Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Çalışkan M, Bochkov YA, Kreiner-Møller E, et al. Rhinovirus wheezing illness and genetic risk of childhood-onset asthma. N Engl J Med 2013;368:1398-407. DOI: 10.1056/NEJMoa1211592

Supplementary Appendix

Supplement to: Çalışkan M, Bochkov YA, Kreiner-Møller E, Bønnelykke K, Stein MM, Du G, Bisgaard H, Jackson DJ, Gern JE, Lemanske RF, Jr., Nicolae DL, and Ober C. **Rhinovirus wheezing illness and genetic risk of childhood onset asthma.**

Table of Contents

Supplementary Methods

Phenotype descriptions in the COAST cohort

Children were diagnosed with asthma if they fulfilled at least one of the following criteria: (1) physician-diagnosed asthma; (2) frequent albuterol use for coughing or wheezing episodes as prescribed by a physician; (3) use of a prescribed daily controller medication; (4) an implemented step-up plan, including use of albuterol or inhaled corticosteroids during illness as prescribed by a physician; (5) use of prednisone for an asthma exacerbation. 200 of the COAST children were tested for asthma at ages six and/or eight years. A diagnosis at either age was considered in this study as the 'asthma' phenotype.

Allergic sensitization was determined by measuring both allergen-specific IgE (*in vitro* IgE) to nine allergens and skin prick test (SPT) reactivity to 12 allergens, as described[1]. Children were classified as atopic if they had at least one positive *in vitro* IgE or at least one positive SPT. A respiratory wheezing illness in the first 3 years of life was defined as: (1) physician-diagnosed wheezing at an office visit; (2) an illness for which the child was prescribed short- or long-acting β-agonists and/or controller medications; or (3) an illness with the following specific diagnoses: bronchiolitis, wheezing illness, reactive airway disease, asthma, or asthma exacerbation. The presence and timing of viral wheezing illnesses during the first three years of life were assessed in 214 children prospectively using nasopharyngeal mucus samples and PCR-based viral diagnostics as previously described[1]. Human rhinovirus (HRV) and respiratory syncytial virus (RSV) wheezing illnesses were defined as the

presence of the specific virus in nasal secretions during a wheezing illness in the first three years of life, as described[1]. Number of HRV and RSV wheezing illnesses was defined as the number of times the specific virus was detected in nasal secretions during distinct wheezing episodes in years 1-3.

Phenotype descriptions in the COPSAC cohort

COPSAC children were diagnosed with asthma if they fulfilled the following four criteria: (1) recurrent episodes (at least five episodes within 6 months; each episode lasting at least three consecutive days) of troublesome lung symptoms that were recorded in diaries; (2) symptoms (e.g., exercise induced symptoms; prolonged nocturnal cough; persistent cough not due to common cold, and symptoms causing wakening at night) typical of asthma based on doctors' interviews of the parents at the clinical research unit; (3) intermittent use of an inhaled β-agonists; and (4) response to a three-month course of inhaled corticosteroids and relapse when stopping treatment[2]. Asthma diagnosis was made at age seven years. The presence and timing of viral wheezing illnesses during the first three years of life were assessed prospectively using nasopharyngeal aspirates and PCR-based viral diagnostics, as previously described[3].

Genetic association and interaction studies

Each SNP was tested for association with each phenotype in the COAST children using a logistic or linear regression model depending on whether the outcome variable was binary or continuous, respectively, as implemented in JMP (version 10) statistical software (SAS Institute Inc.). The model included the

phenotype of interest as the outcome variable and genotype as an explanatory variable. Genotypes were coded as 0, 1, 2 'doses' of the minor allele under an additive genetic model. To test for interactions between 17q21 SNP genotypes and viral (HRV or RSV) wheezing illnesses, viral wheezing illness and a viral wheezing illness*genotype interaction term were included as covariates. In addition, "the number of wheezing episodes in which other (i.e., non-HRV or non-RSV) virus was present" was included as a covariate to assure that observed effects were specific to the virus being tested. In the COPSAC children, rs7216389 was tested for interaction with HRV or RSV wheezing illnesses on asthma risk as described above for COAST children using R statistical environment (http://www.r-project.org/). Combined interaction p-values in the COAST and COPSAC cohorts were obtained using MetaP software (http://compute1.lsrc.duke.edu/softwares/MetaP/metap.php).

