Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Hunninghake GM, Cho MH, Tesfaigzi Y, et al. *MMP12*, lung function, and COPD in high-risk populations. N Engl J Med 2009;361:2599-608. DOI: 10.1056/NEJMoa0904006.

Supplementary Appendix:

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: *MMP12*, Lung Function and COPD in High-Risk Populations

Gary M. Hunninghake, M.D., M.P.H., Michael H. Cho M.D., M.P.H., Yohannes Tesfaigzi Ph.D., Manuel E. Soto-Quiros, M.D., Ph.D., Lydiana Avila, M.D., Jessica Lasky-Su Sc.D, Chris Stidley, Ph.D., Erik Melén, M.D., Ph.D., Cilla Söderhäll, Ph.D., Jenny Hallberg, Ph.D., Inger Kull, R.N., Ph.D., Juha Kere, M.D., Ph.D., Magnus Svartengren, M.D., Ph.D., Göran Pershagen, M.D., Ph.D., Magnus Wickman, M.D., Ph.D., Christoph Lange Ph.D, Dawn L. Demeo M.D., M.P.H., Craig P. Hersh M.D., M.P.H., Barbara J. Klanderman, Ph.D, Benjamin A. Raby M.D., M.P.H., David Sparrow, D.Sc., Steven D. Shapiro M.D., Edwin K. Silverman M.D. Ph.D., Augusto A. Litonjua M.D., M.P.H., Scott T. Weiss M.D., M.S., Juan C. Celedón, M.D., Dr.P.H.

Subjects and Methods

The Genetics of Asthma in Costa Rica (GACRS):

The protocols for subject recruitment and data collection for GACRS have been previously described in detail.[1] From February of 2001 to March of 2005, short questionnaires were sent to the parents of 9,054 children ages 6 to 14 years who were enrolled in 95 schools in Costa Rica. Of the 9,054 questionnaires distributed, 5,355 (59.1%) were returned. Children were eligible for inclusion in the study if they had asthma (defined as physician-diagnosed asthma and at least two respiratory symptoms [wheezing, cough, or dyspnea] or a history of asthma attacks in the previous year) and high probability of having at least 6 great-grandparents born in the Central Valley of Costa Rica.[2] Of the 5,355 children whose parents returned the screening questionnaires, 1,947 (36.4%) had asthma. Of these 1,947 children, 439 (22.5%) unrelated children were willing to participate in the study along with their parents (in whom only genotypic information was collected) and had Central Valley ancestry. Of the 439 participating families in GACRS, 13 were excluded based on DNA quality control and 9 were excluded because of Mendelian inconsistencies, leaving 417 families. Written parental consent was obtained for participating children, from whom written assent was also obtained. The study was approved by the Institutional Review Boards of the Hospital Nacional de Niños (San José, Costa Rica) and Brigham and Women's Hospital (Boston, Massachusetts).

Spirometry was conducted with a Survey Tach Spirometer (Warren E. Collins, Braintree,

MA) following American Thoracic Society recommendations.[3] Height was measured to the nearest half inch. Subjects were instructed to avoid use of short-acting bronchodilators for at least four hours before testing. Spirometric maneuvers were performed with subjects seated and wearing a noseclip. As many as eight attempts were made to obtain three acceptable spirometric measures. The best FEV_1 was selected for data analysis.

The Childhood Asthma Management Program (CAMP):

CAMP was a multicenter clinical trial of the effects of anti-inflammatory medications in children with mild to moderate asthma. All recruited children had asthma defined by symptoms greater than 2 times per week, the use of an inhaled bronchodilator at least twice weekly or the use of daily medication for asthma, and airway responsiveness to methacholine $\leq$ 12.5 mg/ml. Children with severe asthma or other clinically significant conditions were excluded. Of the 1,041 children enrolled in the original clinical trial, 968 children and 1,518 of their parents contributed DNA samples. This analysis was restricted to families of white (non-Hispanic) children because of small sample size for other ethnic groups. Of the 483 white families in CAMP, 13 were removed from this analysis because of Mendelian inconsistencies, leaving 470 families (and 503 children with asthma [probands]). Data, including spirometric measurements, were collected at baseline and during the course of the four-year clinical trial, and blood samples were collected at baseline.[4] Written informed consent was obtained from parents of participating children. CAMP was approved by the Institutional Review Boards of Brigham and Women's Hospital and the other participating centers.

