Supplementary Content

Protective effects of breastfeeding on respiratory symptoms in infants with 17q21 asthma risk variants

Olga Gorlanova, Sabina Illi, Antoaneta A. Toncheva, Jakob Usemann, Philipp Latzin, Michael Kabesch, Jean-Charles Dalphin, Roger Lauener, Juha R. Pekkanen, Erika Von Mutius, Josef Riedler, Claudia Kuehni, Martin Röösli, Urs Frey on behalf of the BILD and PASTURE study groups

eMETHODS

The BILD cohort

Data was obtained from the ongoing prospective Basel-Bern Infant Lung Development (BILD) birth cohort and was collected from June 1999 to October 2015 with 1 year follow-up postpartum in Switzerland

Genotyping and per-individual quality-control filtering

Initial data (n=630) included n=620 DNA samples obtained from umbilical cord blood and n=10 DNA samples from buccal swabs for infants for whom no DNA cord blood samples were taken. Genome-wide genotyping was performed using Illumina HumanOmniExpress Bead Chips (Illumina Inc., San Diego, USA), as described.(1)

We selected five major tagging SNPs at the locus 17q21 associated with asthma: rs7216389, rs4795405, rs8079416, rs8065126 and rs3902025 (sTable 1 in the Supplement) based on data derived from the MAGIC/ ISAAC II cohort (n=1446 children from German and Austrian

sTable 1: SNPs characteristics, the BILD cohort (n=368)

SNP	Chr	Gene/ nearest gene		Call rate (%)	Risk allele	Homozygous non-risk genotypes No. (%)	Heterozygous genotype No. (%)	Homozygous risk genotypes No. (%)	*P*-value (HWE)
rs3902025	17	*GSDMA*	utr-variant-5-prime	100	T	71 (19%)	182 (50%)	115 (31%)	>0.99
rs4795405	17	*ORMDL3/ LRRC3C*	Putative promoter/ intron-variant	100	C	70 (19%)	181 (49%)	117 (32%)	>0.99
rs7216389	17	*GSDMB*	intron-variant	100	T	94 (26%)	184 (50%)	90 (25%)	>0.99
rs8065126	17	*LRRC3C*	intron-variant	100	C	56 (15%)	167 (45%)	147 (40%)	0.51
rs8079416	17	*LRRC3C*	intron-variant	100	C	104 (28%)	184 (50%)	80 (22%)	>0.99

Abbreviations: BILD, Basel-Bern Infant Lung Development birth cohort; Chr, chromosome; *GSDMA*, Human *Gasdermin A* gene locus; *LRRC3C*, *Leucine Rich Repeat Containing 3C*; *GSDMB*, Human *Gasdermin B* gene locus; *ORMDL3,* Human *ORMDL sphingolipid biosynthesis regulator 3* locus; HWE, Hardy-Weinberg equilibrium.

background).(2, 3) For the purposes of our study, either the major tagging SNP from the

respective bin was analysed (rs3902025), or a proxy in high LD was (rs7216389 instead of rs11557467; rs4795405 instead of rs8076131; rs8079416 instead of rs3894194; rs8065126 instead of rs8076474). Linkage disequilibrium (LD) between SNPs was illustrated using pair-wise r^2 and shown in the sTable 2.

Using autosomal markers together with genotype information from the HapMap project (III phase 2 release), we conducted the principle component analysis to identify ancestry. We excluded all unexpected duplicates and related individuals based on the identity-by-descent, individuals with sex mismatch, call rate <97% and heterozygosity rate ±3 standard beyond the mean. In total, 368 unrelated term infants of European ancestry with complete weekly information on respiratory symptoms contributed to the primary analysis with 19,252 observed person-weeks. For the secondary analysis, we included 252 unrelated infants with complete information on wheeze with 13,137 person-weeks. All computations were performed using PLINK, version 1.07 and R, version 3.3.1.(4)

sTable 2: Pairwise linkage disequilibrium[a] (r^2), the BILD cohort

	rs8079416	rs4795405	rs3902025	rs8065126
rs7216389	0.61	0.81	0.50	0.48
rs8065126	0.61	0.63	0.82	
rs3902025	0.75	0.67		
rs4795405	0.56			

Abbreviations: BILD, Basel-Bern Infant Lung Development birth cohort.
[a] Pairwise linkage disequilibrium was calculated using R version 3.3.1.

