ERJ Express. Published on July 9, 2008 as doi: 10.1183/09031936.00065308

Genetic variants of microsomal epoxide hydrolase and glutamate-cysteine ligase in COPD

Running Title: EPHX1 and GCL variation in COPD

Sally Chappell[1], Leslie Daly[2], Kevin Morgan[1], Tamar Guetta-Baranes[1], Josep Roca[3], Roberto Rabinovich[3], Juzer Lotya[2],Ann B. Millar[4], Seamas C. Donnelly[5], Vera Keatings[6], William MacNee[7], Jan Stolk[8], Pieter S. Hiemstra[8], Massimo Miniati[9], Simonetta Monti[9] Clare M. O'Connor[5,10] and Noor Kalsheker[1,10].

[1] The University of Nottingham. Division of Clinical Chemistry, Molecular Medical Sciences, Institute of Genetics, University Hospital, Queens Medical Centre, Nottingham, NG7 2UH, UK

[2] University College Dublin. UCD School of Public Health & Population Science, UCD, Belfield, Dublin 4, IRELAND

[3] Hospital Clinico y Provincial de Barcelona. Service de Pneumologia, Hospital Clinic, Villarroel, 170, 08036, Barcelona, SPAIN.

[4] University of Bristol. Lung Research Group, Department of Clinical Science at North Bristol, Southmead Hospital, Westbury on Trym, Bristol BS10 5NB, UK

[5] University College Dublin. UCD School of Medicine and Medical Science, The Conway Institute, UCD, Belfield, Dublin 4, IRELAND

[6] Letterkenny General Hospital, Letterkenny, Co. Donegal, Ireland

Copyright 2008 by the European Respiratory Society.

[7] ELEGI Colt Laboratories, MRC Centre for Inflammation Research, Level 2, Room C2.29, The Queen's Medical Research Institute, 47 Little France Crescent, Edinburgh EH16 4TJ.

[8] Leiden University Medical Center, Department of Pulmonology (C3-P), Albinusdreef 2, P.O. Box 9600, 2300 RC Leiden, THE NETHERLANDS

[9] CNR Institute of Clinical Physiology, Via G. Moruzzi 1-56124, Pisa, ITALY

[10] Joint senior authors.

Corresponding author:

Professor Noor Kalsheker, Division of Clinical Chemistry, University Hospital, Nottingham, NG7 2UH, UK.

Tel. +44 115 8230720

Fax. +44 115 8230722

Email: noor.kalsheker@nottingham.ac.uk

The work was supported by the European Union 5th Framework programme under contract QLG1-CT-2001-01012 (COPD GENE SCAN project) and is protected as intellectual property.

SCD is supported by Science Foundation Ireland (SFI)
JL is supported by a grant from the Health Research Board (Ireland)

Abstract

The genetic factors that contribute to the development of Chronic Obstructive Pulmonary Disease (COPD) are poorly understood. Many candidate genes have been proposed, including enzymes which protect the lung against oxidative stress such as microsomal epoxide hydrolase (EPHX1) and glutamate-cysteine ligase (GCL). To date, most reported findings have been for EPHX1, particularly in relation to functional variants associated with fast and slow metabolism of epoxide intermediates. We aimed to identify any association of variation in these genes with COPD susceptibility or severity.

We genotyped 1,017 Caucasian COPD patients and 912 non-diseased age and sex matched smoking controls for six single nucleotide polymorphisms (SNPs) in EPHX1 (including the fast and slow variants and associated haplotypes), and eight SNPs in the two genes encoding GCL. GCL is a rate-limiting enzyme in the synthesis of glutathione, a major contributor to anti-oxidant protection in the lung.

We found no association of variation in EPHX1 or GCL with susceptibility to COPD or disease severity.

This is the largest reported study to date and is well powered to detect associations that have been previously suggested. Our data indicate that these genetic variants are unlikely to be related to susceptibility or disease severity in COPD in Caucasians.

Key words: Chronic obstructive pulmonary disease, genetic susceptibility, glutamate-cysteine ligase, microsomal epoxide hydrolase,.

Introduction

Chronic obstructive pulmonary disease (COPD) is the sixth leading cause of morbidity and mortality in the western world and the prevalence of the disease is increasing as the population ages [1]. Cigarette smoking is the major environmental contributor to the disease in western societies yet only about 15% of cigarette smokers develop the disease, indicating that genetic factors play a part in determining susceptibility. With the notable exception of severe alpha$_1$-antitrypsin (AAT) deficiency, these genetic factors are poorly understood.