Spirometry was performed according to American Thoracic Society recommendations[3] with a volume-displaced spirometer. Testing was performed at least 4 hr after the participant's last use of a short-acting bronchodilator, and at least 24 hr after the participant's use of a long-acting bronchodilator.

The Children [Barn in Swedish], Allergy, Milieu, Stockholm, Epidemiological Survey (BAMSE) Study:

Subject recruitment and baseline characteristics of participants in the BAMSE Swedish birth cohort study have been previously described.[5, 6] In brief, between 1994 and 1996, 4,089 newborn infants were recruited and questionnaire data was obtained. Subjects were recruited from central and northwestern parts of Stockholm including both urban and suburban districts. Questionnaires on lifestyle and symptoms related to asthma and other allergic diseases were administered to the parents of participants when they were 1, 2, 4 and 8 years of age. The response rates at ages 1, 2, 4, and 8 years were 96%, 94%, 92% and 84%, respectively. At 8 years of age, all children were invited for clinical testing, including lung function measurements (see below). Blood samples were obtained from 2,480 children and DNA was extracted from 2,033 samples (82%) after exclusion of samples with too little blood, lack of questionnaire data, or lack of parental consent for genetic analyses. At age 8 years, asthma was defined as parental report of physician-diagnosed asthma and one or more episodes of wheezing in the preceding 12 months.

Pulmonary function tests were performed with a spirometer (2200 Pulmonary Function

Laboratory; Sensormedics, Anaheim, CA, USA) at age 8 years. Measurements were performed in accordance with ATS guidelines[3] with subjects seated and wearing a noseclip. The best FEV_1 value was used for analysis. Of the 2,033 children with DNA available for analyses, 1,427 (70%) completed acceptable spirometric measurements and were included in this study. The study was approved by the ethical committee at Karolinska Institutet, and parental consent for genetic analyses was obtained for all children included in this report.

The Boston Early-Onset COPD Study (eoCOPD):

The baseline characteristics and recruitment of probands and family members enrolled in the Boston Early-Onset COPD Study have been reported previously.[7] In brief, ascertainment criteria for probands in the Boston Early-Onset COPD Study included an FEV_1 <40% of predicted, age <53 years, and absence of severe alpha-1-antitrypsin deficiency. All available first-degree and older second-degree relatives were invited to participate in the study. This analysis includes 127 probands and 822 additional family members (for a total of 949 individuals). Participants gave written informed consent and the protocol was approved by the Human Research Committee of Brigham and Women's Hospital in Boston.

Spirometry was performed with a Survey Tach Spirometer (Warren E. Collins, Braintree, MA).[7] Subjects were instructed to abstain from inhaled bronchodilator use for 4 h before testing, unless significant respiratory symptoms necessitated bronchodilator use. Spirometry was performed in accordance with ATS specifications;[3] we report the best FEV_1 value from the best-test effort.

The National Emphysema Treatment Trial (NETT)

The design, methods, and outcomes in NETT have been described.[8, 9] Briefly, 1,218 subjects with severe COPD were enrolled into a randomized trial evaluating lung volume reduction surgery (LVRS) versus optimal medical therapy. Enrollment criteria included an $FEV_1 \leq 45\%$ predicted[10] and bilateral emphysema on computed tomographic (CT) imaging of the chest. This analysis includes 378 white subjects with COPD who were enrolled in the NETT Genetics Ancillary Study, had blood samples available for genotyping. After providing written informed consent, these NETT participants provided a blood sample for DNA extraction for genetic studies of COPD. The study was approved by the institutional review boards at participating NETT centers.

Pulmonary function was measured at the 17 clinical centers, in accordance with ATS guidelines,[3] prior to randomization but after pulmonary rehabilitation.

