Susceptibility to Chronic Mucus Hypersecretion, a Genome Wide Association Study

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

| Citation          | Dijkstra, A. E., J. Smolonska, M. van den Berge, C. Wijmenga, P. Zanen, M. A. Luinge, M. Platteel, et al. 2014. “Susceptibility to Chronic Mucus Hypersecretion, a Genome Wide Association Study.” PLoS ONE 9 (4): e91621. doi:10.1371/journal.pone.0091621. http://dx.doi.org/10.1371/journal.pone.0091621. |
|-------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Published Version | doi:10.1371/journal.pone.0091621                                                                                                                                                                      |
| Citable link      | http://nrs.harvard.edu/urn-3:HUL.InstRepos:12152914                                                                                                                                                   |
| Terms of Use      | This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA|
Susceptibility to Chronic Mucus Hypersecretion, a Genome Wide Association Study

Akkelies E. Dijkstra1,2, Joanna Smolonska2,3, Maarten van den Berge1,2, Ciska Wijmenga3, Pieter Zanen4, Marjan A. Luinge5, Mathieu Platteau3, Jan-Willem Lammers4, Magnus Dahlback6, Kerrie Tosh7, Pieter S. Hiemstra8, Peter J. Sterk9, Avi Spira10, Jorgen Vestbo11,12, Borge G. Nordestgaard13,14, Marianne Benn14, Sune F. Nielsen13, Morten Dahl13, W. Monique Verschueren16, H. Susan J. Picavet16, Henriette A. Smit17, Michael Owsijewitsch18,19, Hans U. Kauczor18,19, Harry J. de Koning20, Eva Nizankowska-Mogilnicka21, Filip Mejza21, Pawel Nastalek21, Cleo C. van Diemen3, Michael H. Cho22,23,24, Edwin K. Silverman22,23,24, James D. Crapo25,26, Terri H. Beaty26, David A. Lomas27, Per Bakke26, Amund Gulsvik28, Yohan Bosse29, M. A. Obeidat30, Daan W. Loth31,32, Lies Lahousse31,33, Fernando Rivadeneira34,35, Andre G. Uitterlinden34,35, Andre Hofman31,35, Bruno H. Stricker31,32, Guy G. Brusselle31,33, Cornelia M. van Duijn36, Uilke Brouwer2,5, Gerard H. Koppelman2,37, Judith M. Vonk2,38, Martijn C. Nawijn2,5, Harry J. M. Groen1, Wim Timens5, H. Marie Boezen2,38, Harry J. de Koning20, Eva Nizankowska-Mogilnicka21, Filip Mejza21, Pawel Nastalek21, Cleo C. van Diemen3, Michael H. Cho22,23,24, Edwin K. Silverman22,23,24, James D. Crapo25,26, Terri H. Beaty26, David A. Lomas27, Per Bakke26, Amund Gulsvik28, Yohan Bosse29, M. A. Obeidat30, Daan W. Loth31,32, Lies Lahousse31,33, Fernando Rivadeneira34,35, Andre G. Uitterlinden34,35, Andre Hofman31,35, Bruno H. Stricker31,32, Guy G. Brusselle31,33, Cornelia M. van Duijn36, Uilke Brouwer2,5, Gerard H. Koppelman2,37, Judith M. Vonk2,38, Martijn C. Nawijn2,5, Harry J. M. Groen1, Wim Timens5, H. Marie Boezen2,38, Dirkje S. Postma1,2, the LifeLines Cohort study1

Abstract

Background: Chronic mucus hypersecretion (CMH) is associated with an increased frequency of respiratory infections, excess lung function decline, and increased hospitalisation and mortality rates in the general population. It is associated with smoking, but it is unknown why only a minority of smokers develops CMH. A plausible explanation for this phenomenon is a predisposing genetic constitution. Therefore, we performed a genome wide association (GWA) study of CMH in Caucasian populations.

Methods: GWA analysis was performed in the NELSON-study using the Illumina 610 array, followed by replication and meta-analysis in 11 additional cohorts. In total 2,704 subjects with, and 7,624 subjects without CMH were included, all current or former heavy smokers (≥20 pack-years). Additional studies were performed to test the functional relevance of the most significant single nucleotide polymorphism (SNP).

Results: A strong consistency with CMH, consistent across all cohorts, was observed with rs6577641 (p = 4.25 x 10^-6, OR = 1.17), located in intron 9 of the special AT-rich sequence-binding protein 1 locus (SATB1) on chromosome 3. The risk allele (G) was associated with higher mRNA expression of SATB1 (4.3 x 10^-3) in lung tissue. Presence of CMH was associated with increased SATB1 mRNA expression in bronchial biopsies from COPD patients. SATB1 expression was induced during differentiation of primary human bronchial epithelial cells in culture.

Conclusions: Our findings, that SNP rs6577641 is associated with CMH in multiple cohorts and is a cis-eQTL for SATB1, together with our additional observation that SATB1 expression increases during epithelial differentiation provide suggestive evidence that SATB1 is a gene that affects CMH.
Introduction
The secretion of mucus is a normal part of the airway defense against inhaled noxious particles and substances. Chronic mucus hypersecretion (CMH) is a condition of overproduction of mucus and defined as the presence of sputum production during at least three months in two consecutive years without any explaining origin whereas airway obstruction is not a prerequisite [1]. Smoking is a risk factor for CMH, i.e. the prevalence of CMH in the general population is reported to be 7.4% in current smokers, 3.7% in ex-smokers and 2.4% in never smokers [2]. CMH is the key presenting symptom in chronic bronchitis, one of the three main sub-groups of chronic obstructive pulmonary disease (COPD), a complex disease characterized by the presence of incompletely reversible and generally progressive airflow limitation [3]. Moreover, CMH is a risk factor for the development of COPD [4,5]. Worldwide, COPD affected 65 million people in 2004 and more than 3 million people died of COPD in 2005, representing 5% of all deaths. It is predicted that COPD will be the third leading cause of death worldwide in 2030 [6]. COPD markedly reduces quality of life and is responsible for high healthcare costs. For instance, the combined (direct and indirect) yearly costs of COPD and asthma in the United States of America were projected at $63 billion in 2008 [7]. CMH is not only associated with COPD but also with an increased duration and frequency of respiratory infections, excess decline in forced expiratory volume in 1 second (FEV1) and increased hospitalization and mortality rates in the general population [4,5,8,9].

It is not known why only a minority of all smokers develops CMH, yet a plausible explanation is the presence of a genetic predisposition for CMH, as evidenced by familial aggregation of mucus overproduction and higher prevalence of CMH in monozygotic than in dizygotic twins [10–12]. Little is known about the identity of the genes that predispose to CMH. One publication suggested that CTLH is associated with chronic bronchitis in COPD [13].

The aim of our study was to identify genetic factors for CMH, thereby obtaining a better insight into the origins of this disorder.
Materials and Methods

Ethics Statement

The Dutch ministry of health and the Medical Ethics Committee of the hospital approved the study protocol for all Dutch centers. Ethics approval and written informed consent was obtained from all participants in all studies participating. For detailed information, see Supplement S1.

