightness, shortness of breath) in the last 12 months?"; and (3) "Has your child had wheezing or whistling in their chest at any time in the last 12 months?" or "Has your child had breathing difficulties (chest tightness, shortness of breath) in the last 12 months?" This definition was used for BAMSE and PIAMA. Because information on doctor-diagnosed asthma was not available in INMA, for this cohort, the asthma definition included a positive response to the 2 remaining questions.

Molecular data

DNA methylation assays at *CHI3L1* CpG sites. Epigenome-wide analysis scans of paired whole peripheral blood DNA methylation samples at birth and age 4 years in INMA and at age 4 and 8 in BAMSE and PIAMA were generated by using the Illumina Infinium HumanMethylation450 BeadChip assays (Illumina, San Diego, Calif). All samples were quality checked, as described in the Methods section in this article's Online Repository. The Illumina 450K included 485,577 assays. During processing, the probes on sex chromosomes, the probes that mapped on multiloci, 65 SNP assays, and the probes containing SNPs at the target CpG sites with a

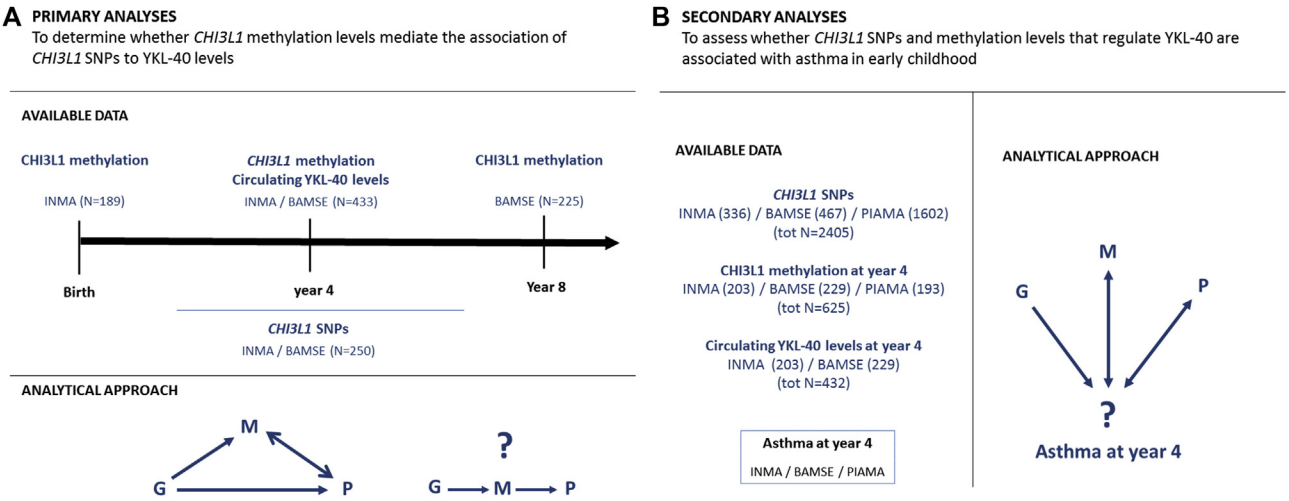

FIG 1. Study design for the primary **(A)** and secondary **(B)** analyses. *G*, *CHI3L1* genetic variation; *M*, methylation levels at *CHI3L1* CpG sites; *P*, circulating YKL-40 levels.

TABLE I. Numbers and basic characteristics of participants included in this study

	INMA	BAMSE	PIAMA
Data available for primary analyses*			
No.	203	230	
Male sex, no. (%)	102 (50%)	126 (55%)	
Age (mo) at the 4-y survey, mean (SD)	53.2 (2.3)	47.9 (1.7)	
Body mass index at 4 y, mean (SD)	16.2 (1.5)	16.2 (1.3)	
Asthma at 4 y, no. (%)	33 (16%)	88 (38%)	
YKL-40 levels at 4 y, no.	203	230	
Methylation data at 4 y, no.	203	230	
Methylation data at birth, no.	189	—	
Methylation data at 8 y, no.	—	225	
GWAS data, no.	172	78	
Data available for association analyses with asthma, secondary analyses†			
GWAS data, no. asthma/no. total	58/336	171/467	69/1602
Male sex, no. (%)	179 (53)	260 (56)	800 (50)
Age (mo) at the 4-y survey, mean (SD)	51.4 (3.3)	48.1 (1.9)	48.3 (1.1)
Methylation data at 4 y, no. asthma/no. total	33/203	88/229	13/193
YKL-40 levels at 4 y, no. asthma/no. total	33/203	88/229	NA

NA, Not available.
*Children with complete data on *CHI3L1* methylation and YKL-40 levels at age 4 years.
†Children with complete data on asthma, covariates, and at least 1 of the 3 biomarkers at age 4 years (ie, *CHI3L1* SNPs, *CHI3L1* methylation, or YKL-40 levels).

minor allele frequency of greater than 10% were excluded. This resulted in a total of 439,306 CpG sites. Among them, 14 CpG sites were annotated to the *CHI3L1* gene and analyzed in the present study (see Table E1 in this article's Online Repository at www.jacionline.org). Two of them (cg17014757 and cg03625911) have probes that include known SNPs with minor allele frequencies of greater than 5% in European populations, as reported by Chen et al.[37] The probe of cg17014757 includes rs10399931 and the probe of cg03625911 includes rs7515776, which, according to the CEU panel of the 1000 Genomes project, is in nearly complete LD with rs10399805. These 2 CpG sites were included in statistical analyses, but associations of their methylation levels with overlapped SNPs or other SNPs in LD with the overlapped SNPs (rs10399931 and rs4950928 for cg17014757 and rs2886117, rs10399805, and rs7542294 for cg03625911) were identified in tables to caution interpretation.

Methylation levels (β values) at a given CpG site were derived from the ratio of the methylated probe intensity and overall intensity (sum of methylated and unmethylated probe intensities). Thus β is equal to M/(U + M+ α), where *M* is intensity of the methylated probe, *U* is the intensity of the unmethylated probe, and *α* is the constant offset with the default value of 100. The intensity has been corrected for type I and type II probe differences and normalized by using the "dasen" method in the wateRmelon R package.[38] Results from analyses on methylation levels were also confirmed after adjustment for blood cell–type composition, as predicted by the Houseman algorithm,[39] and, in a subset of 116 INMA participants at year 4, after adjustment for available differential cell counts obtained directly from microscopic inspection of blood smears. The methylation signal of the top 5 CpG sites obtained by using the Illumina microarray was also validated by means of pyrosequencing in whole-blood DNA from 96 subjects participating in the PIAMA study (see the Methods section, Fig E4, and Table E11 in this article's Online Repository at www.jacionline.org).

YKL-40 measurements. Circulating YKL-40 levels were measured in serum (INMA) and plasma (BAMSE) samples at age 4 years by using a commercially available ELISA kit (Quantikine Human CHI3L1 immunoassay; R&D Systems, Abingdon, United Kingdom).