To correct for multiple testing, the effective number of independent tests given the patterns of pairwise LD was calculated as implemented in the matrix spectral decomposition program[4]. The five SNPs at the 17q21 locus reduced to 1.23 independent variables, generating a significance threshold of 0.04 to maintain a 5% Type I error rate. Odds Ratios (OR) or incidence risk rates were estimated using unconditional maximum likelihood method and 95% Confidence Intervals (CI) were calculated using normal approximation.

HRV16 stimulation of PBMCs, and DNA and RNA extraction

Twenty ml of blood was drawn from each Chicago participant. Between 13 and 16 blood samples were processed on each of seven days of collection. To minimize technical variables that could affect gene expression, one person

(M.C.) processed all samples immediately after collection and in the order that the blood was drawn. PBMCs were isolated from whole blood samples using a Ficoll-Paque separation protocol. After PBMC isolation, cells were counted on a hemocytometer and then resuspended in media, bringing cell concentration to 4 x 10^6 cells/ml. From each subject, 4 x 10^6 cells were treated with media alone for 24 hours and 4 x 10^6 cells were treated with media containing HRV16 for 24 hours. The multiplicity of infection was 10 plaque-forming units per cell. The remaining cells from each subject were used for DNA extraction.

DNA was extracted on the day of sample collection using QIAamp DNA Blood Mini Kit (Qiagen); concentrations were measured on a Nanodrop ND-100 Spectrophotometer (NanoDrop Technologies). Total RNA was extracted after 24-hour incubation, using the RNeasy Plus Mini Kit (Qiagen); concentrations were measured on a Nanodrop ND-100 Spectrophotometer (NanoDrop Technologies) and quality was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies).

Analysis of gene expression data

Among five genes in the 17q21 locus, *IKZF3*, *GSDMB*, and *ORMDL3* genes were successfully amplified in unstimulated and HRV-stimulated PBMCs. Each qPCR reaction was run in duplicate and only samples with a coefficient of variation (CV) less than 0.02 were included in the analyses. After these exclusions, the sample sizes were 96, 97, and 97 for *IKZF3*, *GSDMB*, and *ORMDL3*, respectively. Expression levels of each gene were normalized against *POLR2C* expression within each individual and calibrated against a control

sample, cDNA extracted from T cells from a single individual was run as a control on each plate, using the Delta Delta Ct method[5]. Each data point was divided by the mean gene expression level in unstimulated cells in order to set the mean of unstimulated gene expression to 1. The effects of potential covariates (age, gender, ethnicity, virus batch, processing day, qPCR plate number, cDNA plate number, PBMC count, asthma status) on gene expression levels were explored and the effects two significant covariates, processing day and qPCR plate effect, were regressed out of the gene expression data prior to analysis.

Each 17q21 gene was tested for differential expression between unstimulated and HRV-stimulated cells using a paired t-test. Associations between 17q21 SNPs and gene expression were tested under an additive genetic model. All statistical analyses were performed as implemented in JMP (version 10) statistical software (SAS Institute Inc.).

Figure S1. Genes at the 17q21 locus and patterns of pairwise LD between SNPs in this study. (A) Location of the 17q21 asthma associated locus and the genomic organization of five genes within the association interval; black rectangles are exons and arrows above each gene show direction of transcription. (B) SNPs included in this study and patterns of pairwise LD (r^2) observed in the COAST cohort. Value for r^2 (x100) are shown, solid black boxes are $r^2 = 1$.