The Lovelace Smokers Cohort (Lovelace)

The design and methods for the Lovelace Smokers cohort have been previously described.[11] In brief, study subjects were drawn from eligible participants in a cohort study in NewMexico (Lovelace Smokers' Cohort), between March 2001 and July 2006. The catchment area for this cohort was Albuquerque, NM and its surrounding communities, comprising a population of approximately 700,000 persons. Most

participants were recruited through newspaper or television advertisements and were paid a small stipend for their participation. Subjects were included in the study if they were ages 40 to 75 years, former or current smokers with a minimum smoking history of 20 pack-years on initial screening, and able to understand English. Information related to demographics, respiratory diseases, and smoking was obtained by self-report from all study participants via a questionnaire. This analysis is limited to the 1,487 non-Hispanic white subjects due to relatively small sample size and concerns about population stratification for Hispanic participants. This study was approved by the Western Institutional Review Board (Olympia, WA).

Spirometry was obtained on all subjects by certified and registered respiratory therapists strictly adhering to the 1994 American Thoracic Society (ATS) guidelines.[12] All tests were conducted at Lovelace Scientific Resources (Albuquerque, NM).

The Normative Aging Study (NAS):

Subjects in this study were part of the Veterans Administration (VA) Normative Aging Study, a longitudinal study established in 1963, details of which have been previously described.[13] Briefly, 2,280 men from the greater Boston area, ages 21 to 80 years, enrolled in the study after an initial health screening determined that they were free of known chronic medical conditions. Men with past or current chronic conditions such as coronary heart disease, hypertension, chronic lung disease, asthma, and diabetes were excluded. Since enrollment, the participants have undergone comprehensive clinical

examinations at 5-year intervals for those < 52 years old and at 3-year intervals for those > 52 years old. Genotypic data were present on a total of 1,483 individuals, of whom 1,468 had repeated measures of pulmonary function and are included in this analysis. The study protocol was approved by the Human Studies Subcommittee of the Department of Veterans Affairs Medical Center and the Institutional Review Board of the Brigham and Women's Hospital.

Pulmonary function tests were performed as previously reported.[14] Briefly, a water-filled recording spirometer was used to obtain measures of FEV_1, with values adjusted by body temperature and pressure. These spirometric tests were performed in accordance with American Thoracic Society guidelines.[3]

Genotyping

In GACRS, genotyped markers were selected using a linkage disequilibrium (LD)-tagging algorithm (Tagger)[15] for *MMP12* and its 10-kb flanks in CEPH (Centre d'etude du polymorphisme humain) trios from HapMap. The seventeen SNPs genotyped capture $\geq$ 98 % of the HapMap SNPs in *MMP12* in CEPH trios at an $r^2 \geq 0.8$ for a minor allele frequency (MAF) of 0.1. The same SNPs were genotyped in CAMP as in GACRS, to allow comparison of LD patterns between these ethnically distinct groups. In GACRS and CAMP, 16 SNPs were genotyped using GoldenGate assays on the Illumina BeadStation 500 system (San Diego, CA)[16] and one SNP (rs2276109) was genotyped

with a Taqman genotyping assay (Applied Biosystems, Foster City, CA). In eoCOPD, NETT, and NAS, two SNPs with significant findings in Costa Rica (rs737693 and rs2276109) were genotyped with Taqman; one of these two SNPs (rs2276109) was genotyped with Taqman in two cohorts (BAMSE and Lovelace).

To test whether our findings could be explained by LD between variants in *MMP12* and those in the adjacent gene for matrix metalloproteinase 3 (*MMP3* [stromelysin], see **Figure 1** of this appendix), we genotyped four SNPs in *MMP3* (rs639752, rs650108, rs679620, and rs673163) in GACRS using the mass spectrometry analysis implemented in SEQUENOM (Sequenom, Inc., San Diego, CA). These SNPs and also those genotyped in *MMP12* capture 100 % of the SNPs in *MMP3* in CEPH trios at an $r^2 \geq 0.8$ for SNPs with MAF ≥ 0.1.

The quality of the genotypic data in each cohort was assessed by several methods. In Illumina, Sequenom, and Taqman, duplicate genotyping was performed on approximately 5% of the samples to assess genotype reproducibility. Only one discordant genotype was detected with Taqman genotyping in NAS. In BAMSE, there was >99% concordance with control samples. The genotypic pass rate was >97% for all loci in all platforms.

