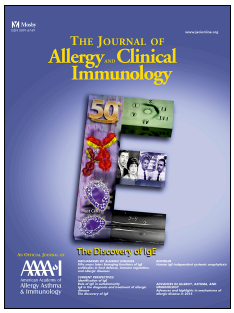
Accepted Manuscript



6q12 and 11p14 variants are associated with postnatal exhaled nitric oxide and respiratory symptoms

Oliver Fuchs, MD PhD, Olga Gorlanova, MD, Philipp Latzin, MD PhD, Anne Schmidt, MD, Maximilian Schieck, PhD, Antoaneta A. Toncheva, PhD, Sven Michel, PhD, Vincent D. Gaertner, BSc, Michael Kabesch, MD, Urs Frey, MD PhD

PII: S0091-6749(17)30046-5

DOI: 10.1016/j.jaci.2016.11.048

Reference: YMAI 12588

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 10 July 2015

Revised Date: 4 November 2016

Accepted Date: 14 November 2016

Please cite this article as: Fuchs O, Gorlanova O, Latzin P, Schmidt A, Schieck M, Toncheva AA, Michel S, Gaertner VD, Kabesch M, Frey U, 6q12 and 11p14 variants are associated with postnatal exhaled nitric oxide and respiratory symptoms, *Journal of Allergy and Clinical Immunology* (2017), doi: 10.1016/j.jaci.2016.11.048.

6q12 and 11p14 variants are associated with postnatal exhaled nitric oxide and respiratory symptoms

Oliver Fuchs, MD PhD[1,2,3*], Olga Gorlanova, MD[1*], Philipp Latzin, MD PhD[1,3], Anne Schmidt, MD[1], Maximilian Schieck, PhD[4,5], Antoaneta A. Toncheva, PhD[4], Sven Michel, PhD[4], Vincent D. Gaertner, BSc[4], Michael Kabesch, MD[4,6**] and Urs Frey, MD PhD[1**]

[1] University Children's Hospital (UKBB), University of Basel, Basel, Switzerland

[2] Dr von Hauner Children's Hospital, Ludwig Maximilians University, Munich, Germany, and Comprehensive Pneumology Center Munich (CPC-M), Germany; Member of the German Center for Lung Research (DZL)

[3] Division of Respiratory Medicine, Department of Pediatrics, Inselspital, University of Bern, Bern, Switzerland

[4] Department of Pediatric Pneumology and Allergy, University Children's Hospital Regensburg (KUNO), Regensburg, Germany

[5] Institute of Human Genetics, Hannover Medical School, Hannover, Germany

[6] Department of Pediatric Pneumology, Allergology and Neonatology, Hannover Medical School, Hannover, Germany, and Biomedical Research in Endstage and Obstructive Lung Disease Hannover (BREATH), Germany; Member of the German Center for Lung Research (DZL)

[*, **] Both authors contributed equally

Corresponding author and address for reprint requests:

Urs Frey, MD PhD

University Children's Hospital (UKBB)

Spitalstrasse 33, 4056 Basel, Switzerland

Phone: +41 61 704 19 00

Fax : +41 61 704 12 53

Email: urs.frey@ukbb.ch

Authors' Contributions

Conception and design: OF, OG, PL, MK, and UF. Acquisition of data, analysis and interpretation: OF, OG, PL, AS, MS, AAT, SM, VDG, MK, and UF. Drafting the manuscript for important intellectual content, final approval of the manuscript: OF, OG, PL, AS, MS, AAT, SM, VDG, MK, and UF. UF is the Prinicipal Investigator of the BILD Cohort.

Funding Sources

This work was supported by the Swiss National Science Foundation (3200-B0-12099, to OF, PL, AS, UF), the European Respiratory Society (Long-Term Research Fellowship 675 to OF) and the Austrian, German and Swiss Pediatric Respiratory Society (Training Scholarships to OF and AS).

Running Head

Genetic variants and exhaled nitric oxide in infants

Key Words

Infants, exhaled nitric oxide, genetics, airway inflammation, genome-wide association study, wheeze, asthma

Article word count: 4,103/3,500 excl. abstract, abbreviations and acknowledgments

Abstract word count: 244/250

Key messages

- Genetic determinants at 6q12 and 11p14 are associated with higher postnatal eNO levels in healthy term-born infants while there were no associations on a genome-wide level with known genetic determinants of eNO in later life.
- While genetic determinants at 6q12 may act in "*trans*" and appear to subsequently be associated with less and less extensive respiratory symptoms during the first year of life in our study population, those at 11p14 seem to impact *NO synthase 3* expression in lung epithelial cells.
- Our findings imply potential differences in the relation between eNO metabolism and lung disease in early infancy in comparison to later childhood.

Capsule Summary

The identification of novel genetic determinants of infant eNO, at previously unknown loci 6q12 and 11p14, may implicate that postnatal eNO metabolism in healthy infants prior to first viral infections and sensitization is related to mechanisms other than those associated with asthma, atopy or increased risk thereof later in life.

DEFINITION OF ABBREVIATIONS

ANO3	Human anoctamin 3 gene locus
BILD	Basel/Bern Infant Lung Development (Cohort)
bp	base pair
BPD	Bronchopulmonary dysplasia
cM	CentiMorgan
CEU	Utah residents with ancestry from northern and western Europe
CTCF	11-zinc finger protein or CCCTC- binding factor
DENND1B	DENN/MADD domain containing 1B
ECR	Evolutionary conserved genomic region
EGF(R)	Epithelial growth factor (receptor)
(e)NO	(Exhaled) nitric oxide
ETS	Environmental tobacco smoke
EYS	Human eyes shut homolog gene locus
GxE	Gene-environment interaction
GxG	Gene-gene interaction
GSDMB	Human gasdermin B gene locus
GWAS	Genome-wide association study
HWE	Hardy-Weinberg equilibrium
IBD	Identity-by-descent
IRR	Incidence risk ratio
LD	Linkage disequilibrium
LUZP2	Human leucine zipper protein 2 gene locus
LYRM9	Human LYR motif containing 9 gene locus

96	MAF	Minor allele frequency
97	Mb	Megabase
98	*MUC15*	Human mucin 15 gene locus
99	*NOS*	Human nitric oxide synthase gene locus
100	OLR	Online repository
101	OR	Odds ratio
102	PC(A)	Principal component (analysis)
103	p_{int}	Interaction p-value
104	QC	Quality control
105	QQ	Quantile-quantile plot
106	QTL	Quantitative trait locus
107	rs	Reference SNP cluster code (dbSNP)
108	SD	Standard deviation
109	*SLC5A12*	Human solute carrier family 5
110		(sodium/monocarboxylate cotransporter) member 12 gene
111		locus
112	SNP	Single-nucleotide polymorphism
113	TFBS	Transcription factors binding sites
114	UCB	Umbilical cord blood
115	VEGF	Vascular endothelial growth factor
116	V’E	Minute ventilation
117	V’NO	Nitric oxide output
118	V_T	Tidal volume
119	95% CI	95% confidence interval

ABSTRACT

Background: Exhaled nitric oxide (eNO) is a biomarker of airway inflammation and seems to precede respiratory symptoms, such as asthma in childhood. Identifying genetic determinants of postnatal eNO may aid in unraveling the role of eNO in epithelial function or airway inflammation and disease.

Objective: To identify genetic determinants of early postnatal eNO and subsequent respiratory symptoms during the first year of life.

Methods: Within a population-based birth cohort, eNO was measured in healthy term infants aged 5 weeks during quiet tidal breathing in unsedated sleep. We assessed associations of single-nucleotide polymorphisms with eNO in a genome-wide association study, and subsequent symptoms of lower respiratory tract infections during the first year of life; and asked if this was modified by prenatal and early-life environmental factors.

Results: We identified so far unknown determinants of infant eNO: rs208515 ($p=3.3 \times 10^{-8}$) located at 6q12, probably acting in "*trans*", and explaining 10.3% of eNO variance; and furthermore rs1441519 ($p=1.6 \times 10^{-6}$) at 11p14, potentially impacting *NO synthase 3 (NOS3)* expression as shown by *in-vitro* functional analyses. Moreover, the 6q12 locus was inversely associated with subsequent respiratory symptoms ($p<0.05$) and time to recovery after first respiratory symptoms during the first year of life ($p<0.05$).

Conclusion: The identification of novel genetic determinants of infant eNO may implicate that postnatal eNO metabolism in healthy infants prior to first viral infections and sensitization is related to mechanisms other than those associated with asthma, atopy or increased risk thereof later in life.

INTRODUCTION

Exhaled nitric oxide (eNO) is an important biomarker of airway inflammation. Low levels of outdoor air pollution and environmental tobacco smoke (ETS) affect eNO levels in newborns.(1, 2) Moreover, eNO seems to also be an early marker for future respiratory morbidity in infants(3) and to precede transient early but not persistent wheeze in infants.(4) It may thus help to differentiate early forms of infant wheeze, at least in individuals at risk. Later, eNO serves as an established indicator for allergic airway inflammation, aiding in discriminating childhood wheeze and asthma phenotypes(5-7) and in identifying exacerbation risk or asthma control.(8, 9)

Interestingly, respiratory epithelial-derived NO is also important for cell signaling. There is an emerging role of NO in airway development and pulmonary angiogenesis both in rodent models(10, 11) and for premature children at risk for bronchopulmonary dysplasia (BPD).(12, 13) In the context of lung epithelial-mesenchymal crosstalk,(14) NO has been shown to modulate recovery from lung injury and to attenuate transition of epithelial cells to myofibroblasts.(15) In rodents, this has important implications for epithelial function, structural maintenance, and lung remodelling in the postnatal phase.(16) In humans however, the role of postnatal eNO as a marker of epithelial function or for future respiratory morbidity irrespective of underlying risk remains unclear. Thus, before it may be used as a marker for "FeNOtyping",(6) it is important to better understand putative genetic determinants of postnatal eNO.

Several lung cell types are able to produce NO from L-arginine through the action of nitric oxide synthase isoforms (NOSs).(17, 18) Previous work in infants at risk of atopy(19) and following a candidate-gene approach in older children(20-23) has suggested that variants in genes such as *DENN/MADD domain containing 1B* (*DENND1B*) and *NOS2* may affect eNO levels. Only recently, a large genome-wide association study (GWAS) in individuals not at increased risk for atopy or asthma identified the loci *LYR motif containing 9* (*LYRM9*), *NOS2*,

and one near *gasdermin B* (*GSDMB*) at 17q21 to be associated with eNO.(24) All these latter analyses have been performed in older children where genetically determined eNO levels may also be affected by a number of other factors, e.g. concomitant disease and environmental effects. Here, we present a GWAS in a unique, unselected population-based cohort of term-born infants not at increased risk of atopy or asthma, the Basel/Bern Infant Lung Development (BILD) cohort.(25) In the cohort, eNO levels measured shortly after birth, and prior to the impact of environmental factors that may play a role in the pathophysiology of atopy or childhood asthma, were shown to be unrelated to asthma diagnosis at school age.(26) We hypothesized that infant eNO levels could rather be determined by early genetic effects which may precede possible disease development from early in life.

METHODS

A more detailed description of the study population and methods is provided in the online repository (OLR).

Study population, study participants

All data were collected in the ongoing prospective BILD birth cohort of unselected, healthy neonates.(25) The study was approved by all involved ethics committees. All caregivers provided written informed consent for this study.

Measurements and quality control of tidal breathing and nitric oxide

Tidal breathing and mixed eNO were measured at five weeks of age as described previously.(25, 27) Outcome parameters were tidal volume (V_T), minute ventilation ($V'E=V_T$ x respiratory rate), eNO, and NO output (V'NO=eNO concentration x expiratory flow).(27)

Measurement of respiratory symptoms and time to recovery from first respiratory infection during the first year of life

Respiratory symptom data was collected by weekly telephone interviews.(25) Respiratory symptoms were defined as 'cough, wheezing, difficulty breathing and reduced activity during night and day'. Weeks with symptoms were defined as total number of weeks with respiratory symptoms. Time to recovery was defined as total number of weeks that the infant suffered from first respiratory symptoms during the first 6 and 12 months of life.

Genotyping, genetic data quality control

Whole-genome genotyping (Illumina HumanOmniExpress Bead Chips, Illumina Inc., San Diego, USA) was performed in two separate batches in all n=425 study participants (in 2011 for n=329 and in 2013 for additional n=96 children).We checked for a batch effect due to two genotyping time-points. See Figure 1 for details on post-hoc exclusion of individuals and

quality control (QC) of eNO measurements and genotyping data. Principal component analysis (PCA) was performed in order to obtain eigenvectors as covariates for later association tests. The first principal component (PC1) was subsequently included for adjustment.

Statistical analyses and association testing

Analyses were restricted to additive models. No transformation was necessary for eNO or V'NO values. Associations between SNPs and eNO or V'NO were analyzed by adjusted linear regression. The Bonferroni method was chosen to control for multiple testing.

Associations between SNPs and number of weeks with respiratory symptoms and time to recovery were analyzed by adjusted Poisson and logistic regression. For the latter, a binary variable (resolution of symptoms < 2 week or ≥ 2 weeks) was defined. Results are expressed as incidence risk ratio (IRR) and odds ratio (OR), respectively. Gene-gene (GxG) and gene-environment interaction (GxE) were assessed by ANOVA. The corresponding p-value was labeled p_{int}. A p-value <0.05 was considered significant.

Association computations were done in R version 3.0.2 (www.r-project.org) (28) using the GenABEL package (29) and Stata 12.1 (STATA Corporation, College Station, TX, USA) in both the final subset of all n=229 individuals that were left for subsequent analyses as well as for each genotyping batch from either 2011 or 2013, separately. Replicating results for SNPs with significant associations with eNO in children,(21, 24) meta-analyses were performed using the *Metan* package in Stata 11.2 (STATA Corporation, College Station, TX, USA). Regional plots were generated by LocusZoom (http://csg.sph.umich.edu/locuszoom). Permutation analyses were performed by chromosome using PLINK version 1.07.(30) Linkage disequilibrium (LD) was evaluated by PLINK version 1.07(30) based on 1000 Genomes Project (CEU population) data.(31) Regional LD plots were generated in R 3.0.0. Base pair (bp) positions correspond to genome assembly GRCh37.p13 and gene annotations

are according to the NCBI RefSeqGene Project (accessed March 24th 2015).

In-silico functional analyses

Expression (eQTL) and methylation QTLs (mQTL) were analyzed. Phylogenetic analyses were performed by ECR Browser (http://ecrbrowser.dcode.org/). CCCTC- binding factor (CTCF) binding sites and chromatin topological domains were determined using data from CTCFBSDB 2.0 - (http://insulatordb.uthsc.edu/).(32)

Cell culture and gene expression analyses

Expression of genes in the vicinity of associated SNPs on 6q12 (*EYS*) and 11p14 (*ANO3*, *LUZP2*, *MUC15* and *SLC5A12*) was evaluated by qRT-PCR in various cell types. Lung epithelial cells (H358) were cultured and *LUZP2*, *MUC15* and *SLC5A12* were knocked down with two siRNAs per target gene after transfection with lipofectamine RNAiMAX (Invitrogen by ThermoFisher Scientific, Germany). Knockdown efficiency and *NOS3* mRNA levels after 24, 48 and 72 hours were assessed by qRT-PCR. For further details see OLR and OLR Tables S1-S8.

RESULTS

Between 1999 and 2013, the study enrolled 425 eligible infants, of whom n=353 presented for lung function measurements at the age of 5 weeks after gestation. After secondary exclusion, technically acceptable data were obtained from n=344 infants. As indicated in Figure 1, n=229 individuals with data for n=567,864 SNPs were then subjected to analyses. Table 1 presents anthropometric and demographic characteristics and distribution of known and possible covariates among all study participants for whom we had complete datasets on eNO measurements and from genotyping after aforementioned QC. Values for eNO and V'NO were highly correlated ($r=0.71$, $p<0.0001$). We did not find any differences in baseline characteristics between infants for whom data were in- or generally or specifically excluded from subsequent analyses (data not shown).

Calculated genomic inflation factors for eNO and V'NO were below 1.006, showing no evidence for population stratification. While we could not replicate associations of SNPs in the *DENND1B* locus with eNO values in infants at risk of developing allergy,(19) this was the case for SNPs in the *NOS2* locus(21) as well as of variants in *LYRM9* and *GSDMB* loci(24) for their association with eNO values in older children with the same effect direction as previously published(21, 24) however not reaching genome-wide significance (see OLR Table E1 and OLR Figure E1, A-D for forest-plots of respective meta-analyses). However, after adjustment for confounders as well as covariates and correction for multiple testing, we identified different polymorphisms significantly associated with eNO and V'NO. Two SNPs on 6q12 (reference SNP (rs) cluster codes rs208515 and rs208520), were in high LD ($r^2>0.8$) and significantly associated with eNO and V'NO, respectively. The T-allele (risk allele) of rs208515 and the G-allele (risk allele) of rs208520 were associated with increased eNO values (explained variance 10,3% and 10,2%, respectively) as well as increased V'NO values (explained variance 9,9% and 9,9%, respectively). Two variants on 11p14 (rs12223678 and rs1441519) were also in high LD ($r^2>0.8$) and were indicative for associations with eNO and

V'NO. Here, the T-alleles (risk alleles) of rs1441519 and rs12223678 were associated with lower eNO values (explained variance 7,9% and 7,3%, respectively) as well as with lower V'NO values (explained variance 9,0% and 8,7%, respectively). Detailed association data for the total final subset of n=229 individuals are provided in OLR Tables E2A and B, and summary statistics of associated SNPs are listed in OLR Table E3. Figures E2A and B in the OLR display quantile-quantile (QQ) plots and Figures 2A and B Manhattan plots for eNO and V'NO. Figures E3A and B in the OLR as well as OLR Figures E4A and B display regional plots for eNO and V'NO, respectively. The above mentioned top SNPs were also robust with regard to effect direction and of comparable effect size albeit not significant for the smaller batch from 2013 when analyses were performed for each genotyping batch from 2011 or 2013, separately (see OLR Tables E4A and B for associations with eNO). Moreover, they remained also significantly associated with measured outcomes when we performed additional permutation analyses. For the 6q12 top SNPs this was also the case after accounting for multiple testing within the permutation analysis, while for 11p14 hits a trend was observed (see OLR Table E5 for eNO).

In a previous study, no association was found between postnatal eNO levels and respiratory symptoms during the first year of life for the entire study population but for the subgroup of mothers who smoked during pregnancy or who were atopic.(3) We therefore investigated whether the identified genetic determinants were also relevant for n=209/229 children with respiratory symptoms during the first year of life. Table E6A in the OLR displays details on duration of respiratory symptoms in weeks for the study population and across group levels. The T-allele of rs208515 (6q12) was negatively associated with the number of weeks with respiratory symptoms with an IRR of 0.80 (95%CI 0.68-0.95, p=0.009, see Figure 3A). This was independent of symptom severity (data not shown) and remained significant also after adjusting for covariates (sex, having older siblings, nursery care, maternal atopy, maternal smoking during pregnancy, V'E, and eNO) with an adjusted IRR (aIRR) of 0.81 (95%CI

0.69-0.95, p=0.011, see OLR Table E6B). The same association with comparable effect size and direction was found for rs208520. In contrast, we did not find any significant associations between 11p14 variants and the number of weeks with respiratory symptoms (see OLR Table E6B for adjusted and Figure 3B for unadjusted results). Stratifying for maternal atopy or maternal smoking during pregnancy neither modified the association between SNPs and eNO or V'NO, nor between SNPs and subsequent symptoms for either locus. Furthermore, association of SNPs with assessed outcomes was not dependent of each other, i.e. there was no GxG for variants at 6q12 or 11p14 identified to be associated with eNO and V'NO in infants (data not shown).

Lastly, we analyzed wether the effect of identified genetic determinants was also relevant for the time to recovery from first respiratory infections. There was no effect of genetic determinants on the duration of first respiratory infection in the study population with symptoms across the first 12 months of life (n=207) in either univariable or adjusted models (data not shown). Constricting our analyses to children with symptoms during the first 6 months of life (n=130) however, rs208515 or rs208520 were significantly associated with faster recovery (i.e. less than two weeks) in both univariable models (data not shown) and models adjusted for known and possible confounders and covariates (sex, age, having older siblings, nursery care, maternal atopy, maternal smoking during pregnancy, season, being breast-fed at time of first respiratory symptoms, V'E, and eNO) with an aOR=0.34, p=0.008 and an aOR of 0.24, p=0.002, respectively.

As association signals on chromosome 6q12 (rs208515 and rs208520) and chromosome 11p14 (rs12223678 and rs1441519) are located in non-coding regions, we applied *in-silico* and database analyses to search for putative causal SNPs in the respective regions driving the observed signals. The closest RefSeq genes identified on 6q12 were *SLC25A51P1*, *EYS* and *LOC441155* located upstream of rs208515 (446kb, 529kb and 930kb distance respectively), whereas no gene was listed within 1,000kb downstream of rs208515 (see Figure 4A). On

11p14, *LUZP2* was the closest gene upstream of rs12223678 (812kb distance). The closest downstream genes of rs1441519 were *ANO3*, *MUC15* and *SLC5A12* (424kb, 651kb and 759kb distance, respectively, see Figure 4B). Linkage disequilibrium analyses in at least 1,000kb surrounding the associated polymorphisms on 6q12 identified an LD $r^2 \geq 0.8$ for 177 SNPs for rs208515 and 88 SNPs for rs208520. Due to r^2 of 0.82 between rs208515 and rs208520, an overlap in tagged SNPs exists, resulting in 201 unique SNPs in the 6q12 tagging block (OLR Tables E7A and B). None of the tagged SNPs is located in close vicinity to known genes in the 6q12 locus (closest SNP: rs1563929 located 203kb from *SLC25A51P1*). Analysis for 11p14 identified an LD of r^2 0.96 between rs12223678 and rs1441519. Both SNPs are tagging for the identical set of SNPs at $r^2 \geq 0.8$. (n=13 SNPs; OLR Tables E7C and D). None of these 13 SNPs is in close vicinity to genes on the 11p14 locus (closest SNPs: rs11028996 with 802kb distance to upstream *LUZP2* and rs1348169 with 424kb distance to downstream *ANO3*).