GWAS. Genome-wide genotyping had been previously completed for the 3 cohorts (INMA, BAMSE, and PIAMA) by using various platforms (see the Methods section in this article's Online Repository). For the present study, in the INMA cohort we tested 68 SNPs that were genotyped from the genomic region surrounding (±50 kb) *CHI3L1* (hg19: chromosome 1: 203,098,059-203,205,922) by using the HumanOmni1-Quad BeadChip (Illumina). SNPs that were found not to be in Hardy-Weinberg equilibrium by using exact tests[40] were filtered (see the Methods section in this article's Online Repository for GWAS assays). SNPs that were significantly associated with YKL-40 levels in INMA were also tested in BAMSE and PIAMA.

Analytic approach and statistical analysis

For the primary goal, we analyzed interrelationships between multilevel biomarkers (SNPs, methylation at CpG sites, and YKL-40 levels) in INMA and BAMSE, according to a stepwise approach (Fig 1, *A*). Because results of these analyses did not significantly differ between cases and control subjects, they are presented with no stratification by disease. YKL-40 levels were log-transformed and then standardized within each cohort by subtracting the mean and dividing the result by the SD, as done in previous multicohort studies.[41]

TABLE II. Associations between SNPs in the *CHI3L1* genomic region and standardized protein levels of YKL-40 at age 4 years in the INMA cohort (n = 172)

SNP*	Location chromosome 1 (hg19)	Effect allele	Other allele	Effect allele frequency	HWE *P* value†	Linear regression coefficient‡	95% CI	*P* value	R^2§
rs2886117	203168881	A	G	15.7	.77	0.61	0.326 to 0.894	$\mathbf{4 \times 10^{-5}}$	0.10
rs10399931	203156080	T	C	19.2	.33	−0.68	−0.941 to −0.420	$\mathbf{7 \times 10^{-7}}$	0.14
rs10399805	203155998	A	G	15.4	.77	0.62	0.331 to 0.901	$\mathbf{3 \times 10^{-5}}$	0.10
rs4950928	203155882	G	C	16.0	.08	−0.82	−1.108 to −0.536	$\mathbf{6 \times 10^{-8}}$	0.16
rs7542294	203151176	A	G	16.9	.42	0.64	0.360 to 0.914	$\mathbf{1 \times 10^{-5}}$	0.11
rs2791718	203141424	A	C	12.8	.32	0.61	0.286 to 0.924	$\mathbf{2 \times 10^{-4}}$	0.08
rs10920576	203129179	T	C	12.8	.32	0.72	0.406 to 1.034	$\mathbf{1 \times 10^{-5}}$	0.11

Only SNPs significant after Bonferroni correction are shown (data for all SNPs are shown in Table E3). Boldface indicates associations significant after Bonferroni correction.
*SNPs are shown according to position in the gene.
†Hardy-Weinberg equilibrium *P* value.
‡Coefficient from additive models predicting standardized YKL-40 levels.
§Percentage variability in YKL-40 levels explained by the SNP.

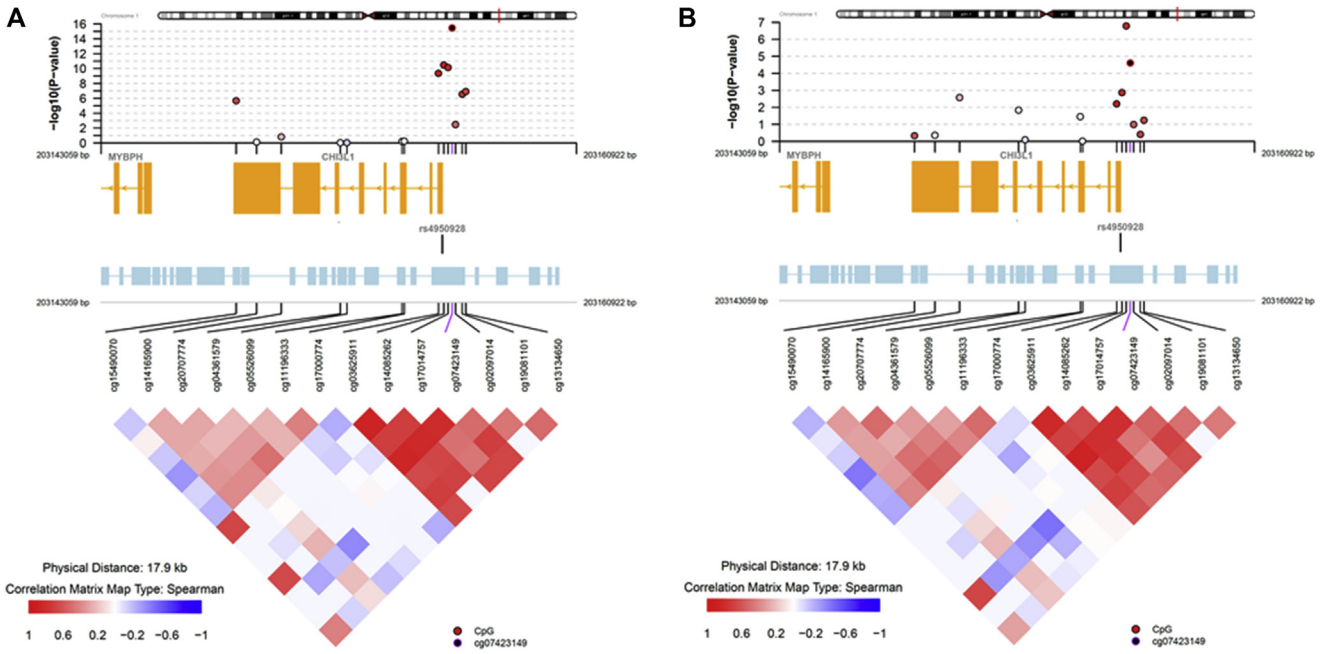

FIG 2. *P* values for association with YKL-40 levels *(top panel)* and correlation matrices *(bottom panel)* for methylation levels at the 14 *CHI3L1* CpG sites in INMA at 4 years of age **(A)** and in BAMSE at 4 years of age **(B)**. The *top panel* of each figure shows the $-\log_{10}$ (*P* value) for the association between YLK-40 levels and DNA methylation at each CpG site (the color of the symbol is the color of the comethylation pattern between that specific CpG and the reference CpG site indicated in black [cg07423149]). The *bottom panel* of each figure shows the correlation between methylation levels at CpG sites, as indicated in the legend. In addition, several annotation tracks are shown (ENSEMBL gene annotation in orange, SNPs from GWASs (rs4950928) in black, and DNAse hypersensitivity regions in *blue*). See Martin et al.[43]

Discovery analyses were completed in INMA with a conservative Bonferroni correction and replication analyses in BAMSE with a 1-tailed α value of .05 (ie, only associations with the same direction of effect between the 2 cohorts were tested for significance). First, associations between *CHI3L1* SNPs and YKL-40 levels were studied. In INMA 68 SNPs were tested in linear regressions predicting YKL-40 protein levels according to additive genetic models. Because INMA and BAMSE used different GWAS platforms, in BAMSE we used genotyped or imputed SNPs for replication of INMA results as appropriate. Second, methylation levels were tested for correlation with YKL-40 levels by using Spearman correlation coefficients. Third, the SNPs and CpG sites that were found to be related to YKL-40 levels were tested for association with each other using robust regressions to reduce the effect of potential outlier observations (see the Methods section in this article's Online Repository for additional information).