Subjects and genotyping

We performed GWA studies in participants of the NELSON-study (n = 3,729), a male population-based lung cancer screening study investigating heavy smokers (≥20 pack-years) [24]. Replication of SNPs with $p \leq 10^{-4}$ was attempted in six cohorts participating in ‘COPD Pathology: Addressing Critical gaps, Early Treatment & diagnosis and Innovative Concepts’ (COPACETIC) and in five non-COPACETIC cohorts. Caucasian subjects with ≥20 pack-years smoking with genotype-, spirometric- and demographic data were included.

An overview of the CMH definitions used in this study is presented in Table 1. A brief description of the included cohorts and details according to the period of data collection, type of population, genotyping platforms and genetic imputation software are presented in Table 2.

| Table 1. Questions used to define chronic mucus hypersecretion in the corresponding cohorts. |
|---|
| **Cohort** | **Question** |
| NELSON [24] | Do you expectorate sputum on the majority of days more than 3 months a year, even when you do not have a cold? |
| Rotterdam [25,26] | Do you expectorate sputum on the majority of days during ≥3 months during the last 2 years? |
| LifeLines [27] | Do you usually expectorate sputum during day or night in winter? If yes: Do you expectorate sputum on the majority of days >3 months a year? |
| Vlagtwedde- Vlaardingen [28,29] | Do you expectorate sputum on the majority of days >3 months a year? |
| Doetinchem [30] | Do you expectorate sputum during winter, day and night, each day for 3 months? |
| Poland [31,32] | Do you usually bring up phlegm from your chest, or do you usually have phlegm in your chest that is difficult to bring up when you don’t have a cold? If yes: Are there months in which you have this phlegm on most days? If yes: Do you bring up this phlegm on most days for as much as 3 months each year? A positive answer to all (3) questions identifies CMH. |
| Heidelberg [33] | Do you expectorate sputum on the majority of days >3 months a year? |
| GLUCOLD [15] | Do you expectorate sputum immediately after getting up on the majority of days in winter >3 months a year? |
| Rucphen [29] | Do you expectorate sputum during day or night in winter? If yes: Do you have expectation on the majority of days >3 months a year? |
| ECLIPSE [34] | Do you usually bring up phlegm from your chest on getting up, first thing in the morning, during the rest of the day or at night, on most days for 3 consecutive months or more during the year? |
| COPDGene [35] | Do you usually bring up phlegm from your chest on getting up, first thing in the morning, during the rest of the day or at night, on most days for 3 consecutive months or more during the year? |
| Norway [36,37] | Do you usually bring up phlegm from your chest on getting up, first thing in the morning, during the rest of the day or at night, on most days for 3 consecutive months or more during the year? |

Materials and Methods

Ethics Statement

The Dutch ministry of health and the Medical Ethics Committee of the hospital approved the study protocol for all Dutch centers. Ethics approval and written informed consent was obtained from all participants in all studies participating. For detailed information, see Supplement S1.

Subjects and genotyping

We performed GWA studies in participants of the NELSON-study (n = 3,729), a male population-based lung cancer screening study investigating heavy smokers (≥20 pack-years) [24]. Replication of SNPs with $p \leq 10^{-4}$ was attempted in six cohorts participating in ‘COPD Pathology: Addressing Critical gaps, Early Treatment & diagnosis and Innovative Concepts’ (COPACETIC) and in five non-COPACETIC cohorts. Caucasian subjects with ≥20 pack-years smoking with genotype-, spirometric- and demographic data were included.

An overview of the CMH definitions used in this study is presented in Table 1. A brief description of the included cohorts and details according to the period of data collection, type of population, genotyping platforms and genetic imputation software are presented in Table 2.

| Table 2. Overview of populations. |
|---|
| **Study** | **Data Collection** | **Type of population** | **Genotyping platform** | **Imputation software** |
| NELSON | 2004/2005 | general population | Illumina Quad 610 | NA |
| GLUCOLD | 2005/2006 | COPD case | Illumina Veracode | NA |
| Vlagtwedde Vlaardingen | 1989/1990 | general population | Illumina Veracode | NA |
| Doetinchem | 1998/2002 | general population | Illumina Veracode | NA |
| Poland | 2005/2006 | general population | Illumina Veracode | NA |
| Heidelberg | 2004/2005 | general population | Illumina Veracode | NA |
| Rucphen | 2002 | Family based COPD on a doctor diagnosis | Illumina Veracode | NA |
| Rotterdam | 2002/2008 | general population | Illumina 550K | MaCH |
| LifeLines | 2008/2010 | general population | Illumina Human CytoSNP-12 | BEAGLE v3.1.0 |
| COPDGene | 2008/2009 | COPD case/control (stage I–IV) | Illumina Human Omni1-Quad | MaCH |
| ECLIPSE | 2005/2007 | COPD case/control (stage II–IV) | Illumina Human HAP 550 V3 | MaCH |
| Norway | 2003/2005 | COPD case/control (stage II–IV) | Illumina Human HAP 550 V1, V3, and DUO | MaCH |

Populations and corresponding period of data collection, type of population, genotyping platform and software used for imputation.

NA = not applicable.

doi:10.1371/journal.pone.0091621.t001
doi:10.1371/journal.pone.0091621.t002
We searched for SNPs associated with CMH by using a two-stage strategy followed by a replication stage and meta-analysis (Figure 1).

Statistical analysis

General characteristics of CMH-cases and controls were compared using Student’s t- and Mann-Whitney-U tests for continuous variables and \( \chi^2 \) tests for dichotomous variables with SPSS 20.0. Sample and SNP quality control (QC), regression- and meta-analysis were performed with PLINK 1.07 [25]. QC criteria are described in Supplement S1.

Logistic regression analysis under an additive model was used to identify SNPs associated with CMH. SNPs with a p-value \( < 10^{-4} \) were included for replication. When two SNPs were in strong linkage disequilibrium \( r^2 \geq 0.8 \), the SNP with the lowest p-value was further analyzed.

SNPs in COPACETIC cohorts and in LifeLines were analyzed using logistic regression with adjustment for sex and smoking (ex-/current smoking). In LifeLines, imputed SNPs with an info-score \(<0.3\) (imputation quality score) were removed. SNPs in non-COPACETIC cohorts were analyzed by the cohort investigators using the same model.

Meta-analysis was performed on SNPs across NELSON and the 11 replication cohorts. The Cochran’s Q test was used to test for heterogeneity in the meta-analysis.

We performed multivariate logistic regression analysis, adjusted for pack-years and lung function, to associate CMH with the risk allele of rs6577641 in the identification cohort.

Functional relevance of SATB1 and rs6577641, our highest ranked-SNP

We performed 4 functional studies with the identified top-SNP. Details on their methods are given in Supplement S1.