Statistical Analysis

In all cohorts, pre-bronchodilator FEV_1 was analyzed as a continuous variable. COPD was defined using modified Global Initiative for Chronic Obstructive Lung Disease (GOLD) Stage II criteria[17] (an FEV_1/FVC ratio < 0.7 and an FEV_1 < 80% of predicted[10])

based on results from pre-bronchodilator spirometry [as post-bronchodilator spirometry was not available in NAS]). For the analyses of COPD, we examined age-of-onset of COPD in NAS and COPD *per se* in eoCOPD and Lovelace. In NAS, subjects were considered to have COPD at the first time point where they developed GOLD criteria specified above, provided that they had an FEV_1 <80% of predicted on all subsequent spirometries. In NAS, analyses of age-of-onset of COPD are presented for both all subjects and after stratifying by current smoking status (current smokers vs. former and never smokers), as current smoking had been previously associated with accelerated lung function decline in a study by Fletcher et al.[18]

Analyses of pre-bronchodilator FEV_1 (hereafter called FEV_1) and COPD in all cohorts were adjusted for age, gender, height, and smoke exposure (including time-varying covariates where appropriate). Smoke exposure variables included binary variables for environmental tobacco smoke exposure (ETS [in the GACRS, CAMP, and BAMSE, as $\underline{\geq}$95% of children were nonsmokers]) and current smoking (in eoCOPD, Lovelace, and NAS [there were no current smokers in NETT]), and a quantitative variable (pack-years [in eoCOPD, NETT, Lovelace, and NAS]).

In family-based studies, Hardy-Weinberg equilibrium (HWE) was tested in parental data by a χ^2 goodness-of-fit test, and deviations from Mendelian inheritance were tested with PedCheck.[19] Genotypes inconsistent with Mendelian transmission within families were removed from family-based association analyses. Estimates of measures of LD -D' and r^2- were obtained using Haploview v3.11.[20]

Family-based association analyses of FEV_1 were performed in GACRS and eoCOPD with the FBAT statistic implemented in PBAT v3.6,[21] and linear regression analyses of FEV_1 were performed in the population-based cohorts (BAMSE, NETT, and Lovelace) using SASv9.1.2 (SAS Institute, Cary, NC). The family-based association analysis of repeated measures of FEV_1 in CAMP was conducted with FBAT-PC.[22] Population based analyses in BAMSE, NETT, and Lovelace were performed in SAS. Longitudinal analyses of FEV_1 in NAS were performed using generalized linear mixed models implemented in the Proc Mixed Procedure in SAS. Survival analyses of COPD onset were performed in NAS using Cox proportional hazards models. No variable in our model violated the proportional hazards assumption. For COPD, family-based association analyses were performed in eoCOPD using the FBAT statistic implemented in PBAT v3.6.[21], and case-control association analyses were conducted in Lovelace using logistic regression.

For the initial analysis of FEV_1 in GACRS, additive, dominant, and recessive genetic models were evaluated. Given results in GACRS, all subsequent analyses were done under a dominant genetic model. P values from all tests in the replication cohorts were one-sided, since all associations in these cohorts were in the same direction as in GACRS. Combined P values for all cohorts were calculated using two methods. We first calculated a P value using Fisher's method, as follows:(described in 23)

$$\chi^2_{2L} \sim -2\ln\left(\prod_{i=1}^{L} p_i\right) \quad (L = \text{\# p values})$$

We then calculated a P value using a weighted Z-score method to account for differences in sample size among study cohorts by initially calculating individual study weights[24]

$$z_{wi} = z_i \sqrt{Ni / Ntotal}$$

and then summing across the individual study weighted Z-scores.

To address multiple testing, we present P values for all cohorts and for all cohorts except the ones originally assessed for each outcome (GACRS for FEV_1 and NAS for COPD). We estimated the population risk (PAR) of COPD attributable to *MMP12* genotypes in NAS using the following formula:[25]

$$PAR = pd[(RR-1)/RR]x100,$$

where "pd" is the prevalence of *MMP12* genotypes among subjects with COPD and "RR" is the adjusted hazard ratio[26] obtained from Cox proportional hazards regression. To evaluate the discriminatory effect of the genotype,[27] we used pooled logistic regression,[28] incorporating the same variables in the Cox proportional hazards model (baseline age, height, and time-dependent variables current smoking and prior smoking) and pooling observations in 5-year periods. We compared C-statistics between models with and without the genotype of interest using a non-parametric approach.[29] All analyses were performed with SAS; the nonparametric ROC curve macro of DeLong et. al.[29] was provided by the SAS knowledgebase at: http://support.sas.com/kb/25/017.html.