Multivariable regression models with backward stepwise variable selection were used to determine what SNPs and CpG sites were independently related to YKL-40 protein levels. A final mediation analysis was completed in INMA by using the R package "mediation"[42] to estimate the effects by the aforementioned independent SNPs on YKL-40 that were mediated by means of methylation. For this analysis, we used both methylation levels at the CpG site (cg07423149) that was identified by using the above backward stepwise selection, as well as the first principal component obtained by using a principal component analysis (PCA) on methylation levels at all 5 CpG sites that were associated with YKL-40.

TABLE III. Spearman correlation coefficients between methylation levels at CpG sites at various ages and YKL-40 levels in INMA and BAMSE

		INMA methylation at 4 y (n = 203)		BAMSE methylation at 4 y (n = 230)		INMA methylation at birth (n = 189)		BAMSE methylation at 8 y (n = 225)	
CpG sites	Position chromosome 1 (hg19)	Spearman correlation with YKL-40 at 4 y	*P* value	Spearman correlation with YKL-40 at 4 y	*P* value	Spearman correlation with YKL-40 at 4 y	*P* value	Spearman correlation with YKL-40 at 4 y	*P* value
cg13134650	**203156765**	**−0.36**	**1×10^{-7}**	**−0.13**	**.06**	**−0.25**	**5×10^{-4}**	**−0.22**	**7×10^{-4}**
cg19081101	203156626	−0.35	3×10^{-7}	−0.06	.38	−0.15	.04	0.01	.90
cg02097014	203156375	−0.21	.003	−0.11	.10	−0.09	.24	−0.12	.08
cg07423149	**203156247**	**−0.53**	**3×10^{-16}**	**−0.27**	**2×10^{-5}**	**−0.30**	**3×10^{-5}**	**−0.26**	**6×10^{-5}**
cg17014757*	**203156098**	**−0.44**	**7×10^{-11}**	**−0.34**	**2×10^{-7}**	**−0.31**	**2×10^{-5}**	**−0.35**	**4×10^{-8}**
cg14085262	**203155939**	**−0.44**	**3×10^{-11}**	**−0.21**	**.001**	**−0.24**	**7×10^{-4}**	**−0.22**	**7×10^{-4}**
cg03625911†	**203155738**	**−0.42**	**4×10^{-10}**	**−0.18**	**.006**	**−0.21**	**.004**	**−0.19**	**.005**
cg17000774	203154457	−0.04	.57	0.00	.97	−0.04	.62	0.03	.65
cg11196333	203154371	−0.04	.62	−0.14	.04	−0.08	.26	−0.16	.01
cg05526099	203152296	0.00	.95	−0.02	.80	−0.05	.51	0.02	.77
cg04361579	203152047	0.00	.95	−0.16	.01	−0.06	.44	0.00	.99
cg20707774	203149828	−0.10	.15	−0.20	.003	−0.04	.56	−0.11	.09
cg14165900	203148901	−0.02	.74	0.05	.44	0.04	.61	0.00	.97
cg15490070	203148132	−0.33	2×10^{-6}	−0.05	.47	−0.13	.08	−0.05	.49

Boldface indicates correlations significant both in the discovery population (INMA, 4 years) after Bonferroni correction and in the replication population (BAMSE, 4 years; 1-sided $\alpha = .05$). Associations in boldface were confirmed after adjustment for blood cell composition.
*Probe overlaps known SNPs with minor allele frequencies of greater than 10% (rs10399931).
†Probe overlaps known SNPs with minor allele frequencies of greater than 10% (rs7515776).

For our secondary analyses (Fig 1, *B*), we performed separate multiple logistic regression models testing the association of SNPs, CpG methylation levels, and YKL-40 levels with asthma, with adjustment for sex, age, and body mass index at 4 years because these demographic factors can affect methylation and YKL-40 levels, as well as asthma risk. To test for nonlinearity of effects, YKL-40 levels were also used as tertiles, and asthma risks were compared between subjects having medium and high YKL-40 levels and subjects having low YKL-40 levels.

RESULTS

Table I shows characteristics of children included in the primary and secondary analyses. Table E2 in this article's Online Repository at www.jacionline.org compares demographic characteristics of participants from the 3 cohorts who were or were not included in the present study.

Primary analyses: Relation between *CHI3L1* SNPs, methylation, and YKL-40 levels

SNPs and YKL-40 protein levels. Among the 68 tested SNPs, after Bonferroni correction, 7 were found to be significantly associated with YKL-40 levels in INMA (see Table E3 in this article's Online Repository at www.jacionline.org for complete analysis and Table II for significant associations). Overall, minor alleles were associated with higher YKL-40 levels, with the exceptions of rs10399931 and rs4950928. These 7 SNPs explained individually up to 16% of the variability in YKL-40 levels ($R^2 = 0.08$-0.16), and they did not belong to a single block of LD. The LD matrix of *CHI3L1* SNPs in INMA is shown in Fig E1 in this article's Online Repository at www.jacionline.org.

In BAMSE associations with YKL-40 levels were replicated for 5 (rs2886117, rs10399931, rs10399805, rs4950928, and rs7542294) of the 7 SNPs that provided significant signals in INMA (see Table E4 in this article's Online Repository at www.jacionline.org).

DNA methylation and YKL-40 protein levels. Descriptive statistics of methylation levels at the 14 CpG sites are shown in Table E5 in this article's Online Repository at www.jacionline.org, their distributions are shown in Fig E2 in this article's Online Repository at www.jacionline.org, and their correlation matrices are shown in Fig 2 and Fig E4 in this article's Online Repository at www.jacionline.org. The methylation levels at the first 7 CpG sites covering the 5′ region to the first exon of the gene correlated with each other. This correlation pattern was found at all ages: birth (INMA), age 4 years (INMA and BAMSE), and age 8 years (BAMSE, Fig 2 and see E4).