1) whether rs6577641 is an eQTL, by analyzing the association of SATB1 expression levels with rs6577641 genotypes in lung tissue from three independent cohorts recruited from Laval University, University of British Columbia, and University of Groningen as described previously [14];

2) CMH-associated mRNA expression in airway wall biopsies from 77 COPD participants in the GLUCOLD-study [15];

3) the association of homozygous genotypes for rs6577641 with a) immunohistochemical staining (IHC) for SATB1 and b) the fraction of mucus positivity on bronchial tissue explanted from COPD or lung cancer subjects that underwent lung surgery;

4) SATB1 expression levels during mucociliary differentiation of primary bronchial epithelial cells cultured at air-liquid interface [26].

Results

Populations

Characteristics of the identification and replication populations are presented in Table 3. Subjects with CMH were more often current smokers and had worse lung function, except for populations including subjects with COPD only.

Identification analysis

After QC, 492,700 SNPs and 2,512 individuals (717 CMH cases, 1,795 controls) from the NELSON study remained. Logistic regression analysis was performed including these individuals.
supplemented with 590 additional healthy controls, adjusting for center. The **QQ-plot** provided no evidence of population stratification (**CMH**) in our identification cohort (p = 1.0185). 77 SNPs were associated with CMH with a p-value < 10^{-4}; CMH was associated with current smoking in our identification cohort (p<0.001). Therefore, we performed a second GWA adjusting for center and current/ex-smoking (717 CMH-cases, 1,795 controls). The QQ-plot showed no evidence of population stratification (**CMH**) in our identification cohort (p = 1.0056). We observed 64 SNPs with a p-value < 10^{-4}. Genome wide association for CMH ordered by chromosome is shown in the Manhattan plot. Figure 2 shows QQ-plots (A, C) and genome wide association signals for CMH ordered by chromosome (Manhattan-plots, B and D) of these sequential analyses. We identified 36 SNPs associated with CMH with a p-value < 10^{-4} in both analyses Table 4. Of these, 32 SNPs were included for replication and 4 SNPs were removed because they were in strong linkage disequilibrium (r^2 > 0.8) with another associated SNP.

### Replication of associated SNPs

Genotyping of SNP rs4775569 failed in the COPACETIC populations, and was removed for further analysis. CMH-associated top-SNPs for each cohort are presented in Table 5, with a complete overview in Table 6. When applying Bonferroni correction in the meta-analysis (p = 1.61 x 10^{-6} for 31 SNPs), we found a strong association with one SNP:

- rs6577641, a SNP located on chromosome 3 in intron 9 of the **special AT-rich sequence-binding protein 1** locus (**SATB1**) (combined p-value = 4.25 x 10^{-6}, OR = 1.17; 1.10–1.26).

The **SATB1** SNP rs6577641 had the lowest p-value for association with CMH in the meta-analysis. Figure 3 shows the forest plot of rs6577641 in the identification and replication cohorts and meta-analysis.

We assessed the percentage of subjects with CMH in each genotyping group for rs6577641 in NELSON-total and stratified for current and ex smokers (Figure 4). Multivariate logistic regression analysis, corrected for pack-years and FEV1%predicted, showed that CMH was significantly associated with the number of G-alleles in the 1,385 current smokers (reference = AA: heterozygous mutant (AG) p = 0.001; OR = 1.50, homozygous mutant (GG) p = 0.001; OR = 1.80) but not in 1,127 ex-smokers (reference = AA: heterozygous mutant (AG) p = 0.380; OR = 1.18, homozygous mutant (GG) p = 0.143; OR = 1.42).

### Table 3. Demographic and clinical characteristics of CMH-cases and -controls with ≥20 pack-years, present in the meta-analysis.

| Population | CMH | N    | Population % | Female, % | Age, yrs (SD) | Pack-years (range) | Current smoking, % | FEV1 %, pred. (SD) | FEV1/FVC, % (SD) |
|------------|-----|------|--------------|-----------|---------------|--------------------|-------------------|-------------------|------------------|
| NELSON     | Control | 1,795 | 71.5 | 0 | 60.2 (5.3) | 34 (21-156) | 47.5 | 100.3 (17.2) | 72.9 (8.7) |
| NELSON     | Case | 717 | 28.5 | 0 | 60.4 (5.6) | 39 (21-140) | 74.2 | 93.5 (20.0) | 69.2 (11.0) |
| Rotterdam  | Control | 1,043 | 84.1 | 46.1 | 68.0 (9.3) | 45 (20-149) | 40.1 | 92.4 (23.5) | 72.8 (8.7) |
| Rotterdam  | Case | 197 | 15.9 | 43.7 | 72.0 (8.4) | 40 (20-168) | 45.2 | 85.0 (26.9) | 68.0 (11.1) |
| Lifelines  | Control | 1,431 | 88.1 | 80.1 | 52.9 (9.2) | 27 (20-100) | 56.4 | 98.2 (15.6) | 72.4 (8.2) |
| Lifelines  | Case | 193 | 11.5 | 46.9 | 53.2 (9.9) | 29 (20-97) | 75.4 | 90.5 (18.0) | 68.3 (11.3) |
| Vlagtwedde-Vlaardingen* | Control | 234 | 82.4 | 27.4 | 52.9 (10.1) | 29 (20-128) | 51.7 | 94.5 (12.1) | 76.6 (4.5) |
| Vlagtwedde-Vlaardingen* | Case | 50 | 17.6 | 18 | 53.4 (10.5) | 33 (22-83) | 68 | 86.7 (18.6) | 71.0 (8.9) |
| Doetinchem | Control | 250 | 80.6 | 37.2 | 54.7 (8.8) | 30 (20-90) | 55.6 | 94.8 (17.6) | 71.5 (9.9) |
| Doetinchem | Case | 60 | 19.4 | 36.7 | 56.4 (7.7) | 33 (20-72) | 68.3 | 89.1 (19.6) | 69.3 (11.4) |
| Poland     | Control | 97 | 85.1 | 22.7 | 56.7 (10.5) | 30 (20-116) | 52.6 | 96.4 (21.4) | 72.5 (0.5) |
| Poland     | Case | 17 | 14.9 | 11.8 | 55.8 (9.4) | 35 (22-86) | 82.4 | 93.5 (24.0) | 69.2 (13.1) |
| Heidelberg | Control | 608 | 84.2 | 35.7 | 58.1 (5.2) | 37 (23-138) | 54.3 | 96.4 (17.6) | 78.9 (9.7) |
| Heidelberg | Case | 114 | 15.8 | 29.8 | 58.0 (5.2) | 37 (23-91) | 91.2 | 86.2 (21.5) | 75.3 (10.6) |
| GLUCOLD**  | Control | 48 | 55.2 | 8.3 | 62.6 (7.6) | 46 (21-182) | 62.5 | 63.4 (9.8) | 50.4 (9.1) |
| GLUCOLD**  | Case | 39 | 44.8 | 20.5 | 59.6 (7.4) | 40 (22-83) | 61.5 | 63.9 (8.8) | 53.1 (7.8) |
| Rucphen**  | Control | 28 | 53.4 | 46.4 | 66.5 (7.9) | 42 (21-120) | 57.1 | 74.5 (15.7) | 57.2 (7.8) |
| Rucphen**  | Case | 24 | 46.2 | 41.7 | 62.2 (10.5) | 43 (21-100) | 70.8 | 70.2 (21.6) | 53.1 (9.7) |
| ECLIPSE**  | Control | 961 | 62 | 37.5 | 64.1 (6.7) | 53 (21-205) | 28.1 | 48.0 (15.7) | 44.5 (11.3) |
| ECLIPSE**  | Case | 590 | 38 | 24.1 | 62.9 (7.4) | 54 (22-220) | 47.5 | 46.2 (15.5) | 44.3 (11.7) |
| COPDGene   | Control | 628 | 71.8 | 53.5 | 63.1 (8.6) | 50 (21-173) | 28.2 | 75.0 (28.3) | 63.7 (17.6) |
| COPDGene   | Case | 247 | 28.2 | 40.5 | 61.9 (8.4) | 54 (21-237) | 50.2 | 60.4 (27.4) | 54.6 (17.9) |
| Norway     | Control | 501 | 52.4 | 44.9 | 61.5 (10.3) | 34 (20-130) | 46.9 | 71.7 (24.2) | 64.6 (15.7) |
| Norway     | Case | 456 | 47.6 | 55.1 | 64.1 (10.1) | 39 (20-119) | 59 | 56.5 (24.4) | 55.0 (17.3) |

