ERJ Express. Published on June 18, 2010 as doi: 10.1183/09031936.00091709

Candidate Genes for Chronic Obstructive Pulmonary Disease in two large data sets

Per S. Bakke[1#*], Guohua Zhu[2#], Amund Gulsvik[1], Xiangyang Kong[2], Alvar G.N. Agusti[3], Peter M.A. Calverley[4], Claudio F. Donner[5], Robert D. Levy [6] ,Barry J. Make[7], Peter D. Paré[6], Stephen I. Rennard[8], Jørgen Vestbo [9], Emiel F.M. Wouters[10], Wayne Anderson[2], David A. Lomas[11], Edwin K. Silverman[12], Sreekumar G. Pillai[2]

[1]Department of Thoracic Medicine, Haukeland UniversityHospital and Institute of Medicine, University of Bergen, Bergen, Norway; [2]GlaxoSmithKline R&D, Research Triangle Park, NC, USA; Current address, Roche Pharmaceuticals, Nutley, NJ [3]Hospital Universitari Son Dureta, Fundación Caubet-Cimera and Ciber Enfermedades Respiratorias, Spain, [4]University of Liverpool, U.K., [5]Division of Pulmonary Disease, S. Maugeri Foundation, Veruno (NO), Italy; [6] Division of Respiratory Medicine and James Hogg iCAPTURE Centre for Cardiovascular and Pulmonary Research, University of British Columbia, St. Paul's Hospital, Vancouver, Canada, [7]National Jewish Medical and Research Centre, Denver, Colorado, USA, [8]University of Nebraska, Omaha, USA, [9]Department of Cardiology and Respiratory Medicine, Hvidovre Hospital, Copenhagen, Denmark and University of Manchester, Manchester, UK, [10]Department of Respiratory Medicine, University Hospital Maastricht, Netherlands; [11]Department of Medicine, University of Cambridge, Cambridge, U.K'; [12] The Channing Laboratory and Pulmonary and Critical Care Division, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

# Per S. Bakke and Guohua Zhu contributed equally to the paper and are both joint first authors

*** Correspondence and requests for reprints**:

Per S. Bakke.
[1]Department of Thoracic Medicine, Haukeland University Hospital, N-5021 Bergen, Norway
Phone: +47 55973245
FAX: +47 55975149

Copyright 2010 by the European Respiratory Society.

Abstract

Lack of reproducibility of findings has been a criticism of genetic association studies in complex diseases like chronic obstructive pulmonary disease (COPD).

We selected 257 polymorphisms of 16 genes with reported or potential relationshipsto COPD and genotyped these variants in a case-control study which included 953 COPD cases and 956 control subjects. We explored the association of these polymorphisms to three COPD phenotypes: a COPD binary phenotype and two quantitative traits (post bronchodilator FEV_1 in percent predicted and FEV_1/FVC). The polymorphisms significantly associated to these phenotypes in this first study were tested in a second, family based, study that included 635 pedigrees with 1910 individuals.

Significant associations to the binary COPD phenotype in both populations were seen for STAT1 (rs13010343) and NFKBIB/SIRT2 (rs2241704) ($p< 0.05$). SNP rs17467825 and rs1155563 of the GC gene were significantly associated with FEV_1 in percent predicted and FEV_1/FVC, respectively in both populations ($p< 0.05$).

This study has replicated associations to COPD phenotypes in the STAT1, NFKBIB/SIRT2, and GC genes in two independent populations, the associations of the former two genes representing novel findings.

Abstract word count: 179

Key words: genetic association; replication; COPD; smoking;

Introduction

The prevalence of Chronic Obstructive Pulmonary Disease (COPD) in Western Europe is approximately 10% (1), and COPD is expected to be the third most significant cause of death world-wide by 2020 (2). The most important risk factor for COPD is smoking and there is a dose-response relationship between smoking exposure and reduced lung function, although there is a substantial heterogeneity in the degree of lung function impairment(3). Only a subset of smokers develops an accelerated rate of decline in lung function that leads to COPD. In addition there appears to be familial clustering of both impaired lung function and COPD (4). These insights suggest that susceptibility to COPD may be influenced by genetic factors. The only well-established genetic cause of COPD, α1-antitrypsin deficiency, is present in only 1-2% of individuals with COPD.(5). A number of studies have been performed to find other genetic susceptibility factors for COPD. So far hundreds of candidate genes have been tested.

Until now it has been difficult to replicate genetic findings from one COPD study to another (6). There may be several explanations for this lack of reproducibility, including small sample sizes, lack of Hardy-Weinberg equilibrium, poor phenotype characterisation of the COPD cases, and geneticheterogeneity (7). It is now recommended that the findings of a genetic association should be replicated in another sample before being published (7); nevertheless, only a minority of genetic COPD studies meets this requirement.

To obtain further insight into the genetic basis of COPD, we replicated the relationships of a number of potential COPD candidate genes in two large, independent and well-characterised populations. We selected 257 single-nucleotide polymorphisms (SNPs) in 16 genes based on reported or potential relationship to COPD. They were analysed in a case-control sample from Bergen, Norway, including 953 COPD cases and 956 controls (8). SNPs with significant associations to COPD were then tested using family-based association analysis in 635

pedigrees with 1910 individuals from the International COPD Genetics Network (ICGN), which is the largest family-based COPD collection reported to date (8).

Methods

Study Subjects

The Norwegian case-control study initially included 953 cases; 189 were recruitedfrom two community studies (9), while the rest were recruited from a registry at Haukeland University Hospital, Bergen. The study also included 956 controls; 735 were recruited from the two community studies while the 221 were volunteers. The inclusion criteria for COPD cases was a post-bronchodilator FEV_1 < 80% predicted and FEV_1/FVC < 0.7. The controls were selected based on post-bronchodilator FEV_1 > 80% predicted and FEV_1/FVC> 0.7. Both cases and controls were Caucasians with at least 2.5 pack-years of smoking history (current or ex smokers).

In the multi-center ICGN study, subjects with known COPD were recruited as probands, and siblings and available parents were ascertained through the probands (8, 10). The probands were recruited from pulmonary and medical clinics and hospital admissions. Inclusion criteria for probands were airflow limitation (post-bronchodilator FEV_1< 60% predicted and FEV_1/VC (vital capacity) < 90% predicted) at a relatively early age (45 to 65 years), a ≥ 5 pack-year smoking history, and at least one eligible sibling (with ≥ 5 pack-year smoking history). COPD in siblings was defined by a post-bronchodilator FEV_1< 80% predicted and FEV_1/VC< 90% predicted. 1910 Caucasian individuals from 635 pedigrees were included in the family-based association analysis.

