ARTICLE INFO

Handling editor: Xavier Querol

Keywords:
Genetics
Genome wide
Interaction
Occupational exposure
Respiratory symptoms

ABSTRACT

Respiratory symptoms are important indicators of respiratory diseases. Both genetic and environmental factors contribute to respiratory symptoms development but less is known about gene-environment interactions. We aimed to assess interactions between single nucleotide polymorphisms (SNPs) and occupational exposures on respiratory symptoms cough, dyspnea and phlegm. As identification cohort LifeLines I (n = 7976 subjects) was used. Job-specific exposure was estimated using the ALOHA + job exposure matrix. SNP-by-occupational exposure interactions on respiratory symptoms were tested using logistic regression adjusted for gender, age, and current smoking. SNP-by-exposure interactions with a p-value < 10^{-4} were tested for replication in two independent cohorts: LifeLines II (n = 5260) and the Vlagtwedde-Vlaardingen cohort (n = 1529). The interaction estimates of the replication cohorts were meta-analyzed using PLINK. Replication was achieved when the meta-analysis p-value was < 0.05 and the interaction effect had the same direction as in the identification cohort. Additionally, we assessed whether replicated SNPs associated with gene expression by analyzing if they were cis-acting expression quantitative trait loci (eQTL) in lung tissue. In the replication meta-analysis, sixteen out of 477 identified SNP-by-occupational exposure interactions had a p-value < 0.05 and 9 of these interactions had the same direction as in the identification cohort. Several identified loci were plausible candidates for respiratory symptoms, such as TMPRSS9, SERPINH1, TOX3, and ARHGAP18. Three replicated SNPs were cis-eQTLs for FCER1A, CHN1, and TIMM13 in lung tissue. Taken together, this genome-wide SNP-by-occupational exposure interaction study in relation to cough, dyspnea, and phlegm identified several suggestive susceptibility genes. Further research should determine if these genes are true susceptibility loci for respiratory symptoms in relation to occupational exposures.

Abbreviations: COPD, Chronic Obstructive Pulmonary Disease; eQTL, Expression Quantitative Trait Locus; GWAS, Genome Wide Association Study; GWIS, Genome Wide Interaction Study; ISCO-88, International Standard Classification of Occupations version 1988; JEM, Job Exposure Matrix; SNP, Single Nucleotide Polymorphism

⁎ Corresponding author at: University of Groningen, University Medical Center Groningen, Department of Epidemiology, Hanzeplein 1, Groningen 9700 RB, the Netherlands.

E-mail address: h.m.boezen@umcg.nl (H.M. Boezen).

https://doi.org/10.1016/j.envint.2018.11.017

Received 9 July 2018; Received in revised form 5 November 2018; Accepted 8 November 2018

Available online 16 November 2018

0160-4120/ © 2018 Elsevier Ltd. All rights reserved.
1. Introduction

Occupational exposures to respirable particles and gases have been associated with low levels of lung function, and with high risk for respiratory symptoms and chronic obstructive pulmonary disease (COPD) (de Jong et al., 2014a; de Jong et al., 2014b; de Jong et al., 2014c; Marchetti et al., 2014). Respiratory symptoms in adults are related to impaired quality of life regardless of a diagnosis of asthma and COPD (Voll-Aanerud et al., 2010). Previous candidate gene approaches have shown that genetic variants may affect the level of lung function (Bosse, 2012) and the risk for respiratory symptoms (Cox et al., 2010). For example polymorphisms in ADAM33 were associated with decline in lung function and genetic polymorphisms in TLR4 were associated with work-related respiratory symptoms (Cho et al., 2011; van Diemen et al., 2005). These candidate gene studies were driven by hypotheses relying on known biological pathways. In contrast, a genome-wide association study (GWAS) can identify genetic loci associated with a phenotype across the entire genome in a hypothesis free manner. However, many susceptibility loci identified in a single GWAS do not replicate consistently. This failure could be because environmental factors trigger the development of symptoms only in subjects with a specific genotype. More recently, hypothesis free genome-wide interaction studies (GWIS) have been performed to identify genetic loci that affect the susceptibility to known harmful exposures (de Jong et al., 2015). We have recently performed a GWAS on respiratory symptoms, and found no significant associations between genetic loci and respiratory symptoms including cough, dyspnea and phlegm (Zeng et al., 2017). An explanation for the negative results is that genetic determinants may play an important role only in combination with specific environmental exposures such as occupational exposure to dust, gases and fumes, pesticides, solvents, or metals. Therefore, it is necessary to investigate genetic susceptibility to environmental exposure in relation to respiratory symptoms using a GWIS.