CMH = Chronic mucus hypersecretion; *lung function is based on FEV1/FVC; **all individuals in this cohort have COPD; *based on lung function of 700 subjects who returned for follow-up study 4 years later.
doi:10.1371/journal.pone.0091621.t003
Functional relevance of SATB1 and rs6577641

1) Transcriptional regulation of SATB1 mRNA expression

We analyzed the association of SATB1 expression levels in lung tissue with rs6577641 genotype in 3 independent data sets of the Universities of Groningen, Laval and UBC [14]. A cis-acting effect of rs6577641 on SATB1 expression was identified and present in all three datasets (n = 1,095), with the same direction of effect across all three SATB1 probes on the array. The (susceptibility) G allele increased expression, the (protective) A allele reduced expression (p = 4.3 × 10^{-9}) in the meta-analysis across the three datasets and across all three SATB1 probes measured (Table 7).

2) SATB1 mRNA expression and CMH

We compared SATB1 expression in baseline airway wall biopsies of COPD patients with (n = 38) and without (n = 39) CMH in GLUCOLD [15]. CMH was significantly associated with SATB1 expression levels (corrected for ex-/current smoking; p = 0.0045; Figure 5). After stratification, the same direction of effect was present in ex- and current smokers. However, this association reached statistical significance in current smokers (p = 0.021) and not in ex-smokers (p = 0.132), probably due to a difference in power as 46 subjects were current smokers versus 33 ex-smokers.

3) Genotype related protein expression and mucus positivity in bronchial epithelium

SATB1 protein expression has previously been observed in IHS analysis of bronchial epithelial cells [16]. Therefore, we stained SATB1 on paraffin embedded lung tissue biopsies of individuals from the Groningen population contributing to the eQTL analysis. We observed clear nuclear staining for SATB1 in bronchial epithelial cells. No significant difference for % of strong positive, positive and weak positive cells was observed between the protective (AA, n = 9) and risk (GG, n = 14) rs6577641 genotypes (11.8±6.5 versus 12.7±6.9, p = 0.74).

We determined whether the fraction of mucus positive bronchial epithelium was different in subjects with different homozygous rs6577641 genotypes and performed PAS-staining on tissue biopsies from the same cohort. We observed no significant difference between individuals with the homozygous protective (AA, n = 10) and risk (GG, n = 7) alleles (19.7±11.9 versus 14.3±9.6, p = 0.34).

4) SATB1 expression levels during bronchial epithelial cell mucociliary differentiation

We investigated whether SATB1 expression was induced during mucociliary differentiation of primary human bronchial epithelial
chronic mucus hypersecretion (CMH), we aimed to identify
end of the 45-day ALI culture period.

Cells. SATB1 expression was induced over time (Figure 6), with an
up to 45 days. ALI culture of HBE cells induced mucociliary

Table 4. SNPs associated with CMH with a p-value<10^{-4}, present in GWAS-I and in GWAS-II, in the NELSON identification cohort.

| Chromosome | SNP       | Base pair position | p-value GWAS I | p-value GWAS II |
|------------|-----------|--------------------|----------------|-----------------|
| 2          | rs6735868 | 103580209          | 1.11 × 10^{-5} | 1.08 × 10^{-5}  |
| 3          | rs1387089 | 1940922            | 7.94 × 10^{-5} | 4.56 × 10^{-5}  |
| 4          | rs1488757 | 1981567            | 2.17 × 10^{-5} | 1.16 × 10^{-5}  |
| 5          | rs6577641 | 18397849           | 6.83 × 10^{-5} | 2.57 × 10^{-5}  |
| 6          | rs4306981 | 7992421            | 9.74 × 10^{-5} | 5.18 × 10^{-5}  |
| 7          | rs4242562 | 115475287          | 7.66 × 10^{-5} | 5.13 × 10^{-5}  |
| 8          | rs7836298 | 115504434          | 1.03 × 10^{-5} | 4.37 × 10^{-5}  |
| 8          | rs7823554*| 115553109          | 6.05 × 10^{-5} | 5.22 × 10^{-5}  |
| 8          | rs7836963*| 115568426          | 5.52 × 10^{-5} | 4.24 × 10^{-5}  |
| 8          | rs16886291| 115711436          | 3.54 × 10^{-5} | 2.09 × 10^{-5}  |
| 8          | rs10098746| 125838127          | 8.47 × 10^{-5} | 4.34 × 10^{-5}  |
| 8          | rs7831595 | 144974963          | 3.08 × 10^{-5} | 2.32 × 10^{-5}  |
| 9          | rs4842047 | 138816796          | 2.63 × 10^{-5} | 4.51 × 10^{-5}  |
| 10         | rs943189  | 22842590           | 5.57 × 10^{-5} | 6.33 × 10^{-5}  |
| 10         | rs1026531 | 22379184           | 2.76 × 10^{-5} | 8.55 × 10^{-5}  |
| 10         | rs1894307*| 12005720           | 9.04 × 10^{-6} | 7.18 × 10^{-6}  |
| 12         | rs2255953 | 12010736           | 1.13 × 10^{-5} | 4.33 × 10^{-5}  |
| 12         | rs2855708 | 12013572           | 6.47 × 10^{-5} | 3.97 × 10^{-5}  |
| 12         | rs10879509*| 73242131          | 6.98 × 10^{-6} | 4.44 × 10^{-5}  |
| 12         | rs4760851 | 73284781           | 4.85 × 10^{-6} | 2.29 × 10^{-5}  |
| 12         | rs952394  | 73441110           | 4.18 × 10^{-5} | 4.22 × 10^{-5}  |
| 12         | rs12822199| 75458164           | 4.82 × 10^{-5} | 8.58 × 10^{-5}  |
| 12         | rs1379963 | 75493882           | 1.18 × 10^{-5} | 2.20 × 10^{-5}  |
| 12         | rs1795669 | 76273692           | 8.01 × 10^{-5} | 7.86 × 10^{-5}  |
| 13         | rs9578362 | 21882381           | 4.28 × 10^{-5} | 7.99 × 10^{-5}  |
| 13         | rs1211304 | 50381016           | 9.96 × 10^{-5} | 1.12 × 10^{-5}  |
| 14         | rs992745  | 27810995           | 7.67 × 10^{-5} | 2.99 × 10^{-5}  |
| 15         | rs754661  | 26934277           | 4.54 × 10^{-5} | 2.88 × 10^{-5}  |
| 15         | rs4775569 | 46850317           | 4.20 × 10^{-5} | 1.72 × 10^{-5}  |
| 15         | rs13333521| 19904082           | 5.08 × 10^{-5} | 2.50 × 10^{-5}  |
| 17         | rs11652469| 49565797           | 1.13 × 10^{-5} | 3.80 × 10^{-5}  |
| 18         | rs8086262 | 69227590           | 1.15 × 10^{-5} | 2.53 × 10^{-5}  |
| 20         | rs4815628 | 3891896            | 4.17 × 10^{-5} | 2.15 × 10^{-5}  |
| 21         | rs2032257 | 27774870           | 3.97 × 10^{-5} | 5.39 × 10^{-5}  |
| 22         | rs1009147 | 30088257           | 8.41 × 10^{-5} | 4.51 × 10^{-5}  |
| 22         | rs1005239 | 47687170           | 9.86 × 10^{-5} | 8.67 × 10^{-5}  |