Given that over 95% of those who develop COPD are smokers and oxidative stress is thought to be important in the pathogenesis of the disease, genetic variation in enzymes that protect the lung against smoke-induced oxidative stress has been a significant focus of study. Microsomal epoxide hydrolase (EPHX1), an enzyme involved in the first-pass metabolism of epoxide intermediates, has received particular attention as two functional variants of the gene, which confer slow and fast metabolic activity, have been identified [2]. The slow variant decreases activity by about 40-50% whilst the fast variant increases activity by about 25%. Since the initial report of an association between increased susceptibility to COPD and His-His homozygosity at the "slow" variant locus (Tyr113His in exon 3) [3], subsequent studies have variously supported this association [4-6], or have reported an association with disease severity rather than susceptibility [7], have observed no association between this polymorphism and disease susceptibility or severity [8-10] or have reported a protective effect of this polymorphism [11]. Similar disparities in study outcomes are reported for the "fast" variant (His139Arg in exon 4)

with possession of the minor variant reported as being associated with protection [12,13] or having no effect [3,11].

The lack of replication and inconsistency in these studies most likely reflect a range of issues that can produce false-positive or false-negative outcomes in case-control studies of complex diseases where relatively weak associations between single genes and disease are expected [14]. These include lack of power due to insufficient numbers of cases and controls, poor matching of cases and controls, particularly for smoking [15] and genetic heterogeneity in and between different study populations. A recent study has also shown that, when using restriction fragment length polymorphism (RFLP) techniques, genotyping error arising from the presence of a SNP at Lys119 near the exon 3 'slow' polymorphism can lead to significant overestimation of His113 homozygotes [10] with resulting lack of Hardy-Weinberg equilibrium in study populations. Indeed, in a meta-analysis of previous studies where populations that were not in HW equilibrium were omitted, His113 homozygosity is reported as being protective against development of COPD [11].

To address many of these issues, we have collected samples from a cohort of 1,017 Caucasian COPD patients and 912 non-diseased control smokers, matched for ethnicity, age and sex which represents the largest collection reported to date in a candidate gene approach* [16]. By matching cases and controls from each centre in Europe we have minimised the risk of population stratification. Recent data from genome-wide

* Numbers here differ from those in the original reference because one subject coded as a COPD case was actually a control

association studies have shown that in the British population at least there is very little evidence for stratification in nearly 3000 controls analysed to date; a relatively small number of genes showed significant differences in allele frequencies thus suggesting that the population is relatively homogeneous and not as variable as previously thought [17]. In this study we have used our resource to evaluate the relationship between EPHX1 variants with respect to both the occurrence and severity of COPD. We also assessed the association between genetic variation in glutamate-cysteine ligase (GCL, also known as gamma glutamyl–cysteine synthetase) and risk of developing COPD. This is the rate-limiting enzyme in the synthesis of glutathione, a central contributor to the lung's anti-oxidant defence system, and variation in the catalytic subunit of the gene has recently been shown to be associated with lung disease severity in Cystic Fibrosis [18]. To date genetic variation in this gene has only been reported in one small study of COPD where no association was found [19].

Materials and Methods

Subjects

COPD cases and control subjects were recruited at six European centres. Numbers from each centre were: Barcelona 70 controls and 138 cases; Bristol 152 controls and 129 cases; Dublin 195 controls and 196 cases; Edinburgh 81 controls and 168 cases; Leiden 216 controls and 188 cases and Pisa 198 controls and 198 cases. Approval for the study was obtained from the appropriate committees at each recruitment centre. Informed consent was obtained from all subjects. Criteria for patient recruitment were a firm clinical diagnosis of stable COPD; airflow limitation as indicated by $FEV_1 \leq 70\%$ normal predicted values and FEV_1/ FVC < 70%; no significant reversibility on bronchodilation and a smoking history of ≥ 20 pack years. Patients were excluded from the study if they had an established diagnosis of asthma, lung cancer, a history of atopy, known AAT deficiency or a serum AAT level of less than 1.0g/L. They were also excluded if they had had an acute exacerbation in the 4 weeks preceding assessment for the study.

Control subjects were recruited at each centre to match COPD patients for ethnicity, age, gender and smoking history. Exclusion criteria were as described for cases and also included a family history of COPD. Only individuals with no evidence of airflow obstruction (FEV_1 and FVC ≥ 80% and FEV_1/ FVC > 70%) were included in the control group. Only white Caucasians were recruited for case and controls. Complete matching between cases and controls was not achieved, but this was taken into account during the analysis.

We have previously not found any evidence for population stratification in this sample set, in keeping with the data obtained from a large sample of the British population [16, 17].