We have previously shown that genetic susceptibility is associated with lung function level (FEV1) in the context of occupational exposure to biological dust, mineral dust and gases and fumes, and that several of these novel identified susceptibility loci were cis-acting expression (mRNA) quantitative trait loci (eQTLs) in lung tissue (de Jong et al., 2015). Whether such gene-by-occupational exposure interactions exist in relation to respiratory symptoms is largely unknown.

The aim of the current GWIS was to identify susceptibility loci for occupational exposures leading to respiratory symptoms. We have previously shown that genetic susceptibility is associated with lung function level (FEV1) in the context of occupational exposure to biological dust, mineral dust and gases and fumes, and that several of these novel identified susceptibility loci were cis-acting expression (mRNA) quantitative trait loci (eQTLs) in lung tissue (de Jong et al., 2015). Whether such gene-by-occupational exposure interactions exist in relation to respiratory symptoms is largely unknown.

The aim of the current GWIS was to identify susceptibility loci for several types of occupational exposures in relation to the respiratory symptoms cough, dyspnea, and phlegm in a general population cohort. We used two independent cohorts to confirm our findings. Findings from this study may shed light on the genetic susceptibility to specific exposures leading to respiratory symptoms.

2. Material and methods

2.1. Identification cohort

Individuals with genotype data from the first data release of the LifeLines cohort study (LifeLines I) were included as the identification cohort (n = 7976). The LifeLines cohort study is a general population-based cohort including subjects from the three Northern provinces of the Netherlands (de Jong et al., 2015; Scholtens et al., 2015). Participants completed a standardized questionnaire, underwent a medical examination, and provided written informed consent. The study protocol was approved by the local medical ethics committee.

2.2. Replication cohorts

To verify our initial findings, we included 5260 subjects from the second data release of the LifeLines cohort study (LifeLines II) and 1529 subjects with full data on genotypes and covariates on the last survey in 1989/1990 from the Vlagtwedde-Vlaardingen cohort, a prospective general population based cohort (Figarska et al., 2012; van Diemen et al., 2005). Both the identification and replication cohorts included only Caucasian individuals of Dutch descent.

2.3. Genotyping and quality control

Blood samples of all subjects included in the identification and replication cohorts were genotyped using IlluminaCytoSNP-12 arrays. SNPs that fulfilled the quality control criteria were included: genotype call-rate ≥ 95%, minor allele frequency ≥ 1%, and Hardy-Weinberg equilibrium cut-off p-value ≥ 10^-4. A total of 227,981 genotyped SNPs were included in the identification analysis. Non-Caucasian samples and first-degree relatives were excluded.

2.4. Respiratory symptoms

Cough, dyspnea, and phlegm were defined by a standardized questionnaire from the European Community Respiratory Health Survey in LifeLines I and LifeLines II (Burney et al., 1994) and the British Medical Research Council respiratory questionnaire in Vlagtwedde-Vlaardingen cohort (Jarvis, 2002). The definition of cough, dyspnea and phlegm was described in detail as follows and just like as mentioned in our previous study (Zeng et al., 2017). Cough was defined as at least one positive answer to the questions: “do you usually cough first thing in the morning in the winter?” or “do you usually cough during the day, or at night, in winter?”. Dyspnea was defined as a positive answer to the question: “are you troubled by shortness of breath when hurrying on level ground or walking up a slight hill or stairs at normal pace?”. Phlegm was defined as at least one positive answer to the questions: “do you usually bring up any phlegm from your chest first thing in the morning in winter?” or “do you usually bring up any phlegm from your chest during the day, or at night, in winter?”

