Supplementary material

Granularity of *SERPINA1* alleles by DNA sequencing in CanCOLD

Nisha Gupta, MD; Nathalie Gaudreault, BSc; Sébastien Thériault, MD; Pei Zhi Li, MSc; Cyndi Henry, MSc; Miranda Kirby, PhD; François Maltais, MD; Wan Tan, MD; Jean Bourbeau, MD; Yohan Bossé, PhD; for the Canadian Respiratory Research Network, for the CanCOLD study group

Methods

Pulmonary function tests

Spirometry and plethysmography were performed according to the American Thoracic Society guidelines [1-3]. For spirometry, FEV1, FVC, and FEV1/FVC were reported. Whole-body plethysmography was also performed for the measurement of functional residual capacity (FRC), residual volume (RV), total lung capacity (TLC), and diffusing capacity of the lung for carbon monoxide (DLCO).

CT image acquisition and analysis

CT images were acquired at multiple sites using various CT system models calibrated similarly with the participant supine at suspended full inspiration and full expiration from the apex to the base of the lung. The CT parameters for image acquisition in the participants investigated were as follows: 100 kVp, 50 mAs, 0.5 s gantry rotation, pitch of 1.375, 1.00–1.25 mm slice thickness, and an intermediate reconstruction kernel (GE: Standard; Siemens: b35; Philips: B) was used for quantitative analysis. All CT image analysis was performed by VIDA Diagnostics, Inc. (Coralville, IA, USA). CT emphysema was measured in the full-inspiration CT images using the percentage of low attenuation areas below -950 Hounsfield units (LAA-950) [4, 5].

DNA sequencing of *SERPINA1*

DNA was extracted from 200-400 µL of frozen whole blood using QIAamp® DNA Blood Mini kit (Qiagen). The DNA quality and concentration were assessed by the UV absorbance ratio 260 nm/280 nm and UV absorbance 260 nm, respectively. Every sample was diluted to a final concentration of 50 ng/µL. The DNA sequences of the coding regions (i.e. exons 2 to 5) of the *SERPINA1* gene were obtained by Sanger sequencing for all subjects. Primer sequences to evaluate the selected regions of

SERPINA1 are provided in **Table S4**. PCR was performed in a final volume of 25 µL containing 100 ng of genomic DNA, 1 U of HotStarTaq DNA polymerase (Qiagen), PCR buffer 1X, Q-Solution 1X, 160 µM of each dNTP and 0.2 µM of each primer. The PCR reaction was carried out on the GeneAmp® PCR system 9700 (Applied Biosystems) with the following cycling conditions: 15 minutes at 95°C, 35 PCR amplification cycles (15 seconds at 94°C, 30 seconds at 59°C (exons 2, 3 and 4) or 60°C (exon 5), and 60 seconds at 72°C), and 7 minutes at 72°C. The sequencing reaction was then performed using standard procedures and the product was run on the ABI 3730xl DNA Analyzer (Applied Biosystems). Sequencing files were assembled and analyzed using the EMBL-EBI Clustal Omega Multiple Alignment Tool (http://www.ebi.ac.uk/Tools/msa/clustalo).

Allele-specific PCR for allelic background determination

For identified variants with unknown allelic background, allele-specific PCR (AS-PCR) was performed in order to amplify and sequence each allele independently. For each variant, two forward AS-primers and one common reverse primer were designed based on the heterozygosity of a single polymorphism. Furthermore, the resulting amplicon was designed in order to include all polymorphisms identified in a single DNA sample of interest. For example, rs201318727 (Ile340Val), rs1303 (Glu376Asp) and rs709932 (Arg101His) were all found heterozygote in the same DNA sample. The design was based on rs1303 heterozygosity and the amplicon spanned 4,598 bp in order to include the other two polymorphisms. Each allele-specific amplicon was then sequenced using primers located near each polymorphism to be sequenced. **Table S5** shows the AS-PCR primers and the polymorphism used to design the AS-PCR. Primers were purchased from Integrated DNA Technologies, Iowa, USA. PCR cycling conditions are indicated in **Table S6**. DNA sequencing was performed as indicated above.