*SNP not selected for replication because of strong linkage disequilibrium with another SNP.
doi:10.1371/journal.pone.0091621.t004

(SATB1) expression was induced over time (Figure 6), with an
approximately 8-fold increased expression from the start to the end of the 45-day ALI culture period.

Discussion

Since not every ex- or current heavy smoker suffers from chronic mucus hypersecretion (CMH), we aimed to identify
genetic variants conferring susceptibility to CMH. Therefore, we performed the first GWA study on CMH, the key presenting
symptom in chronic bronchitis. CMH was associated with 36 SNPs at the p<10^{-4} significance level in the identification cohort. In the meta-analysis combining our identification and replication
cohorts, strong association was observed with rs6577641, a SNP located on chromosome 3 in intron 9 of SATB1. Although the association of rs6577641 with CMH did not reach conventional
genome-wide significance, its effect was in the same direction and was significant (4.25×10^{-5}) at nominal levels (1.61×10^{-5}) across
eleven study populations, showing the robustness of this finding. The detected odds ratio for this SNP suggests an additional risk of

Genetic Influence of SATB1 on Airway Disease
### Table 5. Meta-analysis of top SNPs associated with CMH in replication cohorts, in identification and replication cohorts and corresponding direction of effect in all cohorts and associated feature and gene(s).

| Chr | SNP       | Base pair position | Minor allele | MAF   | p-value | OR   | Direction of effect per cohort* | p-value | OR   | Q | Close(st) gene(s) |
|-----|-----------|--------------------|--------------|-------|--------|------|-------------------------------|---------|------|---|------------------|
| 3   | rs6577641 | 18397849 G         | G            | 0.400 | 5.01E-03 | 1.12 | ++++++++0+0+                  | 4.25E-06 | 1.17 | 6.20E-01 SATB1<sup>+</sup> |
| 3   | rs1488757 | 1981567 G         | G            | 0.109 | 2.34E-01 | 0.92 | -0-0-0-0-0                   | 1.10E-03 | 0.83 | 1.55E-01 LOC727810 and CNTN4 |
| 12  | rs2855708 | 12013572 G        | G            | 0.273 | 2.18E-01 | 1.06 | +0+0++++0+                   | 1.20E-03 | 1.13 | 1.76E-01 ETV6<sup>7</sup> |
| 14  | rs922745  | 27810095 G       | G            | 0.234 | 2.94E-01 | 0.95 | ++++++++0000                | 2.74E-03 | 0.89 | 4.59E-02 LOC728755<sup>7</sup> |
| 4   | rs4306981 | 79924121 G        | G            | 0.307 | 3.37E-01 | 1.04 | ++++++++0000                | 1.38E-03 | 1.12 | 5.19E-02 PAQR3 and ARD1B |
| 12  | rs1795669 | 76273692 A        | A            | 0.059 | 2.83E-01 | 1.09 | ++++++++0000               | 2.90E-03 | 1.22 | 1.77E-01 LOC10055033 and LOC100550330 |
| 9   | rs4820447 | 137556617 A       | A            | 0.303 | 3.88E-01 | 0.96 | -0-XX+X-0-00              | 3.44E-03 | 0.89 | 3.03E-01 CAMPSAP1 and UBAC1 |
| 13  | rs95788362| 21882381 A        | A            | 0.402 | 8.05E-01 | 1.01 | +-----0-00                  | 3.61E-03 | 0.91 | 2.88E-02 LOC6500794 and GRK6PS |
| 12  | rs2255593 | 12010736 G        | G            | 0.212 | 5.31E-01 | 0.97 | ++-X+0-0+                 | 5.12E-03 | 1.13 | 4.54E-02 ETV6<sup>7</sup> |
| 15  | rs754661  | 26934277 G        | G            | 0.405 | 5.45E-01 | 0.96 | -00X++0+                   | 6.29E-03 | 0.91 | 1.08E-01 GABRB3<sup>8</sup> |
| 8   | rs16886291| 11571436 A        | A            | 0.127 | 5.01E-03 | 1.12 | ++++++++00+                | 5.41E-03 | 0.86 | 1.55E-01 hCG_1644355 and TRPS1 |

MAF = minor allele frequency in NELSON;  
*Direction of effect per cohort: each sign reflects one cohort, direction of effect is presented by: ++=(OR>1.05), −= (OR<0.95), 0=(0.95<OR<1.05) and x = (missing result); cohorts are presented in the same order as in Table 2; OR is odds ratio; Q is p-value for heterogeneity;  
<sup>7</sup> means corresponding SNP is located in an intron in this gene.  
doi:10.1371/journal.pone.0091621.t005
### Table 6.