Correlations between methylation levels at CpG sites and YKL-40 levels for INMA and BAMSE are shown in Table III. In INMA serum YKL-40 levels at age 4 years correlated significantly with methylation levels at age 4 years at 8 CpG sites after Bonferroni correction ($P = 3 \times 10^{-16}$ to .003). All correlations were negative; that is, the higher the methylation, the lower the YKL-40 level. Inverse correlations between methylation and YKL-40 levels at age 4 years were replicated in BAMSE for 5 of these CpG sites (cg13134650, cg07423149, cg17014757, cg14085262, and cg03625911; Fig 2 and Table III).[43] In addition, methylation levels at these 5 CpG sites, as measured at birth in INMA and at age 8 years in BAMSE, also correlated with YKL-40 levels at age 4 years (Table III and Fig E3 in this article's Online Repository at www.jacionline.org). These CpG associations with YKL-40 levels were confirmed after adjustment for blood cell composition (data not shown).

At each of these 5 CpG sites, methylation levels obtained by using the microarray and those obtained by means of pyrosequencing were strongly correlated (Spearman correlation coefficients = 0.60-0.92, see Fig E4).

SNPs and DNA methylation. Next, we tested associations between the 7 SNPs and the 5 CpG sites that were found to be related to YKL-40 levels in the aforementioned analyses (see Fig E5 in this article's Online Repository at www.jacionline.org for a

TABLE IV. Associations between the 7 SNPs and methylation levels at the 5 CpG sites that were significantly related to YKL-40 levels in the INMA cohort (n = 172 [methylation levels at 4 years] at top and n = 161 [methylation levels at birth] at bottom)

	cg13134650				cg07423149			
SNP (effect allele)	Regression coefficient‡	95% CI	*P* value	R^2§	Regression coefficient‡	95% CI	*P* value	R^2§
Year 4								
rs2886117 (A)	−1.01	−1.83 to −0.19	.02	0.03	−3.59	−5.34 to −1.84	$\mathbf{8 \times 10^{-5}}$	0.09
rs10399931 (T)	2.42	1.74 to 3.09	$\mathbf{4 \times 10^{-11}}$	0.20	6.52	5.13 to 7.90	$\mathbf{7 \times 10^{-17}}$	0.31
rs10399805 (A)	−1.04	−1.86 to −0.22	.01	0.03	−3.76	−5.50 to −2.01	$\mathbf{3 \times 10^{-5}}$	0.09
rs4950928 (G)	2.49	1.71 to 3.26	$\mathbf{2 \times 10^{-9}}$	0.17	6.66	5.06 to 8.25	$\mathbf{4 \times 10^{-14}}$	0.26
rs7542294 (A)	−1.22	−2.01 to −0.42	**.003**	0.05	−3.54	−5.25 to −1.83	$\mathbf{7 \times 10^{-5}}$	0.09
rs2791718 (A)	−1.09	−2.01 to −0.17	.02	0.03	−3.86	−5.79 to −1.92	$\mathbf{1 \times 10^{-4}}$	0.08
rs10920576 (T)	−1.23	−2.14 to −0.31	.009	0.04	−4.02	−5.94 to −2.10	$\mathbf{6 \times 10^{-5}}$	0.09
Birth								
rs2886117 (A)	−0.55	−1.43 to 0.33	.22	0.01	−1.88	−3.19 to −0.57	**.005**	0.04
rs10399931 (T)	2.02	1.25 to 2.79	$\mathbf{7 \times 10^{-7}}$	0.12	4.75	3.68 to 5.81	$\mathbf{2 \times 10^{-15}}$	0.29
rs10399805 (A)	−0.59	−1.47 to 0.30	.19	0.01	−1.91	−3.22 to −0.60	**.005**	0.04
rs4950928 (G)	1.80	0.91 to 2.68	$\mathbf{9 \times 10^{-5}}$	0.08	4.61	3.42 to 5.79	$\mathbf{2 \times 10^{-12}}$	0.23
rs7542294 (A)	−0.66	−1.53 to 0.21	.13	0.01	−1.97	−3.25 to −0.70	**.003**	0.05
rs2791718 (A)	−0.97	−1.94 to 0.01	.05	0.02	−2.10	−3.55 to −0.65	**.005**	0.05
rs10920576 (T)	−1.17	−2.14 to −0.19	.02	0.03	−2.17	−3.61 to −0.72	**.004**	0.05

Boldface indicates associations significant after Bonferroni correction.
*Probe overlaps known SNPs with minor allele frequencies of greater than 10% (rs10399931).
†Probe overlaps known SNPs with minor allele frequencies of greater than 10% (rs7515776, which is in nearly complete LD with rs10399805).
‡Coefficient from additive models predicting percentage methylation levels.
§Percentage variability in methylation levels explained by the SNP.
‖Association with an SNP overlapped by the CpG probe (or with an SNP in strong LD [$r^2 > 0.7$] with the overlapped SNP).

map of the CpG genomic locations). In INMA 4-year methylation levels at 4 of 5 CpG sites were significantly associated with all 7 SNPs after Bonferroni correction (Table IV), with alleles linked to lower YKL-40 levels being associated with higher methylation levels. These genetic associations were remarkably similar when methylation levels from cord blood were analyzed (Table IV). Based on the effect estimates and the percentage variability in methylation levels that they explained, the SNPs rs10399931 and rs4950928, which are in strong LD, showed the strongest associations for all 5 CpG sites. These 2 SNPs were also associated with methylation levels at all 5 CpG sites in BAMSE at both age 4 and 8 years (see Table E6 in this article's Online Repository at www.jacionline.org).

All the aforementioned 5 CpG sites showed moderate-to-strong correlation of methylation levels between birth and 4 years in INMA (Spearman correlation coefficients = 0.35-0.85) and between 4 and 8 years in BAMSE (Spearman correlation coefficients = 0.44-0.88, see Fig E6 in this article's Online Repository at www.jacionline.org), which is in line with the possibility of a consistent genetic control of their methylation levels from birth to age 8 years.

These 5 CpG sites also showed significant associations with *CHI3L1* SNPs and significant correlations between their methylation levels at 4 and 8 years of age in PIAMA (see Fig E6, *C*, and Table E7 in this article's Online Repository at www.jacionline.org).

Multivariable and mediation analyses. Given the aforementioned associations, we conducted multivariable analyses to identify independent effects by SNPs and CpG sites on YKL-40 levels. Among the 7 SNPs, in INMA the final backward stepwise regression model included rs4950928 and rs7542294, whereas the final model in BAMSE included rs10399931 and rs10399805. Of note, rs4950928 and rs7542294 are in strong LD with rs10399931 and rs10399805, respectively (see Fig E1), indicating that in both cohorts genetic influences on YKL-40 levels can be driven by these 2 groups of SNPs in LD. Among the 5 CpG sites, final stepwise models predicting YKL-40 levels included only cg07423149 in INMA and only cg17014757 in BAMSE, but it should be noted that methylation levels at the 5 CpG sites were strongly correlated with each other in both INMA (Fig 2; Spearman correlation coefficients = 0.64-0.87) and BAMSE (Spearman correlation coefficients = 0.51-0.83, see Fig E3), making it difficult to determine whether a single CpG site or rather a global regional methylation profile was driving the association with YKL-40 levels. To evaluate the components of methylation profiles across these 5 highly correlated CpG sites, we conducted a PCA and found that the first component explained up to 79% and up to 75% of variance in INMA and BAMSE, respectively.