2.5. Occupational exposure

Self-reported job title and descriptions were coded according to the International Standard Classification of Occupations version 1988 (ISCO-88) (International Labour Organization, 1990; de Jong et al., 2015). These four-digit codes were used to estimate job-specific exposure to biological dust, mineral dust, gases and fumes, pesticides, aromatic solvents, chlorinated solvents, other solvents (i.e. non-aromatic solvents such as formaldehyde and dimethylformamide), and metals using the ALOHA+ Job Exposure Matrix (JEM) (de Jong et al., 2014a). To maximize the study power in this interaction study “no exposure” was recoded to “0”, and “low exposure” and “high exposure” were recoded to “1”.

2.6. Statistical analysis

After quality control, genome-wide interaction analyses were performed on three kinds of respiratory symptoms and eight types of occupational exposure. Effects of gene-environment interactions on the respiratory symptoms cough, dyspnea, and phlegm were tested by including the SNP, the occupational exposure and the SNP-by-occupational exposure interaction term in a logistic regression model adjusted for sex, age, and current smoking (no/yes) in the software package PLINK (version 1.07) (Purcell et al., 2007). Each of the eight different occupational exposures was tested separately. SNPs with a SNP-by-occupational exposure interaction p-value < 10^-4 were tested for replication in the two independent cohorts (Liao et al., 2013; Gref et al., 2017; Wu et al., 2014). The interaction estimates of the two replication cohorts were meta-analyzed using PLINK. Replication was achieved when the p-value of the meta-analysis was < 0.05 and the interaction effect was in the same direction as in the identification cohort. SNP annotation was performed using HaploReg version 4.1 (Broad Institute) (Ward and Kellis, 2012).
2.7. Gene expression analysis

To assess the functional relevance of the replicated SNPs in the GWAS analysis we investigated whether the replicated SNPs were cis-acting eQTLs in a lung tissue database established by the lung eQTL consortium (Hao et al., 2012). Lung tissue was collected from patients who underwent lung resectional surgery at three participating sites; University of Groningen, Laval University, and University of British Columbia (Hao et al., 2012). DNA samples were genotyped with Illumina Human1M-Duo BeadChip arrays, and gene expression profiles were obtained using a custom Affymetrix array (GEO accession number GPL10379 and GSE23546). A total of 1087 subjects with no missing data were included in the analysis. The association between the SNPs and gene expression levels (log transformed) were tested in each cohort separately by linear regression analysis adjusted for disease status, age, gender, smoking status and a cohort specific number of principal components, followed by a meta-analysis of the results of the 3 cohorts. A cis-eQTL was defined as a SNP that was significantly associated with expression levels of a gene within a 1 Mb (in both directions) distance of that SNP, with a p-value below the Bonferroni corrected threshold (p = 0.05/number of probe sets within the 2 Mb window) in the meta-analysis.

3. Results

Characteristics of the study populations and the prevalence of respiratory symptoms and occupational exposures are shown in Table 1. The prevalence of smoking, respiratory symptoms, and occupational exposures was higher in the historical Vlagtwedde-Vlaardingen cohort compared to the current LifeLines I and LifeLines II cohorts.

3.1. SNP-by-occupational exposures interactions

An overview of the number of SNPs selected in the identification and replication analyses is presented in Table 2. The genomic inflation factor for the identification sample suggests little population stratification for SNP-by-occupational exposures interaction models such as