Phenotyping

Three phenotypes were defined: 1. The binary COPD phenotype defined according to the criteria above; 2. post bronchodilator FEV_1 in percent predicted; 3. FEV_1/(F)VC.

Candidate Genes Selection and Genotyping

Sixteen candidate genes were selected for analyses based on their potential biological relevance to pathways that may cause COPD, or their proximity to genes with a known relation to COPD (see Table E1 for details of each gene).

. Eight of the 16 genes had previously been shown to be associated with COPD (GC, GSTP1, HDAC2, HDAC5, HMOX1, IL11, JAK3). Genotyping in the two cohorts was performed with the Illumina array-based custom SNP genotyping platform. The selection of the candidate genes and the analyses were performed prior to the genome wide association study that has recently been published using a subset of these subjects (11), which used Illumina's HumanHap550 genotyping BeadChip for the Bergen cohort and Sequenom's iPLEX SNP genotyping protocol developed for measurement with the MassARRAY mass spectrometer for the ICGN study. The six SNPs rs13010343, rs1609181, rs802372, rs10278590, rs8065686, rs4802898, were also present at the Illumina Human Hap550 chip from the genome wide association study (11).

For the selection of SNPs, linkage disequilibrium (LD) bins were established using an r^2 threshold of 0.8. The tagging SNP selection was based on HapMap data for European-Americans (CEU) with a minor allele frequency (MAF) > 5% from the public database (http://www.hapmap.org). Non-synonymous SNPs with any MAF were included. SNPs in these genes were selected and genotyped in both the ICGN family population and the Norwegian case-control population using the Illumina array-based custom SNP genotyping platform. Hardy-Weinberg equilibrium (HWE) was performed for all SNPs in the control data by using the chi-square goodness-of-fit test with SAS software 8.2; HWE for all SNPs was also tested in the family data using PBAT version 3.5 (9). All SNPs (p values > 0.05) were in Hardy-Weinberg equilibrium in both the family data and the case-control data. COPD family data were evaluated for inconsistent Mendelian inheritance using the PedCheckprogram (12). A complete list of the genes, and SNPs tested in each gene is given in table E2

Statistical Methods

In the case-control population, two models were used in the association analysis. A logistic regression model for the COPD binary phenotype and a linear regression model for the quantitative phenotypes (FEV_1 and FEV_1/FVC) with covariates including age, sex and pack-years of smoking. For the quantitative trait analysis, only COPD cases were included. The analyses were done using SAS software 8.2 with an additive genetic model. FBAT version 1.7.3 (13) was used for the family-based single-SNP association analysis of the COPD binary phenotype in the ICGN family study. The analyses of quantitative traits (FEV_1 and FEV_1/VC) were performed with covariates including center, age, sex, height, and pack-years of cigarette smoking using PBAT version 3.5 (13). Biallelic tests were conducted for SNPs using an additive genetic model. The risk allele was determined from the FBAT Z statistic. Haplotype analyses were conducted using the HBAT function of the FBAT program with the use of Monte Carlo sampling for COPD, FEV_1 and FEV_1/VC (14) in the family data. In the case-control data, haplotype analysis was performed using the expectation-maximization algorithm and score tests, implemented in Haplo.stats program (15). The linkage disequilibrium (LD) structure was examined with the program, Haploview (16, 17). We used a p value< 0.05 in both COPD populations to define statistical significance.

We assessed the power to detect significant associations between the genes and the phenotypes based on the following assumptions: We assumed that the allele frequency of disease gene and the marker is 0.1, respectively, and the penetrances for three genotypes of the marker were 0.7, 0.4 and 0.1. At significance level of 0.05, our study had 99.52% and 85.43% power for COPD case-control and family data, respectively, to detect an association.

Results

Study participants

Characteristics of the participants of the two studies are given in table 1. In the case-control study the cases comprised more men and they were older and reported a higher smoking exposure than the controls. The cases had on average a moderate to severe airflow limitation (mean post bronchodilator FEV_1 = 50.3% predicted). In the IGCN study the probands were predominately males, of the same mean age, but with a greater smoking history than the siblings. The probands had on average a severe airflow limitation (mean post bronchodilator FEV_1 = 36.2% predicted).

Table 1. Characteristics of the Norwegian case-control study and the International COPD Genetics Network (ICGN) study

	COPD Case-Control Study		ICGN Family Study	
	Cases	Controls	Probands	Siblings
Subjects, n	953	956	610	1300
Age (± SD), years (± SD)	65.5 (± 10.1)	55.5 (± 9.7)	58.4 (± 5.5)	58.1 (± 9.8)
Female (%)	370 (38.8%)	477 (49.9%)	247 (40.5%)	648 (49.8%)
Post-FEV_1 L (± SD)	1.58 (± 0.71)	3.24 (± 0.74)	1.11 (± 0.44)	2.36 (± 0.98)
Post-FEV_1, % predicted (± SD)	50.3 (± 17.5)	94.1 (± 9.2)	36.3 (± 12.9)	77.5 (± 25.9)
Post-FEV_1/FVC ratio (± SD) [11i]	0.51 (± 0.13)	0.79 (± 0.04)	0.37 (± 0.12)	0.61 (± 0.15)
Pack-years of smoking (± SD)	32.0 (± 18.6)	19.1 (± 13.2)	51.5 (± 26.8)	40.6 (± 24.9)
Current smoking status (%)	457 (48.1%)	407 (42.6%)	207 (33.9%)	661 (50.8%)

[11i]Note: FEV_1/VC was used for the ICGN Family Study, with VC determined by the higher of FVC and SVC

Single marker andhaplotype analysis

The relationships of the single SNP associations with COPD, post bronchodilator FEV_1 in percent predicted, and the ratio of FEV_1 to FVC (FEV_1 to VC in the IGCN study) are shown in the tables 2-4, respectively. Only genes with significant associations in the case-control study are displayed.

The following genes included SNPs that were significantly associated with the binary COPD phenotype in either study: STAT1, GC, MAP3K5, KCND2, RARRES2, HADC5, SIRT2, NFKBIB. Only one SNP in STAT1 (rs13010343) and one SNP in NFKBIB/SIRT2 (rs2241704) were associated with COPD in both studies (table 2). The risk allele was the same in both datasets for rs13010343, and also for rs2241704.