All identified SNPs (n=214) were subjected to further evaluation, whether they locate within regulatory sequences such as enhancers/ super-enhancers and insulators, transcription factors binding sites (TFBS) or evolutionary conserved regions (ECRs) determined in the respective regions as depicted in Figure 4A and B. Indeed, we identified 5 associated SNPs on chromosome 6q12 in binding sites for the CCCTC-binding factor (CTCF), which is specifically allocated to insulator regions in humans and was shown to play an important developmental role (OLR Tables E8A and B).(33-35) Three variants (rs9453663, rs9453664, rs12200490) belong to a sequence that allows cell type-specific binding of CTCF in lung fibroblasts. Polymorphisms rs3899423 and rs4618506 locate in regions where CTCF binds in embryonic stem cells and lymphoblastoid cells, respectively. *In-silico* predictions suggested that all 13 SNPs at 11p14 are located in an area that is part of two cell type-specific (human embryonic stem cells and fetal lung myofibroblasts, IMR-90) chromatin structures called topologically associating domains.

When we investigated whether genes located in the vicinity of the association signals from 6q12 (*EYS*) and 11p14 (*SLC5A12*, *MUC15* and *LUZP2)* were expressed in specific cell types potentially relevant for NO production (OLR Table E9), we found that genes from 11p14, but not from 6q12, were expressed in lung epithelial cells. We hypothesized therefore that 6q12 may have a broader regulatory impact, most likely on *NOS* genes *in trans* while 11p14 genes may influence NO production more directly in lung epithelial cells. Consequently, we performed gene knockdown experiments on the three genes near the 11p14 locus in a lung epithelial cell line (H358) and measured *NOS3* expression in these cells after knockdown (see Figure 5 A and B). Indeed, after an initial rise in *NOS3* expression possibly due to cell transfection stress, gene expression of *NOS3* decreased after knockdown for all three genes, most prominently after *SLC5A12* and *MUC15* knockdown. This suggests that mutations in regulatory elements in the 11p14 region may have an impact on these two genes, in turn affecting NO production via *NOS3*, which is the only NO synthase present in lung epithelial cells, as indicated in OLR Table E9.

DISCUSSION

In this unselected birth cohort,(25) we found two new loci, 6q12 and 11p14, to be associated with NO levels, measured in healthy term-born infants shortly after birth. The 6q12 locus explains a significant proportion of variance at a reasonably high significance level. Furthermore, this locus is also inversely related to the number of weeks with respiratory symptoms and time to recovery after first verified respiratory symptoms during the first year of life, and thus may be of functional relevance.

Our long-distance LD analyses revealed that no polymorphisms in LD with eNO-associated SNPs were located in close vicinity to thus-far identified genes in the respective regions. Therefore, a direct impact of eNO-associated polymorphisms on gene function (e.g. amino acid sequence changes, splicing, promoter activity assays) in either of the two chromosomal loci could not be established.

The most significantly associated SNPs (rs208515 and rs208520) from 6q12 reside in a region that most likely has "*trans* effects" on mechanisms not necessarily located in the vicinity of the original association locus. Indeed, no expression of genes closest to the signal was observed in lung epithelial cells in our initial functional evaluation, and further investigations in this locus will be necessary. Unfortunately, existing eQTL datasets are of no use to study gene expression effects in early childhood, making further targeted experiments necessary. Interestingly, the extensive intergenic regions, where identified SNPs are located, are predicted to harbor regulatory regions. Five of the variants from associated tagging bins at 6q12 are placed in insulator sequences, defined as binding sites for the major human insulator protein CTCF. ChipSeq database derived data revealed that CTCF binds in a cell type-specific manner, and more importantly, in cell types involved in embryotic development and potentially relevant to the course of airway inflammation (lung fibroblasts and lymphoblastoid cells).(33-35)

The second identified locus at 11p14 (rs12223678 and rs1441519) is situated in vicinity of the human *ANO3/MUC15* gene locus. While *ANO3* belongs to a gene family encoding calcium-activated chloride channels,(36) the MUC15 protein belongs to the mucin family, representing glycoproteins, which constitute a part of the physical barrier of the airway epithelium.(36) Remarkably, it also shares structural motifs with the human epithelial growth factor (EGF) receptor (EGFR) and has been shown to modulate EGFR-signaling after EGF stimulation.(37) EGF and homologs regulate several signal cascades(38, 39) that participate in functions such as proliferation, differentiation, mitosis, cell survival, and apoptosis on the transcription level.(38) Additional downstream events encompass glucose metabolism and cell migration.(39)

SLC5A12 was identified as a further candidate gene potentially contributing to the observed association with eNO via *NOS3* expression. This gene is a member of the solute-linked carrier gene family 5 (*SLC5*) that contains the low-affinity Na(+)-coupled lactate transporter.(40) It has been shown that the activity of the SGLT1 (also known as SLC5A1) protein, a member of the *SLC5* gene family as well, is regulated by NOS3.(41) However, how ow exactly these genes on 11p14 are involved in eNO related mechanisms remains unknown and would need to be further investigated. Prediction analysis for the 11p14 variants indicates cell type-specific CTCF-driven formation of topologically associated domains in embryonic stem cells and in fetal lung fibroblasts, potentially orchestrating 11p14 chromatin rearrangements and affecting genes involved in lung tissue development.

Chromatin insulators regulate the effects of enhancers and silencers on a gene promoter(42, 43) and thus might induce or suppress transcription activity. Moreover, insulator activity is mainly driven by DNA conformation and the global chromatin architecture mediated by CTCF,(33) which has been identified as a major regulator in the genome.(33, 34) CTCF establishes boundaries between topologically associating domains in a cell type-specific and developmental fashion.(34, 44)

Importantly, we measured mixed eNO in our study; eNO comprised of nasal and bronchial origin. There is no data supporting that there is a difference between strictly oral or nasal NO (nNO) in term infants.(27, 45) However, such differences could be demonstrated for premature infants very early in life and before any possible significant bacterial colonization and pneumatisation of paranasal sinuses, with much higher levels for nNO, predominantly under iNOS influence,(17) than for bronchial eNO.(46, 47) The link to the already mentioned role of NO metabolism in cell signaling, VEGF-mediated airway development,(10-13) lung epithelial-mesenchymal crosstalk,(14) and epithelial function is striking. These are functional implications and have to be considered with caution. Moreover, further genes adjacent to the most significant SNPs at 6q12 and 11p14 may importantly also play their roles.

Remarkably, we found neither an association with known genetic determinants of eNO levels in children at risk of developing atopy (data not shown)(19) nor in older children(see results and OLR) (22, 24) or adults (data not shown)(20, 23, 48) on a genome-wide level. However, such replications of previously published associations in our cohort on a genome-wide level were also unexpected, as eNO at a later age more typically reflects the impact of factors that may play a role in the pathophysiology of atopy or childhood asthma. We were also unable to replicate our findings in another population, i.e. in study participants of a follow-up of the GABRIEL study, in whom eNO was measured, albeit at an older mean age of 9.9 years (data not shown).(49)

Interestingly, the *ANO3/MUC15* locus at 11p14 has recently been related to atopic eczema in adults ascertained through asthma(36) but the genetic signal in that study was independent from the signal found in our study. Earlier in this study, the role of infant eNO as a marker for later respiratory symptoms was evident for the subgroup of smoking or atopic mothers.(3) This has been corroborated by others.(4) We investigated the potential discrepancy to current findings and found several differences. The population included by *Latzin et al.* was smaller and had higher exposure to ETS (14.0% in (3) vs. 8.7%) or air pollution.(3) Furthermore,

siblings were strictly excluded in the current study. We reproduced the former findings for the smaller population used by *Latzin et al.*(3) We speculate, that discrepancies could be due to sample size differences or differences in environmental exposure.

This study is, to the best of our knowledge, the first effort so far to uncover genetic determinants of eNO levels, which were collected in healthy term infants shortly after birth,(25) according to latest standards and already published as part of a set of normative data, including extensive QC during data collection, and subsequent analyses.(27) Despite being the first effort so far, we are limited by the number of healthy subjects very early in life included to generate sufficient statistical power to detect all variants with at least modest effects and especially with regard to possible replication in independent populations. To our knowledge no such cohort exists, and a decade was required to build up the present study. Still, the analyses within the two different batches may nevertheless serve as discovery and replication samples for our identified determinants of eNO levels in healthy infants shortly after birth. Analysing both batches seperately, they display the same effects, with same effect direction and comparable effect sizes while showing different levels of significance given their unequal numbers of individuals. Moreover, with this GWAS we were able to reach genome-wide significance levels with a significant effect size even with the most conservative method to adjust for multiple testing and with significant results for the identified top SNPs also after permutation. Still, larger longitudinal studies are needed in order to characterize functional implications, as well as further loci with only modest effects that remain to be identified for eNO values in later life.

Conclusion

Limited by small sample size and the lack of a formal replication in an independent population in general, we nevertheless identified novel genetic determinants of most accurately measured postnatal eNO in healthy infants not at increased risk for atopy or

asthma. One identified locus was 6q12 which most likely has "*trans* effects" on mechanisms not necessarily located in the vicinity of the originally associated locus, but for which rs208515 explained up to 10.3% of outcome variance with a high level of significance. In addition, suggestive association in the GWAS was detected for SNPs at chromosome 11p14, for which functional analyses revealed effects on *NOS3* expression. Given the association with respiratory symptoms, these results may provide new insight into possible biological mechanisms underlying eNO levels, their regulation, and their functional relevance within the spectrum of respiratory diseases in young children. The identification of these variants may implicate that postnatal eNO metabolism in healthy infants, not at increased risk for atopy, is rather related to epithelial function prior to first viral infections and postnatal sensitization, than to mechanisms associated with asthma, atopy or increased risk thereof later in life.

ACKNOWLEDGEMENTS

We thank Dr. Tobias Balschun and Tanja Henke from the Department of Clinical Molecular Biology, University Hospital Schleswig-Holstein, Campus Kiel, Germany, for their cooperation and performing the genotyping. We are indebted to our study participants and their families for their participation. We thank Karine Landgren Hugentobler for proofreading the manuscript. Furthermore, we thank the study nurses Monika Graf, Barbara Hofer, Sandra Luescher, Christine Becher, Maja Weber, Anna Padiyath, and the lung function technicians Gisela Wirz and Sharon Schmid for their invaluable assistance and support.

This work was supported by the Swiss National Science Foundation (3200-B0-12099, to OF, PL, AS, UF), the European Respiratory Society (Long-Term Research Fellowship 675 to OF) and the Austrian, German and Swiss Pediatric Respiratory Society (Training Scholarships to OF and AS).

REFERENCES

1. Frey U, Kuehni C, Roiha H, Cernelc M, Reinmann B, Wildhaber JH, et al. Maternal atopic disease modifies effects of prenatal risk factors on exhaled nitric oxide in infants. Am J Respir Crit Care Med. 2004;170(3):260-5.
2. Gabriele C, Asgarali R, Jaddoe VW, Hofman A, Moll HA, de Jongste JC. Smoke exposure, airway symptoms and exhaled nitric oxide in infants: the Generation R study. Eur Respir J. 2008;32(2):307-13.
3. Latzin P, Kuehni CE, Baldwin DN, Roiha HL, Casaulta C, Frey U. Elevated exhaled nitric oxide in newborns of atopic mothers precedes respiratory symptoms. Am J Respir Crit Care Med. 2006;174(12):1292-8.
4. Chawes BL, Buchvald F, Bischoff AL, Loland L, Hermansen M, Halkjaer LB, et al. Elevated exhaled nitric oxide in high-risk neonates precedes transient early but not persistent wheeze. Am J Respir Crit Care Med. 2010;182(2):138-42.
5. Pijnenburg MW, De Jongste JC. Exhaled nitric oxide in childhood asthma: a review. Clin Exp Allergy. 2008;38(2):246-59.
6. de Jongste JC. To wheeze or not to wheeze: prospective FENO-typing in early infancy. Am J Respir Crit Care Med. 2006;174(12):1281-2.
7. Singer F, Luchsinger I, Inci D, Knauer N, Latzin P, Wildhaber JH, et al. Exhaled nitric oxide in symptomatic children at preschool age predicts later asthma. Allergy. 2013;68(4):531-8.
8. Pijnenburg MW, Hofhuis W, Hop WC, De Jongste JC. Exhaled nitric oxide predicts asthma relapse in children with clinical asthma remission. Thorax. 2005;60(3):215-8.
9. van der Valk RJ, Baraldi E, Stern G, Frey U, de Jongste JC. Daily exhaled nitric oxide measurements and asthma exacerbations in children. Allergy. 2012;67(2):265-71.
10. Duong-Quy S, Hua-Huy T, Pham H, Tang X, Mercier JC, Baud O, et al. Early inhaled nitric oxide at high dose enhances rat lung development after birth. Nitric oxide : biology and chemistry / official journal of the Nitric Oxide Society. 2014;38:8-16.
11. Young SL, Evans K, Eu JP. Nitric oxide modulates branching morphogenesis in fetal rat lung explants. American journal of physiology Lung cellular and molecular physiology. 2002;282(3):L379-85.
12. Cole FS, Alleyne C, Barks JD, Boyle RJ, Carroll JL, Dokken D, et al. NIH Consensus Development Conference statement: inhaled nitric-oxide therapy for premature infants. Pediatrics. 2011;127(2):363-9.
13. Mercier JC, Olivier P, Loron G, Fontaine R, Maury L, Baud O. Inhaled nitric oxide to prevent bronchopulmonary dysplasia in preterm neonates. Seminars in fetal & neonatal medicine. 2009;14(1):28-34.
14. Knight DA, Holgate ST. The airway epithelium: structural and functional properties in health and disease. Respirology. 2003;8(4):432-46.
15. Vyas-Read S, Shaul PW, Yuhanna IS, Willis BC. Nitric oxide attenuates epithelial-mesenchymal transition in alveolar epithelial cells. American journal of physiology Lung cellular and molecular physiology. 2007;293(1):L212-21.
16. Kanazawa H. Role of vascular endothelial growth factor in the pathogenesis of chronic obstructive pulmonary disease. Medical science monitor : international medical journal of experimental and clinical research. 2007;13(11):RA189-95.

17. Michel T, Feron O. Nitric oxide synthases: which, where, how, and why? The Journal of clinical investigation. 1997;100(9):2146-52.

18. Guo FH, Comhair SA, Zheng S, Dweik RA, Eissa NT, Thomassen MJ, et al. Molecular mechanisms of increased nitric oxide (NO) in asthma: evidence for transcriptional and post-translational regulation of NO synthesis. J Immunol. 2000;164(11):5970-80.

19. Chawes BL, Bischoff AL, Kreiner-Moller E, Buchvald F, Hakonarson H, Bisgaard H. DENND1B gene variants associate with elevated exhaled nitric oxide in healthy high-risk neonates. Pediatr Pulmonol. 2015;50(2):109-17.

20. Dahgam S, Nyberg F, Modig L, Naluai AT, Olin AC. Single nucleotide polymorphisms in the NOS2 and NOS3 genes are associated with exhaled nitric oxide. Journal of medical genetics. 2012;49(3):200-5.

21. Salam MT, Bastain TM, Rappaport EB, Islam T, Berhane K, Gauderman WJ, et al. Genetic variations in nitric oxide synthase and arginase influence exhaled nitric oxide levels in children. Allergy. 2011;66(3):412-9.

22. Salam MT, Byun HM, Lurmann F, Breton CV, Wang X, Eckel SP, et al. Genetic and epigenetic variations in inducible nitric oxide synthase promoter, particulate pollution, and exhaled nitric oxide levels in children. J Allergy Clin Immunol. 2011.

23. Bouzigon E, Monier F, Boussaha M, Le Moual N, Huyvaert H, Matran R, et al. Associations between nitric oxide synthase genes and exhaled NO-related phenotypes according to asthma status. PloS one. 2012;7(5):e36672.

24. van der Valk RJ, Duijts L, Timpson NJ, Salam MT, Standl M, Curtin JA, et al. Fraction of exhaled nitric oxide values in childhood are associated with 17q11.2-q12 and 17q12-q21 variants. J Allergy Clin Immunol. 2014;134(1):46-55.

25. Fuchs O, Latzin P, Kuehni CE, Frey U. Cohort profile: the Bern Infant Lung Development Cohort. International journal of epidemiology. 2012;41(2):366-76.

26. Usemann J, Fuchs O, Anagnostopoulou P, Korten I, Gorlanova O, Röösli M, et al. Predictive value of exhaled nitric oxide in healthy infants for asthma at school age. Eur Respir J. 2016.

27. Fuchs O, Latzin P, Thamrin C, Stern G, Frischknecht P, Singer F, et al. Normative data for lung function and exhaled nitric oxide in unsedated healthy infants. Eur Respir J. 2011;37(5):1208-16.

28. R Development Core Team. R Foundation for Statistical Computing. Vienna, Austria2006.

29. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007;23(10):1294-6.

30. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics. 2007;81(3):559-75.

31. Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491(7422):56-65.

32. Ziebarth JD, Bhattacharya A, Cui Y. CTCFBSDB 2.0: a database for CTCF-binding sites and genome organization. Nucleic acids research. 2013;41(Database issue):D188-94.

33. Phillips JE, Corces VG. CTCF: master weaver of the genome. Cell. 2009;137(7):1194-211.

34. Ong CT, Corces VG. CTCF: an architectural protein bridging genome topology and function. Nature reviews Genetics. 2014;15(4):234-46.

35. Herold M, Bartkuhn M, Renkawitz R. CTCF: insights into insulator function during development. Development. 2012;139(6):1045-57.

36. Dizier MH, Margaritte-Jeannin P, Madore AM, Esparza-Gordillo J, Moffatt M, Corda E, et al. The ANO3/MUC15 locus is associated with eczema in families ascertained through asthma. J Allergy Clin Immunol. 2012;129(6):1547-53 e3.

37. Pallesen LT, Pedersen LR, Petersen TE, Knudsen CR, Rasmussen JT. Characterization of human mucin (MUC15) and identification of ovine and caprine orthologs. Journal of dairy science. 2008;91(12):4477-83.

38. Bodnar RJ. Epidermal Growth Factor and Epidermal Growth Factor Receptor: The Yin and Yang in the Treatment of Cutaneous Wounds and Cancer. Advances in wound care. 2013;2(1):24-9.

39. Roskoski R, Jr. The ErbB/HER family of protein-tyrosine kinases and cancer. Pharmacological research : the official journal of the Italian Pharmacological Society. 2014;79:34-74.

40. Srinivas SR, Gopal E, Zhuang L, Itagaki S, Martin PM, Fei YJ, et al. Cloning and functional identification of slc5a12 as a sodium-coupled low-affinity transporter for monocarboxylates (SMCT2). The Biochemical journal. 2005;392(Pt 3):655-64.

41. Arthur S, Coon S, Kekuda R, Sundaram U. Regulation of sodium glucose co-transporter SGLT1 through altered glycosylation in the intestinal epithelial cells. Biochim Biophys Acta. 2014;1838(5):1208-14.

42. Valenzuela L, Kamakaka RT. Chromatin insulators. Annual review of genetics. 2006;40:107-38.

43. Holwerda S, de Laat W. Chromatin loops, gene positioning, and gene expression. Frontiers in genetics. 2012;3:217.

44. Li Y, Huang W, Niu L, Umbach DM, Covo S, Li L. Characterization of constitutive CTCF/cohesin loci: a possible role in establishing topological domains in mammalian genomes. BMC genomics. 2013;14:553.

45. Franklin PJ, Turner SW, Mutch RC, Stick SM. Measuring exhaled nitric oxide in infants during tidal breathing: methodological issues. Pediatr Pulmonol. 2004;37(1):24-30.

46. Leipala JA, Williams O, Sreekumar S, Cheeseman P, Rafferty GF, Hannam S, et al. Exhaled nitric oxide levels in infants with chronic lung disease. European journal of pediatrics. 2004;163(9):555-8.

47. Williams O, Rafferty GF, Hannam S, Milner AD, Greenough A. Nasal and lower airway levels of nitric oxide in prematurely born infants. Early human development. 2003;72(1):67-73.

48. Bouzigon E, Nadif R, Thompson EE, Concas MP, Kuldanek S, Du G, et al. A Common variant in RAB27A gene is associated with fractional exhaled nitric oxide levels in adults. Clin Exp Allergy. 2014.

624 49. Fuchs O, Genuneit J, Latzin P, Buchele G, Horak E, Loss G, et al. Farming
625 environments and childhood atopy, wheeze, lung function, and exhaled nitric oxide. J
626 Allergy Clin Immunol. 2012;130(2):382-8 e6.