Additional Results

Genotyping

All SNPs were in HWE in parental genotypes in family-based studies (Costa Rica, CAMP, and eoCOPD), and in genotypes of all subjects in BAMSE, NETT, Lovelace, and NAS.

Lung Function

The results of the analyses of association between SNP rs737693 and FEV_1 were similar to those for SNP rs2276109 and are shown in **Table 2** of this appendix.

COPD

After adjustment for age, height, and time-dependent covariates of current smoking and pack-years, the minor allele of rs737693 was associated with a 39% reduction in COPD onset (hazard ratio=0.61, 95% confidence interval [CI]=0.43-0.88, P= 0.008) in all subjects. Reciprocally, the absence of this allele (homozygosity for the major allele) was associated with a 64% increment in the risk of developing COPD.

The PAR of COPD due to homozygosity for the major allele of SNP rs737693 was 31% (95% CI=10%-46%).

ACKNOWLEDGEMENTS FOR STATISTICAL ANALYSES

All of the statistical analyses of data for GACRS, CAMP, eoCOPD, NETT, and NAS were performed by Gary M. Hunninghake and Michael M. Cho at the Channing Laboratory (with assistance from Jessica Lasky-Su and Christoph Lange, and under the supervision of Edwin K. Silverman and Juan C. Celedón) under appropriate policies for human subjects' protection for each of the five studies. Data from the BAMSE Study was analyzed by Jenny Hallberg and Erik Melén (under the supervision of Magnus Wickman) at the Karolinska Institutet, Stockholm under BAMSE's policies for human subjects' protection. Data from the Lovelace Smokers Cohort was analyzed by Chris Stidley (under the supervision of Yohannes Tesfaigzi) at the Lovelace Respiratory Research Institute in Albuquerque (NM) under Lovelace's policies for human subjects' protection.

Table 1. Pairwise (r2) linkage disequilibrium (LD) between SNP genotyped in both GACRS (Costa Rica) and CAMP.

Pairwise (r^2) linkage disequilibrium (LD) patterns for SNPs in parents of asthmatic children in GACRS																	
rs586701	rs737693	rs654600	rs674546	rs660727	rs484171	rs686375	rs505770	rs476391	rs651159	rs652438	rs632009	rs2276109	rs1277718	rs501371	rs636648	rs1940935	
1.00	0.48	0.97	0.94	0.38	0.42	0.39	0.38	0.38	0.42	0.38	0.05	0.47	0.41	0.41	0.38	0.03	rs586701
	1.00	0.46	0.46	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.98	0.01	0.01	0.01	0.02	rs737693
		1.00	0.95	0.41	0.45	0.42	0.41	0.42	0.45	0.42	0.05	0.46	0.40	0.40	0.36	0.04	rs654600
			1.00	0.39	0.43	0.46	0.45	0.39	0.43	0.39	0.05	0.45	0.40	0.40	0.36	0.04	rs674546
				1.00	0.91	0.86	0.87	1.00	0.91	1.00	0.02	0.01	0.84	0.84	0.93	0.02	rs660727
					1.00	0.94	0.93	0.91	1.00	0.91	0.02	0.01	0.92	0.92	0.85	0.01	rs484171
						1.00	1.00	0.86	0.94	0.86	0.02	0.01	0.87	0.87	0.80	0.01	rs686375
							1.00	0.87	0.93	0.87	0.02	0.01	0.87	0.87	0.81	0.01	rs505770
								1.00	0.91	1.00	0.02	0.01	0.84	0.84	0.93	0.02	rs476391
									1.00	0.91	0.02	0.01	0.84	0.84	0.93	0.02	rs651159
										1.00	0.02	0.01	0.84	0.84	0.93	0.02	rs652438
											1.00	0.02	0.02	0.02	0.02	0.94	rs632009
												1.00	0.01	0.01	0.01	0.02	rs2276109
													1.00	1.00	0.90	0.01	rs1277718
														1.00	0.90	0.01	rs501371
															1.00	0.02	rs636648
																1.00	rs1940935

