Online Data Supplement

Visual Assessment of Chest CT Images is Independently Useful for Genetic Association Analysis in Studies of Chronic Obstructive Pulmonary Disease

Eitan Halper-Stromberg, PhD, Michael H. Cho, MD, Carla Wilson, MS, Dipti Nevrekar, MD, James D. Crapo, MD, George Washko, MD, Raúl San José Estépar, PhD, David A. Lynch, MD, Edwin K. Silverman, MD, PhD, Sonia Leach, PhD, Peter J. Castaldi, MD

Analyst Training and Fleischner Set Scoring:

Visual CT scoring was performed at National Jewish Health between September 2014 and March 2015. Two analysts were trained by Dr. David Lynch to follow the protocols developed by the Fleischner Society. To arrive at a score for emphysema, the lung images were reviewed, and the emphysema score was calculated based on the most advanced grade of emphysema that was present. With regard to panlobular, the Fleischner document specifies that the term panlobular emphysema is best restricted to subjects with alpha-1 antitrypsin deficiency. These subjects were excluded from COPDGene. Visual assessment of airway wall thickening was subjective, based on global assessment of the airways, compared to a set of standard images. When the majority of airways were judged to be thick-walled, airway wall thickening was scored as substantial. When the airway wall thickening was judged borderline, or fewer than half the airways were thick-walled, airway wall thickening was scored as mild. If neither of these conditionS were met, the airway wall thickening was scored as absent.

Training was performed on a random sample of 100 CT image sets previously scored by the Fleischner Society members. The two analysts went on to rescore a set of 600 more subjects previously scored by the Fleischner Society members (total Fleischner set size: 700). The Fleischner society members chose one hundred subjects randomly from within each of six spirometry-defined lung function categories, GOLD 0-4 and GOLD Unclassified ($FEV_1 < 0.8$ and $FEV_1/FVC > 0.7$). An additional randomly selected 100 non-smoking controls were scored. Of these 700, 423 individuals of NHW ancestry were used for genetic association testing. Excluded individuals included 100 non-smoking controls for whom genotyping was not performed, 136 African-Americans, nine individuals who failed quality control (genotype failure, misreported ancestry, mislabeled sample, or relatedness to another individual within COPDgene), and 22 individuals for which consensus was not reached.

Additional Scoring:

At the time of this analysis, 1117 additional NHW subjects with genotyping data had been scored by the two analysts, as part of the ongoing effort to generate visual score data for the full COPDgene cohort, and were therefore included in the genetic association study.

SNP genotyping, imputation, and quality control:

SNP genotyping and imputation has been previously described (1). Briefly, SNPs were genotyped on the Illumina (San Diego, CA) OmniExpress platform with additional SNP genotypes computed by imputation using MaCH and and minimac (2) using 1000 Genomes (3) V3 data as reference. For inclusion, SNPs were required to have a minor allele frequency ≥ 5%, a Hardy-Weinberg equilibrium p-value $\geq 1e10^{-4}$, and a genotyping success rate ≥ 95%. Imputed SNPs were required to have an imputation quality r^2 value > 0.3. For the GWAS, we tested a total of 6,173,964 SNPs across the 22 autosomal chromosomes. For the candidate gene study, we tested 108,511 SNPs across 196 candidate genes. For the prior GWAS lead SNP study we tested 33 SNPs across 19 loci.

Identification of Candidate Genes:

The candidate gene set consisted of SNPs with minor allele frequency ≥ 5% located within 100kb of 184 candidate genes from a published review of COPD genetics(4). We used genes listed in tables 1 and 2 from the review, excluding 3 genes: 2 X chromosome genes (*IL13RA1* and *TIMP1*), and 1 gene on chromosome 8 (*DEFB4A*) for which we had no SNP data. Additionally, an identifier appeared in the review that did not uniquely identify a single gene, *HLA*, which we converted to six genes for inclusion on our list: *HLA-DPA1*, *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DRB1*, and *HLA-DRA* (*HLA-DRA* did not make the final list because we had no high quality SNP data). The boundaries of each candidate gene locus included flanking 100 Kb regions on either side of each gene. Gene boundaries were queried using their HGNC symbols via BioMart and defined using the Ensembl start and end coordinates for H. sapiens build GRCh37 (5). For the lead SNP study we tested 33 SNPs across 19 loci that had been previously identified to reach genome-wide significance for association to emphysema, COPD, or airway disease and which had a minor allele frequency ≥ 1% in our dataset (1, 6–10). . In addition, SNPs within 100kb of twelve genes located near SNPs from the prior GWAS set were also included. These genes are *RIN3, SNRPF, MYO1D, VWA8, HYKK, DLC1, SERPINA10, MIR2054, MAGI2, SERPINE2, NT5C3B*, and *C10orf90*.

Population Stratification:

The first five principal components, computed based upon the full NHW cohort within COPDgene from 35,557 directly genotyped SNPs in linkage equilibrium using Eigensoft 3.0 (11, 12), served as the population stratification variables.

References

1. Cho MH, McDonald M-LN, Zhou X, Mattheisen M, Castaldi PJ, Hersh CP, Demeo DL, Sylvia JS, Ziniti J, Laird NM, Lange C, Litonjua AA, Sparrow D, Casaburi R, Barr RG, Regan EA, Make BJ, Hokanson JE, Lutz S, Dudenkov TM, Farzadegan H, Hetmanski JB, Tal-Singer R, Lomas DA, Bakke P, Gulsvik A, Crapo JD, Silverman EK, Beaty TH. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *lancet Respir Med* 2014;2:214–25.
2. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010;34:816–834.
3. Altshuler D, Lander E, Ambrogio L. A map of human genome variation from population scale sequencing. *Nature* 2010;476:1061–1073.
4. Bosse Y. Updates on the COPD gene list. *Int J Chron Obstruct Pulmon Dis* 2012;doi:10.2147/COPD.S35294.
5. Smedley D, Haider S, Durinck S, Pandini L, Provero P, Allen J, Arnaiz O, Awedh MH, Baldock R, Barbiera G, Bardou P, Beck T, Blake A, Bonierbale M, Brookes AJ, Bucci G, Buetti I, Burge S, Cabau C, Carlson JW, Chelala C, Chrysostomou C, Cittaro D, Collin O, Cordova R, Cutts RJ, Dassi E, Genova A Di, Djari A, *et al.* The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Res* 2015;doi:10.1093/nar/gkv350.
6. Castaldi PJ, Cho MH, Estépar RSJ, McDonald M-LN, Laird N, Beaty TH, Washko G, Crapo JD, Silverman EK. Genome-Wide Association Identifies Regulatory Loci Associated with Distinct Local Histogram Emphysema Patterns. *Ajrccm* 2014;075478:1–45.
7. Cho MH, Castaldi PJ, Hersh CP, Hobbs BD, Barr RG, Tal-Singer R, Bakke P, Gulsvik A, San José Estépar R, Van Beek EJR, Coxson HO, Lynch DA, Washko GR, Laird NM, Crapo JD, Beaty TH, Silverman EK. A Genome-wide Association Study of Emphysema and Airway Quantitative Imaging Phenotypes. *Am J Respir Crit Care Med* 2015;doi:10.1164/rccm.201501-0148OC.
8. Kong X, Cho MH, Anderson W, Coxson HO, Muller N, Washko G, Hoffman EA, Bakke P, Gulsvik A, Lomas DA, Silverman EK, Pillai SG. Genome-wide association study identifies BICD1 as a susceptibility gene for emphysema. *Am J Respir Crit Care Med* 2011;183:43–49.
9. Manichaikul A, Hoffman E a, Smolonska J, Gao W, Cho MH, Baumhauer H, Budoff M, Austin JHM, Washko GR, Carr JJ, Kaufman JD, Pottinger T, Powell C a, Wijmenga C, Zanen P, Groen HJM, Postma DS, Wanner A, Rouhani FN, Brantly ML, Powell R, Smith BM, Rabinowitz D, Raffel LJ, Hinckley Stukovsky KD, Crapo JD, Beaty TH, Hokanson JE, Silverman EK, *et al.* Genome-wide study of percent emphysema on computed tomography in the general population. The Multi-Ethnic Study of Atherosclerosis Lung/SNP Health Association Resource Study. *Am J Respir Crit Care Med* 2014;189:408–18.
10. Wilk JB, Shrine NRG, Loehr LR, Zhao JH, Manichaikul A, Lopez LM, Smith AV, Heckbert SR,