FIGURES

Figure 1. Flowchart of excluded study participants due to general exclusion criteria or quality control

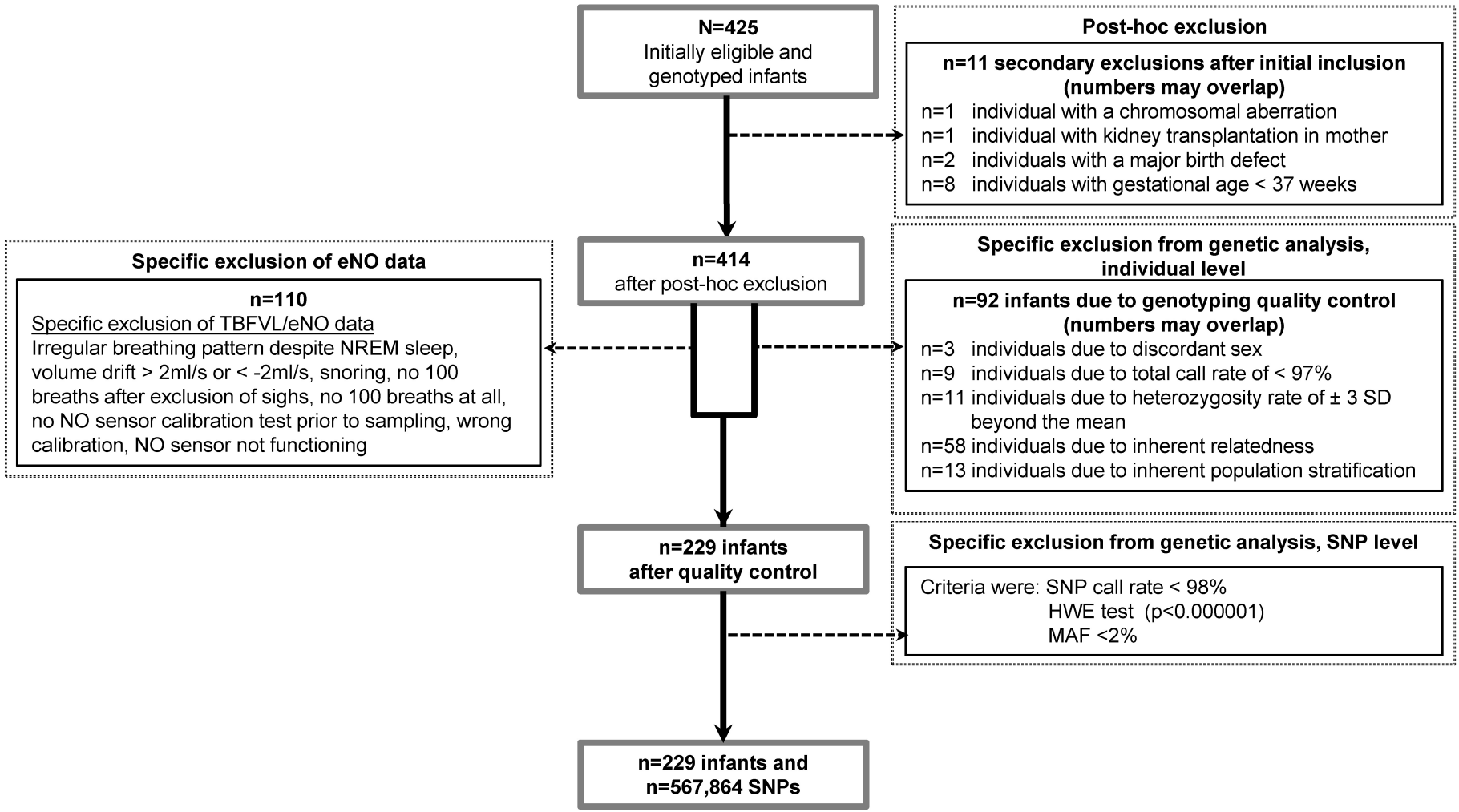
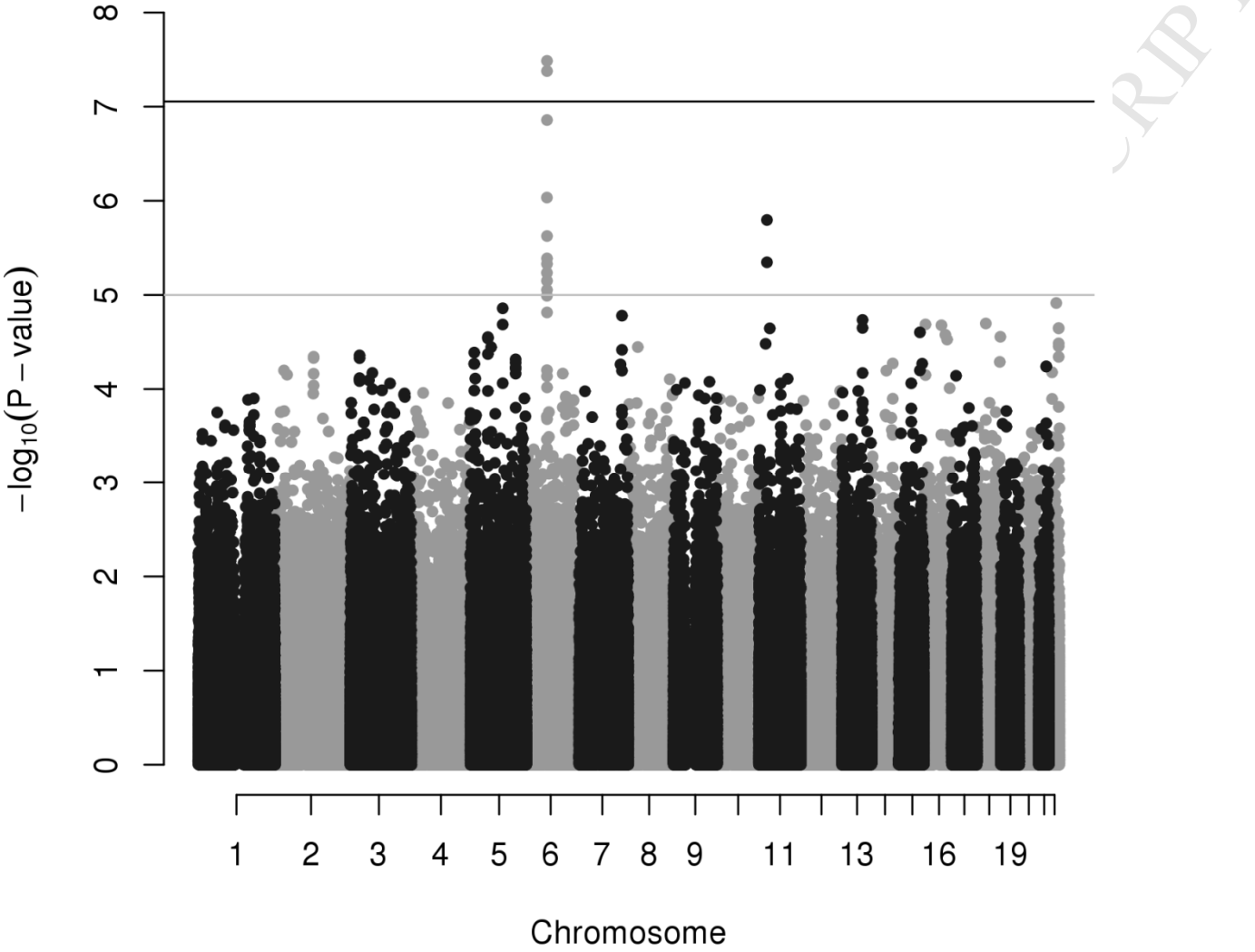


Figure 2, A and B. Manhattan plots displaying the associations of eNO (A) and V'NO values (B), final models

A.



B.

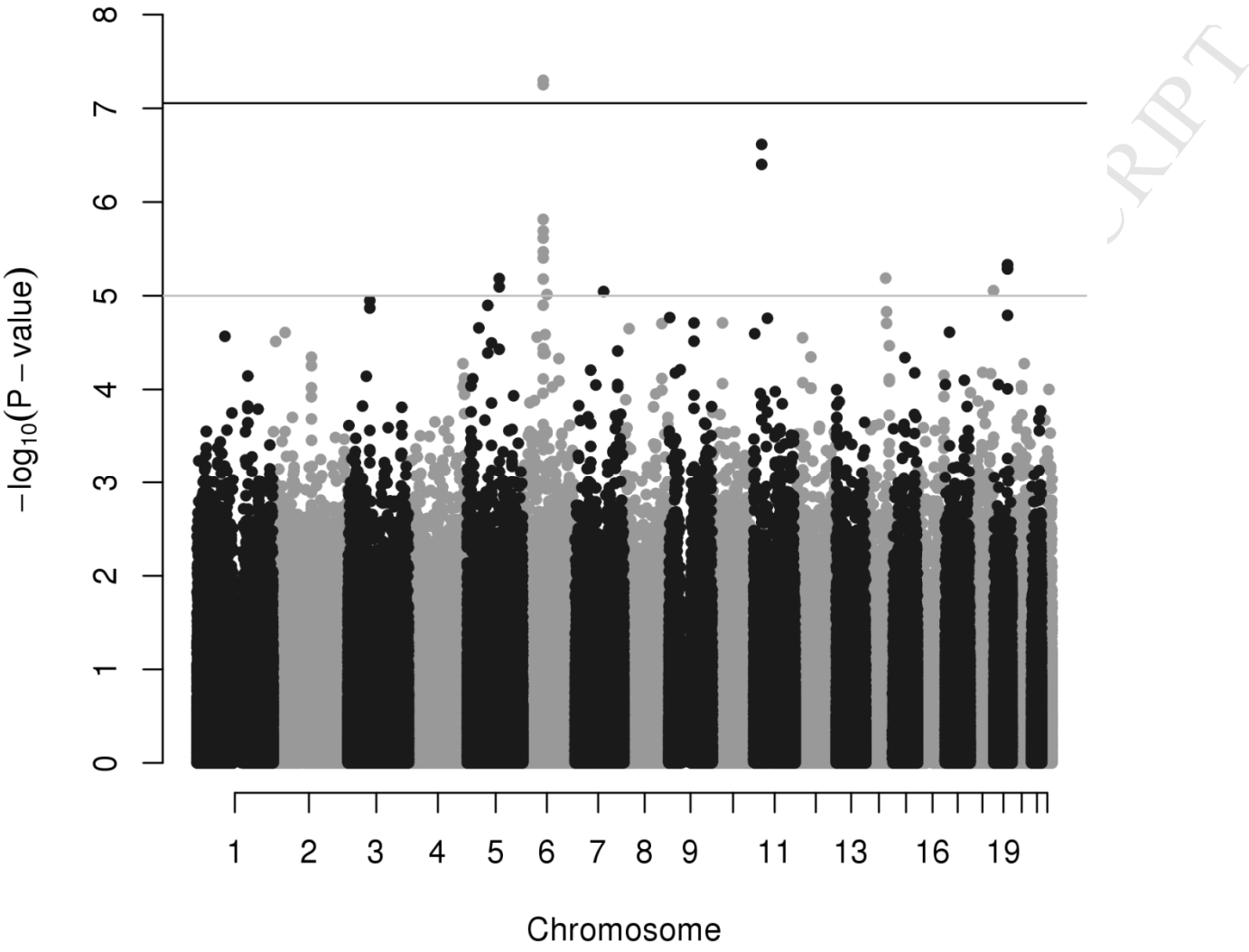
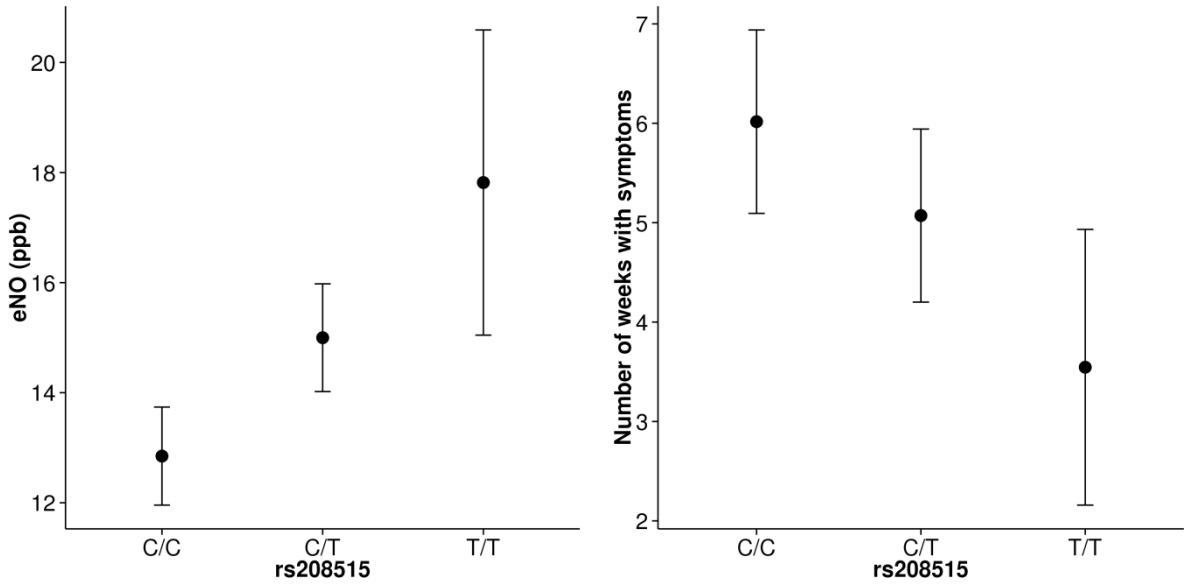


Figure 3, A and B: Association of the 6q12 (A) and of the 11p14 (B) loci with eNO and respiratory symptoms during the first year of life across genotypes

A.



B.

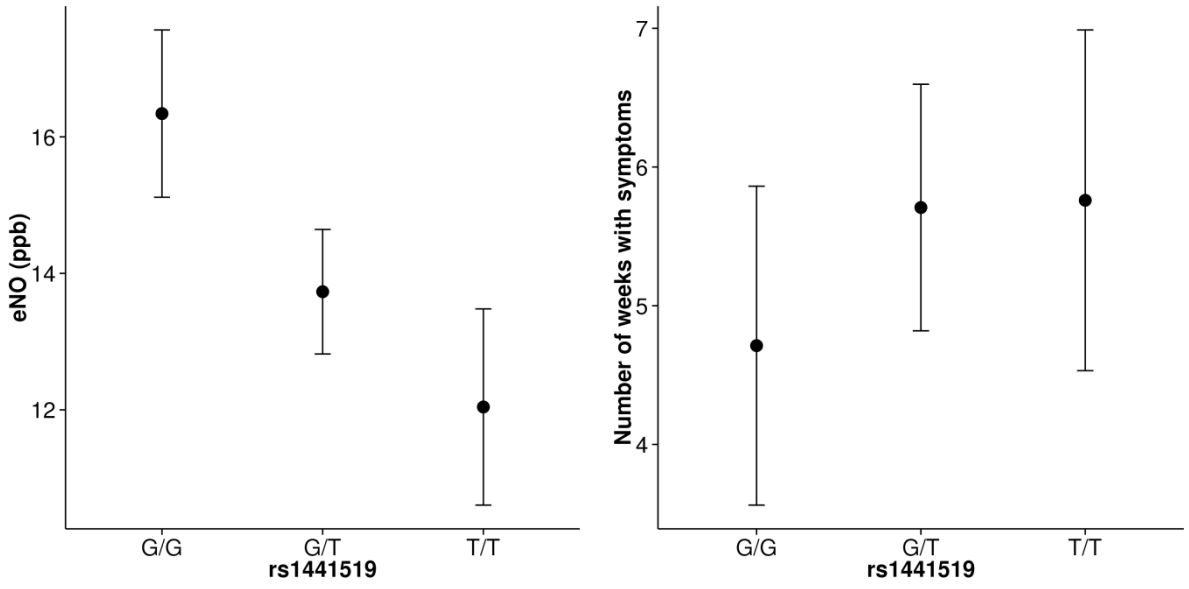


Figure 4, A and B: Fine mapping of eNO association signals on chromosomes 6 and 11

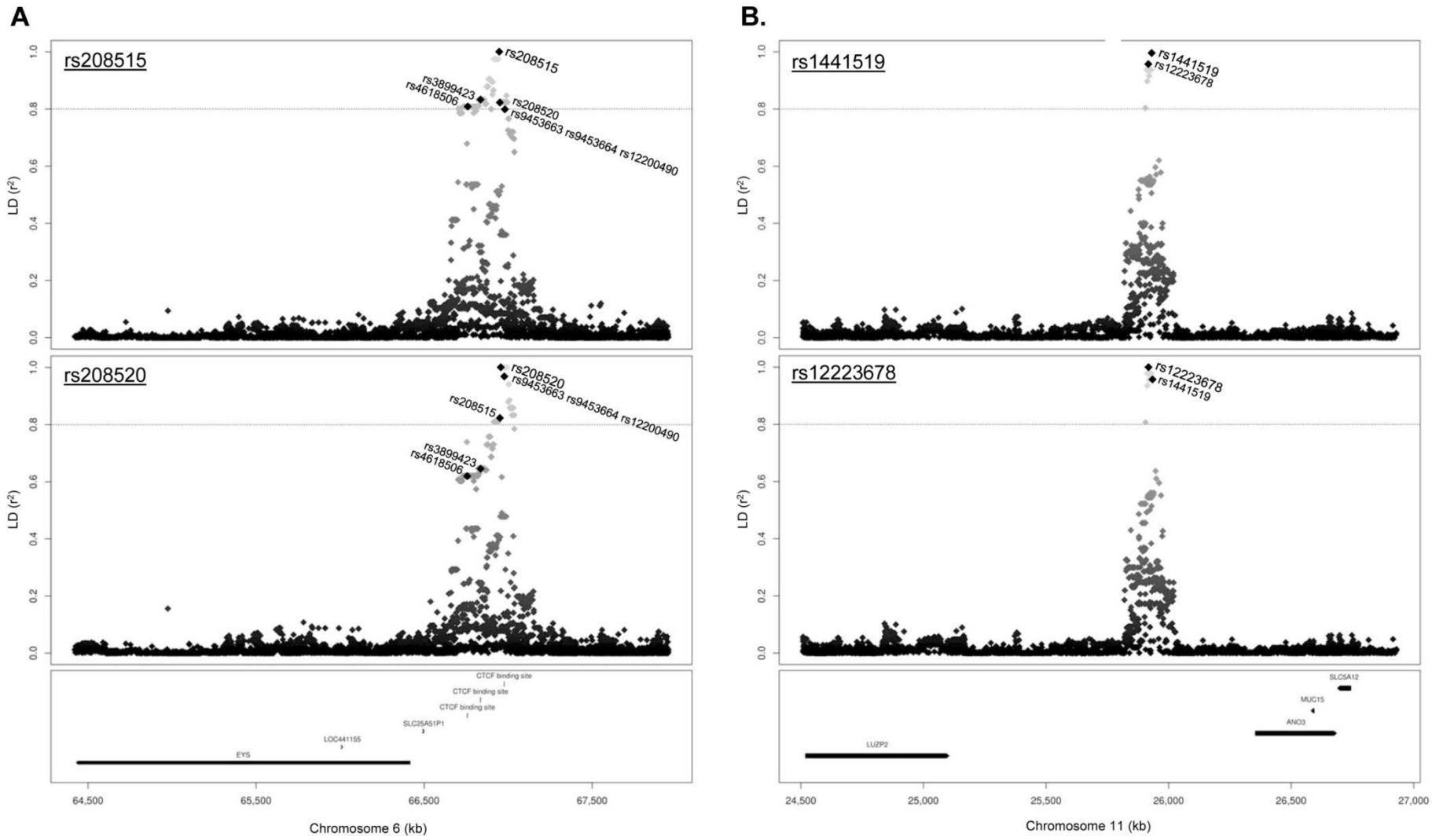


Figure 5, A, and B: mRNA Expression of *NOS3* in epithelial cells before and after knock down of 11p14 genes

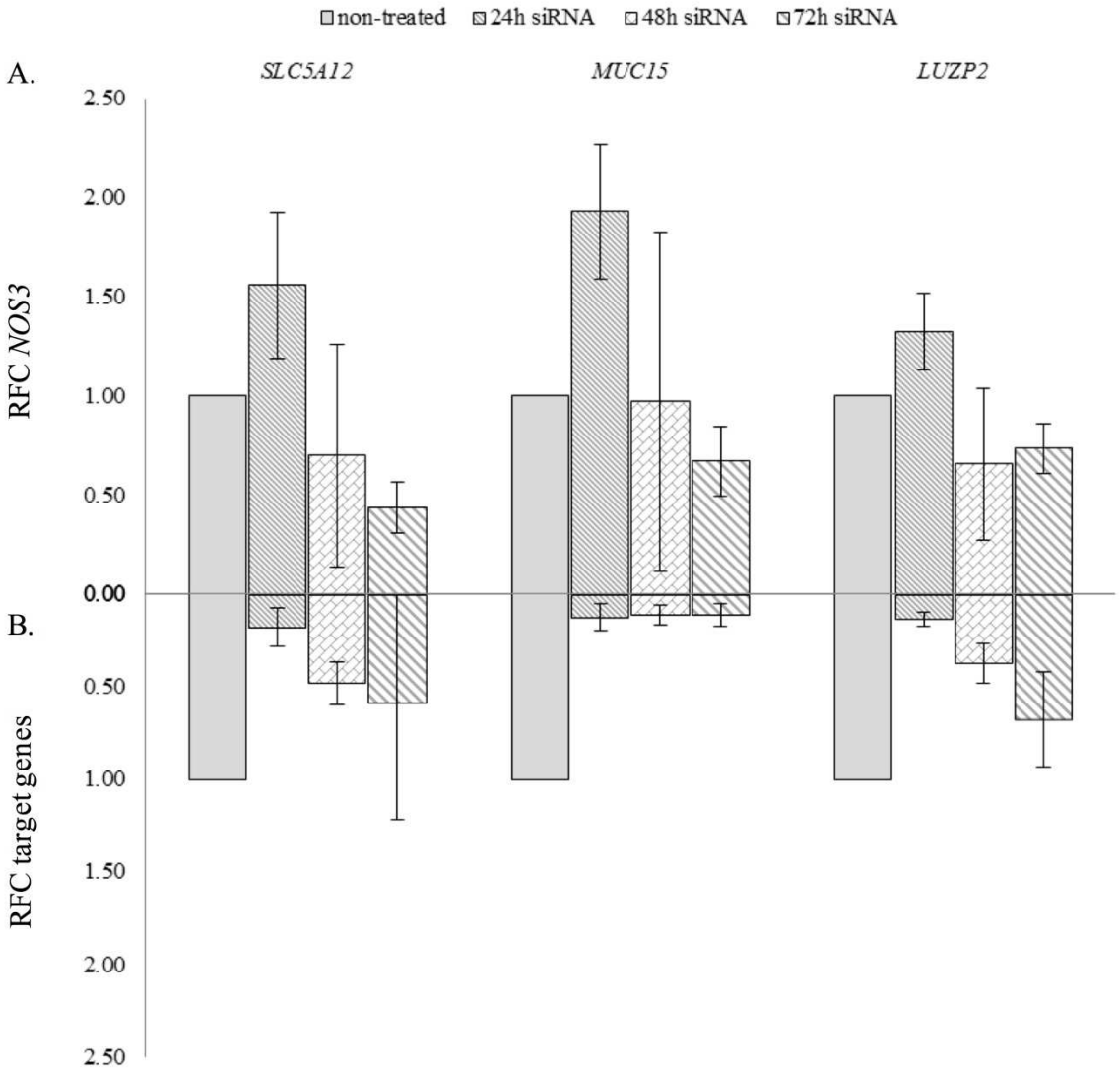


FIGURE LEGENDS

Figure 1. Flowchart of excluded study participants due to general exclusion criteria or quality control.