Pairwise (r^2) linkage disequilibrium (LD) patterns for SNP in parents of non-Hispanic white children in CAMP

rs586701	rs737693	rs654600	rs674546	rs660727	rs484171	rs686375	rs505770	rs476391	rs651159	rs652438	rs632009	rs2276109	rs1277718	rs501371	rs636648	rs1940935	
1.00	0.63	0.99	0.98	0.27	0.27	0.26	0.26	0.27	0.26	0.27	0.11	0.63	0.27	0.27	0.28	0.11	rs586701
	1.00	0.63	0.62	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.07	0.97	0.01	0.01	0.01	0.07	rs737693
		1.00	0.98	0.27	0.27	0.26	0.26	0.27	0.26	0.27	0.10	0.63	0.27	0.27	0.28	0.11	rs654600
			1.00	0.27	0.27	0.28	0.28	0.26	0.25	0.27	0.11	0.62	0.27	0.27	0.27	0.11	rs674546
				1.00	0.99	0.95	0.95	1.00	0.98	1.00	0.02	0.01	0.95	0.95	0.96	0.03	rs660727
					1.00	0.96	0.95	0.99	1.00	0.99	0.02	0.01	0.96	0.96	0.97	0.03	rs484171
						1.00	0.99	0.95	0.95	0.95	0.02	0.01	0.92	0.92	0.93	0.03	rs686375
							1.00	0.95	0.94	0.95	0.02	0.01	0.93	0.93	0.92	0.03	rs505770
								1.00	0.98	1.00	0.02	0.01	0.95	0.95	0.96	0.03	rs476391
									1.00	0.98	0.02	0.01	0.95	0.95	0.96	0.03	rs651159
										1.00	0.02	0.01	0.95	0.95	0.96	0.03	rs652438
											1.00	0.07	0.02	0.02	0.02	0.94	rs632009
												1.00	0.01	0.01	0.01	0.07	rs2276109
													1.00	1.00	0.99	0.03	rs1277718
														1.00	0.99	0.03	rs501371
															1.00	0.03	rs636648
																1.00	rs1940935

Table 2. Results of Analyses of Association between a SNP in the 3' genomic region of *MMP12* (rs737693) and FEV_1 and COPD in All Study Cohorts*

FEV_1

rs number	Alleles	Location	GACRS	CAMP	BAMSE	eoCOPD	NETT	Lovelace	NAS	Combined Overall†	Combined Replication‡
rs737693	A>T	3' genomic	(+) 0.001	(+) 0.42 (+) 0.48 On P/N§ (+) 0.01 On ICS‖	-	(+) 0.003	(+) 0.16	-	(+) 0.005	5×10^{-6} (3×10^{-6})	6×10^{-4} (2×10^{-4})

COPD

rs number	Alleles	Location	NAS	eoCOPD	Lovelace	Combined Overall†
rs737693	A>T	3' genomic	(-) 0.01	(-) 0.006	-	3×10^{-4} (2×10^{-4})

*GACRS=nuclear families of asthmatic children in the Genetics of Asthma in Costa Rica Study. CAMP= nuclear families of asthmatic children in the Childhood Asthma Management Program. BAMSE= children in the BAMSE (Children [Barn in Swedish], Allergy, Milieu, Stockholm, Epidemiological Survey) Birth Cohort Study. eoCOPD= families of subjects with severe COPD in the Boston Early-Onset COPD Study. NETT=subjects in the Genetics Ancillary Study of the National Emphysema Treatment Trial. Lovelace=subjects in the Lovelace Smokers Cohort. NAS=subjects in the longitudinal Normative Aging Study.

† Overall combined P values were obtained using Fisher's method. P values from a weighted Z-score method are in parenthesis.

‡ Combined P values were again obtained (as described above) after excluding data from Costa Rica.

§ Children randomized to receive either inhaled placebo or inhaled nedocromil. Children randomized to receive inhaled corticosteroid (budesonide).

(+/-) Sign before the P value refers to the direction of association between the minor allele and pre-bronchodilator FEV_1 or COPD risk.