Smolonska J, Tang W, Loth DW, Curjuric I, Hui J, Cho MH, Latourelle JC, Henry AP, Aldrich M, Bakke P, Beaty TH, Bentley AR, Borecki IB, Brusselle GG, Burkart KM, Chen TH, Couper D, Crapo JD, Davies G, Dupuis J, Franceschini N, *et al.* Genome-wide association studies identify CHRNA5/3 and HTR4 in the development of airflow obstruction. *Am J Respir Crit Care Med* 2012;186:622–632.

11. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet* 2006;2:2074–2093.

12. Price AL, Price AL, Patterson NJ, Patterson NJ, Plenge RM, Plenge RM, Weinblatt ME, Weinblatt ME, Shadick N a, Shadick N a, Reich D, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38:904–9.

Table E1: Significant Association Results Between Emphysema Phenotypes, both LAA-950 and Visual, and Lead SNPs From Previous COPD and Emphysema GWAS

						Visual emphysema		**LAA-950**	
Locus	Nearest gene	Lead SNP	Effect Allele	MAF	Discovery phenotypes	Effect (SE)	P value	Effect (SE)	P value
1q41	*TGFB2*	rs1690789	C	0.50	Emphysema[1]	-0.19 (0.06)	9.5E-4	-1.19 (0.37)	1.4E-3
1q41	*TGFB2*	rs4846479	G	0.26	COPD[2]	0.16 (0.07)	1.6E-2	0.60 (0.43)	1.6E-1
4q31	*HHIP*	rs13141641	T	0.42	Emphysema[1,2] COPD[3]	0.14 (0.06)	1.7E-2	1.03 (0.38)	6.2E-3
8p22	*DLC1*	rs75200691	T	0.12	Emphysema[2]	0.05 (0.09)	5.5E-1	1.24 (0.57)	3.1E-2
11q22	*MMP3/12*	rs626750	G	0.18	Emphysema[1] COPD[3]	0.18 (0.07)	1.5E-2	1.31 (0.49)	7.0E-3
14q32	*RIN3*	rs754388	C	0.18	COPD[3]	0.20 (0.08)	9.0E-3	0.76 (0.49)	1.2E-1
15q25	*CHRNA3*	rs12914385	C	0.41	Emphysema[1,2] COPD[3]	-0.20 (0.06)	3.9E-4	-0.80 (0.38)	3.4E-2
17q11	*MYO1D*	rs379123	T	0.41	Emphysema[1]	-0.22 (0.06)	2.0E-4	-1.44 (0.38)	1.6E-4
19q13	*CYP2A6*	rs56113850	C	0.38	Emphysema[1]	0.14 (0.07)	3.6E-2	1.27 (0.43)	3.1E-3

Definition of abbreviations: LAA-950 = Percentage of low attenuation areas below -950 Hounsfield units. MAF = minor allele frequency. SE = standard error.

1 Castaldi PJ, Cho MH, Estépar RSJ, et al. Genome-Wide Association Identifies Regulatory Loci Associated with Distinct Local Histogram Emphysema Patterns. *Ajrccm* 2014;075478:1–45

2 Cho MH, Castaldi PJ, Hersh CP, et al. A Genome-wide Association Study of Emphysema and Airway Quantitative Imaging Phenotypes. *Am J Respir Crit Care Med* [Internet] 2015;Available from: http://dx.doi.org/10.1164/rccm.201501-0148OC

3 Cho MH, McDonald M-LN, Zhou X, et al. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *lancet Respir Med* [Internet] 2014;2(3):214–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24621683

P values were calculated using linear regression. For visual emphysema, the scale ranged from 0-5. For LAA-950, the scale ranged from 0 to 100 while observed values ranged from 0 to 58.Thirty three lead SNPs across nineteen loci from previous GWAS were tested. Seven loci were nominally significant by association to visual emphysema and seven loci were nominally significant by LAA-950. One lead SNP per locus is shown for lead SNPs in LD at $R^2 > 0.3$. Only the *TGFB2* locus contains multiple lead SNPs failing to meet this LD threshold. Results for the full list of lead SNPs are included in the supplement, Table E12.

Table E2: Significant Association Results Between Airway Phenotypes, both Semi-Automated and Visual, and Lead SNPs From Previous COPD and Emphysema GWAS

						Visual airway wall thickness		**Pi10**		**WAP**	
Locus	**Nearest gene**	**Lead SNP**	**EA**	**MAF**	**Discovery phenotypes**	Effect (SE)	P value	Effect (SE)	P value	Effect (SE)	P value
1q41	*TGFB2*	rs4846479	G	0.26	COPD[1]	0.06 (0.03)	3.5E-2	-0.01 (0.00)	2.1E-2	0.03 (0.13)	8.1E-1
1q41	*TGFB2*	rs1690789	C	0.50	Emphysema[2]	-0.06 (0.03)	2.4E-2	0.01 (0.00)	4.3E-2	-0.02 (0.11)	8.8E-1
2q36	*SERPINE2*	rs734556	T	0.33	Airway[3]	0.06 (0.03)	3.7E-2	0.00 (0.00)	9.9E-1	0.18 (0.12)	1.1E-1
4q28	*MIR2054*	rs142200419	T	0.01	Airway[4]	0.25 (0.12)	3.7E-2	0.02 (0.02)	4.0E-1	1.35 (0.50)	7.5E-3
5q32	*HTR4*	rs7733088	G	0.40	Airway[5]	0.05 (0.03)	4.9E-2	0.01 (0.00)	5.3E-2	0.21 (0.11)	5.5E-2
11q22	*MMP3/12*	rs626750	G	0.18	COPD[1] Emphysema[2]	0.09 (0.03)	7.0E-3	0.00 (0.01)	9.9E-1	0.18 (0.14)	2.1E-1
14q32	*RIN3*	rs754388	C	0.18	COPD[1]	0.11 (0.03)	8.2E-4	-0.00 (0.01)	8.9E-1	0.07 (0.15)	6.4E-1
15q25	*CHRNA5*	rs17486278	A	0.37	Emphysema[2,4]	-0.01 (0.03)	6.0E-1	0.00 (0.00)	3.1E-1	0.30 (0.11)	8.5E-3

Definition of abbreviations: MAF = minor allele frequency. EA = Effect Allele. Pi10 = square root of wall area for airways standardized at an internal perimeter of 10 mm as calculated by VIDA software. SE = standard error. WAP = wall area percent of airways for segmental (third generation) bronchi as calculated by VIDA software.