SNP Mapping and Haplotypes:

DNA sequencing to identify SNPs was carried out using an approach previously described [16]. Primers were designed for PCR so that the exons and flanking regions of the genes could be screened. These SNPs were then genotyped using Taqman assays in 291 independent Caucasian samples from a previous study of COPD to obtain more robust data on the frequencies of SNPs. All subjects had given informed consent, and the study was approved by the ethics committee. Where SNPs were in full linkage disequilibrium in the sequenced samples a single representative SNP was chosen for genotyping. Where the minor allele frequency of a SNP fell below 5% in the larger group it was removed from further analyses.

Genotyping of Study Samples:

Genotyping of the study population for 6 SNPs in the EPHX1 gene and for 8 SNPs in the GCL genes (5 in the GCLC and 3 in the GCLM subunits) was carried out at Geneservices Ltd using Taqman assays. Primer and probe sequences are available on request. As a quality control measure, the 44 samples of known genotype from sequencing were included. These genotypes were unknown to Geneservices but were known at source. When discrepancies were noted the analysis was repeated. This resulted in 100% concordance for all the assays.

Statistical Analysis:

Analysis of Genetic variation in population groups:

Each of the SNPs in the three genes were analysed for Hardy-Weinberg equilibrium (HWE) using PROC ALLELE in SAS/Genetics Release 9.1. To examine linkage disequilibrium, the correlation coefficient between SNP pairs within each gene in patients and controls was calculated using the same program. Analysis of EPHX1 allele and genotype frequencies were performed using PROC ALLELE and PROC CASECONTROL in SAS/Genetics. As matching was not completely achieved on recruitment, allele and genotype frequencies were adjusted for any residual confounding due to age, sex, smoking and centre using logistic regression. This also gave an increase in power by accounting for the matched design. Interactions of age, sex and smoking with centre were included as previously described [16]. The size of the study gave it the ability to detect small differences between cases and controls. For instance over the SNPs investigated the (rare-type) allele frequency in controls ranged from a low of 7.4% to a high of 26.6% (both SNPs being in the GCL gene - see Table 4). Using these extremes as examples and taking a 1% two-sided level of significance, the study had a power of 95% to detect an allele frequency of 11.4% or greater in the cases compared to the 7.4%, and a frequency of 32.8% or greater compared to the 26.6%. The corresponding detectable percentages for a lower frequency in cases than controls are 4.2% and 20.8%.

Haplotype Analysis

For full analysis of the relationships between COPD and EPHX1 and GCL variants, we used a staged, haplotype-based approach as described previously [16]. We first identified the SNP groupings on which to perform the haplotype analysis. To allow for any combination we screened all the possible SNP groups in the EPHX1 gene and in the modifier and catalytic GCL genes for differences between cases and controls. This was done by performing an omnibus test [22] as implemented in SAS/Genetics PROC HAPLOTYPE and employing a *P*-value of less than 0.01 as a cut-off value for statistical significance. Exact *P*-values based on 10,000 permutations were generated. We excluded extremely rare haplotypes, occurring at a frequency of <1% within each of these groups. Individual haplotypes were compared between COPD cases and controls using the pseudo chi-square test implemented in PROC HAPLOTYPE. A 1% level of significance was taken.

Results

Characteristics of the Study Population

The characteristics of the control and COPD groups recruited to the study are summarised in Table 1. Targeted recruitment of controls yielded groups with similar age and gender profiles, although complete matching of control and COPD groups was not achieved as a high proportion of smokers over the age of 65 had evidence of some obstruction in pulmonary function and were therefore excluded from recruitment as controls. This incomplete matching was adjusted for in the analysis of the data. Within the COPD group those recruited were almost equally distributed across the Global

Obstructive Lung Disease (GOLD) severity categories of moderate, severe and very severe (Supplemental Table 1S).

SNP Mapping:

Microsomal epoxide hydrolase (EPHX1): 25 SNPs were identified in the EPHX1 gene on sequencing (Supplemental Table 2S). Two of these (rs45616640 and rs45505095) had a minor allele frequency of less than 5% so were excluded from further analysis. Two further SNPs (rs45467394 and rs4149227) were unsuitable for Taqman assay design. Complete linkage disequilibrium was observed between SNPs at sites –699/-613/-362 and –290/-200 in the 5' flanking region leaving 19 SNPs with poor LD occurring at a frequency of >5%. As screening for 19 SNPs would result in a large number of haplotypes occurring at a frequency lower than 1%, thus reducing the power of the study, a selection of 6 SNPs was made based on either previous associations or with potential functional variants. The SNPs chosen included three that fully covered variation in the promoter region (rs2854450, rs2854451 and rs3753658), the 'slow' and 'fast' functional SNPs (rs1051740 and rs2234922) that have been shown to affect activity of EPHX1, and one SNP (rs4653695) from the 3' untranslated end of the gene.