Table 2.

Single SNP **Association Analysis with COPD**. SNPs with P-values ≤ 0.05 in the COPD case-control study are listed. Bold SNPs represent significant replication in both populations.

		COPD Case Control Study			ICGN Family Study		
Gene	SNP Name	Risk Allele	Odd Ratio (95% CI)	P value	Risk Allele	Fam #	P value
STAT1,	RS12468579	G	1.18 (1.01-1.37)	0.0307	G	220	0.5505
STAT1	**RS13010343**	**G**	**1.24 (1.01-1.55)**	**0.0483**	**G**	**113**	**0.0371**
KCND2	RS1861064	G	0.79 (0.62-0.99)	0.0469	G	123	0.3622
KCND2	RS1609181	T	0.69 (0.51-0.95)	0.0225	T	61	0.4720
KCND2	RS802358	A	0.79 (0.63-0.99)	0.0432	A	123	0.5845
KCND2	RS802372	C	0.73 (0.54-0.97)	0.0311	C	77	0.8047
KCND2	RS4730979	G	1.22 (1.03-1.44)	0.0162	A	194	0.1882
RARRES2,	RS10278590	G	1,17 (1.01-1.37)	0.0380	T	217	0.6587
HDAC5	RS8065686	C	1.48 (1.12-1.95)	0.0051	T	75	0.6582
SIRT2	RS4802998	A	1.17(1.01-1.37)	0.0402	A	205	0.8746
NFKBIB	**RS2241704**	**T**	**0.81 (0.67-0.96)**	**0.0214**	**T**	**172**	**0.0442**

SIRT2							

* Fam # = number of informative families.

The genes significantly associated with post-bronchodilator FEV_1 percent predicted were STAT1, GC, AGER, MAP3K5, KCND2, JAK3, SIRT2, NFKBIB/SIRT2, IL11, HMOX1) in either study. Only GC (rs17467825) was replicated in both studies (table 3), with negative regression coefficients indicating a risk for impaired FEV1.

Table 3. Single SNP **Association with FEV1**. SNPs with P-values ≤ 0.05 in the COPD case-control study are listed. Bold SNPs represent significant replication in both populations

		COPD Case Control Study			ICGN Family Study		
Gene	SNP Name	Risk Allele	β (95 % CI)	P value	Risk Allele	Fam #	P value
NFE2L2	RS2706110	G	-0.071 (-0.141- -0.002)	0.0438	G	228	0.5608
STAT1	RS3771300	C	-0.073 (-0.145- -0.002)	0.0445	C	148	0.3155
STAT1	RS13010343	A	0.113 (0.020- 0.207)	0.0170	A	176	0.1692
STAT1	RS16833177	T	-0.070 (-0.138- -0.003)	0.0401	C	231	0.2745
GC	**RS17467825**	**A**	**-0.067 (-0.127- -0.008)**	**0.0257**	**A**	**298**	**0.0462**
GC	RS2282680	G	-0.069 (-0.129- -0.010)	0.0222	G	297	0.0717
GC	RS4588	C	-0.071 (-0.131- -0.012)	0.0187	C	298	0.0665
GC	RS1155563	T	-0.078 (-0.136- -0.021)	0.0078	T	300	0.1389
AGER	RS3131297	T	0.114 (0.048- 0.179)	0.0006	T	259	0.6431
AGER	RS2269423	T	-0.066 (-0.118- -0.012)	0.0148	T	332	0.7342
AGER	RS204994	A	0.081 (0.014- 0.147)	0.0173	A	272	0.3190
AGER	RS204993	T	0.084 (0.022- 0.146)	0.0078	T	303	0.5362
KCND2	RS7793864	A	-0.104 (-0.205- -0.004)	0.0411	T	123	0.7888

KCND2	RS802340	T	-0.086 (-0.162- -0.010)	0.0263	G	221	0.8958
SIRT2	RS7257949	A	0.105 (0.002- 0.209)	0.0463	A	117	0.4301
SIRT2	RS11879029	G	0.108 (0.002- 0.214)	0.0460	G	118	0.3010
SIRT2	RS7251489	G	0.109 (0.004- 0.214)	0.0416	G	116	0.3912
SIRT2	RS10410544	C	-0.063 (-0.118- -0.009)	0.0217	C	319	0.5244
NFKBIB, SIRT2	RS2053071	G	0.066 (0.012-0.122)	0.0175	G	311	0.9974
IL11	RS4252552	G	0.107 (-0.007- 0.221)	0.0675	G	109	0.0489
HMOX1	RS17885925	C	-0.143 (-0.271- -0.015)	0.0281	C	76	0.3921
HMOX1	RS11912889	A	-0.147 (-0.274- -0.019)	0.0235	A	76	0.3283

* Fam # = number of informative families.
β : regression coefficient.

Regarding the associations with FEV_1/FVC or FEV_1/VC, significant associations were noted for STAT1, GC, AGER, HDAC2, MAP3K5, KCND2, and JAK3 in either study. Only GC (rs1155563) was observed in both samples (table 4), with negative regression coefficients indicating a risk for impaired ratio.

.

Table 4.
Single SNP **Association with FEV1/FVC**[11†]. SNPs with P-values ≤ 0.05 in the COPD case-control study are listed. Bold SNPs representsignificant replication in both populations

		COPD Case Control Study			ICGN Family Study		
Gene	Name	Risk Allele	β(95 % CI)	P value	Risk Allele	Fam#	P value
STAT1	RS13010343	A	0.024 (0.004- 0.044)	0.0178	A	176	0.3393
GC	**RS1155563**	**T**	**-0.014 (-0.0275- -0.002)**	**0.0203**	**T**	**300**	**0.0281**
AGER	RS3131297	T	0.024 (0.0106- 0.039)	0.0006	T	259	0.2529
AGER	RS2269423	T	-0.014 (-0.025- -0.003)	0.0138	T	332	0.4500

AGER	RS204994	A	0.021 (0.007- 0.036)	0.0030	A	272	0.1003
AGER	RS204993	T	0.018 (0.005- 0.032)	0.0063	T	303	0.1651
MAP3K5	RS4363056	T	0.013 (0.002- 0.0254)	0.0228	T	330	0.2536
KCND2	RS802340	T	-0.019 (-0.036- -0.003)	0.0194	G	221	0.3357

[11†]Note: FEV_1/VC was used for the ICGN Family Study, with VC determined by the higher of FVC and SVC.
*Fam#=number of informative families.
β: regression coefficient.