Results

Association of individual genetic variants with AAT serum levels

All the genetic variants identified were tested against the AAT serum levels. The most significant associated variant was rs28929474 causing the Z phenotype (**Table 1**). Mean AAT levels by genotype groups were 1.28±0.18 for GG (no Z allele), 0.87±0.15 for GA (heterozygote for the Z allele), and 0.42±0.32 for AA (homozygote for the Z allele). The second most significant associated variant was rs17580 causing the S phenotype. Mean AAT levels by genotype groups were 1.29±0.19 for AA, 1.10±0.15 for AT, and 0.94±0.16 for TT. The three common missense variants defining the M1 to M4 phenotypes were also associated with AAT levels. For rs6647 (Val213Ala), the mean AAT levels by genotype groups were 1.28±0.18 for TT, 1.25±0.22 for TC, and 1.23±0.25 for CC. For rs709932 (Arg101His), the mean AAT levels by genotype groups were 1.26±0.20 for GG, 1.30±0.19 for GA, and 1.29±0.16 for AA. For rs1303 (Glu376Asp), the mean AAT levels by genotype groups were 1.26±0.21 for AA, 1.29±0.18 for AC, and 1.29±0.16 for CC. Four additional missense variants were associated with lowered AAT levels including rs28931570 (Arg39Cys) causing the I phenotype, rs121912714 (Asp256Val) causing P_{Lowell}, rs61761869 (Pro369Ser) causing $M_{Wurzburg}$, and rs745463238 (Ser14Phe), but the latter occurs in only one individual with $ZS_{Donosti}$.

Deficient alleles across CanCOLD subgroups

We next evaluated the distribution of the 220 deficient alleles across the four CanCOLD subgroups, i.e. controls, at-risk, mild COPD, and moderate-severe COPD. Deficient alleles were found in 7.7% (47 out of 612) of controls, 7.3% (59 out of 808) of individuals at-risk of COPD, 8.9% (64 out of 718) of mild COPD patients, and 8.6% (50 out of 580) of moderate-severe COPD patients. A slight increase of deficient alleles was observed in the COPD subgroups, however this difference was not statistically

significant (**Figure S1A**). Similarly, the distribution of the S and Z alleles across CanCOLD subgroups were not statistically different (**Figures S1B-C**).

Carriers of deficient alleles across CanCOLD subgroups

Finally, we evaluated the distribution of the 210 individuals carrying at least one deficient alleles across CanCOLD subgroups. There were 14.4% (44 out of 306) of carriers in controls, 14.1% (57 out of 404) in at-risk individuals, 17.0% (61 out of 359) in mild COPD patients, and 16.6% (48 out of 290) in moderate-severe COPD patients. A slight increase of carriers of deficient alleles was observed in the COPD subgroups, however the difference was not statistically significant (**Figure S2A**). Similarly, carriers of the S and Z alleles were not statistically different across CanCOLD subgroups (**Figures S2B-C**). When considering those with or without emphysema according to CT imaging, the percentages of AATD allele carriers was 15.2% (129 out of 848) for LAA-950 <5% and 15.5% (54 out of 349) for LAA-950 ≥5%.

References

1. Macintyre N, Crapo RO, Viegi G, Johnson DC, van der Grinten CP, Brusasco V, Burgos F, Casaburi R, Coates A, Enright P, Gustafsson P, Hankinson J, Jensen R, McKay R, Miller MR, Navajas D, Pedersen OF, Pellegrino R, Wanger J. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *Eur Respir J* 2005: 26(4): 720-735.
2. Wanger J, Clausen JL, Coates A, Pedersen OF, Brusasco V, Burgos F, Casaburi R, Crapo R, Enright P, van der Grinten CP, Gustafsson P, Hankinson J, Jensen R, Johnson D, Macintyre N, McKay R, Miller MR, Navajas D, Pellegrino R, Viegi G. Standardisation of the measurement of lung volumes. *Eur Respir J* 2005: 26(3): 511-522.
3. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CP, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J, Force AET. Standardisation of spirometry. *Eur Respir J* 2005: 26(2): 319-338.
4. Gevenois PA, Zanen J, de Maertelaer V, De Vuyst P, Dumortier P, Yernault JC. Macroscopic assessment of pulmonary emphysema by image analysis. *J Clin Pathol* 1995: 48(4): 318-322.
5. Muller NL, Staples CA, Miller RR, Abboud RT. "Density mask". An objective method to quantitate emphysema using computed tomography. *Chest* 1988: 94(4): 782-787.