Glutamate- cysteine ligase (GCL): Six SNPs were identified in the GCL catalytic subunit (GCLC) and six in the modifier subunit (GCLM). The location and nucleotide substitution of these SNPs are indicated in Table 3S of the supplementary material. In the catalytic subunit, two SNPs (rs2100375 and rs1901773) were in complete linkage disequilibrium, so only the rs2100375 was assessed in further analyses. In 291

independent samples, the remaining 5 SNPs were found to contribute to 9 haplotypes that occurred with a frequency of >1% and accounted for 99% of the observed variation in the gene thus providing almost complete coverage of the gene. These 5 SNPs were genotyped in the study population. In the modifier subunit, three SNPs (rs41303970, rs743119 and rs2273406) were in complete linkage disequilibrium, so rs41303970 was chosen to infer information for all three sites in subsequent analyses. SNP rs35267053 had a minor allele frequency of less than 5%, so was not included in further investigations. The three remaining SNPs were found to contribute to 4 haplotypes that occurred with a frequency of >1% and accounted for 99% of the observed variation in the 291 samples screened. All three SNPs were genotyped in the final sample resource.

Analysis of variation in EPHX1 and GCL genes in COPD and Controls:

The allele and genotype frequencies of the six EPHX1 SNPs that were screened in the full study population are shown in Table 2. For ease of presentation, screened SNPs are numbered sequentially 1 to 6 with the gene location relative to the transcription start site, base substitution and the National Centre for Biotechnology Information (NCBI) dbSNP reference number for each SNP indicated. All of the SNPs were within the accepted criteria for Hardy-Weinberg equilibrium and the frequencies of the SNPs were similar in the populations from all six recruitment centres.

No difference in allele frequency was observed between patient and control groups for any individual EPHX1 SNP nor was there any difference (at the 1% level set for acceptance of significance) in SNP genotype frequencies between patients and controls

(Table 2). Similarly, when classified into functional phenotypes on the basis of genotype at the "slow" and "fast" loci as described by Smith and Harrison [3] no difference was observed between controls and COPD patients. Neither were functional phenotypes found to be associated with disease severity within the patient group (Table 4S in the Supplementary Material). We also evaluated the common 6-SNP haplotypes (occurring with a frequency above 1%) associated with the 'slow' and 'fast' functional variants but found no difference in haplotypes between control and COPD groups (Table 3).

As with the EPHX1 gene, the frequencies of the six SNPs screened in the GCLC and GCLM genes were similar in the populations from all six recruitment centres and, with the exception of SNP 1m (C-590T, modifier subunit), all were within the accepted criteria for Hardy-Weinberg equilibrium (HWE). For SNP 1m, the frequency of the minor allele homozygotes was slightly higher than expected (4.3% vs 2.5%) in the control population ($p = 0.002$). This was unlikely to be due to genotyping error as the blinded controls for this assay were accurately assigned. The frequency of the minor allele is less than 5%, and small deviations in observed numbers may cause significant deviation from HWE. We cannot fully exclude the possibility of a copy number variant, which often contributes to deviations from HWE, although we found no evidence for this in preliminary experiments using real-time PCR (data not shown).

No difference in allele frequency was observed between patient and control groups for any individual SNP in either the GCLC or GCLM genes, nor was there any difference in SNP genotype frequencies (Table 4).

Linkage disequilibrium analysis (see Tables 5S and 6S in Supplementary Material) demonstrated strong linkage between SNPs 2c and 5c (rs2100375 and rs524553) in the GCLC subunit and between all 3 SNPs in the modifier subunit in both patient and control groups. No strong linkage was observed between SNPs evaluated in the EPHX1 gene in either patients or controls.

Haplotype analysis of SNP combinations in each of the three genes did not identify any combinations that were related to COPD. Similarly, when evaluated on the basis of disease severity within the patient group, no association between haplotypes and disease severity was observed for any of the three genes evaluated.

Discussion

The EPHX1 gene has been a significant focus for studies on genetic susceptibility in COPD in the past decade. It is a promising candidate gene as it is strongly expressed in the bronchial epithelium and its biological role in detoxifying epoxides is extremely relevant in protecting the lung against smoke-induced damage. In addition, there is significant individual variation in enzymatic activity, and genetic polymorphisms that modify enzymatic activity have been identified [23].