See text for details.

Abbreviations: (e)NO – (exhaled) nitric oxide, IBD – inheritance-by-descend, NREM – non rapid eye movement, SD – standard deviation, PCA – principal component analysis

Figure 2, A and B. Manhattan plots displaying the associations of eNO (A) and V'NO values (B), final models.

Each dot represents the p-value for the association of a SNP with the outcomes eNO and V'NO. For this, the $-\log_{10}$ of each SNP (y-axis) is plotted against the genomic position, i.e. chromosome (x-axis) adjusting for PC1 as well as covariates. These were study length, maternal smoking during pregnancy, maternal history of atopy, age at measurement, as well as V'E for eNO, and study weight, maternal smoking during pregnancy, maternal history of atopy, age at measurement, and V'E for V'NO in the final models for either outcome. The black lines display the genome-wide Bonferroni significance threshold $p<8.8\times10^{-8}$ (0.05/ 567,864). The light grey lines display the suggestive genome-wide significance level $p<1\times10^{-5}$.

Abbreviations: eNO – exhaled nitric oxide, PC1 – principal component 1, SNP – single nucleotide polymorphism, V'E – minute ventilation, V'NO – NO output

Figure 3, A and B: Association of the 6q12 locus (A) and of the 11p14 locus (B) with eNO and respiratory symptoms during the first year of life across genotypes.

Figures display the effect of rs208515 (A) and of rs1441519 (B) on means of eNO on the left (y-axis, eNO in ppb) and weeks with respiratory symptoms during the first year of life on the

right (y-axis in weeks with symptoms) and their respective 95% confidence interval across genotypes (from left to right: homozygous for non-risk allele, heterozygous, homozygous for risk allele).

Abbreviation: (e)NO – (exhaled) nitric oxide, ppb – parts per billion.

Figure 4, A and B: Fine mapping of eNO association signals on chromosomes 6 and 11.

LD analysis identified that no polymorphisms in linkage with the identified target SNPs at 6q12 (A: rs208515 and rs208520) and 11p14 (B: rs12223678 and rs1441519) are located in close vicinity to known genes (LD threshold r2≥0.8). Database search identified that five tagged polymorphisms on chr.6 are located in CTCF binding sites (rs4618506, rs3899423, rs9453663, rs9453664, rs12200490; please note that only CTCF binding sites matching to analyzed SNPs are depicted). *In-silico* prediction proposed that the 11p14 locus contains cell type-specific chromatin structures called topologically associating domains, which encompass the *LUZP2* and *ANO3/MUC15* loci (domains are not specifically depicted due to their vast predicted extend of over 3,000kb in embryonic stem cells and human lung fetal myofibroblasts). The plots were prepared using R version 3.0.0.

Abbreviations: *ANO3* – Human anoctamin 3 gene locus, CTCF – 11-zinc finger protein or CCCTC- binding factor, *EYS* – Human eyes shut homolog gene locus, kb – kilo bases, LD – linkage disequilibrium, *LUZP2* – Human leucine zipper protein 2 gene locus, *MUC15* – Human mucin 15 gene locus, rs – reference SNP cluster code.

Figure 5, A, and B: mRNA Expression of *NOS3* (A) and of target genes (B) in epithelial cells before and after knock-down of 11p14 target genes.

The reference baseline value to evaluate expression level changes of each target gene is the non-treated sample (RFC = 1). Data is given as means ± standard deviation (SD). (A) upper panel: mRNA levels of *NOS3* in before and after knock down of respective target genes. (B)

lower panel: Knock down efficiency: mRNA levels of *SLC5A12*, *MUC15* and *LUZP2* at 24, 48, and 72 hours after knock-down.

Abbreviations: *LUZP2* - human leucine zipper protein 2 gene locus, *MUC15* - human mucin 15 gene locus, RFC – Relative fold change, *SLC5A12* – human solute carrier family 5 (sodium/monocarboxylate cotransporter) member 12.

TABLES

Table 1: Main outcomes, as well as anthropometric and demographic characteristics for n=229 study participants with complete data on eNO measurements and GWAS

Variable	Mean (SD)	Median	IQR	Range
eNO, ppb	14.1 (5.2)	13.9	11.0-17.2	1.8-32.9
V'NO, nL/s	0.63 (0.23)	0.63	0.46-0.78	0.07-1.29
Minute ventilation, ml	1401.0 (251.1)	1401.0	1232-1565	786-2221
Anthropometric characteristics				
Gestational age, weeks	39.8 (1.14)	40.0	39.1-40.9	37.1-42.3
Postnatal age at measurement, days	35.2 (4.65)	34	32-38	25-53
Birth weight, kg	3.4 (0.42)	3.4	3.1-3.7	2.3-4.9
Birth length, cm	49.6 (2.03)	49.6	48.0-51.0	45.0-57.0
Study weight, kg	4.4 (0.52)	4.3	4.0-4.8	3.1-5.8
Study length, cm	54.9 (2.11)	54.8	53.2-56.6	50.0-59.6
	Number (%)			
Demographic and socioeconomic characteristics, distribution of covariates				
Male sex	115 (50.2%)			
Maternal asthma	27 (11.8%)			
Maternal atopy	79 (34.5%)			
Positive maternal skin prick test*	82 (37.8%)			
Maternal smoking during pregnancy	20 (8.7%)			
Paternal smoking during pregnancy	42 (18.4%)			
Parental smoking during pregnancy	46 (20.1%)			
Caesarean section	39 (17%)			

* data available for n=217 mothers only

Abbreviations: eNO – exhaled nitric oxide, GWAS – genome-wide association study, IQR – interquartile range, SD – standard deviation, V'NO – nitric oxide output.

Online Repository (OLR): 6q12 and 11p14 variants are associated with postnatal exhaled nitric oxide and respiratory symptoms

Oliver Fuchs, MD PhD[1,2,3*], Olga Gorlanova, MD[1*], Philipp Latzin, MD PhD[1,3], Anne Schmidt, MD[1], Maximilian Schieck, PhD[4,5], Antoaneta A. Toncheva, PhD[4], Sven Michel, PhD[4], Vincent D. Gaertner, BSc[4], Michael Kabesch, MD[4,6**] and Urs Frey, MD PhD[1**]

[1] University Children's Hospital (UKBB), University of Basel, Basel, Switzerland

[2] Dr von Hauner Children's Hospital, Ludwig Maximilians University, Munich, Germany, and Comprehensive Pneumology Center,Munich (CPC-M), Germany; Member of the German Center for Lung Research (DZL)

[3] Division of Respiratory Medicine, Department of Pediatrics, Inselspital, University of Bern, Bern, Switzerland

[4] Department of Pediatric Pneumology and Allergy, University Children's Hospital Regensburg (KUNO), Regensburg, Germany

[5] Institute of Human Genetics, Hannover Medical School, Hannover, Germany

[6] Department of Pediatric Pneumology, Allergology and Neonatology, Hannover Medical School, Hannover, Germany, and Biomedical Research in Endstage and Obstructive Lung Disease Hannover (BREATH), Germany; Member of the German Center for Lung Research (DZL)

[*, **] Both authors contributed equally

OLR DEFINITION OF ABBREVIATIONS

ATS	American Thoracic Society
BILD	Basel/Bern Infant Lung Development (Cohort)
BTPS	Body temperature and ambient pressure, saturated with water vapor conditions
cat	Catalogue
cM	CentiMorgan
CEU	Utah residents with ancestry from northern and western Europe
CHB	Han Chinese in Beijing, China
CTCF	11-zinc finger protein or CCCTC- binding factor
ECR	Evolutionary conserved genomic region
(e)NO	(Exhaled) nitric oxide
ERS	European Respiratory Society
ETS	Environmental tobacco smoke
GxE	Gene-environment interaction
GxG	Gene-gene interaction
HWE	Hardy-Weinberg equilibrium
IBD	Identity-by-descent
IRR	Incidence risk ratio
JPT	Japanese in Tokyo, Japan
LD	Linkage disequilibrium
MAF	Minor allele frequency
Mb	Megabase

48	NREM	Non-rapid eye moving
49	OLR	Online repository
50	OR	Odds ratio
51	PC(A)	Principal component (analysis)
52	p_{int}	Interaction p-value
53	QC	Quality control
54	QQ	Quantile-quantile plot
55	QTL	quantitative trait locus
56	rs	Reference SNP cluster code (dbSNP)
57	SD	Standard deviation
58	SNP	Single-nucleotid polymorphism
59	UCB	Umbilical cord blood
60	V'E	Minute ventilation
61	V'NO	Nitric oxide output
62	VT	Tidal volume
63	YRI	Yoruba in Ibadan, Nigeria
64	95% CI	95% confidence interval

65

OLR METHODS

This is a more detailed description of the study population and methods as provided in the main document.

Study population, study participants

All data were collected in an ongoing prospective birth cohort of unselected, healthy neonates - recruited antenatally since 1999 in Bern, and since 2011 in Basel, Switzerland, the Bern-Basel Infant Lung Development (BILD) cohort.(1) The following inclusion criteria apply: white ethnicity, term delivery (≥ 37 weeks), no known major birth defects or perinatal disease of the newborn.

Known and potential confounders of lung function (demographic data, sociodemographic status, smoke exposure and parental atopic disease, the latter being defined as either atopic dermatitis, atopic rhinitis, or atopic asthma in either parent) were assessed by interviews using standardized questionnaires.(1)

In addition, a skin prick test, including six common allergens (Dog dander, cat dander, *Dermatophagoides pteronyssinus*, mixed tree pollens, mixed grass pollens, *Alternaria tenuis*, Allergomed AG, Therwil, Switzerland) was performed in a subgroup of mothers. We validated antenatal maternal smoking history by measuring cotinine levels in the first urine of the newborn (gas-liquid chromatography, IST, Lausanne, Switzerland).

The study was approved by both ethics committees of the cantons of Bern and Basel-Stadt, Switzerland (Kantonale Ethikkommission Bern and Kantonale Ethikkomission Basel-Stadt), as well as the research ethics committees of the University Hospital Bern, Switzerland (Inselspital) and of the University Hospital Basel, Switzerland (Universitätskinderspital beider Basel, UKBB). All caregivers provided written informed consent for this study.

Measurements of tidal breathing, nitric oxide

At the age of five weeks, tidal breathing and mixed exhaled nitric oxide (eNO, oral and nasal) were measured in unsedated neonates during behaviourally-defined quiet natural sleep (non-rapid eye moving, NREM phase), using a multiple-breath online measurement technique as described previously.(1, 2) Measurements followed regular feeding of the infants, usually resulting in natural sleep in this age group. They were performed in a supine position with the head midline, via an infant mask (Size 1; Homedica AG, Huenenberg, Switzerland) according to the standards by the European Respiratory Society (ERS) and the American Thoracic Society (ATS) of infant lung function testing (3-5) eNO measurement(6). Flow was measured using an ultrasonic flowmeter (Spiroson®; EcoMedics AG, Duernten, Switzerland). Exhaled nitric oxide was measured online with a rapid-response chemoluminescence analyzer (CLD 77 AM®, CLD 88 sp®; EcoMedics AG, Duernten, Switzerland). For signal collection, a software package (WBreath 3.28; ndd Medical Technologies, Switzerland) was used and tidal flows and derived volumes were converted to body temperature and ambient pressure, and saturated with water vapor (BTPS) conditions. (1, 2) Main outcome parameters were tidal volume (V_T), minute ventilation ($V'E=V_T$ x respiratory rate), eNO, and NO output (V'NO=eNO concentration x corresponding expiratory flow).(2)

Quality control of tidal breathing and eNO measurements

We performed sensor calibration tests prior to each measurement. Contamination by ambient NO was avoided by always using NO-free air for inspiration. For infant tidal breathing and eNO measurements, we used the first 100 regular breaths of tidal breathing during NREM sleep from a total recording of over 10 min, and excluded the first 20 to 30 breaths after mask placement to allow for adjustment of breathing pattern. In addition, sighs, together with 10 breaths before and after a sigh, were excluded to reduce noise.(2) Mean tidal breathing

parameters of flow, volume and flow-volume loop were then calculated according to ERS/ATS standards.(3) Data were included, if no apparent volume drift was present, defined as a change of less than 2 $mL \cdot s^{-1}$.(2) During multiple-breath measurements, eNO was measured breath-by-breath during the third quartile of expiration and mean eNO was calculated over 100 breaths correcting for expiratory flow as described previously.(2) In total, data of n=110 infants had to be excluded due to specific exclusion of data from tidal breathing and eNO measurements (see Figure 1 in main document).

Measurement of respiratory symptoms and time to recovery from first respiratory infection during the first year of life

We collected detailed information about the respiratory symptoms of the children by weekly telephone interviews during the first year of life using a standardized symptom score.(1) Respiratory symptoms were defined as 'cough, wheezing, difficulty breathing and reduced activity during night and day', as observed by their mother. Weeks with symptoms were calculated as a total number of weeks with day or night respiratory symptoms during the first year of life. Furthermore, data on time to recovery was collected as the number of weeks that the infant suffered from respiratory symptoms.

Genotyping, genetic data quality control

Genotyping was performed in two batches in all n=425 study participants (in 2011 for n=329 and in 2013 for an additional n=96 children) either from umbilical cord blood (UCB) samples (n=418) or buccal swabs (n=7) for infants for whom there was no DNA from UCB samples.(1) Genomic DNA was extracted using the QIAamp Maxi Kit (Qiagen AG, Basel, Switzerland) according to the manufacturer's recommendations, quantified by spectrophotometry, and checked for fragmentation and amplification. Whole-genome

genotyping was performed on 20µl of DNA aliquoted to a concentration of 50 ng/µl with Illumina HumanOmniExpress Bead Chips (Illumina Inc., San Diego, USA).

From all n=425 children, some individuals were excluded post-hoc from further analyses: Individuals with a gestational age below 37 weeks (n=8), chromosomal aberration (n=1), major birth defect (n=2), and maternal kidney transplantation (n=1), see Figure 1 in main document. Before entering subsequent analyses, quality control (QC) of eNO measurements was performed for the remaining n=414 individuals (see above) and their genotyping data were subjected to further QC; both on the individual and single-nucleotid polymorphism (SNP) level to assure robustness of association analyses (see Figure 1 in main document).

On the individual level, several steps of QC were done. First, subjects with total call rates of <97% (n=9), as well as individuals with discordantly called sex (n=3) were excluded. In addition, n=11 individuals with heterozygosity rates of beyond ± 3 standard deviations (SD) of the population mean were excluded. Secondly, we checked for inherent relatedness of individuals as well as population stratification for a subset of autosomal SNPs with minor allele frequencies (MAF) ≥ 5% after linkage disequilibrium (LD) pruning (r^2<0.2, window 50 Mb, resulting in 98,154 SNPs. In order to exclude related individuals, an identity-by-descent (IBD) matrix was calculated. Any pairs of individuals with a relatedness measure ($\hat{\pi}$) of $\hat{\pi}$ > 0.185 were considered to be indicative of relatedness. From such pairs, individuals with the lowest eNO quality were excluded from further analyses, these were n=58 study participants in total. If pairs had equal eNO quality, one individual was randomly excluded. Furthermore, in order to correct for population stratification, principal component analysis (PCA) was conducted in the “genome-wide complex trait analysis” (GCTA) tool based on the EIGENSTRAT algorithm.(7) The PCA was performed for all study participants remaining after QC on the individual level, i.e. after exclusion due to relatedness, and in comparison with genotypes of unrelated individuals from the HapMap Project (www.hapmap.org)

including Utah residents with ancestry from northern and western Europe (CEU), Yoruba in Ibadan, Nigeria (YRI), Japanese in Tokio, Japan (JPT) and Han Chinese in Beijing, China (CHB). This step led to further exclusion of n=13 subjects due to inherent population stratification.

Finally, further selection on the SNP level was performed for all individuals after post-hoc exclusion and passing QC on the individual level and for eNO data. Here, criteria for exclusion from further analysis were call rates of <98%, markers failing the Hardy-Weinberg equilibrium (HWE) test ($p<1.0\times10^{-6}$), and low minor-allele frequencies (MAF, <2%). We also checked for a possible batch effect due to the two time-points of genetic analyses (2011 and 2013). For this, we analyzed the distribution of covariates in both batches in order to rule out non-random distribution of confounding factors. Furthermore, a principal component analysis (PCA) was performed on the final subset of n=229 individuals that were left for subsequent analyses in order to obtain eigenvectors as covariates for later association tests. Here, the first principal component (PC1) explained the largest proportion of variance (0.54%, data not shown) and has been subsequently included for adjustment.

Statistical analyses and association testing

After the study-specific QC, a total of 229 individuals were included in the present study. As values for both eNO and V'NO in the final dataset were already normally distributed (data not shown), no further transformation in order to obtain a normal distribution was necessary. We therefore examined untransformed eNO and V'NO values.

Additive genetic models were used for all subsequent analyses. Association between eNO or V'NO values, and each SNP, was investigated by linear regression, adjusting for PC1 as well as known and possible covariates. These were study length, maternal smoking during pregnancy, maternal history of atopy, age at measurement, as well as V'E for eNO, and study

weight, maternal smoking during pregnancy, maternal history of atopy, age at measurement, and V'E for V'NO in the final models for either outcome. The genomic inflation factor (λ) was calculated as the median of the chi-squared test statistics divided by the expected median (0.455) of the chi-squared distribution with one degree of freedom. In order to control for multiple testing, a conservative critical p-value threshold using the Bonferroni method (calculated as significance level α divided by the number of SNPs, i.e. 0.05/567,864, resulting in $p < 8.8\times10^{-8}$), was chosen. The proportion of variance explained by each SNP was obtained as the difference in R^2 between full (including SNP and all covariates) and reduced (including only covariates) models.

Poisson regression with robust standard errors was used for analysis of the association between top SNPs for each identified locus and the total number of weeks with respiratory symptoms during the first year of life. Results are expressed as incidence risk ratio (IRR) and adjusted for known and possible covariates (sex, having older siblings, nursery care, maternal atopy, maternal smoking during pregnancy, V'E, and eNO).

Logistic regression was used for analysis between SNPs and time to recovery from the first respiratory symptoms during the first 6 and 12 months of life. For this, a binary outcome variable was defined dividing children with respiratory symptoms into those with a resolution of symptoms < 2 week or ≥ 2 week. Results are expressed as odds ratio (OR) and adjusted for for known and possible covariates (sex, age, having older siblings, nursery care, maternal atopy, maternal smoking during pregnancy, season, being breast-fed at time of first respiratory symptoms, V'E, and eNO).

To test for interactive effects, we selected SNPs with the p-value below 1×10^{-5} with the above mentioned analyses. Gene-gene and gene-environment interactions between pairs of top SNPs for each identified locus and each top SNP and environmental factors (e.g. maternal smoking during pregnancy and maternal atopy), respectively, were assessed by ANOVA with one

degree of freedom comparing the full model, including the interaction term, with the reduced model without the interaction term. The corresponding p-value was labeled as p_{int}. A p-value of 0.05 was considered significant.

Association computations were done in R version 3.0.2 (www.r-project.org) (8) using the GenABEL package (9) and Stata 12.1 (STATA Corporation, College Station, TX, USA) both for the final subset of all n=229 individuals that were left for subsequent analyses as well as for each genotyping batch from either 2011 or 2013 separately.