Table S1. Clinical characteristics of the CanCOLD participants

Characteristics	All	Healthy controls	At-risk of COPD	Mild COPD	Moderate-severe COPD
Number of patients	1,359	306	404	359	290
Age (years; mean ± SD)	66.6 ± 9.7	66.7 ± 9.4	65.5 ± 9.4	67.9 ± 9.6	66.6 ± 10.2
AAT levels (g/L)	1.27 ± 0.20 [15]	1.24 ± 0.18 [3]	1.28 ± 0.20 [7]	1.25 ± 0.19 [2]	1.31 ± 0.21 [3]
Sex – no. of patients (%)					
Male	769 (56.6)	144 (47.1)	229 (56.7)	237 (66.0)	159 (54.8)
Female	590 (43.4)	162 (52.9)	175 (43.3)	122 (34.0)	131 (45.2)
BMI (kg/m^2)	27.6 ± 5.0 [1]	27.3 ± 4.9	28.0 ± 5.2 [1]	27.0 ± 4.5	27.8 ± 5.4
Smoking status – no. of patients (%)					
Never	473 (34.8)	306 (100.0)	0 (0.0)	114 (31.8)	53 (18.3)
Former	668 (49.2)	0 (0.0)	312 (77.2)	196 (54.6)	160 (55.2)
Current	218 (16.0)	0 (0.0)	92 (22.8)	49 (13.6)	77 (26.6)
Pack-years (ever-smokers)	16.7 ± 22.3 [21]	0.0 ± 0.0	18.8 ± 18.7 [14]	18.3 ± 23.1 [2]	29.9 ± 26.8 [5]
Post-bronchodilator FEV1, L	2.6 ± 0.8	2.7 ± 0.8	2.8 ± 0.8	2.7 ± 0.7	1.8 ± 0.6
Post-bronchodilator FEV1 % predicted	91.4 ± 20.3	101.6 ± 17.1	99.2 ± 16.6	95.6 ± 11.9	64.8 ± 12.2
Post-FEV1/FVC	69.4 ± 10.4	77.5 ± 4.7	76.8 ± 4.5	64.5 ± 4.6	56.7 ± 9.6
FEV1 reversibility, %	5.5 ± 8.2	3.5 ± 5.7	3.3 ± 5.7	6.5 ± 6.2	9.6 ± 12.7
FEF25-75	1.8 ± 1.0	2.4 ± 0.9	2.4 ± 1.0	1.3 ± 0.5	0.7 ± 0.4
DLCO, % predicted	106.8 ± 26.7 [72]	112.9 ± 24.0 [12]	110.8 ± 29.6 [27]	108.8 ± 23.0 [12]	91.8 ± 24.5 [21]
RV, % predicted	120.5 ± 34.7 [74]	109.0 ± 27.9 [15]	110.4 ± 26.9 [32]	122.8 ± 32.1 [14]	143.4 ± 41.6 [13]
FRC, % predicted	110.7 ± 25.6 [67]	105.6 ± 24.7 [14]	103.6 ± 21.1 [27]	114.8 ± 22.9 [14]	120.5 ± 30.8 [12]
TLC, % predicted	118.2 ± 17.5 [66]	115.2 ± 18.8 [13]	115.1 ± 15.1 [27]	125.2 ± 16.4 [13]	117.0 ± 17.9 [13]
LAA-950	4.3 ± 4.6 [182]	3.0 ± 3.1 [60]	3.0 ± 3.2 [46]	5.6 ± 5.0 [31]	6.0 ± 6.1 [45]
PD15	-917.4 ± 21.3 [162]	-911.2 ± 21.7 [52]	-911.3 ± 20.5 [40]	-924.9 ± 17.8 [27]	-922.5 ± 21.8 [43]
Pi10_all	4.0 ± 0.2 [162]	3.9 ± 0.2 [52]	4.0 ± 0.1 [40]	4.0 ± 0.2 [27]	4.0 ± 0.2 [43]
Pi10_leq20	3.9 ± 0.1 [162]	3.8 ± 0.1 [52]	3.9 ± 0.1 [40]	3.9 ± 0.1 [27]	3.9 ± 0.1 [43]
MRC scales (Level 1-5)	1.4 ± 0.6 [137]	1.3 ± 0.6 [18]	1.3 ± 0.6 [46]	1.3 ± 0.5 [29]	1.8 ± 0.8 [44]
MRC scales ≥ 3, n (%)	59 (4.8) [137]	6 (2.1) [18]	10 (2.8) [46]	7 (2.1) [29]	36 (14.6) [44]
CAT score	6.9 ± 6.1 [28]	5.5 ± 4.3 [10]	6.1 ± 5.7 [14]	5.6 ± 4.9 [2]	10.8 ± 7.7 [2]
SGRQ-Total score	13.0 ± 14.5 [297]	8.1 ± 9.6 [242]	9.5 ± 11.9 [42]	9.4 ± 10.3 [12]	22.7 ± 17.6 [1]
Self-reported comorbidities, n (%)					
Hypertension	474 (34.9)	107 (35.0)	129 (31.9)	114 (31.8)	124 (42.8)
CVD (any CVD excluding Hypertension)	387 (28.5)	86 (28.1)	97 (24.0)	92 (25.6)	112 (38.6)
Diabetes	144 (10.6)	29 (9.5)	48 (11.9)	29 (8.1)	39 (13.1)
Asthma	316 (23.3)	46 (15.0)	67 (16.6)	84 (23.4)	119 (41.0)
Respiratory medications, n (%)					