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Number of selected SNPs for cough, dyspnea and phlegm in the identification and replication analyses.</th>
</tr>
</thead>
<tbody>
<tr>
<td>N SNPs Identification (LifeLines I)</td>
<td>N SNPs Replication (Meta-analysis of LifeLines II and Vlagtwedde-Vlaardingen)</td>
</tr>
<tr>
<td>p value &lt; 10^{-4}</td>
<td>p value &lt; 0.05</td>
</tr>
<tr>
<td>Cough</td>
<td>Biological dust</td>
</tr>
<tr>
<td>Total</td>
<td>152</td>
</tr>
<tr>
<td>Biological dust</td>
<td>25</td>
</tr>
<tr>
<td>Mineral dust</td>
<td>14</td>
</tr>
<tr>
<td>Gases and fumes</td>
<td>25</td>
</tr>
<tr>
<td>Pesticides</td>
<td>13</td>
</tr>
<tr>
<td>Aromatic solvents</td>
<td>14</td>
</tr>
<tr>
<td>Chlorinated solvents</td>
<td>22</td>
</tr>
<tr>
<td>Other solvents</td>
<td>16</td>
</tr>
<tr>
<td>Metals</td>
<td>23</td>
</tr>
<tr>
<td>Dyspnea</td>
<td>Total</td>
</tr>
<tr>
<td>Biological dust</td>
<td>14</td>
</tr>
<tr>
<td>Mineral dust</td>
<td>30</td>
</tr>
<tr>
<td>Gases and fumes</td>
<td>16</td>
</tr>
<tr>
<td>Pesticides</td>
<td>27</td>
</tr>
<tr>
<td>Aromatic solvents</td>
<td>28</td>
</tr>
<tr>
<td>Chlorinated solvents</td>
<td>24</td>
</tr>
<tr>
<td>Other solvents</td>
<td>29</td>
</tr>
<tr>
<td>Metals</td>
<td>13</td>
</tr>
<tr>
<td>Phlegm</td>
<td>Total</td>
</tr>
<tr>
<td>Biological dust</td>
<td>20</td>
</tr>
<tr>
<td>Mineral dust</td>
<td>11</td>
</tr>
<tr>
<td>Gases and fumes</td>
<td>23</td>
</tr>
<tr>
<td>Pesticides</td>
<td>6</td>
</tr>
<tr>
<td>Aromatic solvents</td>
<td>29</td>
</tr>
<tr>
<td>Chlorinated solvents</td>
<td>22</td>
</tr>
<tr>
<td>Other solvents</td>
<td>20</td>
</tr>
<tr>
<td>Metals</td>
<td>13</td>
</tr>
</tbody>
</table>

SNP-by-biological dust exposure on cough, SNP-by-mineral dust exposure on dyspnea and SNP-by-aromatic solvent exposure on phlegm (lambda = 0.99, 1.01, 1.02, respectively, online supplement Fig. S1).

3.2. Cough

We identified 152 SNPs that interacted with occupational exposures for cough in the identification cohort (p-values for interaction < 10^{-4}) (Table 2). In total 4 SNP-by-exposure interactions (2.6%) had a p < 0.05 in the meta-analysis of the replication cohorts, of which 2 interactions (i.e. rs11265204*biological dust and rs7588395*metals) were in the same direction as in the identification cohort (Table 2, online supplement Fig. S2, Table 3).

3.3. Dyspnea

We identified 181 SNPs that interacted with the investigated occupational exposures on dyspnea in the identification cohort (p-values for interaction < 10^{-4}) (Table 2). In the meta-analysis of the replication cohorts, interactions with 8 SNPs had a p < 0.05 of which 4 interactions (i.e. rs7252511*mineral dust, rs4380004*mineral dust, rs4756771*mineral dust, and rs9513670*gases and fumes) were in the same direction as in the identification cohort (Table 2, online supplement Fig. S3, Table 4).
SNPs | Chr | A1 | MAF | Closest gene | LifeLines I identification (n = 7976) | LifeLines II replication (n = 5260) | Vlagtwedde-vlaardingen replication (n = 1529) | Meta-analysis replication (n = 6806) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11265204</td>
<td>1</td>
<td>A</td>
<td>0.32</td>
<td>RP11-180D21.3</td>
<td>1.10 6.50 × 10⁻²</td>
<td>1.05 4.54 × 10⁻¹</td>
<td>1.20 2.72 × 10⁻¹</td>
<td>1.08 2.65 × 10⁻¹</td>
</tr>
<tr>
<td>SNP</td>
<td>Biological dust</td>
<td>SPN</td>
<td>1.69 1.01 × 10⁻⁶</td>
<td>1.29 5.80 × 10⁻²</td>
<td>1.78 3.60 × 10⁻²</td>
<td>1.38 9.00 × 10⁻³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP* biological dust</td>
<td>0.66 1.47 × 10⁻⁵</td>
<td>0.82 9.40 × 10⁻²</td>
<td>0.64 5.60 × 10⁻²</td>
<td>0.78 1.90 × 10⁻²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs7758025</td>
<td>2</td>
<td>A</td>
<td>0.23</td>
<td>AC096649.3</td>
<td>1.86 5.26 × 10⁻⁶</td>
<td>1.12 9.30 × 10⁻²</td>
<td>1.01 9.45 × 10⁻¹</td>
<td>1.10 1.17 × 10⁻¹</td>
</tr>
<tr>
<td>SNP</td>
<td>Metals</td>
<td></td>
<td>1.05 4.21 × 10⁻¹</td>
<td>1.50 3.40 × 10⁻²</td>
<td>0.99 9.89 × 10⁻¹</td>
<td>1.34 7.20 × 10⁻²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNP*metals</td>
<td>0.38 2.31 × 10⁻⁵</td>
<td>0.60 5.10 × 10⁻²</td>
<td>0.64 4.10 × 10⁻¹</td>
<td>0.61 3.50 × 10⁻²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A1 = tested allele, MAF = Minor Allele Frequency.