Replicating results for SNPs with significant association with eNO in children,(10, 11) meta-analyses were performed using the *Metan* package in Stata 11.2 (STATA Corporation, College Station, TX, USA). For this, we included a total of four SNPs or their proxies, respectively, for which we had access to data in our final dataset. Since eNO values were log-transformed (ln) in previous studies, we also performed log-transformation (ln) for eNO values from the BILD cohort prior to meta-analysis. New effect estimates and their 95%CI were derived by calculating weighted differences between fixed-effect meta-analysis estimates from published data(10, 11) and estimates from BILD data. We used the I^2 statistic as a measure of heterogeneity between studies; the corresponding p-value was labeled as p_{het}.

Regional plots were generated by LocusZoom (http://csg.sph.umich.edu/locuszoom). Permutation analyses were performed using PLINK version 1.07 by chromosome using the max(T) (*-mperm*) option with 100.000 permutations in total. Linkage disequilibrium (LD) between SNPs was estimated using PLINK version 1.07 (SNP inclusion parameters: minor allele frequency (MAF) ≥0.05; Hardy-Weinberg equilibrium *p*-value ≥0.0001)(12) based on 1000 Genomes Project data.(13) Regional LD plots based on r2-values were prepared using R version 3.0.0. Base pair (bp) positions are given according to genome assembly GRCh37.p13 and gene annotations are according to the NCBI RefSeqGene Project (accessed March 24th 2015).

241

242 In-silico functional analyses

243 Several databases (the Multiple Tissue Human Expression Resource (MuTHER) of the
244 Wellcome Trust, http://www.muther.ac.uk/public.html(14); a database on eQTL-weighted
245 GWAS, http://www.hsph.harvard.edu/liming-liang/eqtl-weighted-gwas/(15); a database for
246 eQTL results derived from lymphoblastoid cell lines from asthmatic children,
247 http://csg.sph.umich.edu/liang/asthma/(16); the eQTL Browser of the National Institute of
248 Health (NIH), http://www.ncbi.nlm.nih.gov /projects/gap/eqtl/index.cgi; the blood eQTL
249 browser, http://genenetwork.nl /bloodeqtlbrowser/(17); the University of Michigan Center for
250 Statistical Genetics skin eQTL database, http://www.sph.umich.edu/csg/junding/eQTL(18);
251 the Genotype-Tissue Expression (GTEx) portal of the BROAD Institute,
252 http://www.gtexportal.org/home/(19); the Bgee Gene Expression Evolution database,
253 http://bgee.unil.ch/(20); as well as further two datasets mentioned in articles (21, 22)) were
254 used to check the relationship of target polymorphisms with published data on expression
255 (eQTL) and methylation QTLs (mQTL). ECR Browser (http://ecrbrowser.dcode.org/) was
256 used to check whether the target polymorphisms are located in evolutionary conserved
257 genomic regions (ECRs). CTCF binding sites and chromatin topological domains were
258 determined using data from CTCFBSDB 2.0 - a database for CTCF binding sites and genome
259 organization (http://insulatordb.uthsc.edu/).(23)

260

261 Cell culture and gene expression analyses

262 a) *Cell culture*

263 H358 lung epithelial cells were cultured at 37°C with 5% CO_2 in RPMI 1640 Glutamax (cat.
264 #61870-010, Gibco by ThermoFisher Scientific) supplemented with 10% heat-inactivated

FBS (Biochrom, FBS Superior, catalogue, cat. #S0615) and penicillin/ streptomycin (5 µg/ml, cat. #P4333-20ML Sigma-Aldrich, Steinheim, Germany).

Twenty-four hours prior to transfection, cells were washed with 1x PBS (cat. #10010023, Gibco by ThermoFisher Scientific) and treated with 6 ml accutase per 75 cm^2 flask (cat. #A6964-100ML, Sigma-Aldrich, Steinheim, Germany) for 5-10 min at 37°C with 5% CO2. Afterwards, 5 ml of cell culture media were added to inhibit the action of the accutase. Cells were then centrifuged for 5 min at 200 x g at room temperature. Supernatant was removed and 10 ml fresh cell culture media were added. Cells were then counted using trypan blue stain 0.4% (cat. #T10282, Molecular probes by ThermoFisher Scientific), Countess cell counting chamber slides (cat. #C10283 Molecular probes by ThermoFisher Scientific), and the Countess II FL automated cell counter (cat. #AMQAF1000, ThermoFisher Scientific). One hundred thousand cells were seeded in each well of the 24-well plate in 450 µl culture media.

b) siRNA transfection

Twenty-four hours later, cells were transfected with 53.5 µl mixture containing the following components:

- Opti-MEM I Reduced Serum Medium (cat. #31985070, ThermoFisher Scientific) - 50 µl
- Lipofectamine RNAiMAX transfection reagent (cat. #13778150, ThermoFisher Scientific) - 1.5 µl
- siRNA1 and siRNA2 working stock - 1 µl of each

Stock solutions of all siRNAs were prepared by adding 200 µl dH_2O to the lyophilized substance. This solution contained 100µM siRNA. A mix of two different siRNAs was used for each target gene to achieve better knock-down efficiency. The following products were used (see Table S1):

Table S1. siRNAs used for transfection

Product	Cat. #/ ID
LUZP2 siRNA 1	s50366
LUZP2 siRNA 1	s50365
MUC15 siRNA 1	s44561
MUC15 siRNA 2	s44563
SLC5A12 siRNA 1	s46170
SLC5A12 siRNA 2	s46171
Silencer Select GAPDH Positive Control siRNA	4390849

The transfection reagents mixture was incubated for 20 min at room temperature, then 53.5 µl was added to the respective wells. Cells were then incubated at 37°C with 5% CO2. Each condition (i.e. transfection, positive control and non-treated cells) was conducted in 4 independent biological replicates. Silencer Select GAPDH Positive Control siRNA was used to evaluate the knock-down efficiency (cat. #4390849, ThermoFisher Scientific). Approximately 15 hours after the transfection 250 µl of fresh cell culture media were added to each well in order to support the cells with nutrients due to high evaporation rate at 37°C. Cells were harvested at hours 24, 48 and 72 after the transfection. Cell culture supernatants were collected and stored at -80°C for further experiments. Silencer Select *GAPDH* Positive Control siRNA (cat. # 4390849, ThermoFisher Scientific) was used as a transfection control to assess the transfection efficiency (see Table S2). *GAPDH* gene was expressed only at rate 18.42% compared to the non-treated cells 24 hours after transfection. The rate was 5.83% and 15.57% after 48 and 72 hours, respectively. Data was derived from 7 independent biological replicates of the respective treated and non-treated samples.

Table S2. Transfection efficiency

Condition	*GAPDH*	SD
Non-treated	100.00%	
24h siRNA	18.42%	0.0659
48h siRNA	5.83%	0.0082
72h siRNA	15.57%	0.1721

Abbreviation: SD – standard deviation

c) Total RNA isolation and cDNA synthesis

Total RNA isolated from the cell pellets with RNeasy Micro Kit (cat. #74004, Qiagen, Hilden, Germany) according to the manufacturer's protocol. RNA concentration was measured by Qubit 3.0 Fluorometer (cat. # Q33216, ThermoFisher Scientific) with Qubit RNA BR Assay Kit (cat. # Q10211, ThermoFisher Scientific) and cDNA was synthesized using 1 µg total RNA per sample with QuantiTect Reverse Transcription Kit (cat. #205313, Qiagen, Hilden, Germany) according to the manufacturers' protocols.

d) Real-time quantitative PCR gene expression analyses - qRT-PCR reactions

Predesigned TaqMan gene expression assays (listed below, see Table S3) were used to detect *EYS*, *ANO3*, *SLC5A12*, *MUC15*, *LUZP2*, *NOS1*, *NOS2*, *NOS3* genes.

Table S3. Used TaqMan gene expression assays

Product	**Cat. #/ ID**
EYS	Hs01650973_m1
LUZP2	Hs01382617_m1
ANO3	Hs00230040_m1
MUC15	Hs00377336_m1
SLC5A12	Hs01054637_m1
NOS1	Hs00167223_m1
NOS2	Hs01075529_m1
NOS3	Hs01574659_m1
GAPDH	Hs02758991_g1

Human *18S rRNA* was used as an endogenous control. Primers and fluorescence-labelled probe (Metabion International AG, Martinsried, Germany) were designed with Vector NTI Advance 10 software (Invitrogen by ThermoFisher Scientific) and are listed below (see Table S4).

Table S4. Primers and fluorescence-labelled probe for human 18S rRNA

Gene	**Forward primer (5'→3')**	**Reverse primer (5'→3')**	**Fluorescent probe (5'→3')**

18S rRNA	AGTCCCTGCCCTTTGTACACA	GATCCGAGGGCCTCACTAAAC	6-FAM – CGCCCGTCGCTACTACCGATTGG – Tamra

326

327 mRNA levels were assessed by quantitative real-time PCR (qRT-PCR) using the StepOne
328 Plus Real-Time PCR System (cat. #4376600, ThermoFisher Scientific). The following PCR
329 reaction mixtures for all predesigned assays were used (see Table S5).

330

331

332 **Table S5. PCR reaction mixtures used**

Reagent	Volume for 1 reaction (µl)
Metabion Universal PCR Master Mix*	10
Respective predesigned assay	1
H_2O	7.35
cDNA diluted 15x	1.65

333

334 For the *18S rRNA* housekeeping gene the master mix contained the following components
335 (see Table S6):

336 **Table S6. Master mix for 18S rRNA real-time PCR**

Reagent	Volume for 1 reaction (µl)
Metabion Universal PCR Master Mix*	10.00
18SrRNA forward primer (25 pmol/µl)	0.12
18SrRNA reverse primer (25 pmol/µl)	0.12
18SrRNA probe (25 pmol/µl)	0.04
H_2O	6.07
Betain 5M (final conc. 0.5 M)†	2.00
cDNA diluted 15x	1.65

337 * mi-real-time Probe Master (ROX+), cat. #mi-E2006L was purchased from Metabion International
338 AG, Martinsried, Germany
339 † Betain 5M cat. #B0300-5VL was purchased from Sigma-Aldrich. (Steinheim, Germany)

340

341 The PCR cycling conditions (see Table S7) are listed below:

342 **Table S7. Cycling conditions for real-time PCR**

Step	Temperature	Time
Initial denaturation and polymerase activation	95 ˚C	2 min
40 cycles:		
Denaturation	95 ˚C	15 sec
Annealing and extension	60 ˚C	1 min

e) Real-time quantitative PCR gene expression analyses - data analysis

A specific fixed threshold adjustment for each target gene was applied (see Table S8) and cycle threshold (C_t) values were obtained using StepOne Software v2.2.2 (ThermoFisher Scientific).

Table S8. Cycle threshold values for target genes

Gene	Threshold
EYS	0.05
LUZP2	0.03
ANO3	0.05
MUC15	0.025
SLC5A12	0.025
NOS1	0.025
NOS2	0.025
NOS3	0.025
GAPDH	0.03
18S rRNA	0.015

Data evaluation was conducted based on the comparative delta-delta C_t method.(24, 25) Delta C_t (ΔC_t) values were determined by subtracting the C_t of *18S rRNA* from the C_t of the target gene for each sample, respectively. The lower the ΔC_t values the higher the gene expression levels. Relative fold changes (RFC) in the expression were calculated for each transfected sample with regard to the non-treated cells.

OLR REFERENCES

E1. Fuchs O, Latzin P, Kuehni CE, Frey U. Cohort profile: the Bern Infant Lung Development Cohort. International journal of epidemiology. 2012;41(2):366-76.

E2. Fuchs O, Latzin P, Thamrin C, Stern G, Frischknecht P, Singer F, et al. Normative data for lung function and exhaled nitric oxide in unsedated healthy infants. Eur Respir J. 2011;37(5):1208-16.

E3. Bates JH, Schmalisch G, Filbrun D, Stocks J. Tidal breath analysis for infant pulmonary function testing. ERS/ATS Task Force on Standards for Infant Respiratory Function Testing. European Respiratory Society/American Thoracic Society. Eur Respir J. 2000;16(6):1180-92.

E4. Frey U, Stocks J, Coates A, Sly P, Bates J. Specifications for equipment used for infant pulmonary function testing. ERS/ATS Task Force on Standards for Infant Respiratory Function Testing. European Respiratory Society/ American Thoracic Society. Eur Respir J. 2000;16(4):731-40.

E5. Frey U, Stocks J, Sly P, Bates J. Specification for signal processing and data handling used for infant pulmonary function testing. ERS/ATS Task Force on Standards for Infant Respiratory Function Testing. European Respiratory Society/American Thoracic Society. Eur Respir J. 2000;16(5):1016-22.

E6. American Thoracic S, European Respiratory S. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. Am J Respir Crit Care Med. 2005;171(8):912-30.

E7. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. American journal of human genetics. 2011;88(1):76-82.

E8. R Development Core Team. R Foundation for Statistical Computing. Vienna, Austria2006.

E9. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007;23(10):1294-6.

E10. Salam MT, Bastain TM, Rappaport EB, Islam T, Berhane K, Gauderman WJ, et al. Genetic variations in nitric oxide synthase and arginase influence exhaled nitric oxide levels in children. Allergy. 2011;66(3):412-9.

E11. van der Valk RJ, Duijts L, Timpson NJ, Salam MT, Standl M, Curtin JA, et al. Fraction of exhaled nitric oxide values in childhood are associated with 17q11.2-q12 and 17q12-q21 variants. J Allergy Clin Immunol. 2014;134(1):46-55.

E12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics. 2007;81(3):559-75.

E13. Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491(7422):56-65.

E14. Grundberg E, Small KS, Hedman AK, Nica AC, Buil A, Keildson S, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nature genetics. 2012;44(10):1084-9.

406 E15. Ma B, Wilker EH, Willis-Owen SA, Byun HM, Wong KC, Motta V, et al. Predicting
407 DNA methylation level across human tissues. Nucleic acids research. 2014;42(6):3515-
408 28.

409 E16. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M. Mapping complex disease
410 traits with global gene expression. Nature reviews Genetics. 2009;10(3):184-94.

411 E17. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al.
412 Systematic identification of trans eQTLs as putative drivers of known disease
413 associations. Nature genetics. 2013;45(10):1238-43.

414 E18. Ding J, Gudjonsson JE, Liang L, Stuart PE, Li Y, Chen W, et al. Gene expression in
415 skin and lymphoblastoid cells: Refined statistical method reveals extensive overlap in
416 cis-eQTL signals. American journal of human genetics. 2010;87(6):779-89.

417 E19. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot
418 analysis: multitissue gene regulation in humans. Science. 2015;348(6235):648-60.

419 E20. Roux J, Rosikiewicz M, Robinson-Rechavi M. What to compare and how: Comparative
420 transcriptomics for Evo-Devo. Journal of experimental zoology Part B, Molecular and
421 developmental evolution. 2015;324(4):372-82.

422 E21. Nica AC, Montgomery SB, Dimas AS, Stranger BE, Beazley C, Barroso I, et al.
423 Candidate causal regulatory effects by integration of expression QTLs with complex
424 trait genetic associations. PLoS genetics. 2010;6(4):e1000895.

425 E22. Dimas AS, Nica AC, Montgomery SB, Stranger BE, Raj T, Buil A, et al. Sex-biased
426 genetic effects on gene regulation in humans. Genome research. 2012;22(12):2368-75.

427 E23. Ziebarth JD, Bhattacharya A, Cui Y. CTCFBSDB 2.0: a database for CTCF-binding
428 sites and genome organization. Nucleic acids research. 2013;41(Database issue):D188-
429 94.

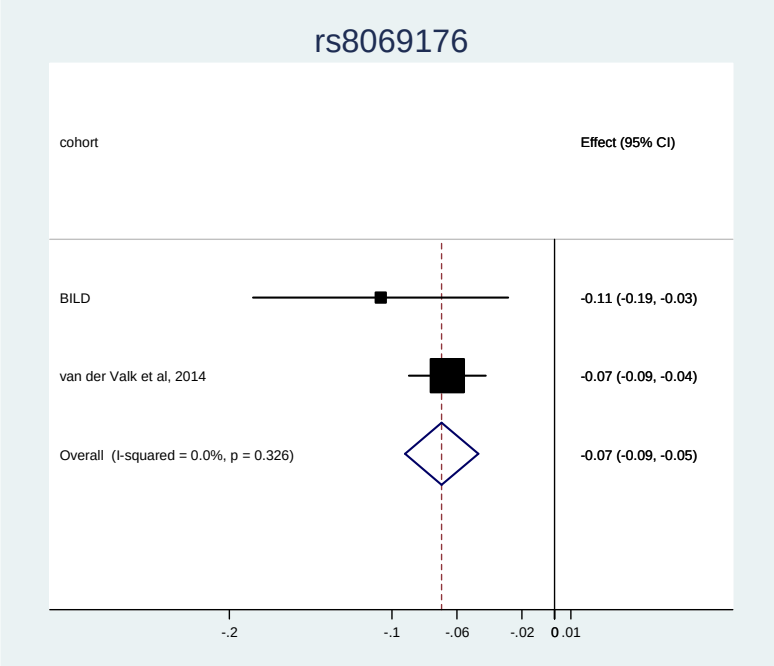
430 E24. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time
431 quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402-8.

432 E25. Toncheva AA, Potaczek DP, Schedel M, Gersting SW, Michel S, Krajnov N, et al.
433 Childhood asthma is associated with mutations and gene expression differences of
434 ORMDL genes that can interact. Allergy. 2015;70(10):1288-99.

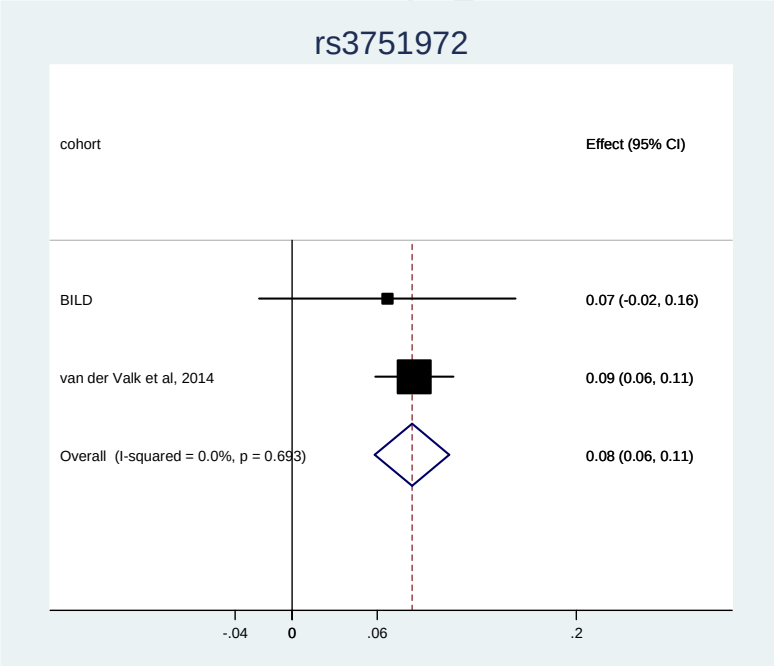
OLR FIGURES

Figure E1, A-D. Forest plots for associations of SNPs known to be associated with eNO in older children for data from the BILD cohort as well is their respective original publication and meta-analyses.

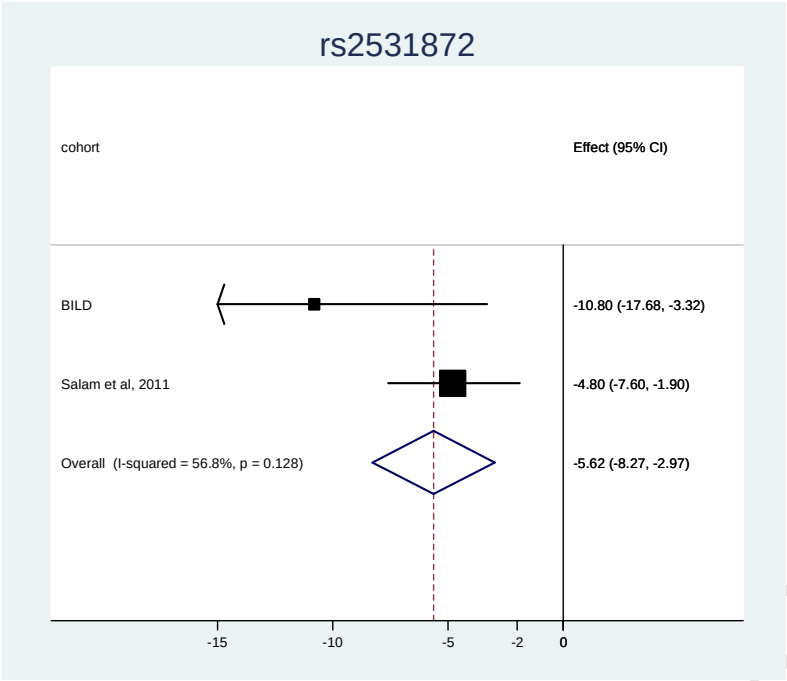
A.



B.



C.



D.