| Chromosome | Base pair position | SNP | Minor allele | Base pair position | SNP | Minor allele | P-value   | OR   | Q       | N       | p-value   | OR   | Q       | Closest gene(s) | Q       |
|------------|--------------------|-----|--------------|--------------------|-----|--------------|-----------|-------|---------|---------|-----------|-------|---------|----------------|---------|
| 3          | 18397849           | rs6577641 | G             | 5.01E-03           | 1.121 | 9.19E-01    | 12        | 4.3E-06 | 1.173   | 6.20E-01 | 1.24E-01  | 9.17E-01 | 1.073   | SATB1*           | 9.17E-01 |
| 18         | 69227500           | rs8086262 | G             | 2.16E-02           | 1.107 | 8.31E-02    | 12        | 1.0E-02 | 1.073   | 1.4E-02  | 9.17E-01  | 1.073   | 1.4E-02  | KCNC2*          | 9.17E-01 |
| 8          | 115473257          | rs1432162 | C             | 6.71E-02           | 1.175 | 6.1E-02     | 12        | 2.7E-02 | 1.175   | 6.1E-02  | 1.175    | 6.1E-02  | 1.175   | KCRC3           | 9.17E-01 |
| 12         | 115302457          | rs1375983 | A             | 8.7E-01            | 1.089 | 6.0E-01     | 11        | 4.7E-01 | 1.089   | 6.0E-01  | 1.089    | 6.0E-01  | 1.089   | KCNT1           | 9.17E-01 |
| 12         | 75493882           | rs13821199 | G             | 9.6E-02           | 1.093 | 5.3E-02     | 12        | 8.0E-02 | 1.093   | 5.3E-02  | 1.093    | 5.3E-02  | 1.093   | LOC220459       | 9.17E-01 |
| 8          | 115473257          | rs1432162 | C             | 5.41E-02           | 1.121 | 7.6E-02     | 12        | 2.4E-03 | 1.121   | 7.6E-02  | 1.121    | 7.6E-02  | 1.121   | KCNC2*          | 9.17E-01 |
| 12         | 75458164           | rs12822199 | A             | 4.24E-02           | 1.107 | 2.1E-02     | 12        | 1.0E-02 | 1.107   | 2.1E-02  | 1.107    | 2.1E-02  | 1.107   | LOC220459       | 9.17E-01 |
| 3          | 1981567            | rs1488757 | G             | 2.34E-01           | 0.923 | 8.44E-01    | 12        | 1.1E-03 | 0.923   | 8.44E-01 | 0.923    | 8.44E-01 | 0.923   | LOC727810       | 9.17E-01 |
| 12         | 75458164           | rs12822199 | A             | 9.6E-02           | 1.093 | 5.3E-02     | 12        | 8.0E-02 | 1.093   | 5.3E-02  | 1.093    | 5.3E-02  | 1.093   | TRPS1           | 9.17E-01 |
| 3          | 1981567            | rs1488757 | G             | 2.34E-01           | 0.923 | 8.44E-01    | 12        | 1.1E-03 | 0.923   | 8.44E-01 | 0.923    | 8.44E-01 | 0.923   | TRPS1           | 9.17E-01 |
| 12         | 75493882           | rs13821199 | G             | 9.6E-02           | 1.093 | 5.3E-02     | 12        | 8.0E-02 | 1.093   | 5.3E-02  | 1.093    | 5.3E-02  | 1.093   | LOC220459       | 9.17E-01 |
| 8          | 115711836          | rs16886291 | A             | 5.45E-01           | 0.963 | 8.77E-01    | 12        | 9.4E-01 | 0.963   | 8.77E-01 | 0.963    | 8.77E-01 | 0.963   | hCG_1644355     | 9.17E-01 |
| 22         | 47687170           | rs1005239 | G             | 8.58E-01           | 0.989 | 9.91E-01    | 12        | 5.1E-02 | 0.989   | 9.91E-01 | 0.989    | 9.91E-01 | 0.989   | NF2*            | 9.17E-01 |
| 12         | 73284781           | rs4760851 | A             | 6.15E-01           | 1.022 | 6.86E-01    | 12        | 2.4E-02 | 1.022   | 6.86E-01 | 1.022    | 6.86E-01 | 1.022   | TRHDE and LOC100128674 | 9.17E-01 |
| 3          | 1940922            | rs1387089 | G             | 6.80E-01           | 0.972 | 4.43E-01    | 11        | 9.1E-02 | 0.972   | 4.43E-01 | 0.972    | 4.43E-01 | 0.972   | EPPK1           | 9.17E-01 |
| 11         | 22379184           | rs11026531 | A             | 9.75E-01           | 1.005 | 1.04E-01    | 12        | 1.2E-02 | 1.005   | 1.04E-01 | 1.005    | 1.04E-01 | 1.005   | SLC17A6*         | 9.17E-01 |

P-value is fixed p-value if p-value for heterogeneity (Q) < 0.005; OR is Odds Ratio; OR is fixed OR if p-value for heterogeneity (Q) < 0.005; and random OR if p-value for heterogeneity (Q) > 0.005; N = number of cohorts; * means that the corresponding SNP is an intron in this gene.

DOI: 10.1371/journal.pone.0091621.t006

Genetic Influence of SATB1 on Airway Disease

PLOS ONE | www.plosone.org 9 April 2014 | Volume 9 | Issue 4 | e91621
that lack of power is the reason for not reaching the level of significance in ex-smokers.

These data strongly suggest that SATB1 plays a role in the susceptibility to CMH in subjects with a history of heavy smoking (≥20 pack-years) within the general population. Moreover, rs6577641 has a cis-eQTL effect on SATB1 lung tissue expression, the risk allele at rs6577641 (G) increasing and the A-allele reducing expression of SATB1 significantly. Additionally, we found a higher SATB1 expression in bronchial biopsies of COPD-patients with CMH. We found no differences between the GG and AA genotypes for protein expression of SATB1 in airway epithelium by IHC in a small sample from our lung tissue registry. Finally, we demonstrate that SATB1 mRNA expression is induced during mucociliary differentiation in ALI cultures of human bronchial epithelial cells of 2 donors supporting our eQTL findings. Interestingly, expression of the mucin gene MUC5AC was also induced during this culture period, with a slightly delayed kinetics compared to SATB1. Together these data strongly suggest that SATB1 is induced during differentiation of bronchial epithelial cells and affects chronic mucus hypersecretion.

The forest plot clearly shows that the effect of SNP rs6577641 is lower in cohorts including COPD patients only (GLUCOLD, Rucphen, COPDGene, ECLIPSE and Norway) than in the other cohorts. Additional meta-analysis of COPD-cohorts and general population based cohorts separately confirmed this (COPD cohorts, combined p-value = 0.236, OR = 1.07 and general population based cohorts, combined p-value = 5.18×10^-2, OR = 1.26). This suggests genetic heterogeneity of CMH in subjects with and without COPD.

The SNP most significantly associated with CMH, rs6577641, is located in an intron of SATB1. SATB1 is a transcription factor and chromatin (re)organizer important for controlling the expression of many genes in a tissue or cell-type specific fashion, for instance in differentiating thymus T-cells [17] or differentiating skin keratinocytes [18]. Expression of SATB1 has been observed in normal human bronchial epithelial cells by immunohistochemistry and lower levels were observed in non-small lung cancer cells [16]. In our study, we also showed the presence of SATB1 in bronchial epithelial cells by IHC staining of lung tissue. However, no significant differences were found between patients homozygous for the protective and risk alleles, for either specific SATB1 staining or for PAS staining, the latter specifically detecting mucus. This inability to detect a genotype effect on protein staining may

The forest plot clearly shows that the effect of SNP rs6577641 is lower in cohorts including COPD patients only (GLUCOLD, Rucphen, COPDGene, ECLIPSE and Norway) than in the other cohorts. Additional meta-analysis of COPD-cohorts and general population based cohorts separately confirmed this (COPD cohorts, combined p-value = 0.236, OR = 1.07 and general population based cohorts, combined p-value = 5.18×10^-2, OR = 1.26). This suggests genetic heterogeneity of CMH in subjects with and without COPD.