1. Cho MH, McDonald M-LN, Zhou X, et al. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *lancet Respir Med* [Internet] 2014;2(3):214–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24621683
2. Castaldi PJ, Cho MH, Estépar RSJ, et al. Genome-Wide Association Identifies Regulatory Loci Associated with Distinct Local Histogram Emphysema Patterns. *Ajrccm* 2014;075478:1–45
3. Dijkstra AE, Postma DS, Ginneken B van, et al. Novel Genes for Airway Wall Thickness Identified with Combined Genome-Wide Association and Expression Analyses. *Am J Respir Crit Care Med* [Internet] 2014;191(5):547–556. Available from: http://dx.doi.org/10.1164/rccm.201405-0840OC
4. Cho MH, Castaldi PJ, Hersh CP, et al. A Genome-wide Association Study of Emphysema and Airway Quantitative Imaging Phenotypes. *Am J Respir Crit Care Med* [Internet] 2015;Available from: http://dx.doi.org/10.1164/rccm.201501-0148OC
5. Wilk JB, Shrine NRG, Loehr LR, et al. Genome-wide association studies identify CHRNA5/3 and HTR4 in the development of airflow obstruction. *Am J Respir Crit Care Med* 2012;186(7):622–632.

P values were calculated using linear regression. For airway wall thickness, the scale ranged from 0-2. For Pi10, the the observed range of values was 3.3 to 4.3. For WAP, the observed range of values was 52 to 75. Thirty three lead SNPs across nineteen loci from previous GWAS were tested. Seven loci were nominally significant by association to at least one airway phenotype, six by visual airway thickness, one by Pi10, and two by WAP. One lead SNP per locus is shown for lead SNPs in LD at $R^2 >$ 0.3. Only the *TGFB2* locus contains multiple lead SNPs failing to meet this LD threshold. Results for the full list of lead SNPs are included in the supplement, Table E12.

Table E3: Nominally Significant lead SNPs by association to LAA-950 and visual emphysema from the candidate gene analysis based upon an FDR < .05 threshold

					Visual emphysema			**LAA-950**		
Locus	**Nearest Gene**	**Lead SNP**	**Effect Allele**	**MAF**	Effect (SE)	P value	q value	Effect (SE)	P value	q value
12p13	*A2M*	rs113350951	C	0.10	-0.3 (0.10)	1.8E-3	3.9E-1	-2.68 (0.64)	2.6E-5	2.2E-2
10p13	*CDC123*	rs2399794	A	0.47	0.24 (0.06)	6.8E-5	5.1E-2	1.71 (0.39)	9.8E-6	9.4E-3
7q22	*CYP3A5*	rs6965239	T	0.13	-0.29 (0.09)	9.2E-4	2.9E-1	-2.67 (0.56)	2.5E-6	4.9E-3
15q25	*IREB2*	rs2656072	G	0.21	0.41 (0.07)	6.3E-9	5.0E-5	2.51 (0.46)	4.5E-8	2.2E-4
11p11	*OR4X1*	rs74824994	G	0.06	-0.3 (0.13)	2.0E-2	6.9E-1	-3.41 (0.84)	5.3E-5	4.1E-2
3p24	*RARB*	rs77815743	G	0.09	-0.32 (0.10)	1.0E-3	3.0E-1	-2.56 (0.64)	6.5E-5	4.8E-2
17q21	*NT5C3B*	rs12452803	A	0.22	-0.23 (0.07)	1.2E-3	3.2E-1	-2.07 (0.45)	5.3E-6	7.7E-3
2q35	*TNS1*	rs3791976	A	0.40	-0.18 (0.06)	2.7E-3	4.4E-1	-1.64 (0.38)	2.0E-5	1.7E-2
14q32	*RIN3*	rs36032836	C	0.18	-0.3 (0.07)	5.7E-5	4.7E-2	-1.91 (0.48)	8.4E-5	6.0E-2
15q23	*THSD4*	rs72742771	G	0.13	-0.36 (0.09)	2.9E-5	2.6E-2	-1.71 (0.56)	2.4E-3	3.6E-1

Definition of abbreviations: MAF = minor allele frequency. LAA-950 = Percentage of low attenuation areas below -950 Hounsfield units. SE = standard error. q value = FDR-based q value calculated from all tests performed for all candidate genes.

P values were calculated using linear regression. For visual emphysema, the scale ranged from 0-5. For LAA-950, the observed range of values was 0 to 58. SNPs at 196 candidate gene loci were tested. Ten loci reached an FDR < .05 by either visual or LAA-950. The 15q25 locus contained four significant genes, *IREB2*, *CHRNA3*, *CHR*NA5, and *HYKK* for both visual and LAA-950. Only one lead SNP from this locus is shown in the table. Lead SNPs were chosen based upon minimum q value at each locus, considering both visual emphysema and LAA-950 q values.

Table E4: Exploring Bias Between Subjects with Complete Visual and Semi-Automated Emphysema Measures and Subjects with Visual Emphysema Measures Only

	Only Visual Emphysema Measures Available	Visual and Semi-Automated Emphysema Measures Available	P value
n	27	1513	
Gender (% Male)	52	50	1.0E+00
Age	64 (8)	63 (8)	4.2E-01
Pack-years	45 (34-63)	42 (30-58)	4.2E-01
FEV1/FVC	0.67 (0.16)	0.64 (0.17)	2.8E-01
FEV1 % of predicted	72 (23)	74 (26)	7.7E-01
Current Smoker	0.19	0.31	2.3E-01

Definitions of abbreviations: FEV1 = forced expiratory volume in one second. FVC = forced vital capacity.

P values were calculated using a chi-squared test for categorical variables (gender and current smoker) and a t-test for continuous variables (age, pack-years, FEV1/FVC, FEV1 % of predicted). Values are mean (sd) or median (interquartile range).

Table E5: Exploring Bias Between Subjects with Complete Visual and Semi-Automated Airway Measures and Subjects with Visual Airway Measures Only

	Only Visual Airway Measures Available	Visual and Semi-Automated Airway Measures Available	P value
n	67	1473	
Gender (% Male)	36	51	2.3E-02
Age	63 (9)	63 (8)	7.3E-01
Pack-years	43 (30-54)	42 (30-58)	2.8E-01
FEV1/FVC	0.65 (0.16)	0.64 (0.17)	4.5E-01
FEV1 % of predicted	73 (26)	74 (26)	7.8E-01
Current Smoker	0.25	0.31	3.8E-01

Definitions of abbreviations: FEV1 = forced expiratory volume in one second. FVC = forced vital capacity.

P values were calculated using a chi-squared test for categorical variables (gender and current smoker) and a t-test for continuous variables (age, pack-years, FEV1/FVC, FEV1 % of predicted). Values are mean (sd) or median (interquartile range).