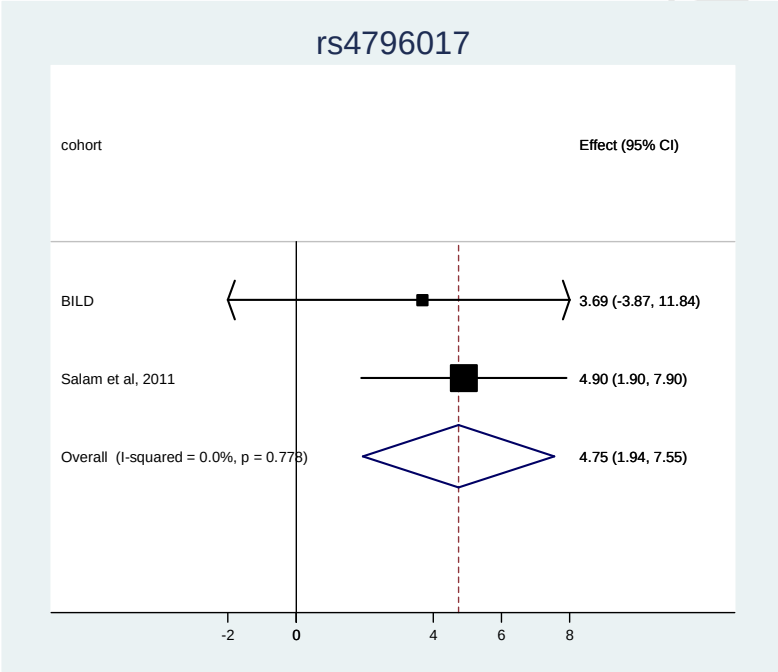
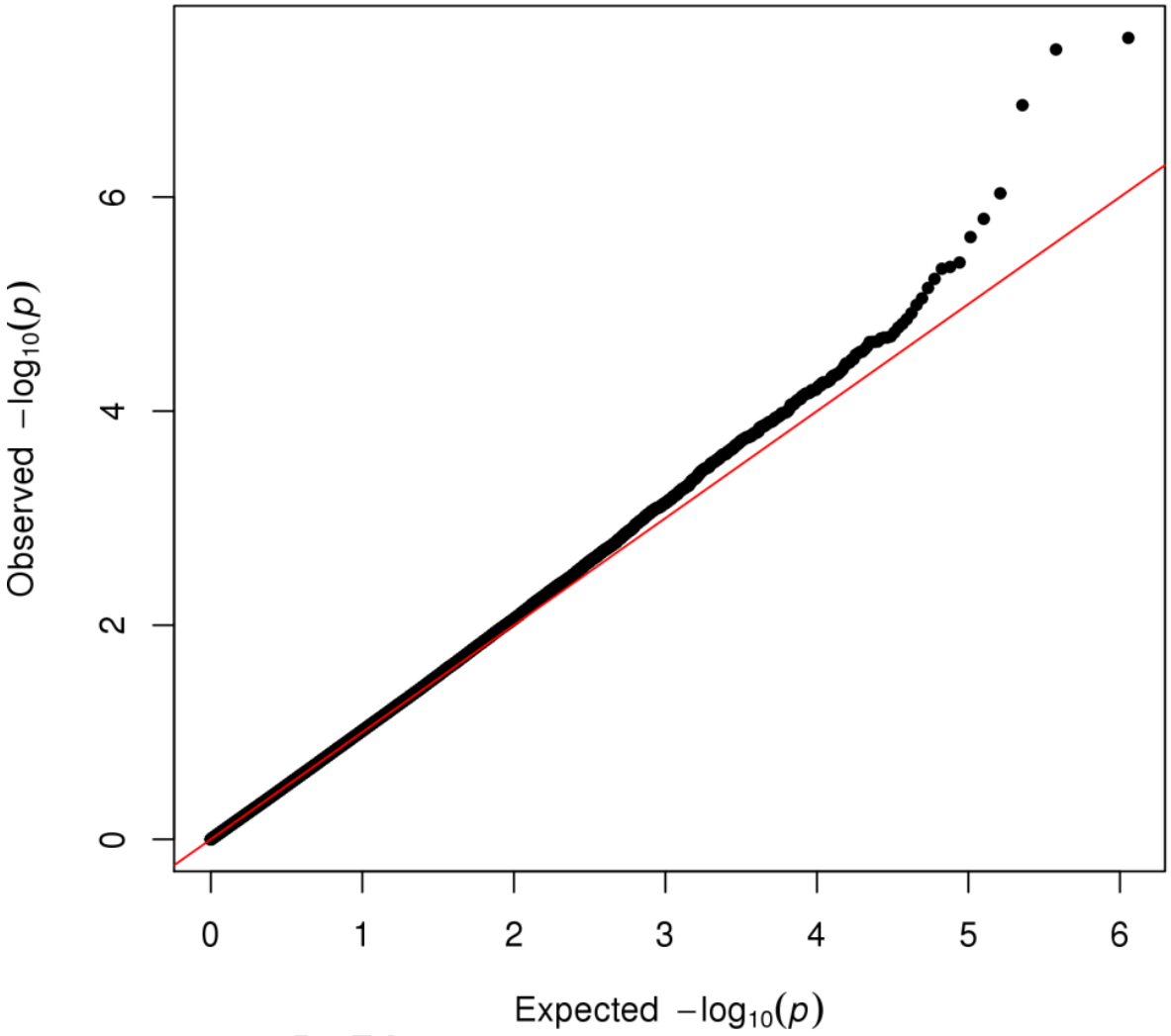


Figure E2, A and B. QQ Plots of 567,864 SNPs for n=229 study participants for eNO (A) and V'NO (B), final models

A.



B.

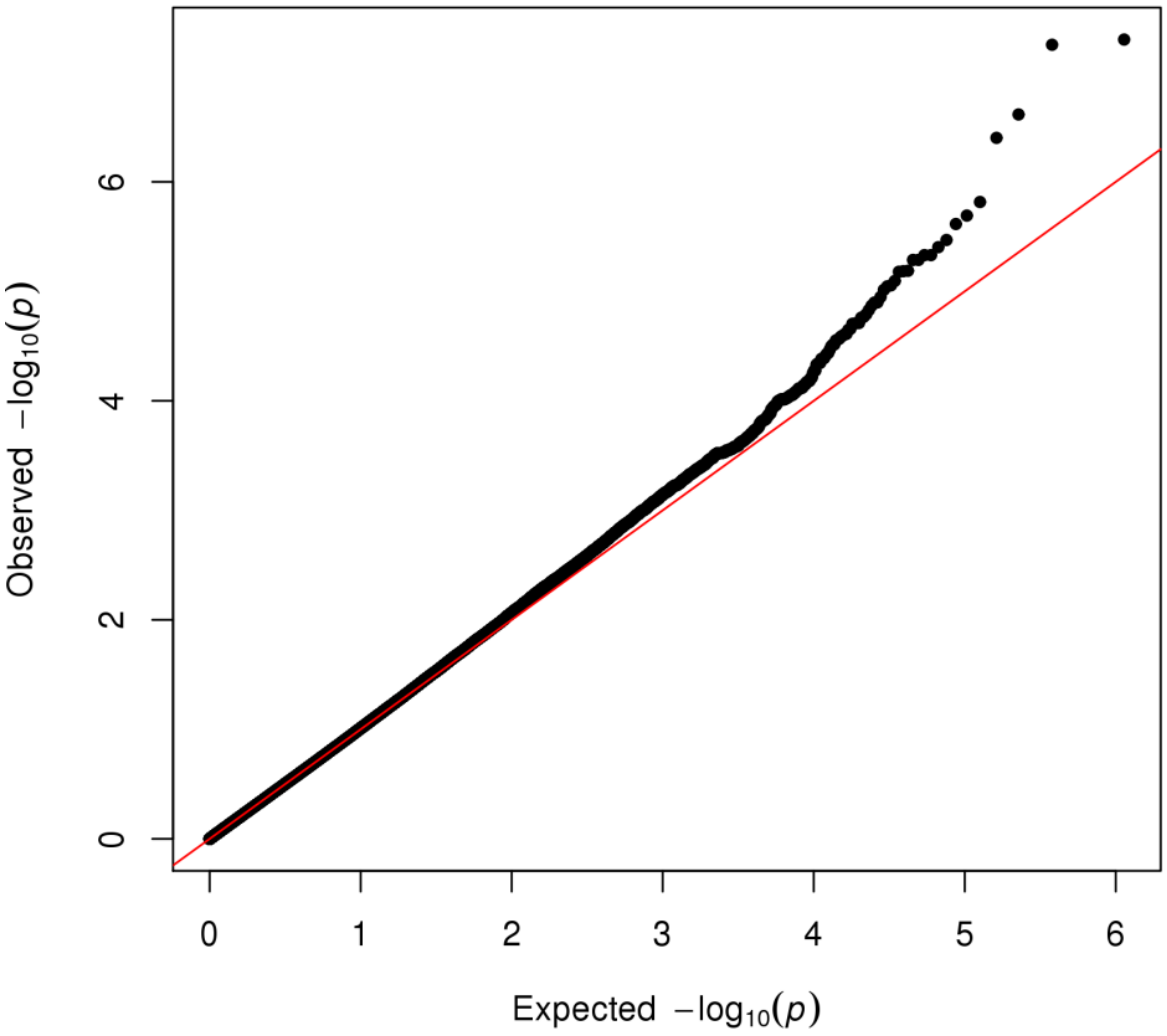
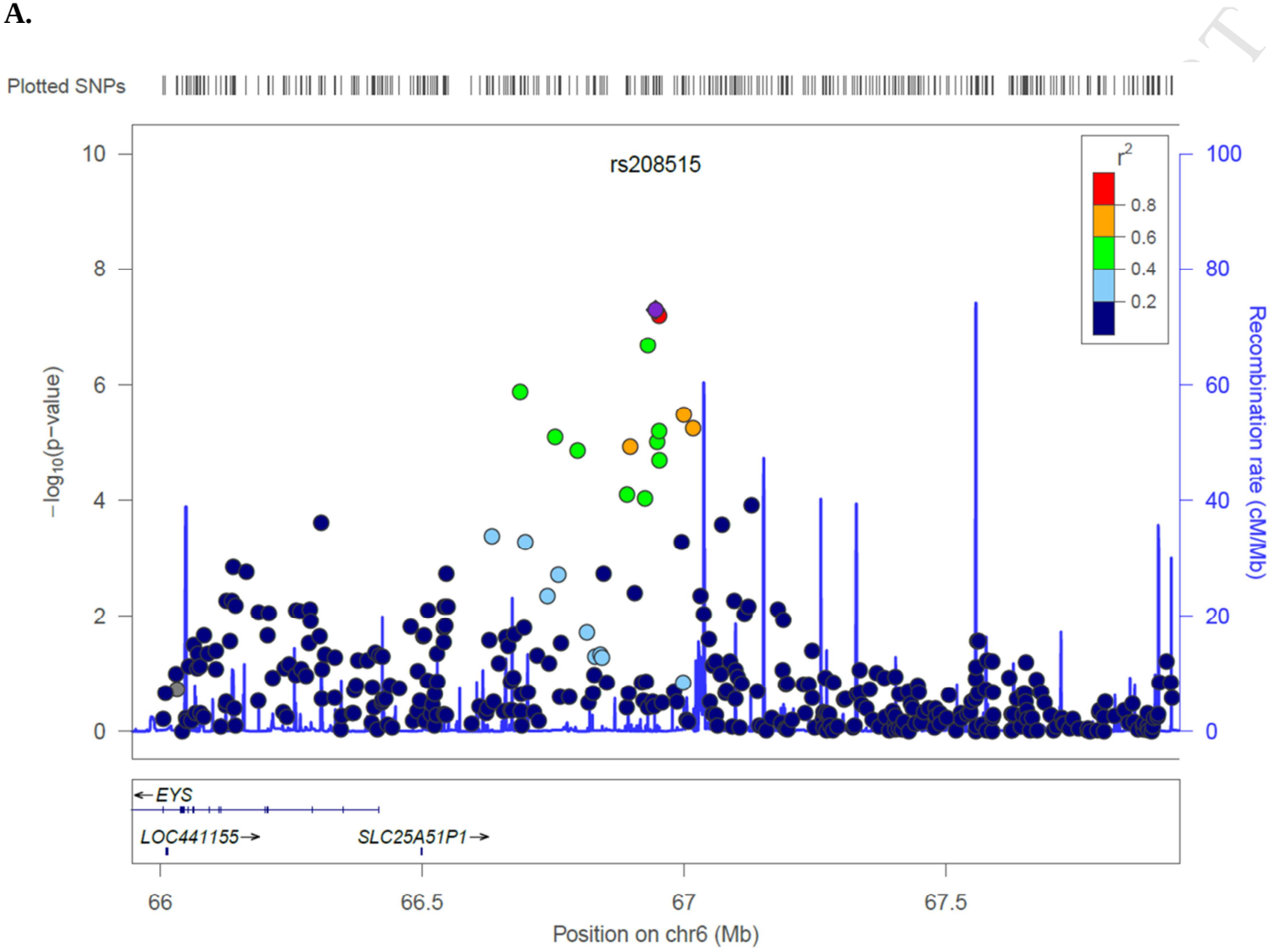


Figure E3, A and B: Regional plots of rs208515 at 6q12 (a) and of rs1441519 at 11p14 (b) for outcome eNO

A.



B.

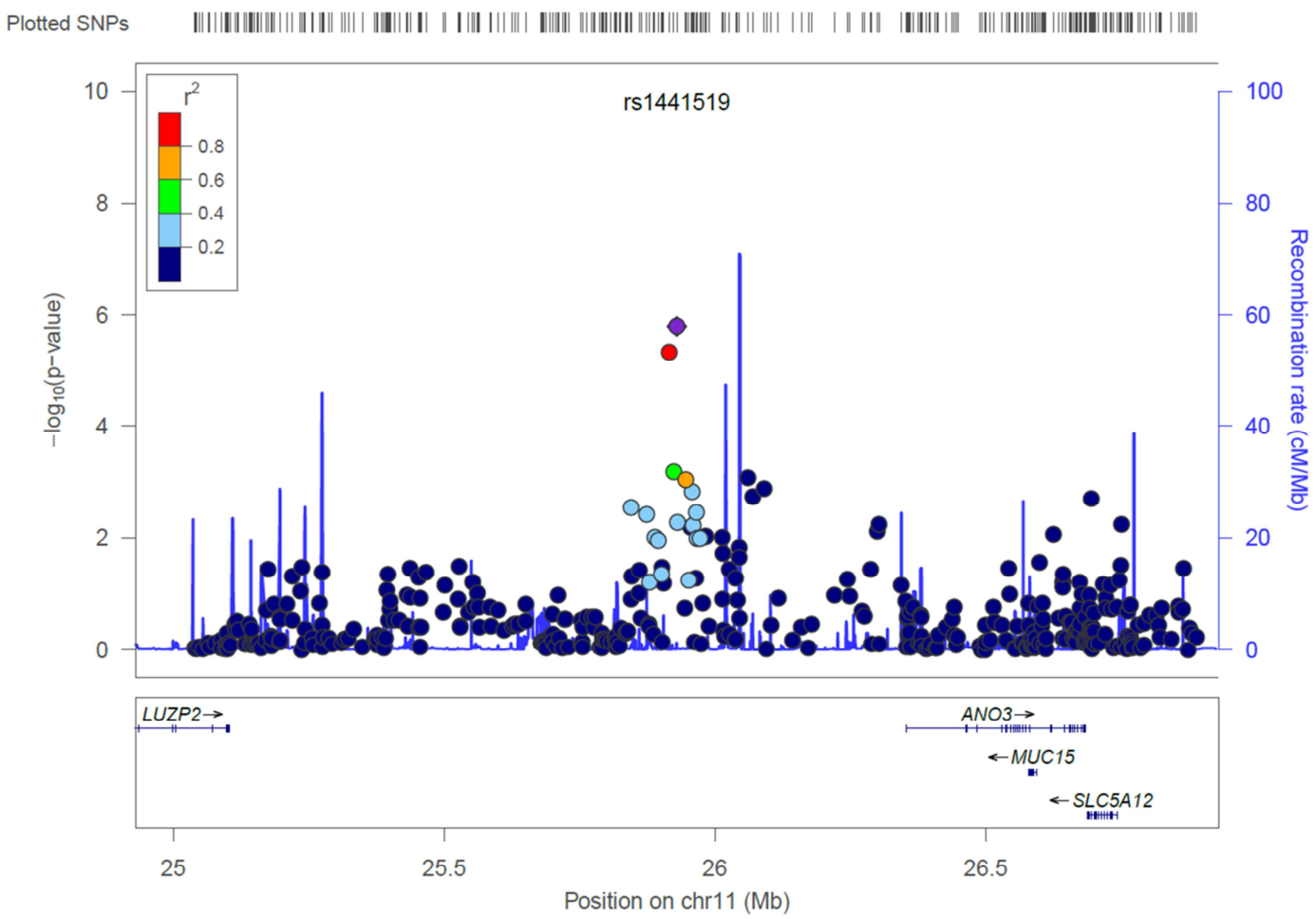
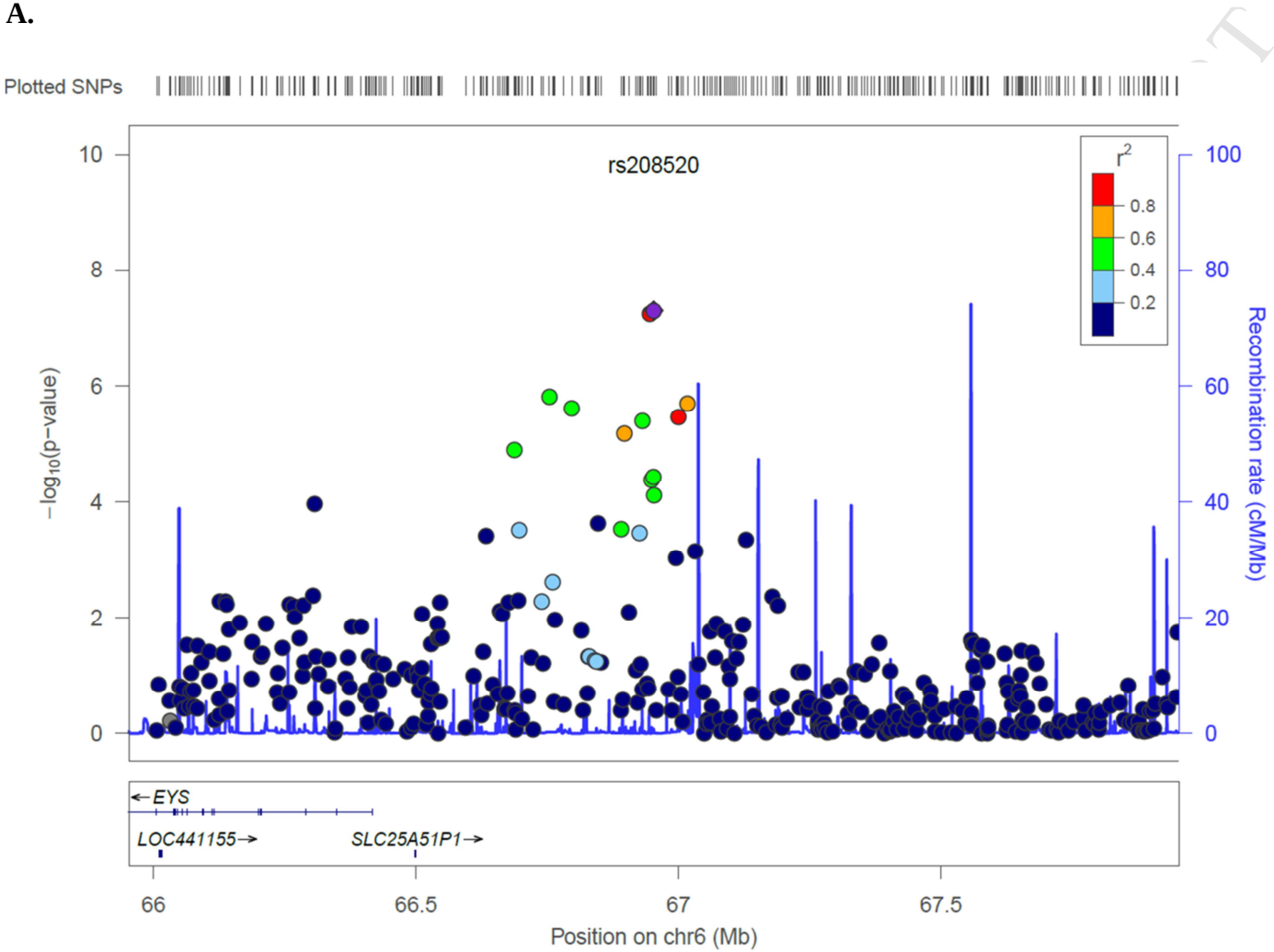
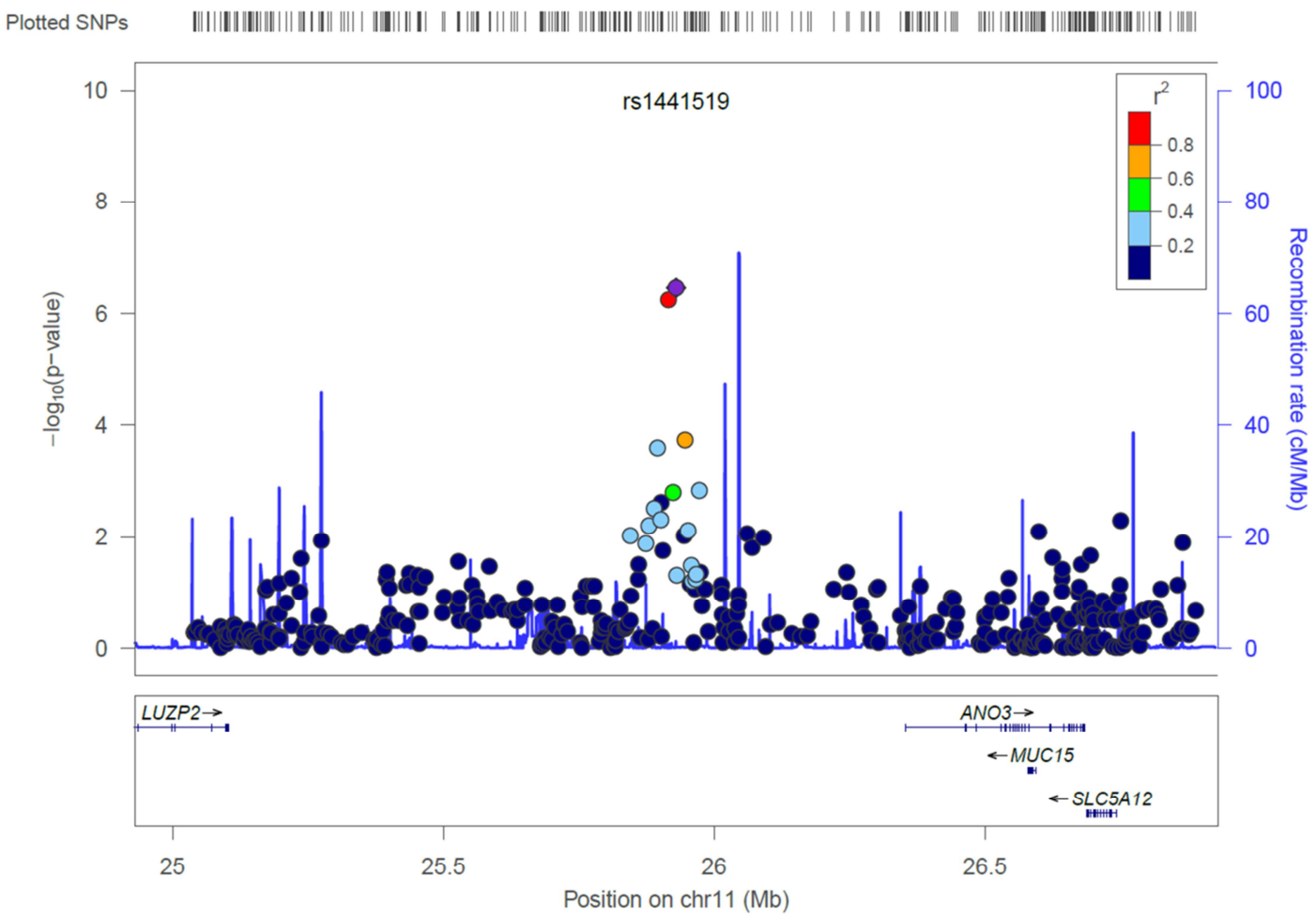

Plotted SNPs
rs1441519
r2
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
LUZP2
ANO3
MUC15
SLC5A12
Position on chr11 (Mb)

Figure E4, A and B: Regional plots of rs208520 at 6q12 (a) and of rs1441519 at 11p14 (b) for outcome V'NO

A.