3.4. Phlegm

We identified 144 SNPs that interacted with the investigated occupational exposures on phlegm in the identification cohort (p-values for interaction < 10⁻⁴) (Table 2). A total of 4 SNP-by-occupational exposure interactions showed a significant association (p < 0.05) in the meta-analysis of the replication cohorts, of which 3 interactions (i.e. rs11291026*biological dust, rs3095611*aromatic solvents, and rs7758025*other solvents) were in the same direction as in the identification cohort (Table 2, online supplement Fig. S4, Table S5).

3.5. Associations with gene expression

Of the 9 replicated SNPs, 6 SNPs were available in the lung tissue database and for 2 other SNPs a SNP in high linkage disequilibrium could be found (i.e. rs11682476 for rs7588395: r² = 0.99 and rs1420546 for rs3095611: r² = 0.93 (20)), one SNP (rs4388004) could not be assessed for its association with gene expression levels in lung tissue. Three SNPs showed cis-eQTL associations with p-values below the Bonferroni corrected threshold (online supplement Table S1 and Table S3). SNP rs11265204 that significantly interacted with exposure to biological dust on cough was associated with expression of FCER1A (p = 1.99 × 10⁻¹⁰) (Fig. 1A). Rs11682476 (proxy SNP for rs7588395 that significantly interacted with exposure to biological dust on cough) was associated with expression of CHN1 (p = 1.79 × 10⁻⁴) (Fig. 1B). SNP rs7252511 that interacted with exposure to mineral dust on dyspnea was significantly associated with expression of TIMM13 (close to TMPRSS9) (p = 1.27 × 10⁻⁸) (Fig. 1C).

4. Discussion

To our knowledge this is the first GWIS on the effects of SNPs-by-occupational exposure interactions on respiratory symptoms. We identified and replicated two SNPs that were associated with cough in subjects with occupational exposure to biological dust and metals, respectively. Four SNPs were identified and replicated in SNP-by-occupational exposure interaction analysis on dyspnea in subjects with exposure to mineral dust, and gases and fumes. Three SNPs were associated with phlegm in subjects with exposure to biological dust, aromatic solvents, and other solvents, respectively. Several identified SNPs were located in genes that have important functions in the lung and may play a role in pathways associated with respiratory symptoms (i.e. TMPRSS9, SERPINH1, TOX3, ARHGAP18, and FCER1A) (online supplement Table S2). In addition, three SNPs were associated with gene expression levels in the lung which indicate their possible functional relevance.

Table 3

Significantly replicated interactions between SNPs and occupational exposures on cough.

Table 4

Significantly replicated interactions between SNPs and occupational exposures on dyspnea.
Table 5
Significantly replicated interactions between SNPs and occupational exposures on phlegm.