Table E6: Exploring Bias Between Subjects with Complete Visual and Semi-Automated Emphysema Measures and Subjects with Semi-Automated Airway Measures Only

	Only Semi-Automated Emphysema Measures Available	Visual and Semi-Automated Emphysema Measures Available	P value
n	5130	1513	
Gender (% Male)	53	50	4.7E-02
Age	62 (9)	63 (8)	2.3E-07
Pack-years	42 (30-59)	42 (30-58)	2.3E-01
FEV1/FVC	0.64 (0.17)	0.64 (0.17)	1.5E-01
FEV1 % of predicted	74 (26)	74 (26)	9.4E-01
Current Smoker	0.42	0.31	2.5E-13

Definitions of abbreviations: FEV1 = forced expiratory volume in one second. FVC = forced vital capacity.

P values were calculated using a chi-squared test for categorical variables (gender and current smoker) and a t-test for continuous variables (age, pack-years, FEV1/FVC, FEV1 % of predicted). Values are mean (sd) or median (interquartile range).

Table E7: Exploring Bias Between Subjects with Complete Visual and Semi-Automated Airway Measures and Subjects with Semi-Automated Airway Measures Only

	Only Semi-Automated Airway Measures Available	Visual and Semi-Automated Airway Measures Available	P value
n	5130	1473	
Gender (% Male)	53	51	1.3E-01
Age	62 (9)	63 (8)	2.6E-07
Pack-years	42 (30-59)	42 (30-58)	3.5E-01
FEV1/FVC	0.64 (0.17)	0.64 (0.17)	1.5E-01
FEV1 % of predicted	74 (26)	74 (26)	9.3E-01
Current Smoker	0.42	0.31	5.3E-13

Definitions of abbreviations: FEV1 = forced expiratory volume in one second. FVC = forced vital capacity.

P values were calculated using a chi-squared test for categorical variables (gender and current smoker) and a t-test for continuous variables (age, pack-years, FEV1/FVC, FEV1 % of predicted). Values are mean (sd) or median (interquartile range).

Table E8: Association Results for Lead SNPs From Previous COPD, Emphysema, and Airway GWAS Adjusting the Visual CT Phenotype by LAA-950 and Vice Versa

					Visual emphysema adjusted by LAA-950		LAA-950 adjusted by visual emphysema	
Locus	Nearest Gene	Lead SNP	Effect Allele	MAF	Effect (SE)	P value	Effect (SE)	P value
1q41	*TGFB2*	rs4846479	G	0.26	0.09 (0.05)	4.5E-02	-0.14 (0.30)	6.4E-01
1q41	*TGFB2*	rs1690789	C	0.50	-0.06 (0.04)	1.5E-01	-0.31 (0.26)	2.4E-01
4q31	*HHIP*	rs13141641	T	0.42	0.02 (0.04)	5.5E-01	0.39 (0.26)	1.4E-01
8p22	*DLC1*	rs75200691	T	0.12	-0.08 (0.06)	1.7E-01	0.99 (0.40)	1.3E-02
11q22	*MMP3/12*	rs626750	G	0.18	0.04 (0.05)	4.7E-01	0.46 (0.34)	1.8E-01
14q32	*RIN3*	rs754388	C	0.18	0.11 (0.05)	3.0E-02	-0.17 (0.34)	6.2E-01
15q25	*CHRNA3*	rs12914385	C	0.41	-0.12 (0.04)	3.6E-03	0.16 (0.26)	5.3E-01
17q11	*MYO1D*	rs379123	T	0.41	-0.06 (0.04)	1.5E-01	-0.42 (0.27)	1.1E-01
19q13	*CYP2A6*	rs56113850	C	0.38	0.00 (0.05)	9.7E-01	0.62 (0.30)	3.8E-02

Definition of abbreviations: MAF = minor allele frequency. LAA-950 = Percentage of low attenuation areas below -950 Hounsfield units. SE = standard error.

SNPs were chosen for inclusion based upon nominal significance to an emphysema phenotype without correlate adjustment (Table E1). P values were calculated using linear regression. For visual emphysema, LAA-950 was included in the regression test as a covariate, and for LAA-950, visual emphysema was included as a covariate. Ten SNPs across three loci remained nominally significant for visual emphysema after correction for LAA-950. Two SNPs across two loci remained nominally significant for LAA-950 after correction for visual emphysema. One lead SNP per locus is shown for lead SNPs in LD at $R^2 > 0.3$. Only the *TGFB2* locus contains multiple lead SNPs failing to meet this LD threshold.

Table E9: Association Results Between Lead SNPs From Previous COPD, Emphysema, and Airway GWAS and Airway Visual Phenotypes adjusted by Semi-Automated Airway Phenotypes and Vice Versa

					Visual airway wall thickening adjusted by Pi10		Visual airway wall thickening adjusted by WAP		Pi10 adjusted by visual airway wall thickening		WAP adjusted by visual airway wall thickening	
Locus	Nearest Gene	Lead SNP	Effect Allele	MAF	Effect (SE)	P value	Effect (SE)	P value	Effect (SE)	P value	Effect (SE)	P value
1q41	*TGFB2*	rs4846479	G	0.26	0.082 (0.028)	3.3E-03	0.056 (0.025)	2.3E-02	-0.014 (0.005)	2.2E-03	-0.098 (0.107)	3.6E-01
1q41	*TGFB2*	rs1690789	C	0.50	-0.073 (0.024)	2.6E-03	-0.053 (0.022)	1.4E-02	0.011 (0.004)	4.5E-03	0.096 (0.094)	3.1E-01
2q36	*SERPINE2*	rs734556	T	0.33	0.06 (0.03)	2.9E-02	0.04 (0.02)	1.2E-01	0 (0)	5.2E-01	0.04 (0.1)	6.6E-01
4q28	*MIR2054*	rs142200419	T	0.01	0.22 (0.11)	5.5E-02	0.11 (0.1)	2.6E-01	0 (0.02)	8.2E-01	0.54 (0.44)	2.2E-01
5q32	*HTR4*	rs7733088	G	0.40	0.04 (0.02)	0.15	0.02 (0.02)	0.29	0.01 (0)	0.16	0.11 (0.1)	0.24
11q22	*MMP3/12*	rs626750	G	0.18	0.089 (0.032)	4.7E-03	0.072 (0.028)	1.1E-02	-0.004 (0.005)	4.0E-01	-0.056 (0.122)	6.5E-01
14q32	*RIN3*	rs754388	C	0.18	0.113 (0.032)	3.7E-04	0.103 (0.028)	2.9E-04	-0.007 (0.005)	2.2E-01	-0.183 (0.124)	1.4E-01
15q25	*CHRNA5*	rs17486278	A	0.37	-0.022 (0.025)	3.8E-01	-0.056 (0.023)	1.4E-02	0.005 (0.004)	2.1E-01	0.374 (0.098)	1.3E-04

Definition of abbreviations: MAF = minor allele frequency. Pi10 = square root of wall area for airways standardized at an internal perimeter of 10 mm as calculated by VIDA software. SE = standard error. WAP = wall area percent of airways for segmental (third generation) bronchi as calculated by VIDA software.