The PASTURE cohort

Study population and phenotype definition

The PASTURE birth cohort study was conducted in rural areas of Austria, Finland, France, Switzerland, and Germany, from January 2002 to October 2006.(5) Pregnant women were recruited in the third trimester of pregnancy. Farming status was determined as living on a farm with livestock, regardless of the amount of time spent on actual farming work. Mothers not living on farms were recruited from the same rural areas, or from towns without industry and with less than 30,000 inhabitants. Half of the women participating in the PASTURE study were exposed to a farm environment. Study protocols were approved by local ethics committees, and informed consent was obtained from all parents. Exclusion criteria were described elsewhere.(5)

Respiratory symptoms were defined as the presence of wheeze or cough in the previous 7 days and coded as a binary variable. Breastfeeding was determined as any breastfeeding in the previous 7 days. Information on respiratory symptoms and breastfeeding was obtained from weekly diaries between 8 and 53 weeks of life. Information on risk factors was collected during the third trimester, at 2 months of age, and at 1, 4, and 6 years of age.(5)

Genotyping and SNPs selection

Genotyping was conducted at the Centre National de Génotypage, Evry, France, as described.(5) In brief, genotyping of SNPs covering the 17q21 region (rs7216389, rs2290400,

and rs8076131) was performed in cord blood using the iPLEX® Gold technology, a MALDI-TOF (Matrix-Assisted-Laser-Desorption/Ionization–Time-Of -Flight-Mass-Spectrometry) system from SEQUENOM.

We selected two variants for the replication analysis: rs7216389 and rs8076131 (the nearest proxy of rs4795405 with r^2=0.92; r^2-value is based on a study by Toncheva et al (6)). The genotype frequencies for these SNPs were similar to those in the BILD cohort (sTable 3 in the Supplement). Both SNPs were in Hardy–Weinberg equilibrium.

sTable 3: SNPs characteristics, , the PASTURE cohort (n=799)

SNP	Chr	Gene/ nearest gene		Call rate (%)	Risk allele	Homozygous non-risk genotypes No. (%)	Heterozygous genotype No. (%)	Homozygous risk genotypes No. (%)	*P*-value (HWE)
rs7216389	17	*GSDMB*	intron-variant	100	T	234 (29.3%)	393 (49.2%)	172 (21.5%)	0.78
rs8076131	17	*ORMDL3*	intron-variant	100	A	201 (25.2%)	399 (49.9%)	199 (24.9%)	>0.99

Abbreviations: PASTURE, Protection against Allergy Study in Rural Environments birth cohort; Chr, chromosome; *GSDMB*, Human *Gasdermin B* gene locus; *ORMDL3*, Human *ORMDL sphingolipid biosynthesis regulator 3* locus; HWE, Hardy-Weinberg equilibrium.

Statistical analysis

Each SNP was analysed using a generalized additive mixed model (GAMM) with binomial distribution under an additive genetic model. We assumed an AR1 (first-order autoregressive) correlation structure. Adjustment covariates were sex, week of age, presence of older siblings, birth weight, mode of delivery, child care (modeled as a weekly time-varying variable: 1- being in child care and 0 – not in child care) , maternal education (defined as low = 9 or less years of schooling, intermediate = 10-12 years of schooling or high = 13 years of schooling or more), maternal atopy (defined as doctor-diagnosed asthma, hay fever or atopic dermatitis), maternal smoking in pregnancy, week of study (includes seasonal variation), farm exposure and study centers. Penalized cubic splines were used to capture the non-linear relationships of respiratory symptoms with week of study (includes seasonal components) and week of age. Gene-environment interaction for each SNP was tested by including into the models the interaction term between breastfeeding and SNP. We also meta-analyzed these SNPs in a combined dataset of discovery and replication cohorts. Since the major outcome - respiratory symptoms - was expressed as an RR and OR in the discovery and replication cohort, respectively, the meta-analysis was performed only for the wheeze outcome. New interaction ORs, their 95%CI and corresponding interaction *P*-value were derived by calculating weighted differences between fixed-effect meta-analysis estimates from the BILD data and estimates from the PASTURE data. We used the I^2 statistic as a measure of heterogeneity between studies; the corresponding *P*-value was labeled as P_{het}.