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Chr</th>
<th>A1</th>
<th>MAF</th>
<th>Closest gene</th>
<th>LifeLines I identification (n = 7976)</th>
<th>LifeLines II replication (n = 5260)</th>
<th>Vlagtwedde-Vlaardingen replication (n = 1529)</th>
<th>Meta-analysis replication (n = 6806)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs12291026 SNP</td>
<td>11</td>
<td>C</td>
<td>0.08</td>
<td>SERPINH1</td>
<td>0.71, 1.71 \times 10^{-3}</td>
<td>0.85, 2.24 \times 10^{-1}</td>
<td>0.90, 7.07 \times 10^{-1}</td>
<td>0.86, 2.07 \times 10^{-1}</td>
</tr>
<tr>
<td>Biological dust SNP</td>
<td>16</td>
<td>A</td>
<td>0.32</td>
<td>TOX3</td>
<td>0.94, 4.61 \times 10^{-1}</td>
<td>1.05, 6.82 \times 10^{-1}</td>
<td>0.95, 7.81 \times 10^{-1}</td>
<td>1.02, 8.24 \times 10^{-1}</td>
</tr>
<tr>
<td>rs3095611 SNP</td>
<td>6</td>
<td>A</td>
<td>0.32</td>
<td>ARHGAP18</td>
<td>2.03, 4.03 \times 10^{-5}</td>
<td>1.45, 6.80 \times 10^{-3}</td>
<td>0.95, 7.81 \times 10^{-1}</td>
<td>1.44, 4.60 \times 10^{-2}</td>
</tr>
<tr>
<td>rs7758025 SNP</td>
<td>16</td>
<td>A</td>
<td>0.32</td>
<td>TIMM13</td>
<td>0.94, 2.73 \times 10^{-1}</td>
<td>0.81, 4.00 \times 10^{-3}</td>
<td>0.81, 1.99 \times 10^{-1}</td>
<td>0.81, 2.00 \times 10^{-3}</td>
</tr>
<tr>
<td>Other solvents SNP</td>
<td>6</td>
<td>A</td>
<td>0.32</td>
<td>SERPINH1</td>
<td>0.89, 4.93 \times 10^{-1}</td>
<td>0.66, 7.80 \times 10^{-2}</td>
<td>0.69, 2.06 \times 10^{-1}</td>
<td>0.68, 3.00 \times 10^{-2}</td>
</tr>
<tr>
<td>Other solvents SNP</td>
<td>11</td>
<td>C</td>
<td>0.08</td>
<td>SERPINH1</td>
<td>2.08, 4.54 \times 10^{-6}</td>
<td>1.65, 2.90 \times 10^{-2}</td>
<td>0.69, 2.06 \times 10^{-1}</td>
<td>1.61, 8.00 \times 10^{-3}</td>
</tr>
</tbody>
</table>

A1 = tested allele, MAF = Minor allele frequency.

Findings of the identification cohort were replicated in two independent cohorts. The occupational exposure definition was the same in all cohorts, and estimated by the ALOHA+ JEM that is specific for population-based studies (Sunyer et al., 1998), and less likely to be affected by recall bias or differential misclassification when compared to self-report (Manneetje and Kromhout, 2003). We investigated a broad range of occupational exposures, including biological dust, mineral dust, gases and fumes, pesticides, aromatic solvents, chlorinated solvents, other solvents, and metals, related to the respiratory symptoms cough, dyspnea and phlegm.

The minor allele of rs7252511, a SNP in TMPRSS9, was associated with a higher risk of dyspnea in subjects exposed to mineral dust. In lung tissue, rs7252511 showed a cis-eQTL association with higher expression of the gene TIMM13, which partially overlaps in antisense orientation with TMPRSS9 (Cal et al., 2003). TMPRSS9, known as serase-1, is a type II transmembrane serine protease and has a unique structure including three tandem serine protease domains (i.e., serase-1, −2, and −3) (Okumura et al., 2006). Serase-1B is known to interact with glycosaminoglycans (GAGs), which maintain lung structure and function, modulate the inflammatory response, influence tissue repair and remodeling, and protect against injury in various respiratory diseases (Okumura et al., 2006). Serase-1B and urokinase-type plasminogen activator (uPA) receptor can efficiently activate uPA, which is involved in epithelial repair and airway remodeling, and is associated with asthma susceptibility, bronchial hyperresponsiveness, and decline in lung function (Barton et al., 2009). Occupational exposure to mineral dust is a well-known risk factor for reduced lung growth and COPD (Cohen et al., 2016; Mehta et al., 2012) of which dyspnea is one of the main symptoms. Our results suggest that subjects carrying the minor allele of this SNP in TMPRSS9 are especially vulnerable to these harmful effects of mineral dust exposure.