The SNP most significantly associated with CMH, rs6577641, is located in an intron of SATB1. SATB1 is a transcription factor and chromatin (re)organizer important for controlling the expression of many genes in a tissue or cell-type specific fashion, for instance in differentiating thymus T-cells [17] or differentiating skin keratinocytes [18]. Expression of SATB1 has been observed in normal human bronchial epithelial cells by immunohistochemistry and lower levels were observed in non-small lung cancer cells [16]. In our study, we also showed the presence of SATB1 in bronchial epithelial cells by IHC staining of lung tissue. However, no significant differences were found between patients homozygous for the protective and risk alleles, for either specific SATB1 staining or for PAS staining, the latter specifically detecting mucus. This inability to detect a genotype effect on protein staining may
be due to lack of power, as we found a large variation in SATB1 and PAS protein expression in the relatively small number of lung tissue samples. Other explanations include possible expression regulation of SATB1 by smoke exposure which could be a dynamic process not readily detected at the protein level by any single-time point analysis such as IHC staining on lung biopsies. Alternatively SATB1 expression levels may vary throughout the lungs or the technique used here is not sensitive enough to detect relatively small differences in protein levels. To further explore the association of SATB1 protein and its underlying regulation, it would be of interest to perform longitudinal investigations on lung tissue samples of subjects with and without CMH, or time series of *in vitro* cultured epithelial cells from donors with a specific genotype and cigarette smoke exposure. This would also allow further studies on epigenetic regulation with methylation, microRNA or histone modifications.

The lack of association between the SATB1 protein and rs6577641 might additionally be due to the location of mucus positive cells in lung tissue. Mucus is produced both by goblet cells and submucosal glands, which we did not investigate further. Normal mucus consists of 97% water and 3% solids including 30% mucins. In case of dysregulation of mucus production, the concentration of solids in mucus may increase up to 15%. A further step therefore could involve investigating mucins/proteins present in mucus, e.g. MUC5AC is predominantly produced by goblet cells in proximal airways and MUC5B by secretory cells throughout the airways and by submucosal glands.

How does SATB1 expression contribute to CMH? SATB1 is known to be a genome organizer, a tissue specific chromatin remodeling protein with a property to modifying chromatin architecture by formation of loops, allowing contact of condensed genomic DNA to regulatory transcription proteins [19]. Thus SATB1 can control gene expression of a series of target genes located within a single locus at a specific chromosomal location [20]. This has for instance been elegantly shown in case of differentiating keratinocytes [18], where Satb1 expression regulates genes located in the keratinocyte-specific loci, leading to adaptation of a specific cell fate of the differentiating keratinocytes. Similarly, a mechanism by which SATB1 could contribute to CMH is the induction of a gene expression program during differentiation of bronchial epithelial cells, leading to adaptation of a cell fate specific for mucus producing cells in the submucosal glands or a goblet cell phenotype in the bronchial epithelium. Involvement of Satb1 in pneumocyte differentiation was previously observed by Baguma et al. in mice [21]. We observed induction of SATB1 expression in bronchial epithelial cells differentiating under ALI culture conditions. Further research will need to test whether a specific gene expression profile is induced by SATB1 expression in differentiating bronchial epithelial cells. SATB1 is also highly expressed in thymocytes, but absent in mature non-activated T cells [22]. Moreover, Satb1 has been shown in mice to be essential for expression of TH2 cells important in the regulation of genes encoding interleukin 4, 5 and 13 [19]. In Satb1-deficient mice, development of thymocytes stopped after the CD4+CD8+ stage with deregulation of many genes [23]. Conversely, in case of excessive SATB1-production an excess of Th2 cells may be formed which all produce IL-13, which may contribute to increased mucus production. Therefore, a putative role of SATB1 in T-cells for the CMH phenotype should not be disregarded.

Strength of our study is the fact that we were able to replicate our findings in different populations, ranging from cohorts consisting of individuals with severe airflow limitation to cohorts mainly consisting of healthy smokers. There are some limitations, e.g. the presence of CMH was not based on actual measurements of the amount of sputum produced but based on questionnaires that were not completely similar in all study cohorts. Underreporting of CMH occurs since those experiencing CMH become accustomed to these symptoms, believing they are smoking related or because they are embarrassed to admit to cough and sputum.

| Table 7. Meta-analysis of the effect of rs6577641 on mRNA expression levels of SATB1 in the lung*. |
|---|
| **Probe Gene Symbol** |
| **Affymetrix Probe ID** |
| **Z-score Groningen** |
| **Z-score Laval** |
| **Z-score UBC** |
| **Z-Score Meta-Analysis** |
| **p-value Meta-Analysis** |
| SATB1 | 100148784_TGI_at | 2.28 | -0.08 | -1.62 | -2.29 | **0.022** |
| SATB1 | 100150253_TGI_at | -0.84 | -0.49 | -1.62 | -1.70 | **0.088** |
| SATB1 | 100305926_TGI_at | -2.81 | -1.38 | -1.46 | -3.26 | **0.001** |

*To assess the effect of the SNP rs6577641 on gene expression, a Kruskal-Wallis test was performed. This test generates a p-value, but does not give a direction of the effect. To assess the direction of the effect, a Spearman’s correlation test was performed. Next, a Z-score was calculated for each center and a meta-analysis performed for each of the three SATB1 probes across all centers. Finally, a meta-analysis for all three SATB1 probes was performed across all centers. This generated a Z-score of -5.87 and a corresponding p-value of 4.3*10^-10, indicating that the susceptibility G allele of the SNP rs6577641 increases SATB1 expression.

doi:10.1371/journal.pone.0091621.t007

Figure 5. Bronchial biopsy mRNA-expression levels of SATB1 in COPD patients with chronic mucus hypersecretion (n = 38) compared to patients without chronic mucus hypersecretion (n = 39).
doi:10.1371/journal.pone.0091621.g005
We demonstrated that \textit{SATB1} mRNA expression is induced during mucociliary differentiation in ALI cultures of HBE cells in a small dataset \((n=2)\). However, these data seem reliable as they are supported by eQTL data from lung tissue. Despite this drawback, we consistently found evidence for association of \textit{SATB1} with CMH in the populations studied, showing the robustness of our finding. Moreover, we corroborated this finding by functional studies in lung tissue, airway wall biopsies of COPD patients and epithelial cultures. More extensive research is needed to investigate which factors induce \textit{SATB1} expression in airway epithelium.

In summary, we performed identification analyses and meta-analyses using data from almost 7,000 participants to identify genes involved in susceptibility for CMH. It is remarkable that we found a genetic association for CMH given this phenotype is partly subjectively determined and not well delineated. Moreover, despite cohort differences to define CMH and severity of airflow limitation, we found consistent effects of SNP rs6577641 on CMH. This confirms that the CMH phenotype, despite the fact that it is self-reported, is a robust phenotype irrespective of the presence or absence of airflow limitation. The association of rs6577641 on chromosome 3 at the \textit{SATB1} locus with CMH was supported by functional studies including gene expression findings, demonstrating \textit{SATB1} to be associated with CMH.