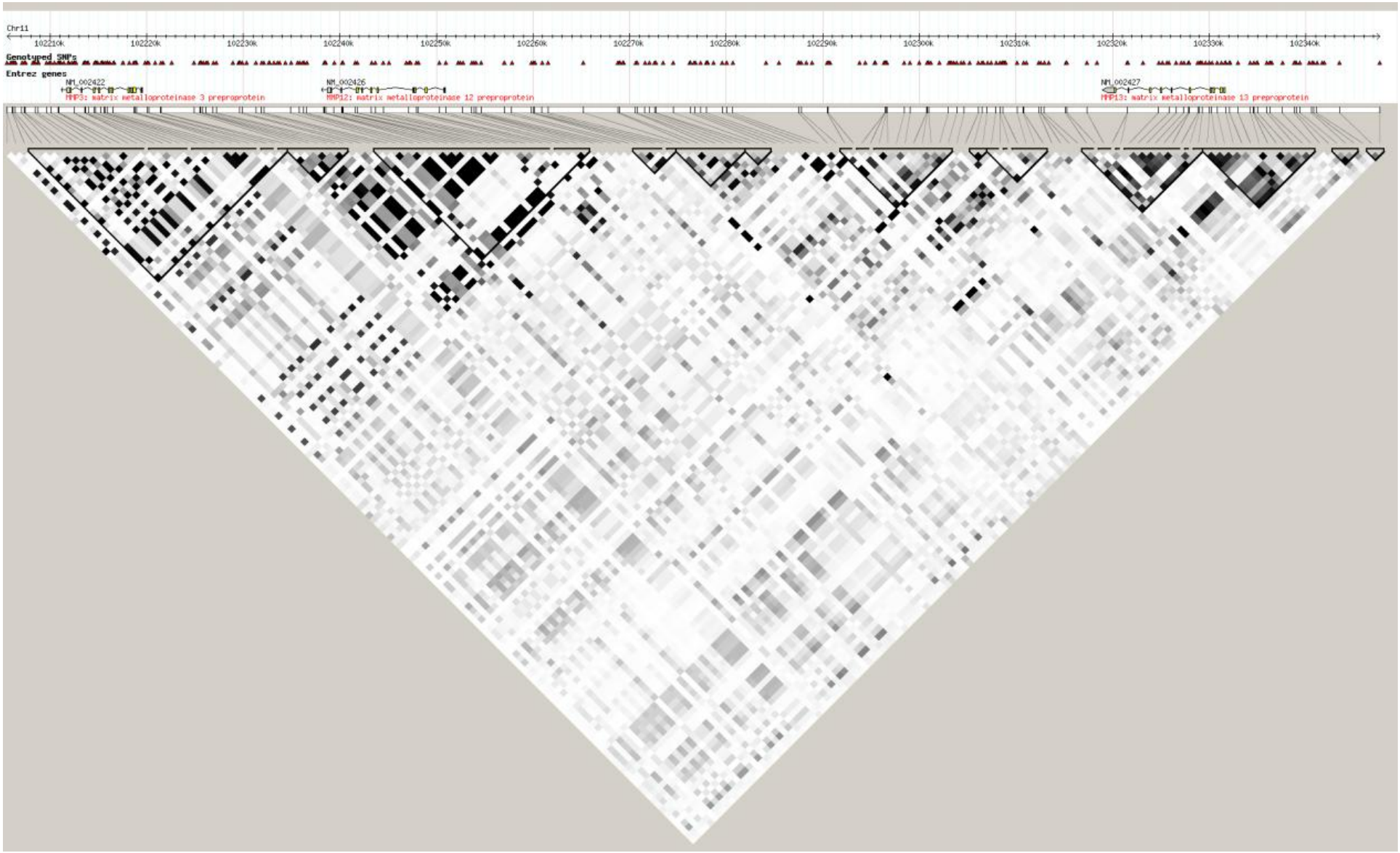
Chr11
102210k
102220k
102230k
102240k
102250k
102260k
102270k
102280k
102290k
102300k
102310k
102320k
102330k
102340k
Genotyped SNPs
Entrez genes
NM_002422
MMP3: matrix metalloproteinase 3 preproprotein
NM_002426
MMP12: matrix metalloproteinase 12 preproprotein
NM_002427
MMP13: matrix metalloproteinase 13 preproprotein

Figure Legends:

Figure 1: This figure represents the pairwise (r^2) linkage disequilibrium (LD) patterns for the genomic region on chromosome 11q22 encompassing *MMP3-MMP13* in CEPH (Centre d'etude du polymorphisme humain), trios in the HapMap project. SNPs in *MMP13* (a gene that lies >68kb 5' to *MMP12*) are not in linkage disequilibrium with *MMP12* SNPs in CEPH trios. This figure was obtained from Haploview (www.broad.mit.edu/mpg/haploview/).[17]