The minor allele of rs12291026, a SNP in SERPINH1, was associated with a higher risk of phlegm in subjects exposed to biological dust. SERPINH1, also known as heat shock protein 47 (HSP47) is a serpin which serves as a human chaperone protein for collagen. SERPINH1 is expressed and resides in the endoplasmic reticulum of procollagen-producing cells and plays a role in collagen biosynthetic process (Bellaye et al., 2014). Collagen is mostly present in tissues in the form of fibers, and is an important marker of fibrosis. Increased collagen in the lung tissue can lead to increased fibrosis, which in turn affects the microenvironments and function of the lungs. SERPINH1 is significantly upregulated during the progression of fibrosis in the lung, liver, and other tissues (Hagiwara et al., 2007). SERPINH1 expression is in alignment with collagen expression. The expression of SERPINH1 and the synthesis of extracellular matrix proteins increased in airway remodeling in asthma (Phipps et al., 2004). Results of animal experiment indicated that the suppression of SERPINH1 expression can abate accumulation of collagens to delay the progression of fibrotic diseases (Taguchi and Razzaque, 2007). Our results suggest that subjects carrying the minor alleles of this SNP in SERPINH1 are more susceptible to phlegm of exposure to biological dust.

Another interesting finding is the association between rs3095611, a SNP in TOX3, and a higher risk of phlegm in subjects exposed to aromatic solvents. TOX3 is a neuronal survival factor that regulates calcium dependent transcription in neurons (Yuan et al., 2009). The minor allele of rs3095611, a SNP in TOX3, was associated with a higher risk of phlegm in subjects exposed to aromatic solvents. TOX3 is a neuronal survival factor that regulates calcium dependent transcription in neurons (Yuan et al., 2009).
et al., 2012; Yuan et al., 2009). Our results indicate there is interaction between TOX3 and exposure to aromatic solvents regarding the risk of phlegm. When aromatic solvents enter the body by inhalation, they can cause respiratory inflammatory, upper respiratory tract damage, and, interestingly, chronic bronchitis, with phlegm as an important symptom (Cakmak et al., 2004; Mcke et al., 2015; Ryu et al., 2013). Our results suggest that subjects carrying the minor alleles of this SNP in TOX3 are more susceptible to the respiratory effects of exposure to aromatic solvents.

Exposure to other solvents (i.e. non-aromatic solvents such as formaldehyde and dimethylformamide) is a risk factor for phlegm in subjects carrying the minor allele of rs7758025 in the ARHGAP18 gene. ARHGAP18 regulates the integrity and morphogenesis of epithelial cells (Maeda et al., 2011; Neisch et al., 2013). The epithelial cells in lungs provide structural integrity, moisten and protect the airways, are a physical barrier against environmental exposures, and prevent infection and tissue injury by action of the mucociliary escalator (Mercer et al., 2006; Vareille et al., 2011). Genetic variation in the ARHGAP18 gene may thus result in an impaired or altered epithelial barrier function, which might lead to increased phlegm production upon exposure to non-aromatic solvents.