B.


Plotted SNPs
rs1441519
r2
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
LUZP2
ANO3
MUC15
SLC5A12
Position on chr11 (Mb)

OLR FIGURE LEGENDS

Figure E1, A-D. Forest plots for associations of SNPs known to be associated with eNO in older children for data from the BILD cohort as well is their respective original publication and meta-analyses.

(A) Original SNP rs8069176 from van der Valk et al. 2014,(11) minor allele (r^2=1.0 with rs2305480); **(B)** original SNP rs3751972 from van der Valk et al. 2014,(11) minor allele; **(C)** original SNP rs2531872 from Salam et al. 2011,(10) minor allele (r^2=0.97 with rs2531864); **(D)** original SNP rs4796017 from Salam et al. 2011,(10) minor allele (r^2=0.86 with rs4569344). 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, (e)NO – (exhaled) nitric oxide, BILD – Basel/Bern Infant Lung Development (Cohort), rs – reference SNP cluster code, p – p-value, SNP – single nucleotid polymorphism.

Figure E2, A and B. QQ plots of 567,864 SNPs for n=229 study participants for eNO (A) and V'NO (B), final models.

Black dots represent observed against expected p-values for SNPs in the final models for each outcome, eNO and V'NO, adjusting for PC1 as well as known and possible confounders. These were study length, maternal smoking during pregnancy, maternal history of atopy, age at measurement, as well as V'E for eNO, and study weight, maternal smoking during pregnancy, maternal history of atopy, age at measurement, and V'E for V'NO in the final models for either outcome. The red line represents the observed against expected p-values under the null distribution.

Abbreviations: (e)NO – (exhaled) nitric oxide, V'NO – NO output, PC – principal component, QQ – quantile-quantile, SNP – single nucleotid polymorphism.

Figure E3, A and B: Regional plots of rs208515/rs208520 at 6q12 (a) and of rs1441519 at 11p14 (b) for outcome eNO.

Figures display association plots of the 6q12 (a) and 11p14 (b) regions for the outcome eNO. For each region, dots represent SNPs plotted with their p-values as ($-\log_{10}$, y-axes on the left) as functions of their genomic position (x-axes). Furthermore, recombination rates from the 1000 Genomes Project (CEU population, www.1000genomes.org) are plotted (y-axes on the right, in cM/Mb) in order to display the local LD structures around the associated SNPs (purple diamonds) for each locus as well as correlated proxies (using scales from blue to red, r^2=0 to 1). The plots were generated using LocusZoom (http://csg.sph.umich.edu/locuszoom/). Abbreviations: *ANO3* – human *anoctamin 3* gene locus (including *MUC15*, the human *mucin 15* gene locus), CEU - Utah residents with ancestry from northern and western Europe, cM – centiMorgan, (e)NO – (exhaled) nitric oxide, *EYS* – human *eyes shut* homolog gene locus, LD – linkage disequilibrium, Mb – Megabase, *MUC15* – human *mucin 15* gene locus, rs – reference SNP cluster code, SNP – single nucleotid polymorphism.

Figure E4, A and B: Regional plots of rs208515 at 6q12 (a) and of rs1441519 at 11p14 (b) for outcome V'NO

Figures display association plots of the 6q12 (a) and 11p14 (b) regions for the outcome V'NO. For each region, dots represent SNPs plotted with their p-values as ($-\log_{10}$, y-axes on the left) as functions of their genomic position (x-axes). Furthermore, recombination rates from the 1000 Genomes Project (CEU population, www.1000genomes.org) are plotted (y-axes on the right, in cM/Mb) in order to display the local LD structures around the associated SNPs (purple diamonds) for each locus as well as correlated proxies (using scales from blue

to red, r^2=0 to 1). The plots were generated using LocusZoom (http://csg.sph.umich.edu/locuszoom/).

Abbreviations: *ANO3* – human *anoctamin 3* gene locus (including *MUC15*, the human *mucin 15* gene locus), CEU - Utah residents with ancestry from northern and western Europe, cM – centiMorgan, *EYS* – human *eyes shut* homolog gene locus, LD – linkage disequilibrium, Mb – Megabase, *MUC15* – human *mucin 15* gene locus, rs – reference SNP cluster code, SNP – single nucleotid polymorphism, V'NO – NO output.

OLR TABLES

Table E1: Associations of SNPs known to be associated with eNO in older children in the BILD cohort, in the respective original publication and respective meta-analyses

			Results for published SNP in BILD cohort[a]					Results for SNP in published original study					Meta-analysis				
SNP	**Chr**	**Gene**	**MA**	**MAF**	**log(β[b])**	**se[b]**	**p-value**	**MA**	**MAF**	**log(β)[b]**	**se[b]**	**p-value**	**log(β)[b]**	**95%CI[b]**	**p-value**	**I^2**	**p_{het}**
van der Valk et al. 2014(11)																	
rs3751972	17	LYRM9	G	0.23	0.067	0.046	0.145	C	0.25	0.086	0.014	2.0×10^{-10}	0.084	0.058; 0.111	3.0×10^{-10}	0	0.693
proxy for rs8069176: rs2305480 (r2=1.00, D'=1.00)	17	ZPBP2-GSDMB	A	0.44	-0.11	0.017	0.040	A	0.43	-0.066	0.012	1.9×10^{-8}	-0.07	-0.09; -0.047	1.6×10^{-9}	0	0.326
Salam et al. 2011(10)																	
SNP	**Chr**	**Gene**	**MA**	**MAF**	**β[c]**	**95%CI[c]**	**p-value**	**MA[d]**	**MAF[d]**	**β[c]**	**95%CI[c]**	**p-value**	**β[c]**	**95%CI[c]**	**p-value**	**I^2**	**p_{het}**
the proxy for rs2531872: rs2531864 (r2=0.97, D'=0.98)	17	5' of NOS2	A	0.35	-10.8	-17.68 ; -3.32	0.0053			-4.8	-7.6, -1.9	0.002	-5.6	-8.27; -2.97	3.2×10^{-5}	56.8%	0.128
proxy for rs4796017: rs4569344 (r2=0.86, D'=0.94)	17	3' of NOS2	C	0.41	3.69	-3.87 ; 11.84	0.349			4.9	1.9; 7.9	0.001	4.8	1.93; 7.6	9.0×10^{-4}	0	0.778

Abbreviations: β - estimate, Chr – chromosome, (95%)CI – (95%) confidence interval, I^2 – heterogeneity, log – natural logarithm, *LYRM9* - human LYR motif containing 9 gene locus, MA – minor allele, MAF – minor allele frequency, *NOS2* – nitric oxide synthase 2 gene locus, se – standard error, p_{het} – p-value for heterogeneity, rs – reference SNP cluster code, SNP – single nucleotid polymorphism, *ZDBP2-GSDMB* – human D site albumin promoter binding protein b-gasdermin B gene locus.

[a]Data obtained from linear regression adjusted for minute ventilation, postnatal age, study length, maternal smoking during pregnancy, maternal history of atopy, and PC1. [b] Results are expressed as a difference in natural log-transformed values per minor allele. [c]Results are expressed as a

percent difference [100 × (exponential of mean -1)] per minor allele with 95% confidence interval (CI). [d]Data on allele and allele frequency (it is shown only by ethnicity) are not reported.

Table E2A: Main associations with eNO values ($p<1x10^{-5}$), additive model

					Adjusted model*		Unadjusted model	
SNPs	**Chr, region**	**Position, bp**	**Effect allele**	**n**	**β, ppb**	**p-value**	**β, ppb**	**p-value**
rs208515	6q12	66945749	T	229	2.53	3.3×10^{-8}	2.36	1.7×10^{-6}
rs208520	6q12	66952828	G	229	2.56	4.2×10^{-8}	2.32	3.9×10^{-6}
rs6903678	6q12	66931208	T	227	2.23	1.4×10^{-7}	2.15	2.4×10^{-6}
rs10944868	6q12	66687508	C	229	2.41	9.2×10^{-7}	2.34	7.7×10^{-6}
rs1441519	11p14	25929959	T	229	-2.09	1.6×10^{-6}	-2.18	2.3×10^{-6}
rs12206488	6q12	66999917	T	229	2.29	2.4×10^{-6}	2.11	4.7×10^{-5}
rs12223678	11p14	25916225	T	229	-2.09	4.5×10^{-6}	-2.20	4.9×10^{-6}
rs1029400	6q12	66952385	T	228	2.05	4.7×10^{-6}	1.92	6.2×10^{-5}
rs6913344	6q12	66754285	T	229	2.03	5.8×10^{-6}	1.90	6.2×10^{-5}
rs6455090	6q12	66948909	C	228	2.02	7.1×10^{-6}	1.88	9.6×10^{-5}
rs208436	6q12	66897470	C	226	1.93	8.8×10^{-6}	1.84	7.3×10^{-5}
rs9453585	6q12	66796668	C	229	1.97	1.0×10^{-5}	1.86	7.9×10^{-5}

Most significant SNPs printed in bold.

Abbreviations: β – regression coefficient; bp – base pairs; chr – chromosome; (e)NO – (exhaled) nitric oxide; n – number of individuals; ppb – parts per billion; rs – reference SNP cluster code; SNP – single nucleotide polymorphism.

*Adjusted for minute ventilation, postnatal age, study length, maternal smoking during pregnancy, maternal history of atopy, and the factor explaining most of the variance during a principal component analysis (PC1, see text).

Table E2B: Main associations with V'NO values ($p<1x10^{-5}$), additive model

SNPs	Chr, region	Position, bp	effect allele	Adjusted model*			Unadjusted model	
				n	β, nL/s	p-value	β, nL/s	p-value
rs208520	6	66952828	G	229	0.11	5.0×10^{-8}	0.11	7.1×10^{-7}
rs208515	6	66945749	T	229	0.11	5.6×10^{-8}	0.11	7.6×10^{-7}
rs1441519	11	25929959	T	229	-0.10	2.4×10^{-7}	-0.09	3.4×10^{-5}
rs12223678	11	25916225	T	229	-0.10	4.0×10^{-7}	-0.10	9.6×10^{-6}
rs6913344	6	66754285	T	229	0.10	1.5×10^{-7}	0.11	1.3×10^{-7}
rs9345809	6	67017543	C	225	0.10	2.0×10^{-6}	0.10	1.3×10^{-5}
rs9453585	6	66796668	C	229	0.09	2.4×10^{-6}	0.11	2.4×10^{-7}
rs12206488	6	66999917	T	229	0.10	3.4×10^{-6}	0.10	2.1×10^{-5}
rs6903678	6	66931208	T	227	0.09	4.0×10^{-6}	0.09	1.5×10^{-5}
rs12669	19	44235535	A	229	-0.10	4.7×10^{-6}	-0.09	7.2×10^{-5}
rs9612	19	44235517	C	229	-0.10	4.7×10^{-6}	-0.09	7.2×10^{-5}
rs346527	19	44259361	C	229	-0.10	5.2×10^{-6}	-0.09	7.7×10^{-5}
rs346545	19	44256386	C	229	-0.10	5.2×10^{-6}	-0.09	7.7×10^{-5}
rs10149105	14	85615940	A	229	-0.10	6.5×10^{-6}	-0.09	1.7×10^{-4}
rs208436	6	66897470	C	226	0.09	6.6×10^{-6}	0.08	8.1×10^{-5}
rs1153996	5	104227705	A	229	-0.09	8.0×10^{-6}	-0.08	1.6×10^{-4}

Most significant SNPs printed in bold.

Abbreviations: β – regression coefficient; bp – base pairs; chr – chromosome; n – number of individuals; nL/s – nanoliter per second; rs – reference SNP cluster code; SNP – single nucleotide polymorphism; V'NO –nitric oxide output.

* Adjusted for minute ventilation, study weight, postnatal age, maternal smoking during pregnancy, maternal history of atopy, and the factor explaining most of the variance during a principal component analysis (PC1, see text).

Table E3. **Summary statistics for top SNPs**

							Genotype frequencies		
SNP	**Chr**	**A**	**B**	**n**	**Call rate**	**MAF**	**AA, n**	**AB, n**	**BB, n**	**p-value HWE**
rs1708611	5	C	T	226	0.99	0.34	106	87	33	0.04
rs1153996	5	G	A	229	1.00	0.34	106	89	34	0.04
rs6913344	6	G	T	229	1.00	0.40	85	107	37	0.78
rs9453585	6	T	C	229	1.00	0.40	84	107	38	0.68
rs208436	6	T	C	226	0.99	0.34	106	87	33	0.04
rs6903678	6	C	T	227	0.99	0.42	80	103	44	0.34
rs208515	6	C	T	229	1.00	0.28	123	84	22	0.19
rs208520	6	T	G	229	1.00	0.26	132	77	20	0.08
rs12206488	6	G	T	229	1.00	0.24	139	72	18	0.07
rs9345809	6	T	C	225	0.98	0.22	146	60	19	0.00
rs742715	6	T	C	229	1.00	0.27	121	93	15	0.74
rs10944868	6	T	C	229	1.00	0.29	114	98	17	0.63
rs9453585	6	T	C	229	1.00	0.40	84	107	38	0.68
rs6455090	6	T	C	228	1.00	0.40	81	110	37	1.00
rs1029400	6	C	T	228	1.00	0.40	81	110	37	1.00
rs12206488	6	G	T	229	1.00	0.24	139	72	18	0.07
rs9345809	6	T	C	225	0.98	0.22	146	60	19	0.00
rs10488529	7	C	T	229	1.00	0.12	173	55	1	0.22
rs12223678	11	C	T	229	1.00	0.43	71	118	40	0.50
rs1441519	11	G	T	229	1.00	0.47	66	113	50	0.89
rs10149105	14	G	A	229	1.00	0.27	122	91	16	1.00
rs8087255	18	C	A	228	1.00	0.09	191	35	2	0.67
rs9612	19	T	C	229	1.00	0.23	139	73	17	0.10
rs12669	19	G	A	229	1.00	0.23	139	73	17	0.10
rs346545	19	T	C	229	1.00	0.24	138	74	17	0.14
rs346527	19	T	C	229	1.00	0.24	138	74	17	0.14

Most significant SNPs printed in bold. Abbreviations: A – reference allele; B – effect allele, chr – chromosome, HWE – Hardy-Weinberg equilibrium, n – number of individuals, rs – reference SNP cluster code, SNP – single nucleotide polymorphism.

Table E4A. **Main associations with eNO values of top SNPs for first batch from 2011, additive model**

					Adjusted model*		Unadjusted model	
SNPs	**Chr, region**	**Position, bp**	**Effect allele**	**N**	**β, ppb**	**p-value**	**β, ppb**	**p-value**
rs208515	6q12	66945749	T	174	2.9	9.6×10^{-8}	2.9	8.1×10^{-7}
rs208520	6q12	66952828	G	174	3.0	4.2×10^{-8}	2.9	9.3×10^{-7}
rs1441519	11p14	25929959	T	174	-2.6	1.3×10^{-6}	-2.7	7.0×10^{-7}
rs12223678	11p14	25916225	T	174	-2.7	8.8×10^{-7}	-2.9	5.0×10^{-7}

Abbreviations: β – regression coefficient; bp – base pairs; chr – chromosome; (e)NO – (exhaled) nitric oxide; n – number of individuals; ppb – parts per billion; rs – reference SNP cluster code; SNP – single nucleotide polymorphism.
*Adjusted for minute ventilation, postnatal age, study length, maternal smoking during pregnancy, and maternal history of atopy.

Table E4B. **Main associations with eNO values of top SNPs for second batch from 2013, additive model**

					Adjusted model*		Unadjusted model	
SNPs	**Chr, region**	**Position, bp**	**Effect allele**	**N**	**β, ppb**	**p-value**	**β, ppb**	**p-value**
rs208515	6q12	66945749	T	55	1.1	0.258	0.8	0.378
rs208520	6q12	66952828	G	55	0.8	0.428	0.5	0.575
rs1441519	11p14	25929959	T	55	-1.0	0.190	-0.8	0.296
rs12223678	11p14	25916225	T	55	-0.7	0.384	-0.5	0.584

Abbreviations: β – regression coefficient; bp – base pairs; chr – chromosome; (e)NO – (exhaled) nitric oxide; n – number of individuals; ppb – parts per billion; rs – reference SNP cluster code; SNP – single nucleotide polymorphism.
*Adjusted for minute ventilation, postnatal age, study length, maternal history of atopy. Due to the small number of smoking mothers in this batch, we did not adjust for maternal smoking during pregnancy for this subanalysis.

Table E5. **Associations with eNO values of top SNPs after permutation**

SNPs	Chr, region	Empirical p-value 1	Empirical p-value 2
rs208515	6q12	$1{,}0\times10^{-4}$	0.004
rs208520	6q12	$1{,}0\times10^{-4}$	0.005
rs1441519	11p14	$1{,}0\times10^{-4}$	0.071
rs12223678	11p14	$1{,}0\times10^{-4}$	0.162

Abbreviations: chr – chromosome; (e)NO – (exhaled) nitric oxide; rs – reference SNP cluster code; SNP – single nucleotide polymorphism.
Empirical p-value 1: p-value for pointwise estimate of an individual SNPs significance
Empirical p-value 2: p-value controlling for tests of other SNPs, thus controlling the so-called familywise error rate, reflecting the chance of such a test statistic given the number of individual tests performed

Table E6A. Distribution of respiratory symptoms in final study population and across risk factors

	Mean (SD)	Interquartile range	Range
Respiratory symptoms, weeks	5.4 (4.7)	2-8	0-21
Maternal atopy, weeks	5.3 (4.7)	1-8	0-17
No maternal atopy, weeks	5.5 (4.6)	2-8	0-21
Maternal smoking, weeks	4.6 (3.8)	2-8	0-14
No maternal smoking, weeks	5.5 (4.7)	2-8	0-21

Abbreviations: SD – Standard deviation. A total of n=20/229 infants did not have any weeks with respiratory symptoms during the first year of life.