References

1. Hunninghake GM, Soto-Quiros ME, Avila L, et al. Sensitization to Ascaris lumbricoides and severity of childhood asthma in Costa Rica. J Allergy Clin Immunol 2007;119:654-61.
2. Escamilla MA, Spesny M, Reus VI, et al. Use of linkage disequilibrium approaches to map genes for bipolar disorder in the Costa Rican population. Am J Med Genet 1996;67:244-53.
3. Standardization of Spirometry, 1994 Update. American Thoracic Society. Am J Respir Crit Care Med 1995;152:1107-36.
4. The Childhood Asthma Management Program (CAMP): design, rationale, and methods. Childhood Asthma Management Program Research Group. Control Clin Trials 1999;20:91-120.
5. Wickman M, Kull I, Pershagen G, Nordvall SL. The BAMSE project: presentation of a prospective longitudinal birth cohort study. Pediatr Allergy Immunol 2002;13 Suppl 15:11-3.
6. Melen E, Nyberg F, Lindgren CM, et al. Interactions between glutathione S-transferase P1, tumor necrosis factor, and traffic-related air pollution for development of childhood allergic disease. Environ Health Perspect 2008;116:1077-84.
7. Silverman EK, Chapman HA, Drazen JM, et al. Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. Risk to relatives for airflow obstruction and chronic bronchitis. Am J Respir Crit Care Med 1998;157:1770-8.

8. Rationale and design of The National Emphysema Treatment Trial: a prospective randomized trial of lung volume reduction surgery. The National Emphysema Treatment Trial Research Group. Chest 1999;116:1750-61.

9. Fishman A, Martinez F, Naunheim K, et al. A randomized trial comparing lung-volume-reduction surgery with medical therapy for severe emphysema. N Engl J Med 2003;348:2059-73.

10. Crapo RO, Morris AH, Gardner RM. Reference spirometric values using techniques and equipment that meet ATS recommendations. Am Rev Respir Dis 1981;123:659-64.

11. Sood A, Stidley CA, Picchi MA, et al. Difference in airflow obstruction between Hispanic and non-Hispanic White female smokers. COPD 2008;5:274-81.

12. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. American Thoracic Society. Am J Respir Crit Care Med 1995;152:S77-121.

13. Bell B, Rose CL, Damon A. The Normative Aging Study: An interdisciplinary and longitudinal study of health and aging. Aging & Human Development 1972; 3: 5-17.

14. Sparrow D, O'Connor G, Colton T, Barry CL, Weiss ST. The relationship of nonspecific bronchial responsiveness to the occurrence of respiratory symptoms and decreased levels of pulmonary function. The Normative Aging Study. Am Rev Respir Dis 1987;135:1255-60.

15. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nat Genet 2005;37:1217-23.

16. Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS. A genome-wide scalable SNP genotyping assay using microarray technology. Nat Genet 2005;37:549-54.

17. Rabe KF, Hurd S, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med 2007;176:532-55.

18. Fletcher C, Peto R. The natural history of chronic airflow obstruction. Br Med J 1977;1:1645-8.

19. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998;63:259-66.

20. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263-5.

21. Lange C, DeMeo D, Silverman EK, Weiss ST, Laird NM. PBAT: tools for family-based association studies. Am J Hum Genet 2004;74:367-9.

22. Lange C, van Steen K, Andrew T, et al. A family-based association test for repeatedly measured quantitative traits adjusting for unknown environmental and/or polygenic effects. Stat Appl Genet Mol Biol 2004;3:Article17.

23. Zaykin DV, Zhivotovsky LA, Westfall PH, Weir BS. Truncated product method for combining P-values. Genet Epidemiol 2002;22:170-85.

24. Lettre G, Jackson AU, Gieger C, et al. Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet 2008;40:584-91.

25. Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait or intervention. Am J Epidemiol 1974;99:325-32.

26. Kenchaiah S, Evans JC, Levy D, et al. Obesity and the risk of heart failure. N Engl J Med 2002;347:305-13.

27. Meigs JB, Shrader P, Sullivan LM, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med 2008;359:2208-19.

28. D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. Stat Med 1990;9:1501-15.

29. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837-45.