The majority of studies on genetic variants of EPHX1 in COPD have evaluated polymorphisms in the coding region of the gene that are associated with 'slow' (Tyr113His) and 'fast' (His139Arg) enzymatic activity, with conflicting results. Many of these studies have limitations with respect to power, with numbers of subjects in the patient and control groups between 60 and 200, and matching for ethnicity, age, gender and smoking vary considerably. In addition, those that analysed polymorphic variation using RFLP techniques likely overestimated the proportion of 'slow' (113His) homozygotes [10].

The present study overcomes many of these limitations by evaluating EPHX1 polymorphisms in patient and control groups with sufficient power to readily replicate previously reported associations. All subjects were Caucasian and any potential effects of population stratification was minimised by recruiting patients and controls from the same populations. In addition, targeted recruitment of controls facilitated enrolment of subjects with a similar profile to patients with respect to gender, age and smoking.

Analysis of EPHX1 polymorphisms showed no relationship between susceptibility to COPD and variation at the Tyr113His or His139Arg loci. Given the significantly greater level of power and case-control matching in the present study it is unlikely that effects of the magnitude observed in previous, less powered, studies have been missed. All EPHX1 SNPs evaluated were in Hardy-Weinberg equilibrium in control populations. This, and the inclusion of previously sequenced samples as quality controls in the genotyping analysis, indicates that our analysis is not affected by genotyping error.

Furthermore, detailed analysis of haplotypes corresponding to the fast (His139Arg) and slow (Tyr113His) variants failed to identify haplotypes that were associated with disease. We also analysed variation in the 5' promoter region of the EPHX1 gene which has been shown to modulate enzyme expression by up to 30% [23] but found no association with disease susceptibility.

Our results are at variance with two adequately-powered recent studies which variously report protective effects of the His139Arg [12] and of the Try113His [11] alleles, though neither study replicated the other's results. In a study of 304 emphysema patients and 441 controls, Hersh et al [12] report a protective effect (odds ratio 0.73, p=0.03) with the 'fast' variant of codon 139 while no association was observed with variation in the 113 codon, By contrast, Brogger et al [11] evaluated the same polymorphisms in 492 Caucasian smokers and ex-smokers (244 COPD patients and 248 controls) and report an protective effect of His/His homozygosity at the 113 locus, with no effect of variation at

the 139 locus. Our study did not replicate either of these observations, nor were they noted in the majority of smaller studies [3-10].

Sequencing of the EPHX1 gene revealed that variation in this gene is high, with 25 polymorphisms being identified. This may possibly explain some of the conflicting results reported in studies of the slow and fast variants as these polymorphisms could contribute to disease susceptibility on a background of other specific SNP variants or haplotypes. This, however, seems unlikely as we have captured the most significant haplotypes corresponding to the fast and slow variants. It may be that variation in other genes play a role as studies on the impact of the slow and fast EPHX1 genetic variants indicate that they may have only a modest impact on *in vivo* activity levels and that post-transcriptional modification may be important [23,25]. It seems unlikely therefore that these particular variants contribute to COPD susceptibility though we cannot exclude the possibility of rare variants playing a role and significantly larger studies would be needed to test this reliably.

The rate limiting enzyme in glutathione synthesis, GCL, is also a potential candidate gene for susceptibility to COPD. Significantly decreased glutathione (GSH) levels are observed in the lungs of GCLm knockout mice [26] and expression of message for this subunit is decreased in alveolar macrophages from healthy smokers [27]. In addition, the presence of GCL protein in the bronchial airways appears to be diminished in smokers and COPD patients compared to non-smoking controls [28, 29]. We genotyped both subunits of this enzyme and investigated the role of the genetic variation observed as a

potential risk factor for disease. As with EPHX1, we found no associations with disease and SNPs in genes encoding either subunit of the GCL enzyme. This confirms the only other reported study of GCL genetic variation in COPD, which looked at a single SNP in the GCLC gene (rs17883901; SNP 1c) in 322 subjects from the Han Chinese population, and failed to find evidence for an association with COPD. Our work has extended these previous findings to include more SNPs throughout the GCLC gene, as well as considering variation in the gene for the modifier subunit.

In conclusion, we have found no evidence for a significant association of polymorphisms in EPHX1 or GCL with susceptibility to COPD in a Caucasian population.