No SNP had replicated associations across the two studies for all the COPD phenotypes.

We also performed haplotype analyses. The results supported the findings of the SNP analyses, but no additional significantly associated genes were identified (results not shown)

Linkage Disequilibrium Analysis

Figures 1-3 show the pair wise LD (r2) values for the SNPs in the genes STAT1, NFKBIB/SIRT2 and GC. The LD structure appeared generally quite similar across these two study populations. Three haplotype blocks were identified in the LD map of the STAT1 gene in both populations with significant SNPs given in bold text. Two haplotype blocks and one haplotype block were revealed for the NFKBIB/SIRT2 gene IGCN and case/control samples, respectively, while 3 haplotype blocks were observed in the two data sets for the GC gene. The 6 SNPs which were also present in the Illumina chip from the GWAS study was not in significant LD (r2 $\geq$ 0.8) with any other SNPs tested in this study.

Discussion

In the present study, we selected 257 polymorphisms in 16 genes with reported or potential relationship to COPD, and genotyped the variants in a case-control COPD study. Significant associations were tested in a family-based COPD study. We detected that STAT1, NFKBIB/SIRT2 and GC genes were associated with COPD-related phenotypes in the two data sets. To our knowledge this is the first study to replicate the same SNPs of the STAT1 and NFKBIB/SIRT2 genes with COPD-related phenotypes in two populations. These genes have multiple potential relationships with established disease mechanisms in COPD.

Regarding the mechanisms by which STAT1 exhibits its functions is unclear. Progression of COPD is associated with increased numbers of $CD8^+$ T lymphocytes and B lymphocytes in the airways supporting a role for these cells in the pathophysiology of COPD (18). It is known that $CD8^+$ T lymphocytes express the chemokine receptor CXCR3 together with IFN –γ (18). The chemokines act as ligands at the CXCR3 receptor, peaking 8-12 hours after stimulation with IFN –γ (19). This pathway is thought to be dependent on STAT1 (20). It is worth noting that Tudhope et al observed that this pathway was dexamethasone resistant (20), a finding in keeping with the limited impact of corticosteroids on the inflammatory profile in COPD (21).

The two genes NKFBIB and SIRT2 are closely located (according to HAPMAP). The SNP RS2241704 is located in the flanking regions of the genes and we cannot distinguish them based on the current data. Hence, the observed association to COPD may work through either gene.

Increasing evidence supports a key role for the transcription factor nuclear factor (NF)-κB in the host response to pneumococcal infection (22). Control of NF-κB activity is achieved through interactions with the IκB family of inhibitors, encoded by the genes NFKBIA,

NFKBIB, and NFKBIE. COPD patients are frequently have lower respiratory tract colonisation with *S. pneumonia* and this chronic infection may play a role in the pathogenesis of COPD. Functional polymorphisms in the NFKBIB gene may affect the host response to the chronic bacterial infection and thereby the chronic airway inflammation.

The SIRT2 gene may be related to degenerative process through deacteylation of the alfa – tubulin (23). The SIRT genes are believed to be involved in the ageing process (24, 25). Emphysema may be regarded as a premature ageing of the lung with loss of elastic fibres (26).

GC codes for the vitamin D binding protein (DBP) that may have several functions. The major function is binding, solubilization and transport of vitamin D and its metabolites (27). It is also reported to augment the chemotactic effect of complement derived molecules on neutrophils (28). Neutrophils play an important role in parenchymal destruction and airway inflammation in COPD. Another important function of DBP is its deglycosylation to DBP macrophage activating factor. The absence of a glycosylated residue at position 420 in GC2 inhibits this conversion. This may be a partial explanation of its protective effect (29). Unlike the GC2-allele, the homozygous GC1F-phenotype is a significant risk factor for the development of COPD (30, 31). Although the GC1F-allele has no effect on the age of onset of COPD, the annual decline in FEV_1 has been reported to be significantly higher in patients with this allele. High resolution CT-parameters show that GC1F-allele carriers suffer from more severe emphysema (28). Black et al. recently analysed serum 25(OH)-vitamin D3 concentration, FEV_1 and FVC of 14091 subjects (age≥20) and demonstrated a significant correlation between these parameters (32).

Several genes were only significantly associated with the COPD-related traits in one of the two studies. This does not necessarily imply that there is no true relationship between these genes and COPD. It especially relates to the genes of which the SNPs reached borderline significance in the other data set, for instance RS2282680 and RS4588 or the GC gene (table 3). This makes it less likely that these associations are false positive. Although the three phenotypes of COPD used in the analyses were identical across the samples (except for VC versus FVC), they are still based solely on spirometry. It is well acknowledged that FEV_1 and FVC far from reflect the whole picture of COPD (33). Hence, the COPD cases from the two populations might differ with respect to other important characteristics of the disease such as degree of emphysema and chronic bronchitis, systemic inflammation, body mass index, respiratory failure and rate of exacerbations. More specific COPD related phenotypes than applied in the present study might have revealed stronger associations to the genes and enhanced the reproducibility of the relationships across the samples.

In both samples the participants were examined only once. Hence, we have no data regarding disease progression and rate of lung function decline. Such data might have further strengthened the associations and reproducibility of the findings.

The major exposure in COPD, active smoking, was taken into account in the analyses based on questionnaire responses. However, distribution of other environmental risk factors such as occupational airborne exposure as well as indoor and outdoor air pollution may differ between the samples and affect the observed associations between the genes and COPD. Passive smoking *in utero* or in early life may cause airway disease in adults (34). The degree of passive smoking and at what time in life it occurs may vary across the samples and potentially influence any gene-environmental interaction effect on COPD.

A common cause of lack of reproducible findings is small sample sizes (35). In our study limited statistical power is not likely to explain lack of replication. The size of the population

in these studies is appropriate to identify even modest effects. Random genotyping errors will cause a non-differential misclassification and bias any association towards nil (36, 37). However, systematic errors may cause a false positive association. Deviation from Hardy Weinberg Equilibrium (HWE) in the control group may be a sign of genotyping error (35). However, in the current study all the SNPs tested were in HWE in both data sets.

Obviously, the possibility exists that there are true differences in genetic determinants for COPD in various populations. Several genetic COPD association studies have seen inconsistent results between Caucasian and Asian populations (38). It is worth noting though that in both the ICGN and the Bergen data sets of the current study all of the participants were self-reported Caucasians.