SNPs with a Bonferroni *P*-value threshold <0.025, and with the same direction of effect in all cohorts, were considered to be replicated. All statistical analyses were conducted with R version 3.3.1 (www.r-project.org).(4)

e**Table 1. Demographic characteristics.**

Characteristic	BILD		PASTURE
	Main sample N=368	Sample with wheeze information[a] N=252	N=799
Anthropometrics			
Boys, No. (%)	182 (49.5%)	127 (50.4%)	397 (49.7%)
Birth weight, mean (SD), kg	3.4 (0.5)	3.4 (0.5)	3.5 (0.5)
Gestational age, mean (SD), week	39.8 (1.1)	39.7 (1.1)	40.0 (1.3)
Respiratory symptoms			
≥1 week, No. (%)	342 (92.9%)	235 (93.3%)	646 (80.9%)
No. of week (total), mean (SD)	5.7 (4.7)	6.1 (4.8)	3.8 (4.0)
Average symptom score, mean (SD)	0.25 (0.2)	0.26 (0.2)	
Wheeze			
≥1 week, No. (%)		69 (27.4%)	282 (35.3%)
No. of week (total), mean (SD)		0.5 (1.0)	1.0 (2.2)
Breastfeeding			
≥1 week, No. (%)	361 (98.1%)	246 (97.6%)	737 (92.2%)
No. of week with breastfeeding, mean (SD)	34.0 (15.1)	34.0 (14.7)	26.5 (18.2)
Risk factors			
Presence of older siblings, No. (%):	191 (51.9%)	127 (50.0%)	521 (65.2%)
Child care ≥1 week, No. (%)	104 (28.3%)	80 (31.8%)	627 (78.5%)
Maternal atopy[b], No. (%)	138 (37.5%)	88 (34.9%)	192 (24.0%)
Maternal smoking during pregnancy, No. (%)	33 (9.0%)	14 (5.6%)	97 (12.1%)
Caesarean section, No. (%)	69 (18.8%)	56 (22.2%)	126 (15.8%)
Educational status, No. (%):			
Low	91 (24.7%)	41 (16.3%)	124 (15.5%)
Middle	114 (31.0%)	82 (32.1%)	350 (43.8%)
High	163 (44.3%)	130 (51.6%)	325 (40.7%)
Farm exposure, No. (%)			390 (48.8%)

Abbreviations: SD, standard deviation; BILD, Basel-Bern Infant Lung Development birth cohort; PASTURE, Protection against Allergy Study in Rural Environments birth cohort.
[a]infants born since 2004.
[b] defined as either atopic asthma, atopic rhinitis, or atopic dermatitis.

90

eTable 2: Association[a] of presence of older siblings and maternal smoking in pregnancy with respiratory symptoms score during the 1st year of life by rs7216389 and rs4795405 genotypes, the BILD cohort

Exposure	Homozygous non-risk genotype	Heterozygous genotype	Homozygous risk genotype	*P*-value interaction[b]
	RR (95% CI)	RR (95% CI)	RR (95% CI)	
Presence of older siblings				
rs7216389_T	1.30 (0.98-1.74)	**1.69 (1.39-2.06)**	**1.70 (1.30-2.23)**	**0.019**
rs4795405_C	1.27 (0.90-1.79)	**1.55 (1.28-1.88)**	**1.77 (1.39-2.25)**	**0.014**
Maternal smoking in pregnancy				
rs7216389_T	1.29 (0.90-1.87)	0.74 (0.49-1.13)	0.84 (0.41-1.73)	0.262
rs4795405_C	1.27 (0.84-1.93)	0.83 (0.57-1.20)	0.69 (0.36-1.33)	0.244