P values were calculated using linear regression. For visual airway wall thickening, Pi10 and WAP were included in the regression test as covariates, and for Pi10 and WAP, visual airway wall thickening was included as a covariate. One lead SNP per locus is shown for lead SNPs in LD at $R^2 > 0.3$. Only the *TGFB2* locus contains multiple lead SNPs failing to meet this LD threshold.

Table E10: Nominally Significant Association Results Between Visual Emphysema Phenotypes and Lead SNPs From Candidate Genes and Previous COPD, Emphysema, and Airway GWAS Shown with Values Based on Ordinal Regression

					Visual emphysema	
Locus	Nearest Gene	Lead SNP	Effect Allele	MAF	Effect (SE)	P value
12p13	*A2M*	rs113350951	C	0.10	-0.3 (0.11)	8.60E-03
10p13	*CDC123*	rs2399794	A	0.47	0.23 (0.07)	7.70E-04
7q22	*CYP3A5*	rs6965239	T	0.13	-0.31 (0.10)	2.20E-03
15q25	*IREB2*	rs2656072	G	0.21	0.50 (0.08)	4.20E-09
11p11	*OR4X1*	rs74824994	G	0.06	-0.30 (0.15)	5.10E-02
3p24	*RARB*	rs77815743	G	0.09	-0.29 (0.12)	1.20E-02
17q21	*TNS1*	rs3791976	A	0.40	-0.21 (0.07)	2.40E-03
2q35	*NT5C3B*	rs12452803	A	0.22	-0.22 (0.08)	7.00E-03
14q32	*RIN3*	rs36032836	C	0.18	-0.32 (0.09)	2.90E-04
15q23	*THSD4*	rs72742771	G	0.13	-0.41 (0.10)	5.10E-05
1q41	*TGFB2*	rs4846479	G	0.26	0.18 (0.08)	2.00E-02
1q41	*TGFB2*	rs1690789	C	0.50	-0.20 (0.07)	3.50E-03
4q31	*HHIP*	rs13141641	T	0.42	0.16 (0.07)	1.60E-02
8p22	*DLC1*	rs75200691	T	0.12	0.06 (0.10)	5.40E-01
11q22	*MMP3/12*	rs626750	G	0.18	0.17 (0.09)	5.60E-02
14q32	*RIN3*	rs754388	C	0.18	0.26 (0.09)	3.70E-03
15q25	*CHRNA3*	rs12914385	C	0.41	-0.25 (0.07)	2.30E-04
17q11	*MYO1D*	rs379123	T	0.41	-0.25 (0.07)	2.60E-04
19q13	*CYP2A6*	rs56113850	C	0.38	0.13 (0.08)	1.00E-01

Definition of abbreviations: MAF = minor allele frequency. SE = standard error.

P values were calculated using ordinal regression. Lead SNPs are included in this table if they were considered significant based upon a linear regression test (Tables E1 and E3). One lead SNP per locus is shown for lead SNPs in LD at $R^2 > 0.3$. Only the *TGFB2* locus contains multiple lead SNPs failing to meet this LD threshold.

Table E11: Nominally significant Association Results Between Airway Phenotypes and Lead SNPs From Previous COPD, Emphysema, and Airway GWAS Shown with Values Based on Ordinal Regression

					Visual airway wall thickening	
Locus	Tested Gene	Lead SNP	Effect Allele	MAF	Effect (SE)	P value
1q41	*TGFB2*	rs4846479	G	0.26	0.17 (0.08)	3.30E-02
1q41	*TGFB2*	rs1690789	C	0.50	-0.15 (0.07)	2.90E-02
2q36	*SERPINE2*	rs734556	T	0.33	0.15 (0.07)	3.40E-02
4q28	*MIR2054*	rs142200419	T	0.01	0.68 (0.32)	3.20E-02
5q32	*HTR4*	rs7733088	G	0.40	0.15 (0.07)	3.10E-02
11q22	*MMP3/12*	rs626750	G	0.18	0.21 (0.09)	1.60E-02
14q32	*RIN3*	rs754388	C	0.18	0.30 (0.09)	9.60E-04
15q25	*CHRNA5*	rs17486278	A	0.37	-0.04 (0.07)	5.60E-01

Definition of abbreviations: MAF = minor allele frequency. SE = standard error.

P values were calculated using ordinal regression. Lead SNPs are included in this table if they were considered significant based upon a linear regression test (Table E2). One lead SNP per locus is shown for lead SNPs in LD at $R^2 > 0.3$. Only the *TGFB2* locus contains multiple lead SNPs failing to meet this LD threshold.

Table E12: Association Results Between All Tested Phenotypes and All 33 Lead SNPs Across 19 Loci From Previous COPD, Emphysema, and Airway GWAS Shown with Values Based on Linear Regression

				Visual emphysema	LAA-950	Visual airway thickening	Pi10	WAP
Locus	Tested Gene	Lead SNP	MAF	P value	P value	P value	P value	P value
1q41	*TGFB2*	rs4846479	0.26	1.60E-02	1.60E-01	3.50E-02	2.10E-02	8.10E-01
1q41	*TGFB2*	rs4846480	0.26	1.60E-02	1.60E-01	3.50E-02	2.10E-02	8.10E-01
1q41	*TGFB2*	rs1690789	0.50	9.50E-04	1.40E-03	2.40E-02	4.20E-02	8.80E-01
2q36	*SERPINE2*	rs734556	0.33	5.00E-01	9.80E-01	3.70E-02	9.90E-01	1.10E-01
4q22	*FAM13A*	rs4416442	0.40	3.00E-01	1.00E-01	2.90E-01	9.60E-01	2.30E-01
4q28	*MIR2054*	rs142200419	0.01	7.80E-01	4.10E-01	3.70E-02	4.00E-01	7.50E-03
4q31	*HHIP*	rs138641402	0.41	9.20E-02	2.20E-02	4.40E-01	5.60E-01	8.80E-01
4q31	*HHIP*	rs13141641	0.42	1.70E-02	6.20E-03	5.30E-01	3.80E-01	6.00E-01
5q32	*HTR4*	rs7733088	0.40	9.00E-01	3.00E-01	4.90E-02	5.30E-02	5.50E-02
7q21	*MAGI2*	rs10251504	0.45	7.40E-01	5.00E-01	6.90E-01	4.80E-01	1.60E-01
8p22	*DLC1*	rs74834049	0.12	6.80E-01	5.40E-02	1.30E-01	6.30E-01	6.10E-01
8p22	*DLC1*	rs75200691	0.12	5.50E-01	3.10E-02	1.40E-01	7.70E-01	5.50E-01
10q26	*C10orf90*	rs7078439	0.31	5.30E-01	7.60E-01	2.10E-01	4.40E-01	7.40E-01
10q26	*C10orf90*	rs10794108	0.27	9.50E-01	7.90E-01	7.50E-02	8.50E-01	2.10E-01
11q22	*MMP3/12*	rs626750	0.18	1.50E-02	7.00E-03	7.00E-03	9.90E-01	2.10E-01
11q22	*MMP3/12*	rs17368582	0.12	1.00E-01	4.10E-02	1.70E-02	5.80E-01	1.30E-01