The potential weaknesses of the study should be acknowledged. First, the case-control population is from a single centre in Norway, and the age and pack-years of smoking were different between the cases and controls. Therefore, we used a logistic regression model for the association analysis, at least partly correcting for age and pack-years of smoking. The fact that the case- control population is from a single centre can be considered as an advantage also, because the population is more homogeneous and the possibility of false positive results by population stratification is minimal. The analysis of population stratification using a set of 257 unlinked SNPs showed no evidence for population stratification. Second, no SNPs were significantly associated with COPD after correction for multiple comparisons. However, the replication in two independent populations and validation by haplotype analysis suggest that the association results are valid. Although various procedures have been studied for correction of multiple testing, including Bonferroni correction and permutation testing, there has not been an ideal statistical framework to deal with raw p values from SNP association analyses (39), especially for replication of a previously reported association result. (40). Here

we are using the replication of the results in an independent cohort to validate our primary findings

In conclusion, we conducted a robust genetic association study and found that variants in the STAT1, NFKBIB/SIRT2 and GC genes are likely to contribute to the susceptibility to COPD. Functional tests need to be performed to find the molecular mechanism that drives the genetic association between COPD phenotypes and the above three genes.

References

1. Buist AS, McBurnie MA, Vollmer WM et al. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. Lancet 2007; 370: 741-750 PubMed .
2. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global burden of disease and risk factors. Washington: The World Bank, 2006
3. Hersh CP, Demeo DL, Lazarus R et al. Genetic association analysis of functional impairment in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2006; 173: 977-984.
4. Seifart C, Plagens A. Genetics of chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis. 2007; 2:541-550
5. Ganrot PO, Laurell CB, Erikson S. Obstructive lung disease and trypsin inhibitors in alpha-1-antitrypsin deficiency. Scand J Clin Lab Invest 1967; 19: 205-208 PubMed .
6. Hersh CP, Demeo DL,, Lange C et al. Attempted replication of reported chronic obstructive pulmonary disease candidate gene associations. Am J Respir Cell Mol Biol. 2005 ;33:71-78.
7. Hall IP, Blakey JD. Genetic association studies in Thorax. A guide to assessing the validity of genetic association studies in respiratory disease. Thorax 2005; 60: 357-359 PubMed PubMed .
8. Zhu G, Warren L, Aponte J et al. The SERPINE2 gene is associated with chronic obstructive pulmonary disease in two populations. Am J Respir Crit Care Med. 2007; 176:167-173.
9. Johannessen A, Omenaas ER, Eide GE, Bakke PS, Gulsvik A. Feasible and simple exclusion criteria for pulmonary reference populations. Thorax 2007; 62: 792-798 PubMed PubMed
10. Patel BD, Coxson HO, Pillai SG, Calverley PM, Donner CF, Make BJ et al. Airway wall thickening and emphysema show independent familial aggregation in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2008; 178: 500-5 PubMed .
11. Pillai SG, Ge D, Zhu G, Kong X, Shianna KV, Need AC, Feng S, Hersh CP, Bakke P. Gulsvik A, Ruppert A, Lødrup Carlsen KC, Roses A, Anderson W, Rennard SI, Lomas DA, Silverman EK, Goldstein DB; ICGN Investigators. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci PLoS Genet. 2009 Mar;5 (3):e1000421. Epub 2009 Mar 20.
12. Lange C, DeMeoD, Silverman EK,Weiss ST, Laird NM. PBAT: tools for family-based association studies. Am J Hum Genet 2004; 74: 367– 369 PubMed PubMed .
13. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998; 63: 259–266 PubMed PubMed .
14. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM. Familybased tests for associating haplotypes with general phenotype data:application to asthma genetics. Genet Epidemiol 2004; 26: 61–69 PubMed PubMed .
15. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet2002;70:425– 434 PubMed PubMed .
16. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis andvisualization of LD and haplotype maps. Bioinformatics 2005;21:263–265.

17. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, *et al.* The structure of haplotype blocks in the human genome. *Science* 2002;296:2225–2229.
18. Hogg JC, Chu F, Utokaparch S et al. The nature of small-airway obstruction in chronic obstructive pulmonary disease. N Engl J Med. 2004; 350: 2645-2653 PubMed .
19. Saetta M, Mariani M, Panina-Bordignon P et al. Increased expression of the chemokine receptor CXCR3 and its ligand CXCL10 in peripheral airways of smokers with chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2002 May 15;165(10 PubMed):1404-1409.
20. Tudhope SJ, Cartley MC, Fenwick PS et al. The role of IkappaB kinase 2, but not activation of NF-kappaB, in the release of CXCR3 ligands from IFN-gamma-stimulated human bronchial epithelial cells. J Immunol. 2007; 179: 6237-6245 PubMed .
21. Hattotuwa KL. Gizycki MJ, Ansari TW, Jeffery PK, Barnes NC. The effects of inhaled fluticasone on airway inflammation in chronic obstructive pulmonary disease: a double-blind, placebo-controlled biopsy study. Am J Respir Crit Care Med. 2002; 165: 1592-1596.
22. Chapman SJ, Khor CC, Vannberg FO et al. IkappaB genetic polymorphisms and invasive pneumococcal disease. Am J Respir Crit Care Med. 2007; 176: 181-187 PubMed .
23. Rajendrasozhan S, Yang SR, Kinnula VL, Rahman I. SIRT1, an anti-inflammatory and antiaging protein, is decreased in lungs of patients with chronic obstructive pulmonary disease. Am J Respir Dis 2008; 177: 861-70 PubMed PubMed .
24. Outeiro TF, Kontopoulos E, Altmann SM, Kufareva I,Strathearn KE, Amore AM, et Sirtuin 2 inhibitors rescue alpha-synucleinmediated toxicity in models of Parkinson's disease. Science 2007; 317:516-519 PubMed PubMed .
25. North BJ, Marshall BL, Borra MT, Denu JM, Verdin E: The human Sir2 ortholog, SIRT2, is an NAD+-dependent tubulin deacetylase. Mol Cell 2003; 11: 437-444 PubMed PubMed .
26. Milne JC, Denu JM. The Sirtuin family: therapeutic targets to treat diseases of aging Curr Opon Chem Biol 2008: 12: 11–17 PubMed
27. Daiger SP, Schanfield MS, Cavalli-Sforza LL. Group-specific component (Gc) proteins bind vitamin D and 25-hydroxyvitamin D. Proc Natl Acad Sci USA 1975; 72: 2076–2080 PubMed .
28. Speeckaert M, Huang G, Delanghe JR, Taes YE. Biological and clinical aspects of the vitamin D binding protein(Gc-globulin) and its polymorphism. Clin Chim Acta. 2006; 372: 33-42 PubMed .
29. Horne Sl, Cockcroft DW, Dosman JA. Possible protective effect against chronic obstructive airways disease by the GC2 allele. Hum Hered. 1990; 40: 173-176 PubMed .
30. Ito I, Nagai S, Hoshino Y et al. Risk and severity of COPD is associated with the group-specific component of serum globulin 1F allele. Chest. 2004; 125: 63-70.
31. IshiiT, Keicho N, Teramoto S et al. Association of Gc-globulin variation with susceptibility to COPD and diffuse panbronchiolitis. Eur Respir J. 2001;18: 753-757.
32. Black PN, Scragg R. Relationship between serum 25-hydroxyvitamin D and pulmonary function in the third national health and nutrition examination survey. Chest 2005; 128: 3792–3798 PubMed PubMed
33. Celli BR. Roger s. Mitchell lecture. Chronic obstructive pulmonary disease phenotypes and clinical relevance. Proc Am Thorac Soc. 2006; 3: 461-465 PubMed .