To evaluate to what extent methylation profiles mediated the relation of *CHI3L1* variation to YKL-40 levels, we completed mediation analyses in INMA (the largest cohort with complete data). In these analyses we tested the 2 SNPs (rs4950928 and rs7542294) and the CpG site (cg07423149) identified by using multivariable analyses plus the first component identified from the PCA. In these analyses cg07423149 mediated 62% of the rs4950928 effects and 43% of the rs7542294 effects on YKL-40 levels (see Table E8 in this article's Online Repository at www.jacionline.org). Corresponding percentages for the overall methylation score were only slightly higher (66% and 46%, respectively). These results indicate that *CHI3L1* methylation profiles might mediate a substantial proportion of the effects of *CHI3L1* genetic variation on YKL-40 levels.

TABLE IV. Continued.

cg17014757*				cg14085262				cg03625911†			
Regression coefficient‡	95% CI	*P* value	R^2§	Regression coefficient‡	95% CI	*P* value	R^2§	Regression coefficient‡	95% CI	*P* value	R^2§
−6.56	−9.41 to −3.71	$\mathbf{1 \times 10^{-5}}$	0.11	−4.26	−5.93 to −2.60	$\mathbf{1 \times 10^{-6}}$	0.13	−2.76‖	−4.09 to −1.43	$\mathbf{6 \times 10^{-5}}$	0.08
15.55‖	14.28 to 16.81	$\mathbf{3 \times 10^{-57}}$	0.70	5.25	3.73 to 6.77	$\mathbf{1 \times 10^{-10}}$	0.20	4.09	2.94 to 5.23	$\mathbf{4 \times 10^{-11}}$	0.21
−6.56	−9.41 to −3.71	$\mathbf{1 \times 10^{-5}}$	0.11	−4.26	−5.93 to −2.58	$\mathbf{1 \times 10^{-6}}$	0.13	−2.75‖	−4.08 to −1.41	$\mathbf{7 \times 10^{-5}}$	0.08
16.16‖	14.56 to 17.75	$\mathbf{2 \times 10^{-46}}$	0.58	4.82	3.07 to 6.57	$\mathbf{2 \times 10^{-7}}$	0.14	4.02	2.70 to 5.35	$\mathbf{1 \times 10^{-8}}$	0.17
−5.00	−7.89 to −2.11	$\mathbf{8 \times 10^{-4}}$	0.07	−3.96	−5.61 to −2.31	$\mathbf{5 \times 10^{-6}}$	0.11	−2.78‖	−4.09 to −1.48	$\mathbf{4 \times 10^{-5}}$	0.09
−5.73	−8.98 to −2.47	$\mathbf{7 \times 10^{-4}}$	0.07	−3.46	−5.38 to −1.55	$\mathbf{5 \times 10^{-4}}$	0.07	−2.54	−4.04 to −1.04	**.001**	0.06
−5.79	−9.04 to −2.54	$\mathbf{6 \times 10^{-4}}$	0.07	−3.59	−5.51 to −1.68	$\mathbf{3 \times 10^{-4}}$	0.07	−2.60	−4.09 to −1.10	$\mathbf{7 \times 10^{-4}}$	0.06
−5.28	−8.04 to −2.52	$\mathbf{2 \times 10^{-4}}$	0.09	−3.31	−4.85 to −1.77	$\mathbf{4 \times 10^{-5}}$	0.09	−1.05‖	−2.18 to 0.09	.07	0.02
14.77‖	13.53 to 16.01	$\mathbf{1 \times 10^{-53}}$	0.71	3.92	2.49 to 5.35	$\mathbf{2 \times 10^{-7}}$	0.14	3.01	2.03 to 3.99	$\mathbf{9 \times 10^{-9}}$	0.17
−5.24	−8.02 to −2.46	$\mathbf{3 \times 10^{-4}}$	0.08	−3.46	−5.00 to −1.93	$\mathbf{2 \times 10^{-5}}$	0.10	−1.16‖	−2.30 to −0.03	.05	0.02
14.72‖	13.08 to 16.36	$\mathbf{2 \times 10^{-39}}$	0.56	3.55	1.93 to 5.16	$\mathbf{2 \times 10^{-5}}$	0.10	3.00	1.90 to 4.09	$\mathbf{2 \times 10^{-7}}$	0.14
−3.91	−6.72 to −1.10	**.007**	0.05	−3.17	−4.69 to −1.64	$\mathbf{6 \times 10^{-5}}$	0.09	−1.33‖	−2.44 to −0.22	.02	0.03
−5.22	−8.34 to −2.10	**.001**	0.07	−3.39	−5.11 to −1.67	$\mathbf{1 \times 10^{-4}}$	0.08	−1.49	−2.75 to −0.22	.02	0.03
−5.64	−8.74 to −2.54	$\mathbf{4 \times 10^{-4}}$	0.08	−3.83	−5.52 to −2.13	$\mathbf{2 \times 10^{-5}}$	0.11	−1.77	−3.03 to −0.51	**.006**	0.04

Secondary analyses: Associations of 3-level biomarkers (SNPs, methylation, and YKL-40) with asthma

We also tested all the markers (ie, variation at 7 SNPs, methylation at 5 CpG sites at age 4 years, and YKL-40 levels at age 4 years) for cross-sectional associations with asthma using the maximum number of available genetic and epigenetic samples in INMA, BAMSE, and PIAMA. Numbers and characteristics of participants included in these analyses are shown in Table I. Table V shows results of association analyses with asthma for each of the biomarkers. Neither the SNPs nor the CpG sites showed consistent associations with asthma in INMA, BAMSE, and PIAMA. The lack of association between methylation at any of the CpG sites and asthma was also confirmed after further adjustment for blood cell composition (data not shown). In contrast, YKL-40 levels were associated, although weakly, with asthma (Table V). In meta-analyses each SD increase in YKL-40 level was associated with a borderline 22% increase in the odds for asthma ($P = .08$), and participants in the highest YKL-40 tertile had 90% increased odds for asthma compared with subjects in the lowest tertile ($P = .03$).

In additional analyses we confirmed the lack of association of genetic and epigenetic markers with asthma when methylation levels at the other 9 CpG sites in *CHI3L1* were analyzed (see Table E9 in this article's Online Repository at www.jacionline.org) and when SNPs and methylation levels were tested for associations with the comorbidity cluster (asthma, rhinitis, and eczema) that was recently described[44] in these cohorts (see Table E10 in this article's Online Repository at www.jacionline.org).