11q22	*MMP3/12*	rs17368659	0.12	9.50E-02	4.60E-02	1.70E-02	5.80E-01	1.20E-01
11q22	*MMP3/12*	rs17368814	0.12	9.50E-02	4.60E-02	1.70E-02	5.80E-01	1.20E-01
12p11	*BICD1*	rs161976	0.46	2.70E-01	1.90E-01	4.10E-01	1.60E-01	5.80E-01
12q23	*SNRPF*	rs7957346	0.44	6.10E-02	7.60E-01	4.00E-01	7.90E-01	4.30E-01
13q14	*VWA8*	rs9590614	0.40	6.30E-01	2.00E-01	8.70E-01	5.10E-01	8.60E-01
14q32	*RIN3*	rs17184313	0.17	1.70E-02	1.90E-01	1.80E-03	5.70E-01	6.60E-01
14q32	*RIN3*	rs754388	0.18	9.00E-03	1.20E-01	8.20E-04	8.90E-01	6.40E-01
14q32	*SERPINA10*	rs45505795	0.03	4.90E-01	3.90E-01	5.70E-01	7.70E-02	3.60E-01
15q25	*HYKK*	rs11852372	0.37	2.70E-03	2.60E-02	3.70E-01	3.30E-01	8.80E-03
15q25	*HYKK*	rs9788721	0.38	2.90E-03	2.70E-02	5.10E-01	3.40E-01	8.90E-03
15q25	*CHRNA5*	rs17486278	0.37	3.00E-03	3.40E-02	6.00E-01	3.00E-01	8.50E-03
15q25	*CHRNA3*	rs12914385	0.41	3.90E-04	3.40E-02	2.50E-01	5.60E-01	1.10E-01
15q25	*CHRNA3*	rs55676755	0.37	1.80E-03	3.00E-02	6.50E-01	3.60E-01	1.10E-02
15q25	*CHRNA3*	rs114205691	0.37	1.90E-03	3.70E-02	7.20E-01	3.70E-01	1.10E-02
17q11	*MYO1D*	rs379123	0.41	2.00E-04	1.60E-04	7.20E-01	7.80E-01	8.10E-01
17q21	*NT5C3B*	rs4796712	0.10	4.70E-01	8.20E-01	9.60E-01	7.40E-01	3.60E-01
19q13	*CYP2A6*	rs56113850	0.38	3.60E-02	3.10E-03	9.60E-01	6.60E-01	7.00E-01

Definition of abbreviations: MAF = minor allele frequency. LAA-950 = Percentage of low attenuation areas below -950 Hounsfield units. Pi10 = square root of wall area for airways standardized at an internal perimeter of 10 mm as calculated by VIDA software. SE = standard error. WAP = wall area percent of airways for segmental (third generation) bronchi as calculated by VIDA software.

Figure E1: Emphysema visual score (left) and LAA-950 (right) local association plot at the CYP3A5 locus (n=1513)

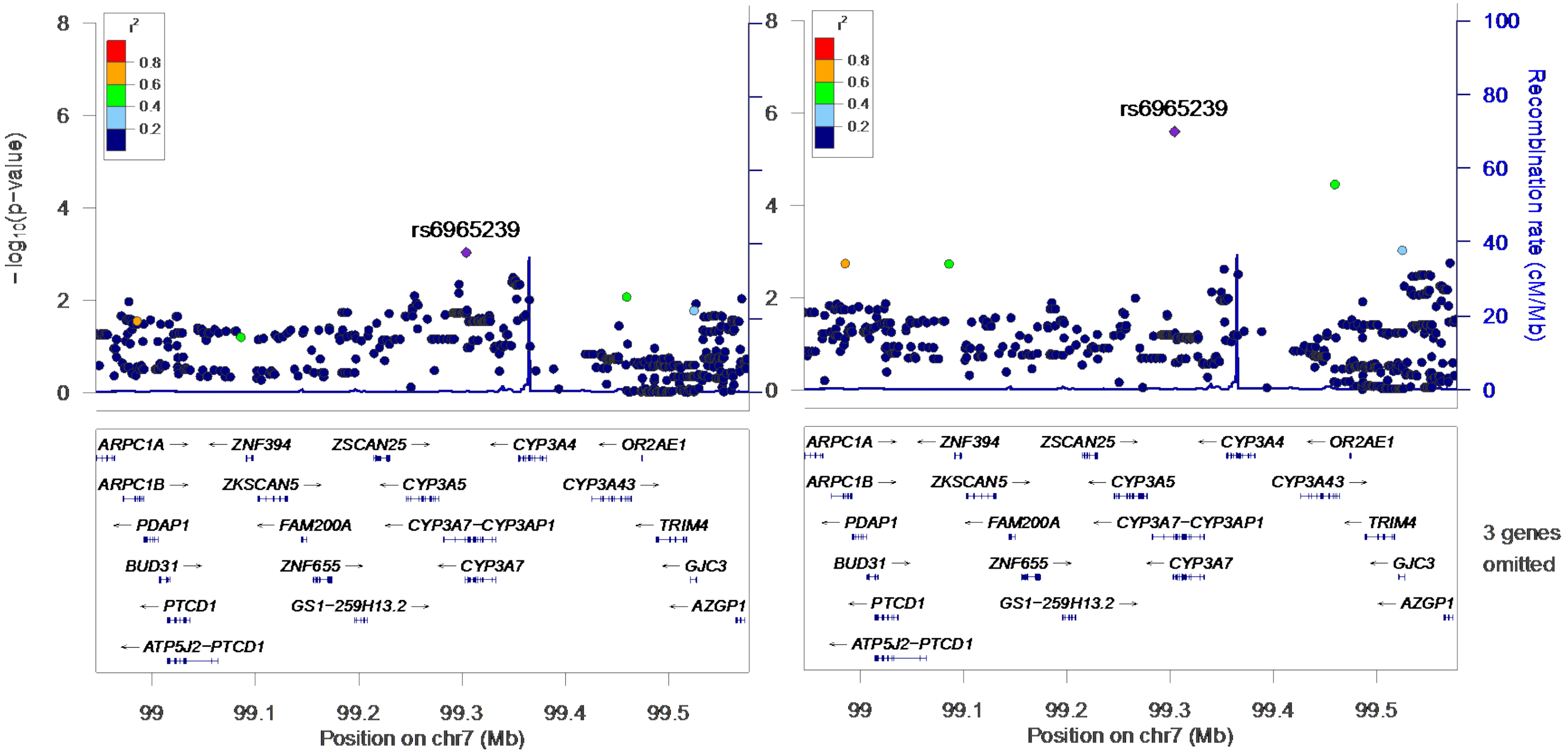

Figure E2: Emphysema visual score (left) and LAA-950 (right) local association plot at the CDC123 locus (n=1513)