Table E6B. Distribution of respiratory symptoms in final study population by genotype

SNP	Mean (SD)			aIRR*	p-value
	Homozygous non-risk genotype	Heterozygous genotype	Homozygous risk genotype		
rs208515	6.0 (5.2)	5.1 (4.0)	3.5 (3.1)	0.81	0.011
rs208520	5.9 (5.1)	5.1 (3.9)	3.5 (3.3)	0.81	0.015
rs1441519	4.7 (4.7)	5.7 (4.8)	5.8 (4.3)	1.05	0.513
rs12223678	4.6 (4.5)	5.9 (4.8)	5.6 (4.3)	1.09	0.293

Abbreviations: aIRR – adjusted incidence risk ratio; eNO – exhaled nitric oxide; rs – reference SNP cluster code; SD – standard deviation; SNP – single nucleotide polymorphism.
*Adjusted for sex, older siblings, nursery care, maternal smoking, maternal atopy, minute ventilation and eNO

Table E7A. Fine mapping of eNO association signals by regional LD analysis for rs208515 (6q12)

rs number	position on chr.6 (bp)	r^2 with rs208515	overlap with rs208520		rs number (continued)	position on chr.6 (bp) (continued)	r^2 with rs208515 (continued)	overlap with rs208520 (continued)
rs1563929	66.702.108	0.80	no		rs1776365	66.834.474	0.83	no
rs12202805	66.703.903	0.80	no		rs3899423	66.835.285	0.83	no
rs12209753	66.703.951	0.80	no		rs3843513	66.836.797	0.83	no
rs12213561	66.705.646	0.80	no		rs3909071	66.836.947	0.83	no
rs12192851	66.705.862	0.80	no		rs851464	66.839.802	0.83	no
rs10944872	66.708.239	0.80	no		rs851465	66.840.293	0.83	no
rs12208310	66.730.108	0.81	no		rs851466	66.840.476	0.83	no
rs12202913	66.734.198	0.81	no		rs10080491	66.847.597	0.83	no
rs12213581	66.735.818	0.81	no		rs1304672	66.847.973	0.83	no
rs12213575	66.735.989	0.81	no		rs4710566	66.848.538	0.83	no
rs10455594	66.737.894	0.81	no		rs866452	66.849.936	0.83	no
rs10455595	66.738.034	0.81	no		rs1100981	66.852.052	0.83	no
rs12204635	66.740.779	0.81	no		rs1233101	66.855.505	0.83	no
rs2126120	66.742.754	0.81	no		rs1233103	66.855.943	0.83	no
rs10455194	66.743.356	0.81	no		rs7747492	66.857.788	0.83	no
rs12194057	66.745.616	0.81	no		rs7739253	66.861.052	0.82	no
rs12201219	66.748.862	0.81	no		rs6932475	66.863.155	0.82	no
rs12199537	66.750.478	0.81	no		rs2351880	66.863.707	0.82	no
rs4386800	66.755.242	0.81	no		rs7740835	66.865.548	0.82	no
rs4618506	66.756.889	0.81	no		rs2208442	66.867.004	0.82	no
rs12189780	66.759.619	0.81	no		rs6937856	66.867.468	0.82	no
rs12202401	66.760.544	0.81	no		rs2797693	66.868.985	0.82	no
rs1351867	66.760.995	0.81	no		rs1885100	66.871.047	0.82	no
rs12213928	66.762.978	0.81	no		rs9351549	66.871.851	0.88	no
rs207740	66.764.132	0.81	no		rs851592	66.875.175	0.88	no
rs176289	66.766.522	0.81	no		rs7753158	66.875.633	0.88	no
rs207077	66.769.855	0.81	no		rs851593	66.875.783	0.88	no
rs207079	66.770.462	0.81	no		rs851594	66.876.695	0.88	no
rs207081	66.773.742	0.81	no		rs1100985	66.878.688	0.88	no
rs207082	66.773.796	0.81	no		rs1100988	66.880.008	0.88	no
rs207083	66.774.216	0.81	no		rs1100989	66.880.532	0.88	no
rs150263	66.777.620	0.81	no		rs1342959	66.883.329	0.90	no
rs207090	66.783.963	0.81	no		rs1342960	66.883.514	0.90	no
rs207119	66.786.408	0.81	no		rs4710314	66.883.656	0.90	no
rs207118	66.788.948	0.81	no		rs2224422	66.883.993	0.90	no
rs207114	66.791.703	0.81	no		rs2153941	66.885.521	0.90	no
rs207112	66.792.092	0.81	no		rs4710315	66.887.594	0.90	no
rs207111	66.792.513	0.81	no		rs1935894	66.891.434	0.90	no
rs207108	66.794.096	0.81	no		rs851600	66.893.754	0.90	no
rs207107	66.794.385	0.81	no		rs851601	66.893.853	0.90	no
rs207103	66.795.568	0.81	no		rs208446	66.903.515	0.89	no
rs207102	66.795.671	0.81	no		rs208447	66.903.905	0.89	no
rs207099	66.796.373	0.81	no		rs208451	66.906.155	0.89	no
rs207096	66.797.287	0.81	no		rs208452	66.906.513	0.85	no
rs207094	66.799.009	0.81	no		rs208453	66.906.841	0.85	no
rs207091	66.799.847	0.81	no		rs208454	66.907.783	0.89	no
rs704519	66.801.136	0.81	no		rs208456	66.907.942	0.89	no
rs851467	66.802.721	0.81	no		rs208459	66.911.055	0.87	no
rs3846803	66.805.743	0.81	no		rs208463	66.913.962	0.87	no
rs7738469	66.807.446	0.81	no		rs208467	66.914.837	0.97	yes
rs851458	66.813.279	0.81	no		rs208469	66.915.891	0.97	yes
rs851459	66.813.989	0.81	no		rs208470	66.916.402	0.97	yes
rs3846808	66.818.604	0.81	no		rs208471	66.916.749	0.97	yes
rs9345764	66.820.597	0.81	no		rs208473	66.917.265	0.97	yes
rs1776361	66.822.270	0.83	no		rs208481	66.922.636	0.97	yes
rs4710311	66.823.223	0.81	no		rs208485	66.925.654	0.97	yes
rs4710564	66.823.478	0.81	no		rs208486	66.926.255	0.97	yes
rs2144081	66.824.972	0.83	no		rs208488	66.926.902	0.97	yes
rs1318604	66.825.278	0.83	no		rs208489	66.927.584	0.97	yes
rs851468	66.825.875	0.83	no		rs208491	66.928.459	0.97	yes
rs1100950	66.827.285	0.83	no		rs208493	66.930.162	0.97	yes
rs7764570	66.827.600	0.81	no		rs208496	66.930.786	0.97	yes

rs7766144	66.827.795	0.83	no	rs208498	66.931.560	0.97	yes

Table E7A. Fine mapping of eNO association signals by regional LD analysis for rs208515 (6q12) - continued

rs number (continued)	position on chr.6 (bp) (continued)	r^2 with rs208515 (continued)	overlap with rs208520 (continued)	rs number (continued)	position on chr.6 (bp) (continued)	r^2 with rs208515 (continued)	overlap with rs208520 (continued)
rs208499	66.932.146	0.97	yes	rs12190316	66.973.365	0.82	yes
rs208509	66.941.706	0.97	yes	rs12194996	66.973.366	0.82	yes
rs208512	66.943.869	1.00	yes	rs72882079	66.973.878	0.82	yes
rs208514	66.945.520	1.00	yes	rs72882086	66.976.671	0.82	yes
rs208515	**66.945.749**	**1.00**	**yes**	rs12189683	66.977.716	0.82	yes
rs208516	66.946.119	1.00	yes	rs12191403	66.978.137	0.82	yes
rs208517	66.947.040	1.00	yes	rs12193164	66.978.962	0.82	yes
rs208518	66.948.501	1.00	yes	rs28684412	66.979.430	0.82	yes
rs208520	**66.952.828**	**0.82**	**yes**	rs72882100	66.981.495	0.82	yes
rs208521	66.953.625	0.82	yes	rs72882102	66.983.408	0.82	yes
rs208523	66.954.195	0.82	yes	rs12190073	66.983.740	0.82	yes
rs421337	66.955.907	0.82	yes	rs12191598	66.983.962	0.82	yes
rs208527	66.957.209	0.82	yes	rs12212063	66.985.148	0.82	yes
rs208529	66.957.549	0.82	yes	rs12192261	66.985.268	0.82	yes
rs208530	66.958.350	0.82	yes	rs12195505	66.986.279	0.82	yes
rs208531	66.960.439	0.82	yes	rs12216109	66.987.488	0.82	yes
rs208532	66.961.280	0.82	yes	rs12199876	66.989.469	0.85	yes
rs208535	66.962.286	0.82	yes	rs12193077	66.989.548	0.82	yes
rs208536	66.963.440	0.82	yes	rs9453671	66.991.208	0.82	yes
rs208537	66.963.544	0.82	yes	rs9453672	66.991.431	0.82	yes
rs208538	66.964.492	0.82	yes	rs72884017	66.992.513	0.82	yes
rs72880045	66.965.232	0.82	yes	rs28773210	66.993.566	0.82	yes
rs9453656	66.965.407	0.82	yes	rs72884025	66.994.185	0.82	yes
rs12190773	66.968.194	0.82	yes	rs72884027	66.995.044	0.82	yes
rs72882065	66.968.566	0.82	yes	rs72884033	66.995.747	0.82	yes
rs12190187	66.973.207	0.82	yes	rs12209225	66.996.443	0.82	yes

LD with surrounding SNPs of identified variants was determined based on the 1,000 genomes catalog and only SNPs with r2≥0.8 are displayed (CEU population n=99; SNP inclusion parameters: MAF≥0.05 and HWE p-value≥0.0001; regions: chr6:64,419,876-67,952,828 and chr11:24,508,516-26,929,959 [GRCh37.p13]). Abbreviations: bp – base pair, CEU - Utah residents with ancestry from northern and western Europe, chr – chromosome, HWE – Hardy-Weinberg equilibrium, eNO – exhaled nitric oxide, LD – linkage disequilibrium, MAF – minor allele frequency, rs – reference SNP cluster code, SNP – single nucleotide polymorphism.

Table E7B. Fine mapping of eNO association signals by regional LD analysis for rs208520 (6q12)

rs number (continued)	position on chr.6 (bp) (continued)	r^2 with rs208520 (continued)	overlap with rs208515 (continued)
rs208467	66.914.837	0.81	yes
rs208469	66.915.891	0.81	yes
rs208470	66.916.402	0.81	yes
rs208471	66.916.749	0.81	yes
rs208473	66.917.265	0.81	yes
rs208481	66.922.636	0.81	yes
rs208485	66.925.654	0.81	yes
rs208486	66.926.255	0.81	yes
rs208488	66.926.902	0.81	yes
rs208489	66.927.584	0.81	yes
rs208491	66.928.459	0.81	yes
rs208493	66.930.162	0.81	yes
rs208496	66.930.786	0.81	yes
rs208498	66.931.560	0.81	yes
rs208499	66.932.146	0.81	yes
rs208509	66.941.706	0.81	yes
rs208512	66.943.869	0.82	yes
rs208514	66.945.520	0.82	yes
rs208515	**66.945.749**	**0.82**	**yes**
rs208516	66.946.119	0.82	yes
rs208517	66.947.040	0.82	yes
rs208518	66.948.501	0.82	yes
rs208520	**66.952.828**	**1.00**	**yes**
rs208521	66.953.625	1.00	yes
rs208523	66.954.195	1.00	yes
rs421337	66.955.907	1.00	yes
rs208527	66.957.209	1.00	yes
rs208529	66.957.549	1.00	yes
rs208530	66.958.350	1.00	yes
rs208531	66.960.439	1.00	yes
rs208532	66.961.280	1.00	yes
rs208535	66.962.286	1.00	yes
rs208536	66.963.440	1.00	yes
rs208537	66.963.544	1.00	yes
rs208538	66.964.492	1.00	yes
rs72880045	66.965.232	1.00	yes
rs9453656	66.965.407	1.00	yes
rs12190773	66.968.194	1.00	yes
rs72882065	66.968.566	1.00	yes
rs12190187	66.973.207	1.00	yes
rs12190316	66.973.365	1.00	yes
rs12194996	66.973.366	1.00	yes
rs72882079	66.973.878	1.00	yes
rs9453663	66.976.077	0.97	no
rs9453664	66.976.079	0.97	no

rs number (continued)	position on chr.6 (bp) (continued)	r^2 with rs208520 (continued)	overlap with rs208515 (continued)
rs12200490	66.976.084	0.97	no
rs72882086	66.976.671	1.00	yes
rs12189683	66.977.716	1.00	yes
rs12191403	66.978.137	1.00	yes
rs12193164	66.978.962	1.00	yes
rs28684412	66.979.430	1.00	yes
rs72882100	66.981.495	1.00	yes
rs72882102	66.983.408	1.00	yes
rs12190073	66.983.740	1.00	yes
rs12191598	66.983.962	1.00	yes
rs12212063	66.985.148	1.00	yes
rs12192261	66.985.268	1.00	yes
rs12195505	66.986.279	1.00	yes
rs12216109	66.987.488	1.00	yes
rs12199876	66.989.469	0.97	yes
rs12193077	66.989.548	1.00	yes
rs9453671	66.991.208	1.00	yes
rs9453672	66.991.431	1.00	yes
rs72884017	66.992.513	1.00	yes
rs28773210	66.993.566	1.00	yes
rs72884025	66.994.185	1.00	yes
rs72884027	66.995.044	1.00	yes
rs72884033	66.995.747	1.00	yes
rs12209225	66.996.443	1.00	yes
rs12206488	66.999.917	0.88	no
rs12204033	67.001.525	0.94	no
rs12202599	67.001.573	0.94	no
rs17644076	67.004.973	0.94	no
rs12216176	67.005.130	0.94	no
rs9345808	67.008.262	0.89	no
rs1468153	67.009.054	0.86	no
rs9363572	67.012.231	0.86	no
rs55813840	67.012.508	0.86	no
rs62414636	67.015.723	0.86	no
rs561301289	67.016.490	0.86	no
rs9345809	67.017.543	0.86	no
rs9354410	67.018.584	0.86	no
rs9354411	67.019.930	0.86	no
rs58958359	67.021.381	0.86	no
rs62414637	67.022.423	0.83	no
rs60604603	67.022.955	0.86	no
rs62414639	67.023.706	0.86	no
rs62414641	67.029.895	0.86	no
rs2352058	67.036.037	0.83	no

LD with surrounding SNPs of identified variants was determined based on the 1,000 genomes catalog and only SNPs with r2≥0.8 are displayed (CEU population n=99; SNP inclusion parameters: MAF≥0.05 and HWE p-value≥0.0001; regions: chr6:64,419,876-67,952,828 and chr11:24,508,516-26,929,959 [GRCh37.p13]). Abbreviations: bp – base pair, CEU - Utah residents with ancestry from northern and western Europe, chr – chromosome, HWE – Hardy-Weinberg equilibrium, eNO – exhaled nitric oxide, LD – linkage disequilibrium, MAF – minor allele frequency, rs – reference SNP cluster code, SNP – single nucleotide polymorphism.

Table E7C. Fine mapping of eNO association signals by regional LD analysis for rs1441519 (11p14)

rs number	position on chr.11 (bp)	r^2 with rs1441519	overlap with rs12223678
rs11028996	25.905.876	0,80	yes
rs1348167	25.912.486	0,94	yes
rs56407206	25.913.261	0,90	yes
rs1542085	25.914.681	0,94	yes
rs12806413	25.915.245	0,94	yes
rs4434958	25.915.996	0,94	yes
rs12223678	**25.916.225**	**0,96**	**yes**
rs4550176	25.917.182	0,96	yes
rs12574733	25.920.888	0,92	yes
rs10834842	25.921.306	0,94	yes
rs12807641	25.924.207	0,94	yes
rs1441519	**25.929.959**	**1,00**	**yes**
rs1348169	25.930.009	0,94	yes

LD with surrounding SNPs of identified variants was determined based on the 1,000 genomes catalog and only SNPs with r2≥0.8 are displayed (CEU population n=99; SNP inclusion parameters: MAF≥0.05 and HWE p-value≥0.0001; regions: chr6:64,419,876-67,952,828 and chr11:24,508,516-26,929,959 [GRCh37.p13]). Abbreviations: bp – base pair, CEU - Utah residents with ancestry from northern and western Europe, chr – chromosome, HWE – Hardy-Weinberg equilibrium, eNO – exhaled nitric oxide, LD – linkage disequilibrium, MAF – minor allele frequency, rs – reference SNP cluster code, SNP – single nucleotide polymorphism.

Table E7D. Fine mapping of eNO association signals by regional LD analysis for rs12223678 (11p14)

rs number	position on chr.11 (bp)	r^2 with rs12223678	overlap with rs1441519
rs11028996	25.905.876	0,81	yes
rs1348167	25.912.486	0,98	yes
rs56407206	25.913.261	0,94	yes
rs1542085	25.914.681	0,98	yes
rs12806413	25.915.245	0,98	yes
rs4434958	25.915.996	0,98	yes
rs12223678	**25.916.225**	**1,00**	**yes**
rs4550176	25.917.182	1,00	yes
rs12574733	25.920.888	0,96	yes
rs10834842	25.921.306	0,98	yes
rs12807641	25.924.207	0,98	yes
rs1441519	**25.929.959**	**0,96**	**yes**
rs1348169	25.930.009	0,98	yes

LD with surrounding SNPs of identified variants was determined based on the 1,000 genomes catalog and only SNPs with r2≥0.8 are displayed (CEU population n=99; SNP inclusion parameters: MAF≥0.05 and HWE p-value≥0.0001; regions: chr6:64,419,876-67,952,828 and chr11:24,508,516-26,929,959 [GRCh37.p13]). Abbreviations: bp – base pair, CEU - Utah residents with ancestry from northern and western Europe, chr – chromosome, HWE – Hardy-Weinberg equilibrium, eNO – exhaled nitric oxide, LD – linkage disequilibrium, MAF – minor allele frequency, rs – reference SNP cluster code, SNP – single nucleotide polymorphism.

Table E8A. **Polymorphisms in high linkage disequilibrium with rs208515 and rs208520 are located in CTCF binding sites**[*]

SNP	Position[‡]	Location of CTCF binding site	Cell type in which CTCF binds to the respective site
rs4618506	chr6:66756889	chr6:66756893-66757236	GM12878 (EBV-transformed lymphoblastoid cell line)
rs3899423	chr6:66835285	chr6:66835142-66835309	H1-hESC (H1 human embryonic stem cells)
rs9453663	chr6:66976077	chr6:66975920-66976110	NHLF (Normal Human Lung Fibroblasts)
rs9453664	chr6:66976079	chr6:66975920-66976110	NHLF (Normal Human Lung Fibroblasts)
rs12200490	chr6:66976084	chr6:66975920-66976110	NHLF (Normal Human Lung Fibroblasts)

Abbreviations: bp – base pairs, chr – chromosome, CTCF – CCCTC-binding factor, EBV – Epstein-Barr virus, rs – reference SNP cluster code, SNP – single nucleotide polymorphism.
*CTCF binding sites were extracted from CTCFBSDB 2.0 - A database for CTCF binding sites and genome organization (http://insulatordb.uthsc.edu/). This data is based on experimentally derived ChipSeq results.
‡Chromosomal positions are provided according to human reference sequence (GRCh37) hg19 assembly.

Table E8B. **Polymorphisms at 6q12 located in CTCF-binding sites are located in Evolutionary Conserved Regions (ECRs)**[*]

SNP	Position**	ECR		
		location	Length	Type of DNA region
rs4618506	chr6:66756889	chr6:66756845-66756933	89 bp	transposons and simple repeats
rs3899423	chr6:66835285	chr6:66833841-66836729	2889 bp	transposons and simple repeats
rs9453663	chr6:66976077	chr6:66975933-66976221	289 bp	intergenic regions
rs9453664	chr6:66976079	chr6:66975933-66976221	289 bp	intergenic regions
rs12200490	chr6:66976084	chr6:66975933-66976221	289 bp	intergenic regions

Abbreviations: bp – base pairs, chr – chromosome, ECRs - Evolutionary Conserved Regions, rs – reference SNP cluster code, SNP – single nucleotide polymorphism.

* ECRs were extracted from ECR Browser (http://ecrbrowser.dcode.org/) and resemble conserved regions in humans and *Macaca mulatta* (Rhesus).

‡Chromosomal positions are provided according to human reference sequence (GRCh37) hg19 assembly.

Table E9. **Gene expression screening for target genes on 6q12 and 11p14, and *NOS1*, *NOS2* and *NOS3* genes in various cell types**

Cell line	Cell type	Description	*EYS*	*ANO3*	*SLC5A12*	*MUC15*	*LUZP2*	*NOS1*	*NOS2*	*NOS3*
Jurkat T	T cell leukemia	T lymphocytes, growing in suspension	-	-	-	-	-	-	+	+
YT	T/NK cell leukemia	T/ NK cells, growing in suspension	-	+	+	+	+	-	-	-
DG-75	Burkitt's lymphoma	B lymphocytes, growing in suspension	-	-	-	-	-	-	-	+
BJAB	Burkitt's lymphoma	B lymphocytes, growing in suspension	+	-	-	-	-	+	-	+
THP-1	Acute monocytic leukemia	Monocytes, growing in suspension	-	-	-	-	-	-	+	-
HL-60	Acute myeloid leukemia	Promyeloblasts, , growing in suspension	-	-	-	-	-	-	-	-
U937	Histiocytic lymphoma	Monocytes, growing in suspension	-	-	-	-	-	-	-	-
HepG2	Hepatocellular carcinoma	Adherent, epithelial-like cells	+	-	-	+	-	-	+	+
HEK293	Embryonal kidney	Adherent, fibroblastoid cells	-	-	+	-	-	-	-	-
H358	Lung carcinoma Bronchioalveolar carcinoma; Non-small cell lung cancer	Adherent, epithelial cells	-	-	+	+	+	-	-	+
H441	Lung carcinoma Papillary adenocarcinoma	Adherent, epithelial cells	-	-	+	+	+	-	-	+
A549	Lung carcinoma	Adherent, epithelial cells	-	-	-	-	-	+	-	-
Primary fibroblasts	Fibroblasts	Fibroblasts	+	+	+	-	+	-	-	-
Peripheral whole blood	-	-	+	-	-	-	-	-	-	-

Abbreviations: ANO3 - human anoctamin 3 gene locus, EYS - human eyes shut homolog gene locus, LUZP2 - human leucine zipper protein 2 gene locus, NOS - human nitric oxide synthase gene locus, MUC15 - human mucin 15 gene locus, SLC5A12 - human solute carrier family 5 (sodium/monocarboxylate cotransporter) member 12 gene locus.