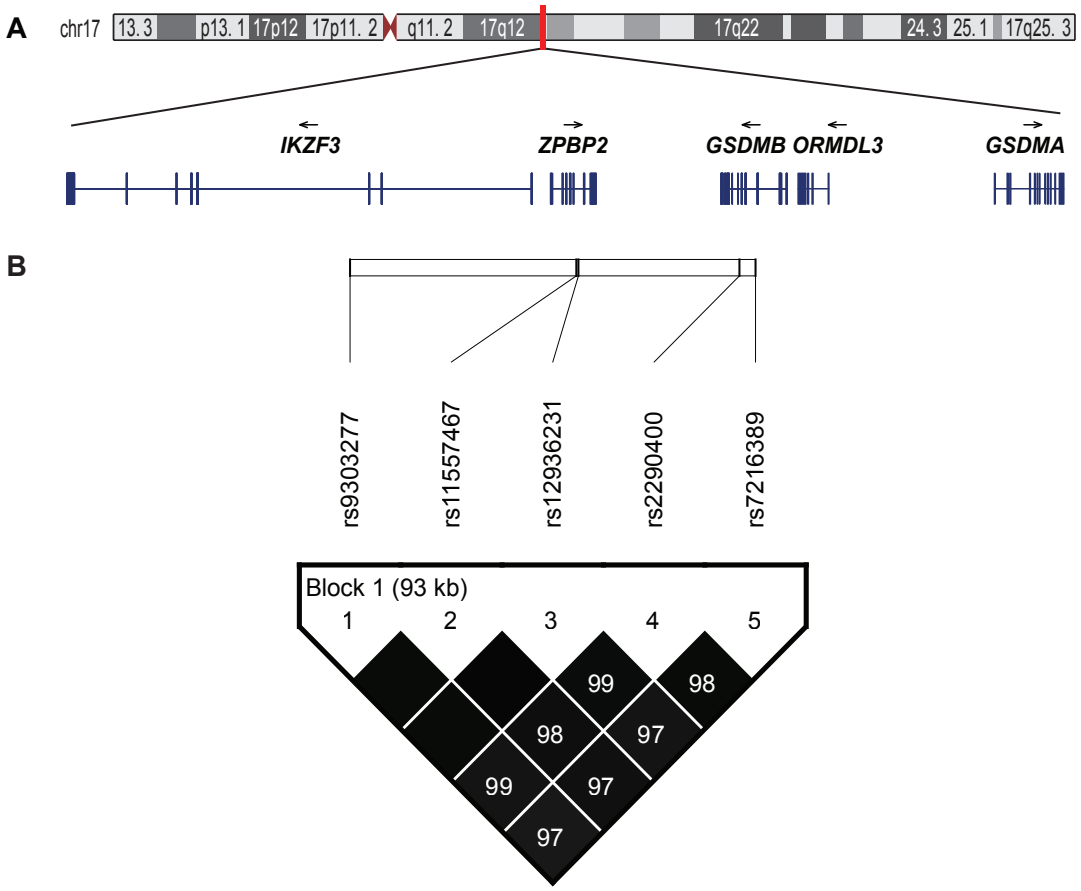

Figure S2. Association between 17q21 genotype and response to HRV stimulation in the Chicago cohort. Difference in gene expression in response to HRV stimulation is stratified by rs7216389 genotype. Sample sizes of each genotype group are shown under the x-axis. Each grey circle corresponds to an individual subject. Green lines show the mean difference in gene expression within each genotype group.