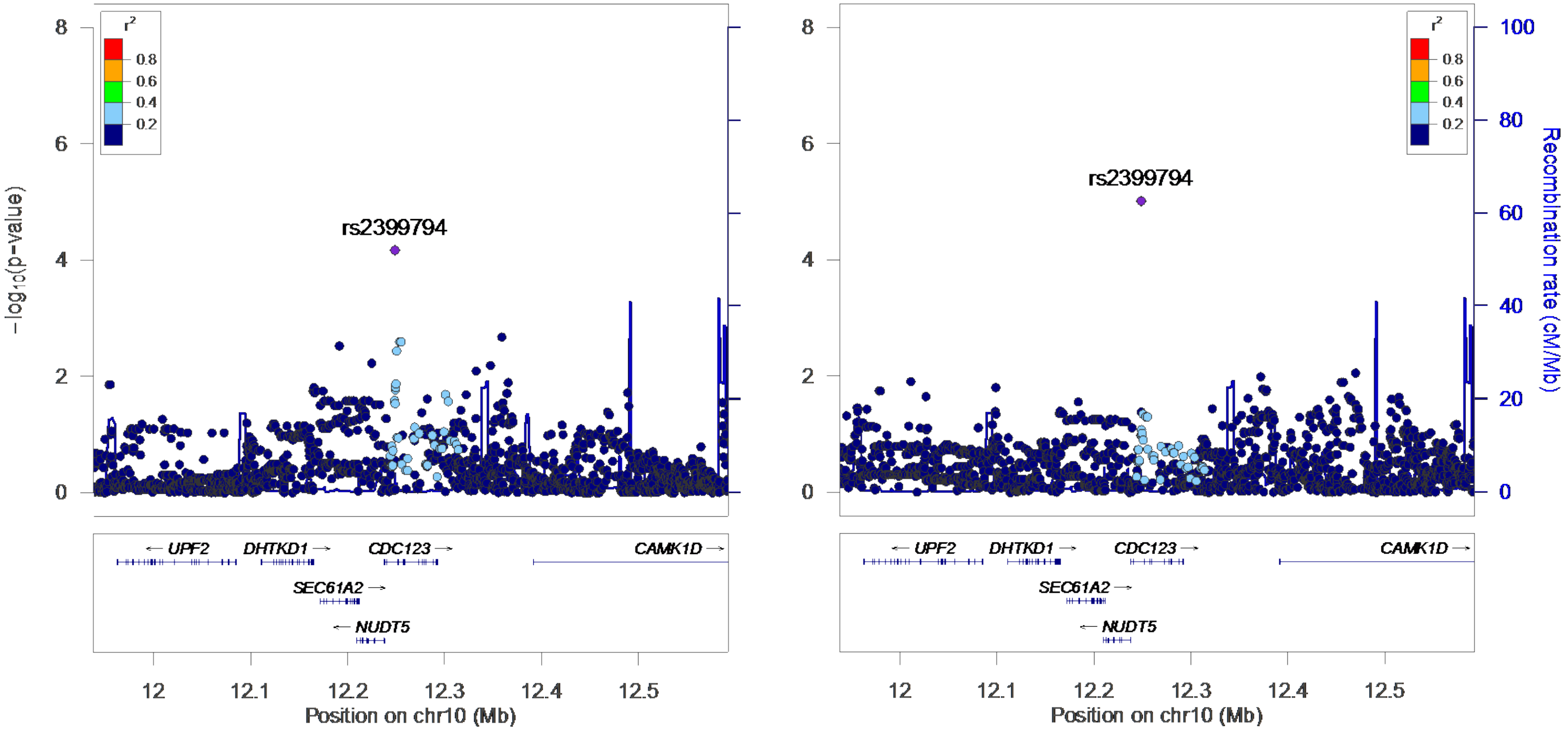

Figure E3: Emphysema visual score (left) and LAA-950 (right) local association plot at the A2M locus (n=1513)

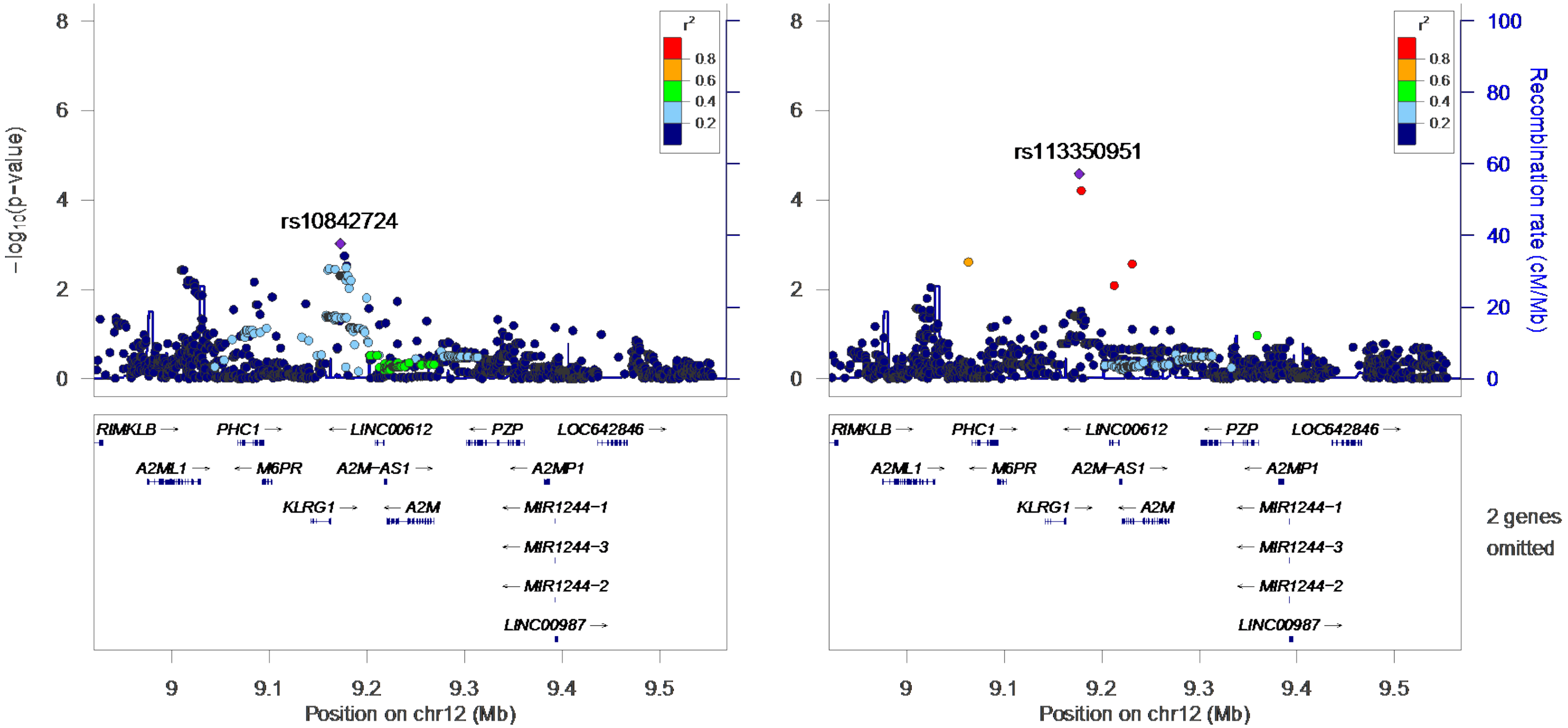

Figure E4: Emphysema visual score (left) LAA-950 (right) local association plot at the OR4X1 locus (n=1513)