DISCUSSION

This is the first study to integrate genetic and epigenetic regulation of YKL-40 levels in childhood. We found consistent interrelationships between *CHI3L1* genetic variants, methylation levels at several *CHI3L1* CpG sites, and circulating YKL-40 levels. These associations were replicated in multiple independent cohorts. Taken together with our mediation analysis, these data indicate that methylation levels mediate part of the known effects of genetic variation on YKL-40 levels. We also observed an association of YKL-40 levels, but not *CHI3L1* SNPs or methylation levels, with asthma.

YKL-40 has been proposed as a potential biomarker for a broad range of diseases.[1] Increased levels of YKL-40 have been consistently found in the blood and airways of adults with asthma, particularly those with severe and persistent disease.[2,3,5-10] The underlying mechanisms of these associations remain largely to be determined. *In vitro* studies have shown that human bronchial epithelial cells express *CHI3L1* and secrete YKL-40 in response to mechanical stress similar to that experienced during bronchoconstriction.[45] Furthermore, YKL-40 levels were increased in bronchoalveolar lavage fluid from patients with asthma on segmental allergen challenge.[46,47] In support of a possible causal involvement of this gene in asthma, *Chi3l1*-null mice showed decreased aeroallergen-induced T_H2 inflammatory responses in their lungs compared with wild-type animals.[10]

Whereas the link between YKL-40 levels and asthma has been consistently reported in adults, results from asthma studies in children have been conflicting. Serum YKL-40 levels at birth and in the first 5 years of life were not significantly associated with asthma at age 6 years in 2 studies, although positive trends were observed in both cases.[11,48] Serum YKL-40 levels were increased in children with therapy-resistant asthma,[12] but no association between YKL-40 levels and asthma severity was found in a subsequent study of 61 asthmatic children.[13] In our larger study we found children with asthma to have increased circulating YKL-40 levels, although this association was relatively weak. This

TABLE V. Associations of *CHI3L1* SNPs, *CHI3L1* methylation levels, and YKL-40 levels with asthma in the study cohorts

	INMA			BAMSE			PIAMA			Meta-analysis		
	adjOR	95% CI	*P* value	adjOR	95% CI	*P* value	adjOR	95% CI	*P* value	adjOR	95% CI	*P* value
SNPs (effect allele), no. = 2405*												
rs2886117 (A)	1.30	0.774-2.174	.32	0.85	0.552-1.294	.44	1.30	0.811-2.080	.28	1.09	0.835-1.431	.516
rs10399931 (T)‡§	0.55	0.305-0.992	.05	1.27	0.919-1.748	.15	0.91	0.590-1.393	.66	1.00	0.792-1.269	.985
rs10399805 (A)	1.40	0.829-2.354	.21	0.75	0.482-1.178	.21	1.12	0.693-1.825	.64	1.02	0.776-1.352	.867
rs4950928 (G)‡§	0.50	0.250-0.986	.05	1.44	1.014-2.042	.04	0.74	0.457-1.193	.22	1.01	0.778-1.314	.933
rs7542294 (A)	1.12	0.658-1.909	.67	0.81	0.540-1.219	.31	1.18	0.754-1.855	.47	1.00	0.767-1.297	.987
rs2791718 (A)‡§	0.66	0.331-1.295	.22	0.68	0.422-1.094	.11	1.42	0.873-2.313	.16	0.90	0.664-1.221	.499
rs10920576 (T)	0.65	0.338-1.253	.20	0.67	0.435-1.046	.08	1.41	0.886-2.252	.15	0.89	0.665-1.181	.411
Methylation levels, n = 625†												
cg13134650‖	0.95	0.837-1.081	.44	1.07	0.966-1.187	.19	1.22	0.967-1.550	.09	1.04	0.965-1.123	.297
cg07423149‖	1.01	0.945-1.075	.81	0.99	0.951-1.029	.59	1.06	0.981-1.150	.14	1.00	0.974-1.036	.796
cg17014757‖	0.98	0.941-1.021	.33	1.01	0.985-1.036	.42	1.05	0.995-1.116	.07	1.01	0.988-1.028	.446
cg14085262‖	1.02	0.952-1.090	.59	0.96	0.926-1.004	.08	1.07	0.987-1.162	.10	0.99	0.960-1.024	.603
cg03625911‖	1.04	0.951-1.128	.42	0.98	0.932-1.026	.37	1.09	0.992-1.200	.07	1.01	0.969-1.047	.714
Standardized YKL-40 levels, n = 432¶												
Effect for 1-SD increase in YLK-40	1.27	0.861-1.860	.23	1.21	0.913-1.591	.19	NA			1.23	0.978-1.534	.077
Medium vs low YKL-40 tertile	2.62	0.927-7.377	.07	1.50	0.761-2.96	.24	NA			1.77	1.004-3.130	.048
High vs low YKL-40 tertile	2.36	0.820-6.773	.11	1.75	0.889-3.427	.11	NA			1.90	1.078-3.364	.026

Complete biomarker data were available for INMA and BAMSE. No YKL-40 levels were available for PIAMA. All models were adjusted for sex, age, and body mass index.
adjOR, Adjusted odds ratio; *NA*, not available.
*No. asthma/no. total: INMA (58/336), BAMSE (171/467), and PIAMA (69/1602).
†No. asthma/no. total: INMA (33/203), BAMSE (88/229), and PIAMA (13/193).
‡Imputed in BAMSE (minimal quality of imputation > 0.77).
§Imputed in PIAMA (minimal quality of imputation > 0.969).
¶No. asthma/no. total: INMA (33/203) and BAMSE (88/229).
‖Odds ratio for CpG sites express effects for 1% increase in their methylation levels.

might be due to the fact that we studied population-based epidemiologic cohorts in which the prevalence of severe asthma is expected to be quite low.

Circulating YKL-40 levels have been previously described to be under strong genetic control. In multivariable analyses we found SNPs from 2 groups of LD to be independently associated with serum YKL-40 levels. Among them, the 2 SNPs rs10399931 and rs4950928 fell within a single LD block and were identified as the strongest protein quantitative trait loci in our study, which is consistent with previous reports.[11,49] They are located in proximity of the transcription start site, and because they are in strong LD, it is difficult to dissect their independent effects on YKL-40 levels and disease risk. Gene reporter assays for *CHI3L1* promoter haplotypes indicated that both SNPs contributed significant cis-regulatory effects on gene expression in Jurkat cells and that the magnitude of these effects was the strongest for rs10399931,[16] which was also the SNP with the strongest effects on methylation levels in our study.

Findings from our study indicate that the effects of *CHI3L1* SNPs on circulating YKL-40 levels might be mediated by methylation levels at *CHI3L1* CpG sites. Several SNPs were associated with methylation levels at 5 CpG sites located within 1 kb of the transcription start site that were, in turn, negatively correlated with YKL-40 levels. It should be noted that the probe for cg17014757 includes rs10399931 and the probe for cg03625911 includes rs7515776. Therefore the strong associations that were found between methylation levels at these CpG sites and the corresponding SNPs (or other SNPs in LD with them) might be due to allele-specific differences in probe hybridization.[37] However, similar genetic associations, although smaller in magnitude than those observed for cg17014757, were found for CpG sites (ie, cg13134650, cg07423149, and cg14085262), the probes of which do not include known SNPs. Mediation analyses that were completed both on cg07423149 and on the methylation principal component supported SNP→CpG→YKL-40 causal models.