SAMA/SABA	46 (3.4)	4 (1.3)	5 (1.2)	17 (4.7)	20 (6.9)
LABA± SAMA/SABA	3 (0.2)	1 (0.3)	0 (0.0)	1 (0.3)	1 (0.3)
LAMA± SAMA/SABA	14 (1.0)	0 (0.0)	2 (0.5)	1 (0.3)	11 (3.8)
LAMA+LABA± SAMA/SABA	84 (6.2)	12 (3.9)	20 (5.0)	26 (7.2)	26 (9.0)
ICS± SAMA/SABA	2 (0.1)	0 (0.0)	0 (0.0)	0 (0.0)	2 (0.7)
LABA+ICS± SAMA/SABA	111 (8.2)	11 (3.6)	20 (5.0)	22 (6.1)	58 (20.0)
LAMA+ICS± SAMA/SABA	4 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	4 (1.4)
LAMA+LABA+ICS± SAMA/SABA	40 (2.9)	0 (0.0)	1 (0.2)	4 (1.1)	35 (12.1)
Any above medications	304 (22.4)	28 (9.2)	48 (11.9)	71 (19.8)	157 (54.3)

The number of missing values is indicated in square brackets. For lung function measurements, the maximum values among all the trials are indicated.

AAT: alpha-1 antitrypsin; BMI, body mass index; CAT, COPD Assessment Test; CVD, cardiovascular diseases; DLCO: diffusion capacity to carbon monoxide; FEF25-75: forced expiratory flow at 25–75%; FEV1: forced expiratory volume in 1 second; FRC: functional residual capacity; FVC: forced vital capacity; ICS, inhaled corticosteroid; LAA-950: percentage of lung voxels below -950 Hounsfield units; LABA, long-acting β-agonist; LAMA, long-acting muscarinic antagonist; MRC, modified British Medical Research Council Questionnaire; PD15: 15th percentile of the CT attenuation histogram; Pi10_all: airway wall thickness of an airway with internal perimeter of 10mm generated using all CT airways; Pi10_leq20: airway wall thickness of a theoretical airway with internal perimeter of 10mm generated using all CT airways with an internal perimeter <=20mm; RV: residual volume; SABA, short-acting β-agonist; SAMA, short-acting muscarinic antagonist; SGRQ, St. George's Respiratory Questionnaire for COPD; TLC: total lung capacity.

Table S2. Clinical characteristics by *SERPINA1* genotyping groups in ever- and never-smokers