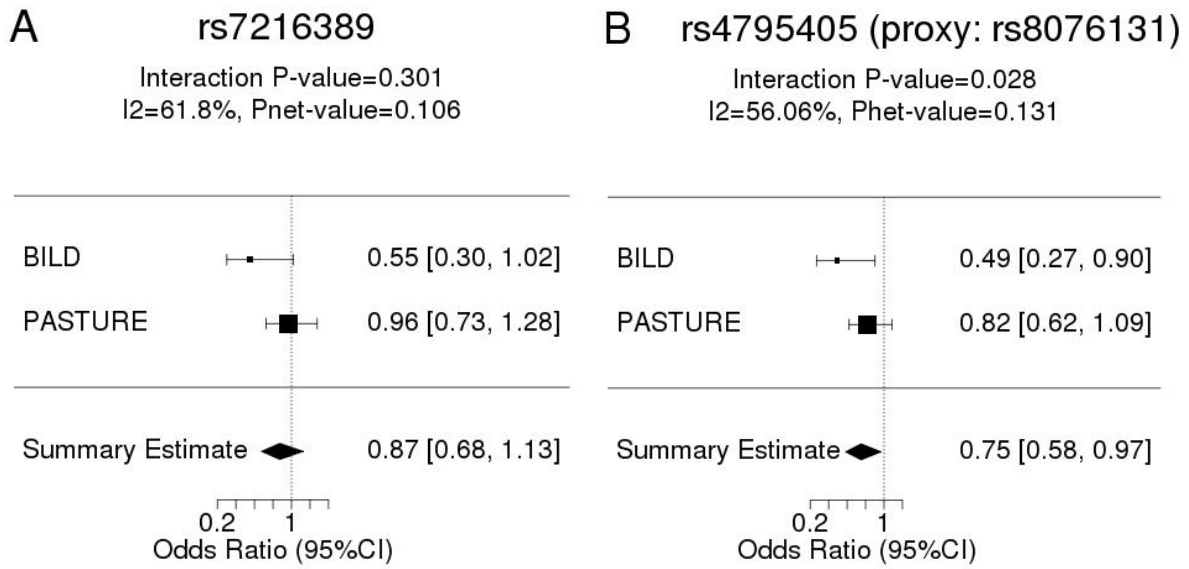
Abbreviations: BILD, Basel-Bern Infant Lung Development birth cohort; RR, risk ratio; CI, confidence interval.
[a] Estimates derived from generalized additive mixed-model with quasi-Poisson distribution; models adjusted week of age (smooth function), sex, birth weight, gestational age, mode of delivery, child care, maternal education, maternal atopy, week of study (smooth function), study centres, breastfeeding and presence of older siblings or maternal smoking in pregnancy .
[b] Interaction was tested by adding the product between the exposures and corresponding SNP in the additive model.
Significant interactions are in boldface.

91

eFigure 1: Forest plots for interactions of top SNPs (additive effect for risk allele) and breastfeeding in relation to wheeze.

(A) Rs7216389, risk allele T **(B)** rs4795405, risk allele C and its proxy rs8076131, risk allele A. For each figure, the size of the squares is proportional to the precision of the estimates for each center, while the horizontal lines indicate their 95% confidence intervals. The dashed vertical lines indicate the overall estimate, whereas the solid ones indicate the null effect.

Abbreviations: 95%CI – 95% confidence interval, BILD – Basel-Bern Infant Lung Development cohort, PASTURE - Protection against Allergy Study in Rural Environments birth cohort, rs – reference SNP cluster code, *P*het – p-value of heterogeneity between studies.



eREFERENCES

1. Fuchs O, Gorlanova O, Latzin P, Schmidt A, Schieck M, Toncheva AA, et al. 6q12 and 11p14 variants are associated with postnatal exhaled nitric oxide levels and respiratory symptoms. *Journal of Allergy and Clinical Immunology* 2017.
2. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007;**448**(7152):470-473.
3. Lluis A, Schedel M, Liu J, Illi S, Depner M, von Mutius E, et al. Asthma-associated polymorphisms in 17q21 influence cord blood ORMDL3 and GSDMA gene expression and IL-17 secretion. *Journal of Allergy and Clinical Immunology* 2011;**127**(6):1587-1594. e1586.
4. Team RC. R: a language and environment for statistical computing. R Core Team, R Foundation for Statistical Computing, Vienna, Austria: ISBN 3-900051-07-0, URL h ttp://www. R-project. org; 2014.
5. Loss GJ, Depner M, Hose AJ, Genuneit J, Karvonen AM, Hyvärinen A, et al. The early development of wheeze. Environmental determinants and genetic susceptibility at 17q21. *American journal of respiratory and critical care medicine* 2016;**193**(8):889-897.
6. Toncheva A, Potaczek D, Schedel M, Gersting S, Michel S, Krajnov N, et al. Childhood asthma is associated with mutations and gene expression differences of ORMDL genes that can interact. *Allergy* 2015;**70**(10):1288-1299.