Our findings are in line with a growing body of evidence that points to significant effects of DNA sequence variants across the genome on methylation levels at nearby and distal CpG sites.[23,25-30] Previous studies have shown that these methylation quantitative trait loci (meQTLs) can in turn affect gene expression[25,50] and are enriched in motifs for DNA-binding factors and DNaseI hypersensitivity regions.[29,51] meQTLs have been shown to be enriched for disease risk variants,[29] and they can ultimately influence disease risk through their effects on methylation patterns.[23,28] Similarly, *CHI3L1* meQTLs could influence asthma risk by affecting methylation, gene expression, and, in turn, YKL-40 levels. However, in contrast with this scenario, in our study neither *CHI3L1* SNPs nor methylation at *CHI3L1* CpG sites were associated with asthma, although YKL-40 levels tended to be higher in children with asthma. These results have 2 alternate explanations: either they represent a true negative finding, which would argue against a causal role of *CHI3L1* variants in childhood asthma, or they are due to possible methodological factors that should be taken into account. First, asthma, particularly in the preschool years, is characterized by a large phenotypic heterogeneity that can affect our ability to capture phenotype-specific genetic associations. Along the same lines, the lack of genetic associations can be explained by *CHI3L1* influencing asthma risk through interactions with environmental factors or through effects on asthma phenotypes that were not included or were underrepresented in our study (eg, adult or severe asthma). It is also possible that our genetic association analyses were underpowered to detect a true signal, although in that

case the magnitude of the *CHI3L1* genetic effects would be expected to be relatively small.

Our study has some limitations. Although we did not find any statistical evidence that interrelationships between *CHI3L1* genetic variation, methylation, and YKL-40 levels differed between cases and control subjects, these stratified analyses had limited sample size, and potential differences by disease status could not be determined conclusively. The methylation array that we used is primarily designed for genomic discovery. Thus information from a large proportion of *CHI3L1* CpG sites and, in turn, analyses on regional methylation profiles could not be included in our study and will need to be addressed in future studies. Finally, we acknowledge that, by using whole blood for methylation studies, we might have missed contributions of methylation profiles from other tissues (eg, from the airways), and our results might have been affected by blood cell composition, particularly in analyses that used cord blood.[52] However, it should be noted that we were able to replicate all the associations between methylation at the 5 CpG sites and serum YKL-40 levels after adjustment for blood cell composition, both as estimated by using the Houseman method and, in a subset of INMA participants, as directly assessed from blood smears.

In conclusion, in multiple independent cohorts we found genetic variation in the *CHI3L1* gene to be related to both methylation levels in nearby CpG sites and circulating YKL-40 levels. Our findings indicate that *CHI3L1* genetic variation can affect circulating YKL-40 levels by regulating its gene methylation profiles.

Pyrosequencing was performed at University Medical Center Groningen in the Laboratory of Obstetrics under the supervision of Dr Torsten Plösch. We thank Dr Xavier Basagaña for his helpful advice on mediation analysis.

Clinical implications: Methylation levels at *CHI3L1* CpG sites mediate part of the effects of *CHI3L1* genetic variation on circulating levels of YKL-40, but they are not associated with childhood asthma.