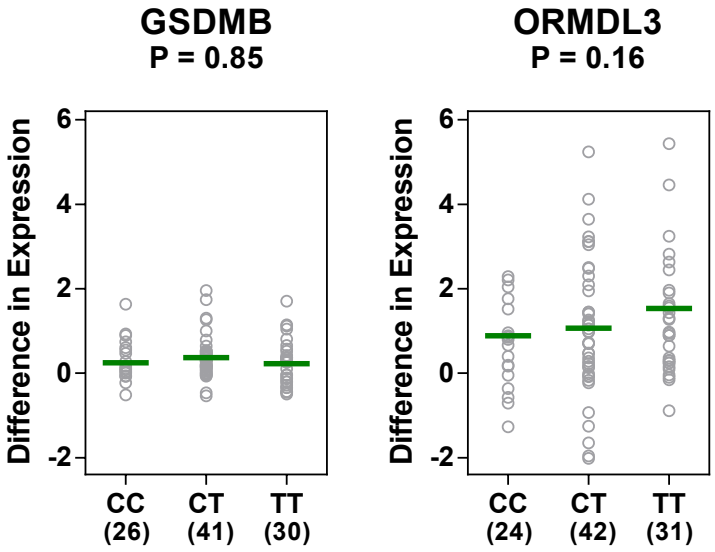

Table S1. Characteristics of the COAST and COPSAC cohorts.

	Gender		Asthma*		HRV wheezing illness (ages 1-3)		RSV wheezing illness (ages 1-3)	
	Female	**Male**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**
COAST	91	123	77	123	67	147	66	148
COPSAC	152	145	45	252	73	224	56	241

***** Asthma diagnosis was made at ages 6 and/or 8 in the COAST children and at age 7 in the COPSAC children.

Table S2. Primer sequences.

Transcript	Direction	Sequence (5' → 3')
POLR2C	Forward	GAGACCTCATCTCCAACAGC
	Reverse	ATAGGCTCGAAGTCTCAGCTC
IKZF3	Forward	GTGGAAAGATGAACTGCGATG
	Reverse	AGAATGTGTCCTAAGATGCCC
ZPBP2	Forward	CTCCAGGCAAGAGATCGAATAG
	Reverse	ACGCAGGTCTGACAAGTTAC
GSDMB	Forward	TCAGCTATAAACACAAGGGCC
	Reverse	CTTAGCGAGGGAGTTTAGCAC
ORMDL3	Forward	ATCCCGTTTGTGAGTGTCC
	Reverse	CTTAGTGTAGAAGCTGGTGAGG
GSDMA	Forward	GCTGTAACCATCCCCAAGG
	Reverse	CCTCTCCTGACTTTTCTCCAG

Table S3. Main and interaction effects of 17q21 SNPs and HRV wheezing illness on asthma risk in the COAST cohort. Note that all odds ratios are relative to the homozygous non-risk genotype with no HRV wheezing illness.

SNP	HRV wheezing illness in the first three years of life (Sample Size)	Sample Size (Asthma YES/NO)	Asthma (P-value)	Odds Ratios (95% CI interval)			17q21-HRV wheezing illness interaction (P-value)
				Homozygous non-risk genotype	Heterozygous genotype	Homozygous risk genotype	
rs9303277	NO (133)	34/99	0.98	1.0	1.2 (0.5-3.1)	1.0 (0.3-3.1)	0.01
	YES (64)	41/23	0.009	2.3 (0.6-9.1)	4.6 (1.7-12.7)	25.8 (5.0-134.0)	
rs11557467	NO (133)	34/99	0.79	1.0	1.2 (0.5-3.1)	0.8 (0.3-2.6)	0.007
	YES (64)	41/23	0.007	2.2 (0.6-8.8)	4.2 (1.5-11.7)	26.4 (5.1-137.2)	
rs12936231	NO (135)	35/100	0.75	1.0	1.1 (0.5-2.7)	0.8(0.3-2.5)	0.01
	YES (65)	23/42	0.01	2.5 (0.7-9.2)	3.9 (1.5-10.7)	24.7 (4.8-126.1)	
rs2290400	NO (135)	35/100	0.65	1.0	1.2 (0.5-2.9)	0.7 (0.2-2.3)	0.007
	YES (65)	23/42	0.009	2.5 (0.7-9.2)	3.7 (1.4-10.2)	26.1 (5.1-133.0)	
rs7216389	NO (135)	35/100	0.70	1.0	1.2 (0.5-2.8)	0.8 (0.2-2.4)	0.004
	YES (65)	23/42	0.006	2.1 (0.5-8.0)	3.9 (1.5-10.7)	26.1 (5.1-133.0)	

Table S4. Main and interaction effects of 17q21 SNPs and RSV wheezing illness on asthma risk in the COAST cohort. Note that all odds ratios are relative to the homozygous non-risk genotype with no RSV wheezing illness.