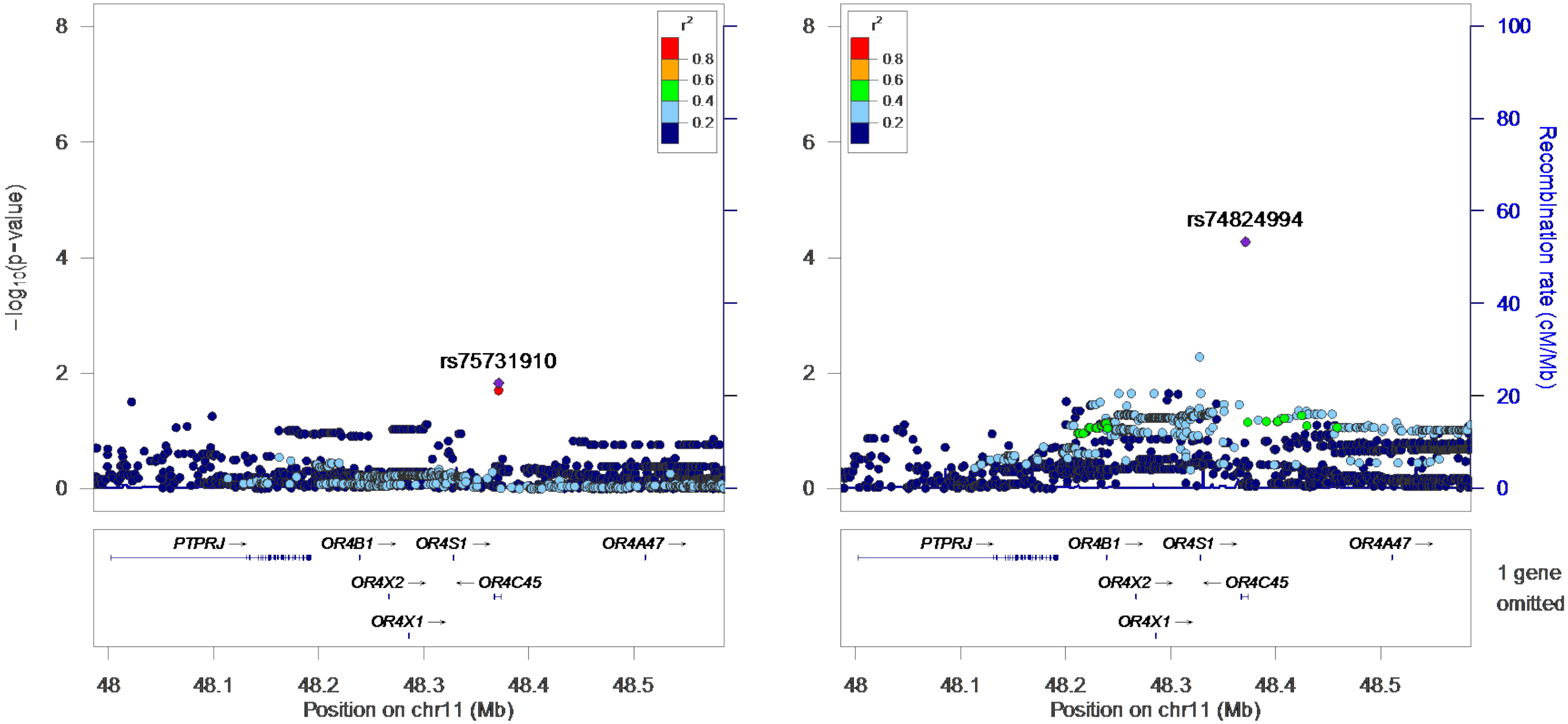

Figure E5: Emphysema visual score (left) and LAA-950 (right) local association plot at the RARB locus (n=1513)

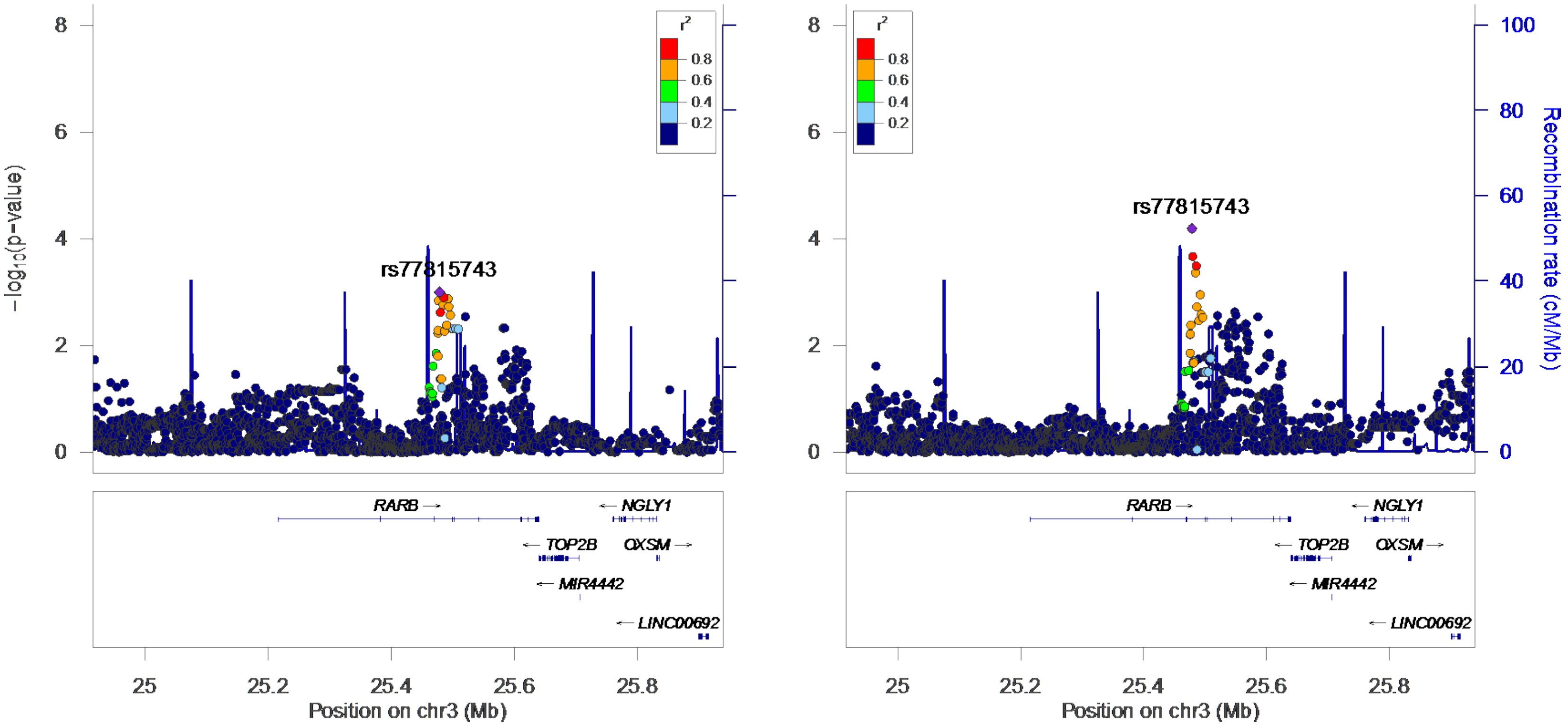

Figure E6: Emphysema visual score (left) and LAA-950 (right) local association plot at the NT5C3B locus (n=1513)