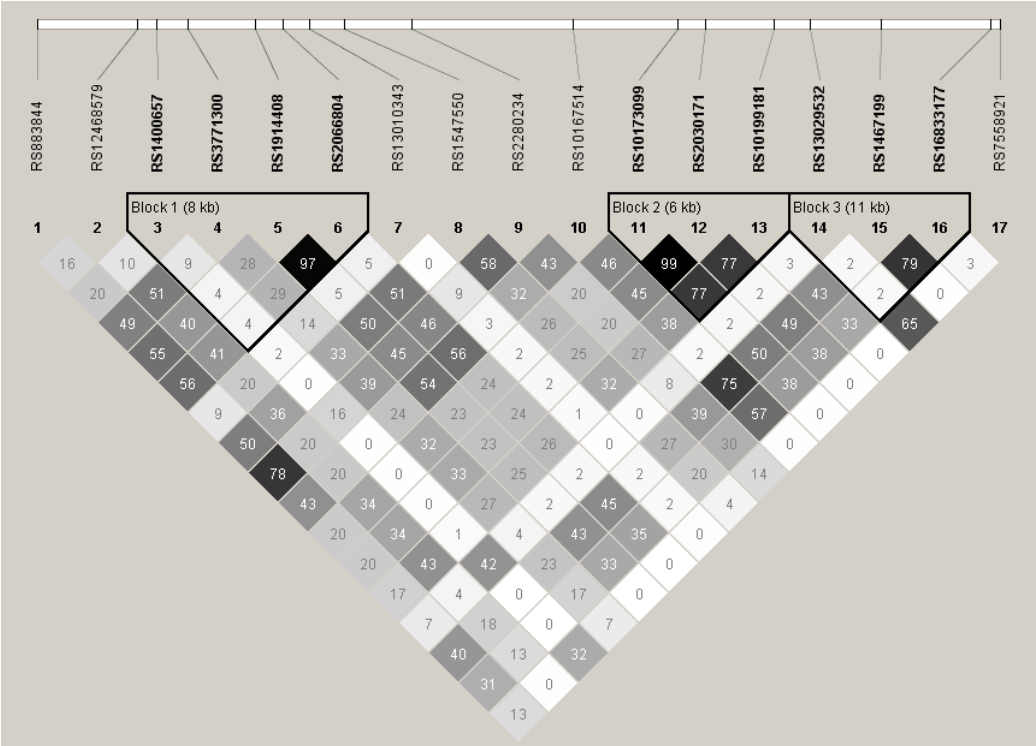
34. Skorge TD, Eagan TM, Eide GE, Gulsvik A, Bakke PS. The adult incidence of asthma and respiratory symptoms by passive smoking in uterus or in childhood. Am J Respir Crit Care Med. 2005; 172: 61-65.
35. Silverman EK, Palmer LJ. Case-control association studies for the genetics of complex respiratory diseases. Am J Respir Cell Mol Biol. 2000; 22: 645-648 PubMed .
36. Rothman KJ, Greenland S. Precision and validity in epidemiological studies. In: Rothman KJ, Greenland S (eds). Modern epidemiology. Lippincott- Raven; Philadelphia 1998, pp115-134.
37. Bakke PS. Designing studies for diagnostic tests. The Clinical Respiratory Journal 2008; 2: 72-75.
38. Molfino N. Current thinking on genetics of chronic obstructive pulmonary disease. Curr Opin Pulm Med. 2007; 13: 107-113.
39. Roeder K, Bacanu SA, Sonpar V, Zhang X, Devlin B. Analysis of singlelocus tests to detect gene/disease associations. Genet Epidemiol 2005;28:207-19 PubMed .
40. Webster PM, Lorimer EG, man SF, Woolf CR, Zamel N. Pulmonary function in identical twins: comparison of nonsmokers and smokers. *Am Rev Respir Dis* 1979;119:223-228.

Figure legends

Figures 1-3. Linkage Disequilibrium (LD) Map in STAT1, NFKBIB/ SIRT2 and GC Genes Region.Pattern of linkage disequilibrium for four genes region in the COPD ICGN family-based population (A) and COPD case-control population (B). Values of r^2 (× 100) are shown. Black squares, $r^2= 1$; white squares, $r^2= 0$; squares in shades of gray, $0< r^2< 1$ (the intensity of the grey is proportional to r^2). Haplotype block structure was estimated with the Haploview program.