Acknowledgements:

We thank Ann Hann (Bristol), Breda Callaghan and Gemma Hogan (Dublin), Joyce Barr (Edinburgh), Clara Kolster-Bijdevaate (Leiden) and Simonetta Monti (Pisa) for assisting with recruitment; G. Catapano E. Fornai and C. Carli for clinical and technical assistance in Pisa; Erika Daly (Dublin) for assistance in data analysis and Dr B. Veldhuisen & JJ Houwing (Leiden) for comments on data analysis.

1. Pauwels RA, Rabe KF. Burden and clinical features of chronic obstructive pulmonary disease (COPD). *Lancet* 2004; 364: 613-20.
2. Hassett C, Aicher L, Sidhu JS, Omiecinski CJ. Human microsomal epoxide hydrolase: genetic polymorphism and functional expression in vitro of amino acid variants. *Hum Mol Genet* 1994; 3: 421-428.
3. Smith CA, Harrison DJ. Association between polymorphism in gene for microsomal epoxide hydrolase and susceptibility to emphysema. *Lancet* 1997; 350: 630-3.
4. Park JY, Chen L, Wadhwa N, Tockman MS. Polymorphisms for microsomal epoxide hydrolase and genetic susceptibility to COPD. *Int J Mol Med* 2005; 15: 443-8.
5. Xiao D, Wang C, Du MJ, Pang BS, Zhang HY, Xiao B, Liu JZ, Weng XZ, Su L, Christiani DC. Relationship between polymorphisms of genes encoding microsomal epoxide hydrolase and glutathione S-transferase P1 and chronic obstructive pulmonary disease. *Chin Med J* 2004; 117: 661-7.
6. Vibhuti A, Arif E, Deepak D, Singh B, Pasha MAQ. Genetic polymorphisms of GSTP1 and mEPHX correlate with oxidative stress markers and lung function in COPD. *Biochem Biophys Res Comm* 2007; 359: 136-142.
7. Yoshikawa M, Hiyama K, Ishioka S, Maeda H, Maeda A, Yamakido M. Microsomal epoxide hydrolase genotypes and chronic obstructive pulmonary disease in Japanese. *Int J Mol Med* 2000; 5: 49-53.
8. Takeyabu K, Yamaguchi E, Suzuki I, Nishimura M, Hizawa N, Kamakami Y. Gene polymorphism for microsomal epoxide hydrolase and susceptibility to emphysema in a Japanese population. *Eur Respir J* 2000; 15: 891-4.

9. Yim JJ, Park GY, Lee CT, Kim YW, Han SK, Shim YS, Yoo CG. Genetic susceptibility to chronic obstructive pulmonary disease in Koreans: combined analysis of polymorphic genotypes for microsomal epoxide hydrolase and glutathione S-transferase M1 and T1. *Thorax* 2000; 55: 121-5.
10. Matheson CM, Raven J, Walters EH, Abramson MJ, Ellis JA. Microsomal Epoxide Hydrolase is not associated with COPD in a community-based sample. *Hum Biol* 2006; 78: 705-17.
11. Brogger J, Steen VM, Eiken HG, Gulsvik A, Bakke P. Genetic association between COPD and polymorphisms in TNF, ADRB2 and EPHX1. *Eur Respir J* 2006; 27: 682-8.
12. Hersh CP, Demeo DL, Lange C, Litonjua AA, Reilly JJ, Kwiatkowski D, Laird N, Sylvia JS, Sparrow D, Speizer FE, Weiss ST, Silverman EK. Attempted replication of reported chronic obstructive pulmonary disease candidate gene associations. *Am J Respir Cell Mol Biol* 2005; 33: 71-8.
13. DeMeo DL, Hersch CP, Hoffman EA, Litonjua AA, Lazarus R, Sparrow D, Benditt JO, Criner G, Make B, Martinez FJ, Scanlon PD, Sciurba FC, Utz JP, Reilly JJ, Silverman EK. Genetic determinants of emphysema distribution in the National Emphysema Treatment Trial. *Am J Respir Crit Care Med* 2007; 176: 42-8.
14. Molfino NA. Genetics of COPD. *Chest* 2004; 125: 1929-40.
15. Lomas DA, Silverman EK. The genetics of chronic obstructive pulmonary disease. *Respir Res* 2001; 2: 20-6.
16. Chappell S, Daly L, Morgan K, Guetta Baranes T, Roca J, Rabinovich R, Millar A, Donnelly SC, Keatings V, Macnee W, Stolk J, Hiemstra P, Miniati M, Monti S,

O'Connor CM, Kalsheker N. Cryptic haplotypes of SERPINA1 confer susceptibility to chronic obstructive pulmonary disease. *Hum Mutat* 2006; 2: 103-109.

17. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared. *Nature* 2007; 447: 661-78.
18. McKone EF, Shao J, Frangolias DD, Keener CL, Shephard CA, Farin FM, Tonelli MR, Pare PD, Sandford AJ, Aitken ML, Kavanagh TJ. Variants in the glutamate-cysteine-ligase gene are associated with cystic fibrosis lung disease. *Am J Respir Crit Care Med* 2006; 174: 415-9.
19. Liu S, Zhou Y, Zong N, Ran P. Genetic analysis of CC16, OGG1, GCLC polymorphisms and susceptibility to COPD. *Respirology* 2007; 12: 29-33.
20. Crapo RO, Morris AH, Gardner RM. *Am Rev Respir Dis* 1981; 123: 659-664.
21. Brand PLP, Quanjer PH, Postma DS, Kerstjens HAM, Koeter GH, Dekhuijzen PNR, Sluiter HJ. Interpretation of bronchodilator response in patients with obstructive airways disease. *Thorax* 1992; 47: 429-436.
22. Zhao JH, Curris D, Sham PC. Model-free analysis and permutation tests for allelic associations. *Hum Hered* 2000; 50: 133-139.
23. Omiecinski CJ, Hassett C, Hosagrahara V. Epoxide hydrolase: Polymorphism and role in toxicology. *Toxicol Lett* 2000; 15: 112-3.
24. Raaka S, Hassett C, Omiencinski CJ. Human microsomal epoxide hydrolase: 5'-flanking region genetic polymorphisms. *Carcinogenesis* 1998; 19: 387-93.
25. Hassett C, Lin J, Carty CL, Laurenzana EM, Omiecinski CJ. Human hepatic microsomal epoxide hydrolase: comparative analysis of polymorphic expression. *Arch Biochem Biophys* 1997; 337: 275-83.

26. Yang Y, Dieter MZ, Chen Y, Shertzer HG, Nebert DE, Dalton TP. Initial characterization of the glutamate-cysteine ligase modifier subunit Gclm(-/-) knockout mouse. Novel model system for a severely compromised oxidative stress response. *J Biol Chem* 2002; 277: 49446-52.
27. Rahman I, MacNee W. Lung glutathione and oxidative stress: implications in cigarette smoke-induced airway disease. *Am J Physiol* 1999; 277: L1067-88.
28. Rahman I, van Schadewijk AA, Hiemstra PS, Stolk J, van Krieken JH, MacNee W, de Boer WI. Localization of gamma-glutamylcysteine synthetase messenger RNA expression in lungs of smokers and patients with chronic obstructive pulmonary disease. *Free Radic Biol Med* 2000; 28: 920-5.
29. Harju T, Kaarteenaho-Wiik R, Soini Y, Sormunen R, Kinnula VL. Diminished immunoreactivity of gamma-glutamylcysteine synthetase in the airways of smokers' lung. *Am J Respir Crit Care Med* 2002; 166: 754-9.

Table 1: Characteristics of Control and COPD subjects recruited to the study

	Control	COPD	*P* -value
Male (%)	63.6%	69.8%	0.0038
Age (yr)	60.8 ± 8.9	65.9 ± 8.2	< 0.0001
Smoking pack years	38.7 ± 17.4	48.7 ± 22.8	< 0.0001
% pred FEV_1	95.4 ± 11.0	43.2± 15.2	< 0.0001
FEV_1/FVC	77.8± 4.9	47.6 ± 12.1	< 0.0001
N	912	1017	

Values indicated are means ± SD

Table 2: EPHX1 – (a) Allele, (b) Genotype and (c) functional phenotype frequencies in Controls and COPD subjects

(a) Minor Allele Frequencies

SNP No	SNP ID	dbSNP reference	Control	COPD	*P* -value[#]	Corrected *P*-value[§]
1	C-399T	rs2854450	0.217	0.206	0.40	0.30
2	G-362A	rs2854451	0.245	0.242	0.82	0.95
3	G-290T	rs3753658	0.124	0.132	0.43	0.63
4	T6651C	rs1051740	0.301	0.294	0.60	0.86
5	A13426G	rs2234922	0.204	0.208	0.76	0.90
6	A20094C	rs4653695	0.107	0.111	0.65	0.91

[#]Crude *P* values; [§]*P* values adjusted by logistic regression for age, gender, centre and smoking