Chronic mucus hypersecretion is a bothersome symptom for many people, it increases in prevalence with aging and affects quality of life, exacerbations of symptoms due to respiratory infections and ultimately increases mortality. The involvement of \textit{SATB1} in CMH offers opportunities to better understand the process leading to CMH, and future development of tailored medicines.

**Supporting Information**

**Supplement S1**

(\text{DOC})

---

**Figure 6.** \textit{SATB1}, \textit{MUC5AC} and \textit{FOXJ1} mRNA expression levels during mucociliary human airway epithelial cell differentiation \((n=2\) donors). Expression of \textit{SATB1}, the identified gene in our study, \textit{MUC5AC} a marker of mucus, and \textit{FOXJ1}, representing ciliated cells in epithelial cell culture on air liquid interface. doi:10.1371/journal.pone.0091621.g006

---

**Acknowledgments**

The authors would like to thank the research staff at Respiratory Health Network Tissue Bank of the FRQS for their valuable assistance in the lung eQTL study.

LifeLines Cohort Study: BZ Alizadeh\(^1\), RA de Boer\(^2\), HM Boezen\(^3\), M Bruinenberg\(^2\), L Franke\(^4\), P van der Harst\(^5\), HL Hillege\(^6\), MM van der Klauw\(^7\), G Navis\(^8\), J Ormel\(^9\), DS Postma\(^9\), JGM Rosmalen\(^9\), JP Slaets\(^9\), H Snieder\(^1\), RP Stolk\(^1\), BHR Wolffenbuttel\(^1\), C Wijmenga\(^1\)

\(^1\)University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen, the Netherlands;

\(^2\)University of Groningen, University Medical Center Groningen, Department of Cardiology, Groningen, the Netherlands;

\(^3\)University of Groningen, University Medical Center Groningen, the LifeLines Cohort Study, Groningen, the Netherlands;

\(^4\)University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, the Netherlands;

\(^5\)University of Groningen, University Medical Center Groningen, Department of Endocrinology, Groningen, the Netherlands;

\(^6\)University of Groningen, University Medical Center Groningen, Department of Internal Medicine, Division of Nephrology, Groningen, the Netherlands;

\(^7\)University of Groningen, University Medical Center Groningen, Interdisciplinary Center of Psychopathology of Emotion Regulation (ICP-E), Department of Psychiatry, Groningen, the Netherlands;

\(^8\)University of Groningen, University Medical Center Groningen, Department of Pulmonology, Groningen, the Netherlands

**Author Contributions**

Conceived and designed the experiments: AED DSP HMB CW JS PZ HG JWL WT MCN MvdB. Performed the experiments: AED JS ML MCN UB MP. Analyzed the data: AED MvdB DL WT MHC MP MB FM M. Obeidat. Contributed reagents/materials/analysis tools: ML MP JS M. Obeidat MvdB DL. Wrote the paper: AED DSP HMB CW JMV GHK PJS EKS MHC MCN. Acquisition of replication cohorts: M. Dahlback KT PSH PSJ AS JV BGN M. Dahl WMV HSP HAS M. Owsijewitsch HUK HJK EN FN CCvD EKS JDC THB DAL PB AG YB LL FR AGU AH BHS GGB CMvD UB GHK SN.
19. Cai C, Zhang HY, Le JJ, Dong JC, Cui Y, et al. (2010) Inflammatory airway disease. Ann J Respir Crit Care Med 182: S77–121.

20. Yasui D, Miyano M, Cai S, Varga-Weisz P, Kohwi-Shigematsu T (2002) SATB1 targets chromatin remodelling to regulate genes over long distances. Nature 419: 641–645.

21. Baqima-Nisharsha M, Angka HE, Isamolou MR, Kabhar B (2007) Microarray analysis of Mif5+/− /Myod1−/− hypoplastic mouse lungs reveals a profile of genes involved in pulmonary differentiation. Histol Histopathol 22: 483–495.

22. Dickinson LA, Joh T, Kohwi Y, Kohwi-Shigematsu T (1992) A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. Cell 70: 631–645.

23. Krangel MS (2007) T cell development: Better living through chromatin. Nat Immunol 8: 687–694.

24. van Klaveren RJ, Oudkerk M, Prokop M, Scholten ET, Nackaerts K, et al. (2009) Management of lung nodules detected by volume CT scanning. N Engl J Med 361: 2221–2229.

25. Soler Artigas M, Loth DW, Wain LV, Ghazib SA, Obiedi M, et al. (2011) Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. Nat Genet 43: 1082–1090.

26. Hofman A, van Duijn CM, Franco OH, Ikram MA, Janssen HL, et al. (2011) The rotterdam study: 2012 objectives and design update. Eur J Epidemiol 26: 657–686.

27. Stolk RP, Rosmalen JG, Postma DS, de Boer RA, Navis G, et al. (2008) Universal risk factors for multifactorial diseases: LifeLines: A three-generation population-based study. Eur J Epidemiol 23: 67–74.

28. Rijker B, Schouten JP, Mensinga TT, Weiss ST, De Vries K, et al. (1993) Factors associated with bronchial responsiveness to histamine in a population sample of adults. Am Rev Respir Dis 147: 1447–1453.

29. van Diemen CC, Postma DS, Vonk JM, Bruineman M, Nohe IM, et al. (2006) Decorin and TGF-beta1 polymorphisms and development of COPD in a general population. Respir Res 7: 89.

30. Verschuren WM, Blokstra A, Picaev HS, Smul HA (2008) Cohort profile: The dutchtemcoh cohort study. Int J Epidemiol 37: 1236–1241.

31. Nizankowska-Mogilnicka E, Mejja F, Buist AS, Vollmer WM, Skucha W, et al. (2007) Prevalence of COPD and tobacco smoking in malopolska region: results from the BOLD study in Poland. Pol Arch Med Wewn 117: 402–410.

32. Lamprecht B, McBurnie MA, Vollmer WM, Gudmundsson G, Weite T, et al. (2011) COPD in never smokers: Results from the population-based burden of obstructive lung disease study. Chest 139: 752–763.

33. Becker N, Delorme S, Kauzkom HU (2008) LUSI: The german component of the european trial on the efficacy of multislice-CT for the early detection of lung cancer. onkologie 31.

34. Vesty J, Anderson W, Coxson HO, Cetin C, Dawber F, et al. (2008) Evaluation of COPD longitudinally to identify predictive surrogate end-points (ECLIPSE). Eur Respir J 31: 869–873.

35. Regan EA, Hokanen JE, Murphy JP, Make B, Lynch DA, et al. (2010) Genetic epidemiology of COPD (COPDGene) study design. COPD 7: 32–43.

36. Cho MH, Castaldi PJ, Wain ES, Siedlinski M, Hersh CP, et al. (2011) A genome-wide association study of COPD identifies a susceptibility locus on chromosome 19p13. Hum Mol Genet.

37. Grydeland TB, Dirksen A, Coxson HO, Pillai SG, Sharma S, et al. (2009) Quantitative computed tomography: Echymema and airway wall thickness by sex, age and smoking. Eur Respir J 34: 838–855.