Figure 1

Figure 1. Linkage Disequilibrium (LD) Map in STAT1, Gene Region
LD plot of STAT1 Gene with ICGN Family-Based Population



LD plot of STAT1 Gene with COPD Case Control Population

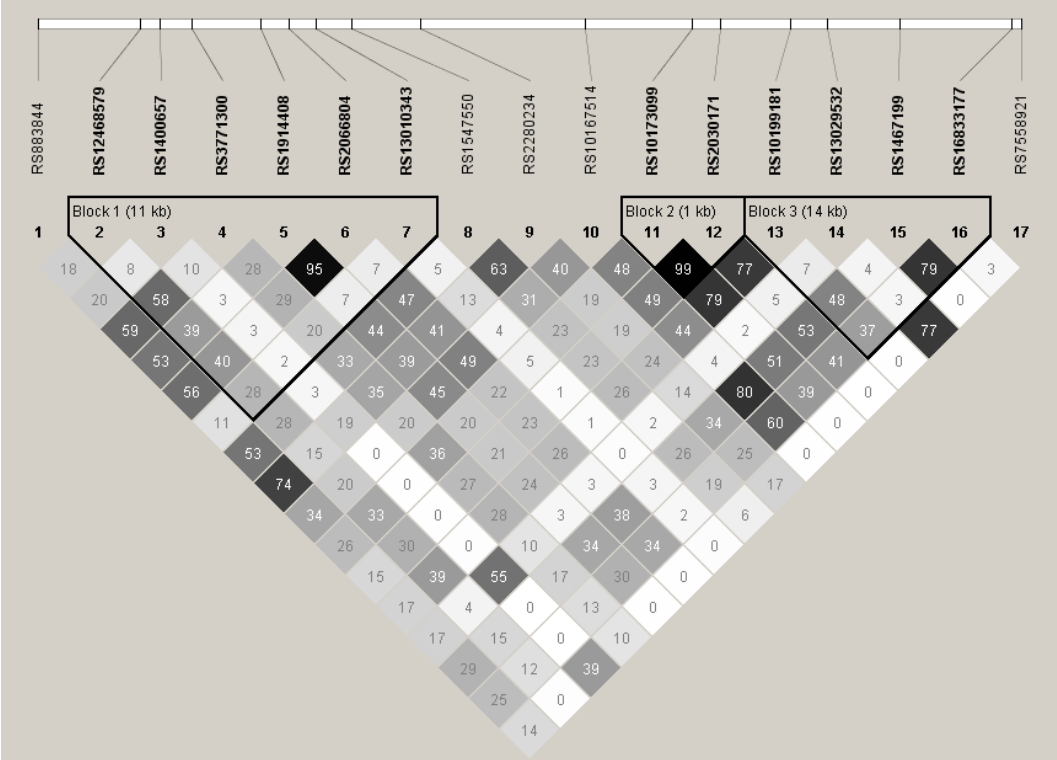
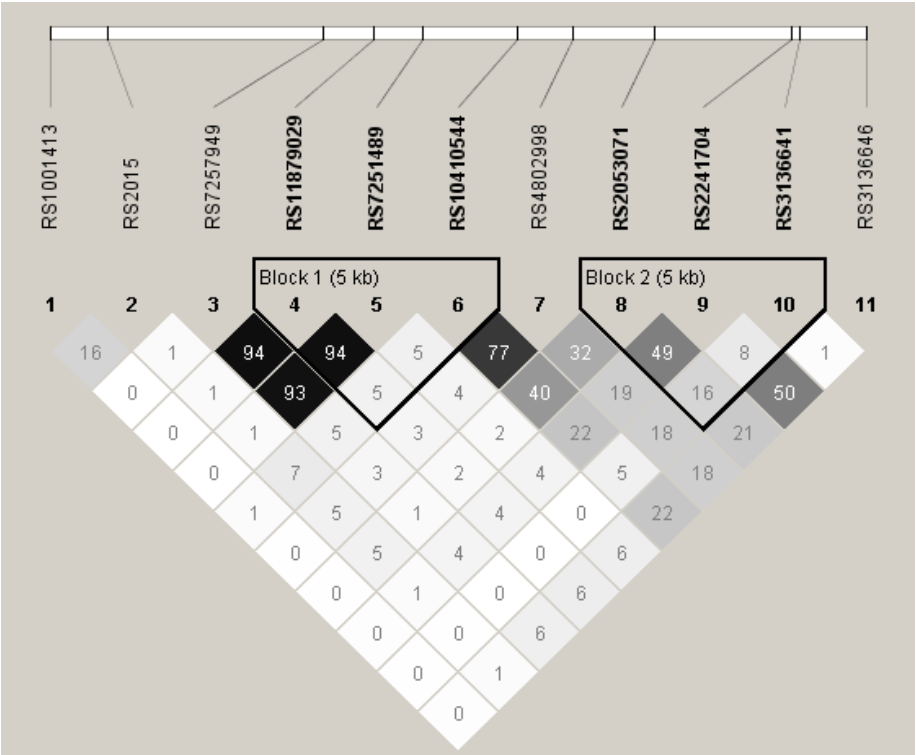