(b) Genotype Frequencies

SNP No	SNP ID	Genotype	Genotype Frequency Control	COPD	*P*-value[#]	Corrected *P*- value[§]
1	C-399T	CC/CT/TT	0.62/0.33/0.05	0.63/0.33/0.04	0.50	0.17
2	G-362A	GG/GA/AA	0.57/0.36/0.06	0.58/0.35/0.07	0.83	0.98
3	G-290T	GG/GT/TT	0.77/0.21/0.02	0.75/0.23/0.02	0.27	0.12
4	T6651C	TT/TC/CC	0.49/0.41/0.10	0.50/0.41/0.09	0.80	0.98
5	A13426G	AA/AG/GG	0.64/0.31/0.05	0.63/0.33/0.04	0.72	0.97
6	A20094C	AA/AC/CC	0.80/0.19/0.01	0.79/0.21/0.01	0.30	0.43

[#]Crude *P* values; [§]*P* values adjusted by logistic regression for age, gender, centre and smoking

(c) Phenotype* frequencies

Phenotype	SNP4/SNP5 Genotype	Control	COPD	*P*-value
Normal	TT/AA or TC/AG	0.433	0.453	0.37
Slow	TC/AA or CC/AA	0.364	0.345	0.41
Fast	TT/AG or TT/GG	0.204	0.201	0.90

*Phenotype Definition:
Normal - No mutation in gene (exon 3:Tyr/Tyr and exon 4: His/His), or heterozygote for both exon 3 (Tyr113/His113) and exon 4 (His139/Arg139) mutations
Slow - One slow (exon 3: Tyr113/His113) allele or two slow alleles.
Fast: at least one fast mutation (exon 4: His139/Arg139 or Arg139/Arg139) and no exon 3 mutation (Tyr113/Tyr113)

Table 3: Haplotypes associated with EPHX1 slow 6651C (113His) and fast 13426G (139Arg) variants with relevant nucleotide substitutions underlined, in COPD and controls. None of the haplotypes were significantly different between COPD and controls. Haplotypes with frequencies less than 1% in both cases and controls are not shown.

	Haplotype	Freq. in Controls	Freq. in COPD	*P*-value
6651C (113His):	CAGCAA	0.0998	0.0955	0.613
	CGGCAA	0.0818	0.0775	0.5974
	TGGCAA	0.0262	0.0349	0.1197
	CGTCAA	0.0195	0.0142	0.2036
13426G (139Arg):	CGGTGA	0.0656	0.0725	0.3944
	TGGTGA	0.0488	0.0357	0.0425
	CAGTGA	0.0247	0.0285	0.3959

Table 4: GCL - (a) Allele and (b) Genotype frequencies in Controls and COPD subjects

(a)

SNP No	SNP	dbSNP reference	Control	COPD	*P*- value[#]	Corrected *P*- value[§]
GCLC (catalytic subunit)						
1c	C-129T	rs17883901	0.074	0.068	0.48	0.67
2c	G24434A	rs2100375	0.279	0.277	0.86	0.65
3c	G36138A	rs16883912	0.102	0.102	0.95	0.41
4c	A37764G	rs1555903	0.091	0.087	0.62	0.58
5c	G39514A	rs524553	0.222	0.231	0.50	0.15
GCLM (modifier subunit)						
1m	C-590T	rs41303970	0.159	0.152	0.62	0.30
2m	C7623T	rs7515191	0.354	0.336	0.26	0.21
3m	G14613T	rs769211	0.266	0.250	0.29	0.25

[#]Crude *P* values; [§]*P* values adjusted by logistic regression for age, gender, centre and smoking

(b)

SNP No	SNP ID	Genotype	Genotype Frequency Control	COPD	*P*- value[#]	Corrected *P*- value[§]
GCLC (catalytic subunit)						
1c	C129T	CC/CT/TT	0.86/0.14/0.004	0.87/0.12/0.01	0.41	0.22
2c	G24434A	GG/GA/AA	0.54/0.37/0.09	0.51/0.42/0.07	0.02	0.12
3c	G36138A	GG/GA/AA	0.80/0.19/0.01	0.81/0.18/0.01	0.57	0.19
4c	A37764G	AA/AG/GG	0.82/0.17/0.004	0.84/0.16/0.01	0.29	0.14
5c	G39514A	GG/GA/AA	0.61/0.34/0.05	0.58/0.39/0.04	0.07	0.05
GCLM(modifier subunit)						
1m	C-590T	CC/CT/TT	0.73/0.23/0.04	0.73/0.24/0.03	0.42	0.31
2m	C7623T	CC/CT/TT	0.43/0.44/0.13	0.44/0.45/0.11	0.34	0.42
3m	G14613T	GG/GT/TT	0.54/0.39/0.07	0.56/0.39/0.06	0.38	0.42

[#]Crude *P* values; [§]*P* values adjusted by logistic regression for age, gender, centre and smoking