REFERENCES

1. Lee CG, Da Silva CA, Dela Cruz CS, Ahangari F, Ma B, Kang MJ, et al. Role of chitin and chitinase/chitinase-like proteins in inflammation, tissue remodeling, and injury. Annu Rev Physiol 2011;73:479-501.
2. Chupp GL, Lee CG, Jarjour N, Shim YM, Holm CT, He S, et al. A chitinase-like protein in the lung and circulation of patients with severe asthma. N Engl J Med 2007;357:2016-27.
3. Tang H, Fang Z, Sun Y, Li B, Shi Z, Chen J, et al. YKL-40 in asthmatic patients, and its correlations with exacerbation, eosinophils and immunoglobulin E. Eur Respir J 2010;35:757-60.
4. James AJ, Reinius LE, Verhoek M, Gomes A, Kupczyk M, Hammar U, et al. Increased YKL-40 and chitotriosidase in asthma and chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2016;193:131-42.
5. Otsuka K, Matsumoto H, Niimi A, Muro S, Ito I, Takeda T, et al. Sputum YKL-40 levels and pathophysiology of asthma and chronic obstructive pulmonary disease. Respiration 2012;83:507-19.
6. Guerra S, Halonen M, Sherrill DL, Venker C, Spangenberg A, Carsin AE, et al. The relation of circulating YKL-40 to levels and decline of lung function in adult life. Respir Med 2013;107:1923-30.
7. Lai T, Chen M, Deng Z, L Y Wu D, Li D, et al. YKL-40 is correlated with FEV1 and the asthma control test (ACT) in asthmatic patients: influence of treatment. BMC Pulm Med 2015;15:1.
8. Duru S, Yuce G, Ulasli SS, Erdem M, Kizilgun M, Kara F, et al. The relationship between serum YKL-40 levels and severity of asthma. Iran J Allergy Asthma Immunol 2013;12:247-53.
9. Saba M, Sharif MR, Akbari H, Nikoueinejad H, Ramazani Jolfaii M. YKL-40 in asthma and its correlation with different clinical parameters. Iran J Allergy Asthma Immunol 2014;13:271-7.
10. Ahangari F, Sood A, Ma B, Takyar S, Schuyler M, Qualls C, et al. Chitinase 3-like-1 regulates both visceral fat accumulation and asthma-like Th2 inflammation. Am J Respir Crit Care Med 2015;191:746-57.
11. Ober C, Tan Z, Sun Y, Possick JD, Pan L, Nicolae R, et al. Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. N Engl J Med 2008;358:1682-91.
12. Konradsen JR, James A, Nordlund B, Reinius LE, Soderhall C, Melen E, et al. The chitinase-like protein YKL-40: a possible biomarker of inflammation and airway remodeling in severe pediatric asthma. J Allergy Clin Immunol 2013; 132:328-35.e5.
13. Santos CB, Davidson J, Covar RA, Spahn JD. The chitinase-like protein YKL-40 is not a useful biomarker for severe persistent asthma in children. Ann Allergy Asthma Immunol 2014;113:263-6.
14. Lee CG, Hartl D, Lee GR, Koller B, Matsuura H, Da Silva CA, et al. Role of breast regression protein 39 (BRP-39)/chitinase 3-like-1 in Th2 and IL-13-induced tissue responses and apoptosis. J Exp Med 2009;206:1149-66.
15. Cunningham J, Basu K, Tavendale R, Palmer CN, Smith H, Mukhopadhyay S. The CHI3L1 rs4950928 polymorphism is associated with asthma-related hospital admissions in children and young adults. Ann Allergy Asthma Immunol 2011; 106:381-6.
16. Verlaan DJ, Ouimet M, Adoue V, Sirois-Gagnon D, Lariviere M, Ge B, et al. Promoter polymorphisms in CHI3L1 are associated with asthma. J Allergy Clin Immunol 2012;130:533-5.
17. Gomez JL, Crisafi GM, Holm CT, Meyers DA, Hawkins GA, Bleecker ER, et al. Genetic variation in chitinase 3-like 1 (CHI3L1) contributes to asthma severity and airway expression of YKL-40. J Allergy Clin Immunol 2015; 136:51-8.e10.
18. Tsai Y, Ko Y, Huang M, Lin M, Wu C, Wang C, et al. CHI3L1 polymorphisms associate with asthma in a Taiwanese population. BMC Med Genet 2014;15:86.
19. Kjaergaard AD, Johansen JS, Nordestgaard BG, Bojesen SE. Genetic variants in CHI3L1 influencing YKL-40 levels: resequencing 900 individuals and genotyping 9000 individuals from the general population. J Med Genet 2013;50:831-7.
20. Rathcke CN, Holmkvist J, Husmoen LL, Hansen T, Pedersen O, Vestergaard H, et al. Association of polymorphisms of the CHI3L1 gene with asthma and atopy: a populations-based study of 6514 Danish adults. PLoS One 2009;4:e6106.
21. Wu AC, Lasky-Su J, Rogers CA, Klanderman BJ, Litonjua A. Polymorphisms of chitinases are not associated with asthma. J Allergy Clin Immunol 2010;125: 754-7, e1-7.
22. Sohn MH, Lee JH, Kim KW, Kim SW, Lee SH, Kim KE, et al. Genetic variation in the promoter region of chitinase 3-like 1 is associated with atopy. Am J Respir Crit Care Med 2009;179:449-56.
23. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat Rev Genet 2011;12:529-41.
24. Gruzieva O, Merid SK, Melen E. An update on epigenetics and childhood respiratory diseases. Paediatr Respir Rev 2014;15:348-54.
25. Zhang D, Cheng L, Badner JA, Chen C, Chen Q, Luo W, et al. Genetic control of individual differences in gene-specific methylation in human brain. Am J Hum Genet 2010;86:411-9.
26. Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E, et al. Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. Nat Genet 2008;40:904-8.
27. Hellman A, Chess A. Extensive sequence-influenced DNA methylation polymorphism in the human genome. Epigenetics Chromatin 2010;3:11.
28. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nat Biotechnol 2013;31:142-7.
29. Shi J, Marconett CN, Duan J, Hyland PL, Li P, Wang Z, et al. Characterizing the genetic basis of methylome diversity in histologically normal human lung tissue. Nat Commun 2014;5:3365.
30. Lemire M, Zaidi SH, Ban M, Ge B, Aissi D, Germain M, et al. Long-range epigenetic regulation is conferred by genetic variation located at thousands of independent loci. Nat Commun 2015;6:6326.
31. Bousquet J, Anto J, Auffray C, Akdis M, Cambon-Thomsen A, Keil T, et al. MeDALL (Mechanisms of the Development of ALLergy): an integrated approach from phenotypes to systems medicine. Allergy 2011;66:596-604.
32. Ribas-Fito N, Ramon R, Ballester F, Grimalt J, Marco A, Olea N, et al. Child health and the environment: the INMA Spanish Study. Paediatr Perinat Epidemiol 2006;20:403-10.
33. Kull I, Melen E, Alm J, Hallberg J, Svartengren M, van Hage M, et al. Breastfeeding in relation to asthma, lung function, and sensitization in young schoolchildren. J Allergy Clin Immunol 2010;125:1013-9.

34. Brunekreef B, Smit J, de Jongste J, Neijens H, Gerritsen J, Postma D, et al. The prevention and incidence of asthma and mite allergy (PIAMA) birth cohort study: design and first results. Pediatr Allergy Immunol 2002; 13(suppl 15):55-60.
35. Wijga AH, Kerkhof M, Gehring U, de Jongste JC, Postma DS, Aalberse RC, et al. Cohort profile: the prevention and incidence of asthma and mite allergy (PIAMA) birth cohort. Int J Epidemiol 2014;43:527-35.
36. 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. Lancet Respir Med 2014;2:131-40.
37. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 2013;8:203-9.
38. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics 2013;14:293.
39. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 2012;13:86.
40. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 2005;76:887-93.
41. Guerra S, Halonen M, Vasquez MM, Spangenberg A, Stern DA, Morgan WJ, et al. Relation between circulating CC16 concentrations, lung function, and development of chronic obstructive pulmonary disease across the lifespan: a prospective study. Lancet Respir Med 2015;3:613-20.
42. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: R package for causal mediation analysis. J Stat Software 2014;59(5).
43. Martin TC, Yet I, Tsai PC, Bell JT. coMET: visualisation of regional epigenome-wide association scan results and DNA co-methylation patterns. BMC Bioinformatics 2015;16:131.
44. Garcia-Aymerich J, Benet M, Saeys Y, Pinart M, Basagana X, Smit HA, et al. Phenotyping asthma, rhinitis and eczema in MeDALL population-based birth cohorts: an allergic comorbidity cluster. Allergy 2015;70:973-84.
45. Park JA, Drazen JM, Tschumperlin DJ. The chitinase-like protein YKL-40 is secreted by airway epithelial cells at base line and in response to compressive mechanical stress. J Biol Chem 2010;285:29817-25.
46. Kuepper M, Bratke K, Virchow JC. Chitinase-like protein and asthma. N Engl J Med 2008;358:1073-5, author reply 5.
47. Gavala ML, Kelly EA, Esnault S, Kukreja S, Evans MD, Bertics PJ, et al. Segmental allergen challenge enhances chitinase activity and levels of CCL18 in mild atopic asthma. Clin Exp Allergy 2013;43:187-97.
48. Usemann J, Frey U, Mack I, Schmidt A, Gorlanova O, Roosli M, et al. CHI3L1 polymorphisms, cord blood YKL-40 levels and later asthma development. BMC Pulm Med 2016;16:81.
49. Kruit A, Grutters JC, Ruven HJ, van Moorsel CC, van den Bosch JM. A CHI3L1 gene polymorphism is associated with serum levels of YKL-40, a novel sarcoidosis marker. Respir Med 2007;101:1563-71.
50. Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol 2011;12:R10.
51. Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schubeler D. Identification of genetic elements that autonomously determine DNA methylation states. Nat Genet 2011;43:1091-7.
52. Yousefi P, Huen K, Quach H, Motwani G, Hubbard A, Eskenazi B, et al. Estimation of blood cellular heterogeneity in newborns and children for epigenome-wide association studies. Environ Mol Mutagen 2015;56:751-8.