Figure 2. Linkage Disequilibrium (LD) Map in NFKBIB/ SIRT2, Gene Region

LD plot of NFKBIB/ SIRT2 Genes with ICGN Family-Based Population



LD plot of SIRT2/NFKBIB Genes with COPD Case Control Population

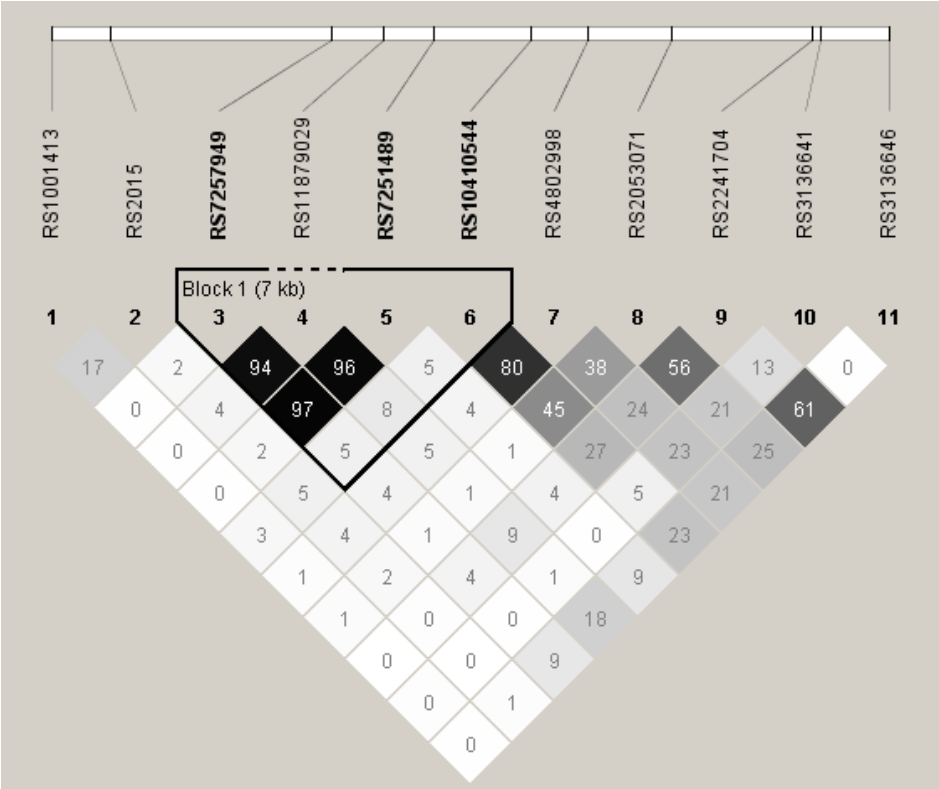
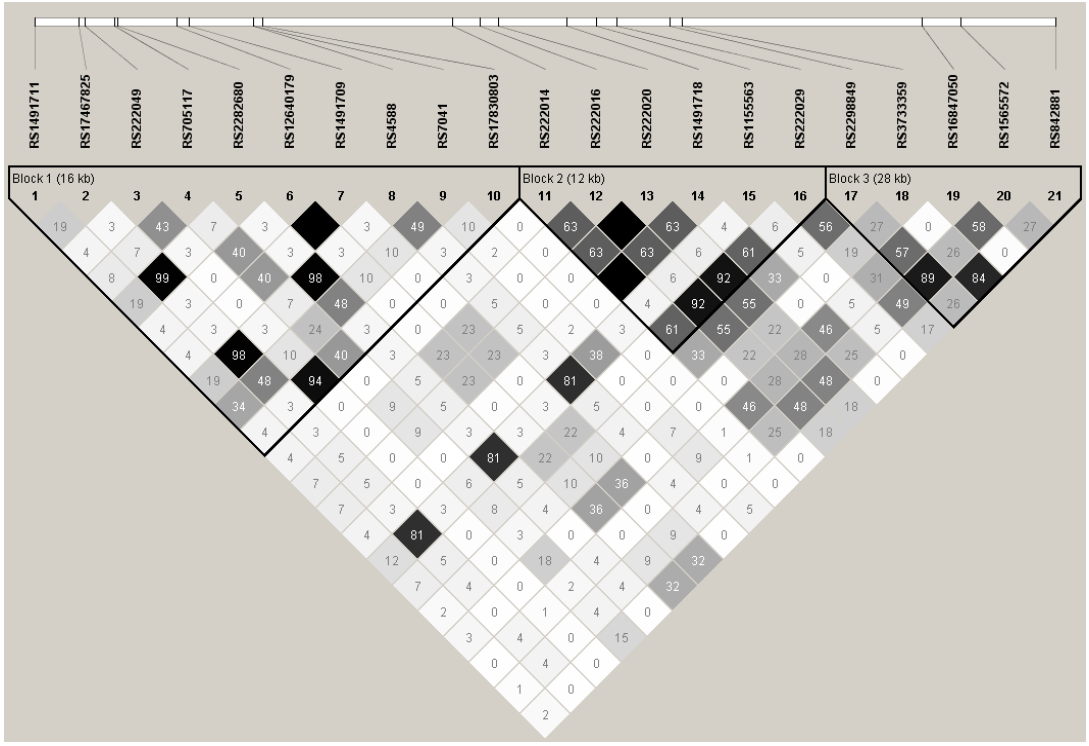
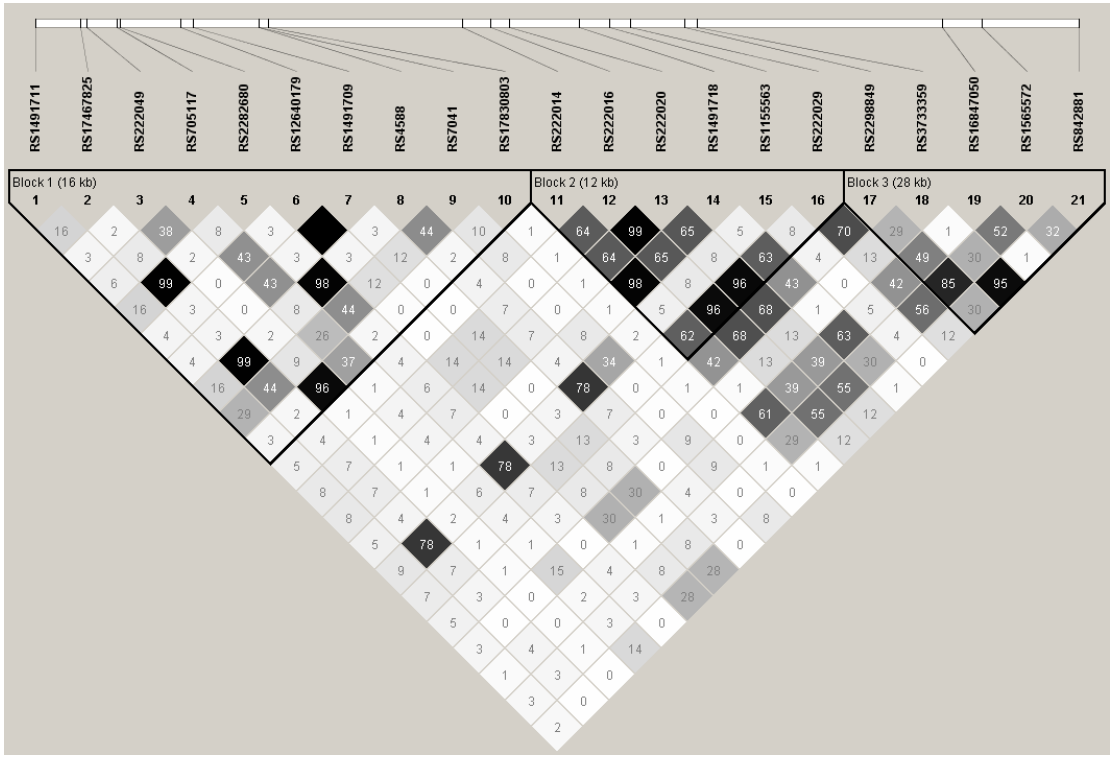


Figure 3. Linkage Disequilibrium (LD) Map in GC Gene Region

LD plot of GC Gene with ICGN Family-Based Population



LD plot of GC Gene with COPD Case Control Population