SNP	RSV wheezing illness in the first three years of life (Sample Size)	Sample Size (Asthma YES/NO)	Asthma (P-value)	Odds Ratios (95% CI interval)			17q21-RSV wheezing illness interaction (P-value)
				Homozygous non-risk genotype	Heterozygous genotype	Homozygous risk genotype	
rs9303277	NO (135)	41/94	0.05	1.0	1.9 (0.7-5.3)	3.0 (1.0-9.3)	0.46
	YES (62)	28/34	0.10	3.5 (1.0-12.5)	4.9 (1.6-14.9)	17.3 (2.9-103.4)	
rs11557467	NO (136)	41/95	0.09	1.0	2.0 (0.7-5.5)	2.6 (0.9-8.0)	0.38
	YES (61)	27/34	0.10	3.9 (1.1-14.2)	4.6 (1.5-14.1)	19.5 (3.3-114.6)	
rs12936231	NO (137)	42/95	0.12	1.0	1.7 (0.6-4.4)	2.4 (0.8-6.9)	0.37
	YES (63)	28/35	0.10	3.3 (1.0-11.4)	4.0 (1.3-11.6)	16.7 (2.9-95.7)	
rs2290400	NO (137)	42/95	0.12	1.0	1.7 (0.6-4.4)	2.3 (0.8-6.7)	0.36
	YES (63)	28/35	0.10	3.3 (1.0-11.4)	4.0 (1.3-11.6)	16.7 (2.9-95.7)	
rs7216389	NO (137)	42/95	0.10	1.0	2.1 (0.8-5.5)	2.4 (0.8-7.0)	0.30
	YES (63)	28/35	0.07	3.0 (0.9-10.4)	4.2 (1.4-12.2)	16.7 (2.9-95.7)	

Table S5. (A) Main and interaction effects of rs7216389 and HRV wheezing illness on asthma risk in the COPSAC cohort. Note that all odds ratios are relative to the homozygous CC genotype with no HRV wheezing illness. **(B) Main and interaction effects of rs7216389 and RSV wheezing illness on asthma risk in the COPSAC cohort.** Note that all odds ratios are relative to the homozygous CC genotype with no RSV wheezing illness.

A

SNP	HRV wheezing illness in the first three years of life (Sample Size)	Odds Ratios (95% CI interval)			17q21-HRV wheezing illness interaction (P-value)
		Homozygous non-risk genotype	Heterozygous genotype	Homozygous risk genotype	
rs7216389	NO (224)	1.0	1.0 (0.4-2.7)	0.7 (0.2-2.4)	0.08
	YES (73)	1.5 (0.3-8.2)	2.5 (0.8-7.6)	3.9 (1.3-11.7)	

B

SNP	RSV wheezing illness in the first three years of life (Sample Size)	Odds Ratios (95% CI interval)			17q21-RSV wheezing illness interaction (P-value)
		Homozygous non-risk genotype	Heterozygous genotype	Homozygous risk genotype	
rs7216389	NO (241)	1.0	1.0 (0.4-2.7)	1.2 (0.4-3.4)	0.29
	YES (56)	1.5 (0.3-8.2)	2.9 (0.9-9.4)	3.4 (1.0-11.9)	

References:

1. Jackson DJ, Gangnon RE, Evans MD, et al. Wheezing rhinovirus illnesses in early life predict asthma development in high-risk children. Am J Respir Crit Care Med 2008;178:667-72.

2. Bisgaard H, Pipper CB, Bonnelykke K. Endotyping early childhood asthma by quantitative symptom assessment. J Allergy Clin Immunol 2011;127:1155,64.e2.

3. Bisgaard H, Hermansen MN, Bonnelykke K, et al. Association of bacteria and viruses with wheezy episodes in young children: prospective birth cohort study. BMJ 2010;341:c4978.

4. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet 2004;74:765-9.

5. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25:402-8.