The cis-eQTL analysis showed two additional associations that are of interest. The first was an association between the intergenic SNP rs11265204 and a lower expression of FCERIA. In our analysis, the interaction between rs11265204 and biological dust was significantly associated with a lower risk to cough. Polymorphisms in FCERIA are associated with IgE levels and allergic sensitization (Chen et al., 2009; Potaczek et al., 2013; Zhou et al., 2012), and therefore, speculations are that FCERIA is associated with asthma via IgE (Gould and Sutton, 2008; Platts-Mills, 2001). Given that cough is a symptom associated with asthma and allergy (Goldsoibel and Chipp, 2010), our results indicating a lower expression of FCERIA and a lower risk of exposure-associated coughing in subjects carrying the minor alleles of rs11265204 seems plausible. The second cis-eQTL is rs11682476 (proxy-SNP for rs7588395) which is associated with lower expression of CHN1. According to our analysis, the interaction between rs7588395 and metal exposure was significantly associated with cough. CHN1 encodes α2-chimaerin which is a non-protein kinase C phorbol ester receptor with Rac-GTPase-activating protein activity and cysteine-rich domain. Ahmed et al demonstrated that cysteine-rich proteins indicate the presence of metal-binding domains, where metals, usually zinc, play a structural role in maintaining the function of these domains, such as dimer formation and DNA-binding (Miyake et al., 2008). In the current study, we found that subjects with the major allele of rs7588395 and exposure to metals had increased prevalence of cough. An important paradigm of this gene is earlier mentioned ARHGAP18 which belongs, like CHN1, to a family of Rho GTPase-activating proteins. Therefore, CHN1 may have a similar function in the lung as ARHGAP18.

Since GWI studies are an explorative approach, we used a more liberal p-value threshold (p < 10^{-4}) for identification of SNPs in the identification cohort to reduce the risk of missing true associations between SNP-by-occupational exposure interactions and respiratory symptoms (Liao et al., 2013; Gref et al., 2017; Wu et al., 2014). When we assessed these interactions in the independent replication cohorts, the total number of significant gene-by-exposure interactions in the replication meta-analysis is less than expected by chance (i.e. 16 out of 477 (approximately 3.4%) had a p-value < 0.05). However, given the putative function of the identified genes and the association with gene expression levels we think the associations may be true findings. Further research is necessary to confirm whether the risk to develop respiratory symptoms upon specific occupational exposures depends on the identified genes. This information may offer insight into the cellular and molecular mechanism underlying the development of respiratory symptoms and this may further indicate potential treatment targets especially for the susceptible subgroups.

5. Conclusion

This paper presents a GWIS on the risk of respiratory symptoms cough, dyspnea and phlegm, which investigated interactions between SNPs and several types of occupational exposures (i.e. dust, gases and fumes, pesticides, solvents, and metals). We identified some plausible candidate genes that may be involved in biological pathways leading to respiratory symptoms, i.e. FCERIA, CHN1, TMPRSS9, SERPINH1, TOX3, and ARHGAP18. The next step should be to determine if the identified genes are true susceptibility loci for respiratory symptoms. Findings from this study may eventually contribute to the understanding of pathways underlying the development of respiratory symptoms, which may lead to the identification of novel therapeutic targets and provide targeted protection measures to environmental exposures, especially for the genetically susceptible population.

Acknowledgements

This study was funded by the Groningen Research Institute for Drug Exploration (GUIDE), University Medical Center Groningen, University of Groningen, the Netherlands. The LifeLines Cohort Study, and generation and management of GWAS genotype data for the LifeLines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes Research Foundation. The Vlagtwedde-Vlaardingen cohort study was supported by the Ministry of Health and Environmental Hygiene of the Netherlands and the Netherlands Asthma Fund (grant 187) and the Netherlands Asthma Fund grant no. 3.2.02.51, the Stichting Astma Bestrijding, BBMRI-NL (Complementation project), and the European Respiratory Society COPD research award 2011 to H.M. Boezen. The lung eQTL study at Laval University was supported by the Fondation de l’Institut universitaire de cardiologie et de pneumologie de Québec, the Respiratory Health Network of the FRQS, the Canadian Institutes of Health Research (MOP - 123369). The funding sources were not involved in study design, in collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the article for publication. The authors would like to thank the staff at the Respiratory Health Network Tissue Bank of the FRQS for their valuable assistance with the lung eQTL dataset at Laval University.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2018.11.017.

References

X. Zeng et al.


Cal, S., Quresada, V., Garabaya, C., Lopez-Otin, C., 2003. Polyserase-4, a human poly-
tion to wheat flour in bakery workers. Ann Allergy Asthma Immunol 107, 57–64.


International Labour Organization, 1990. The revised international standard classi-


International Labour Organization, 1990. The revised international standard classi-


International Labour Organization, 1990. The revised international standard classi-


International Labour Organization, 1990. The revised international standard classi-