Characteristics	Group 1: no deficiency	Group 2: mild deficiency	Group 3: intermediate deficiency	Group 4: severe deficiency	P value*	P value adj. age, sex, and study sites
Ever-smokers	**n=748**	**n=92**	**n=44**	**n=2**		
Age (years; mean ± SD)	66.6 ± 9.7	65.1 ± 9.4	68.6 ± 8.2	61.5 ± 4.9	0.208	-
AAT levels (g/L)	1.33 ± 0.17 [10][a]	1.12 ± 0.14[b]	0.98 ± 0.21 [1][c]	0.41 ± 0.31[d]	**<0.001**	**<0.001**
Sex – no. of patients (%)					0.306	
Male	441 (59.0)	59 (64.1)	27 (61.4)	0 (0.0)		
Female	307 (41.0)	33 (35.9)	17 (38.6)	2 (100.0)		
BMI (kg/m^2)	27.9 ± 5.1 [1]	28.4 ± 5.6	27.1 ± 4.6	28.3 ± 6.5	0.748	0.834
Smoking status – no. of patients (%)					0.909	0.96
Former	561 (75.0)	70 (76.1)	35 (79.5)	2 (100.0)		
Current	187 (25.0)	22 (23.9)	9 (20.5)	0 (0.0)		
Pack-years (ever-smokers)	26.0 ± 23.4 [19]	27.1 ± 22.2 [1]	21.9 ± 20.1 [1]	23.5 ± 30.5	0.621	0.757
Post-bronchodilator FEV1, L	2.5 ± 0.8	2.6 ± 0.8	2.5 ± 0.8	1.8 ± 0.4	0.463	0.854
Post-bronchodilator FEV1 % predicted	88.6 ± 20.9	88.5 ± 20.1	87.8 ± 22.0	72.4 ± 11.4	0.613	0.768
Post-FEV1/FVC	68.0 ± 10.8	67.6 ± 10.1	67.2 ± 13.2	52.2 ± 8.4	0.259	0.138
FEV1 reversibility, %	5.8 ± 8.2	7.1 ± 9.5	5.7 ± 12.3	14.3 ± 18.2	0.544	0.325
FEF25-75	1.7 ± 1.0	1.7 ± 1.0	1.7 ± 1.0	0.5 ± 0.2	0.302	0.387
DLCO, % predicted	103.9 ± 24.7 [42]	104.0 ± 22.7 [7]	100.7 ± 26.2 [4]	73.8 ± 11.4	0.308	0.329
RV, % predicted	123.4 ± 36.2 [42]	129.4 ± 37.0 [7]	122.5 ± 35.6 [3]	164.3 ± 41.9	0.209	0.128
FRC, % predicted	111.6 ± 26.0 [36]	114.6 ± 25.4 [7]	111.1 ± 28.3 [3]	145.3 ± 29.6	0.241	0.125
TLC, % predicted	118.7 ± 17.4 [35]	120.2 ± 15.5 [8]	116.9 ± 17.2 [3]	137.5 ± 19.1	0.354	0.248
LAA-950 (%)	4.6 ± 4.9 [79][a]	4.3 ± 4.2 [15][a]	4.6 ± 5.6 [2][a]	14.7 ± 4.0[b]	**0.034**	**0.010**
PD15	-918.1 ± 21.8 [79]	-917.3 ± 18.9 [15]	-918.6 ± 20.7 [2]	-948.2 ± 10.1	0.205	0.109
Pi10_all	4.0 ± 0.2 [79]	4.0 ± 0.1 [15]	4.0 ± 0.2 [2]	3.9 ± 0.1	0.297	0.713
Pi10_leq20	3.9 ± 0.1 [79]	3.9 ± 0.1 [15]	3.9 ± 0.1 [2]	3.9 ± 0.0	0.643	0.543
MRC scales (Level 1-5)	1.50 ± 0.68 [40]	1.46 ± 0.70 [8]	1.60 ± 0.90 [4]	2.00 ± 0.00	0.435	0.638
MRC scales ≥ 3, n (%)	47 (6.6) [40]	6 (7.1) [8]	6 (15.0) [4]	0 (0.0)	0.256	0.359
CAT score	7.4 ± 6.4 [15]	8.8 ± 7.4 [2]	7.7 ± 8.8	6.5 ± 2.1	0.34	0.226
SGRQ-Total score	14.1 ± 14.7 [37]	17.3 ± 18.4 [3]	16.8 ± 19.6 [5]	17.5 ± 0.8	0.531	0.178
Never-smokers	**n=401**	**n=55**	**n=15**	**n=2**		
Age (years; mean ± SD)	67.1 ± 9.8	64.6 ± 10.1	69.3 ± 10.1	75.0 ± 11.3	0.122	-
AAT levels (g/L)	1.28 ± 0.16 [2][a]	1.07 ± 0.14 [1][b]	0.88 ± 0.15 [1][c]	0.60 ± 0.08[c]	**<0.001**	**<0.001**
Sex – no. of patients (%)					0.843	
Male	206 (51.4)	26 (47.3)	9 (60.0)	1 (50.0)		
Female	195 (48.6)	29 (52.7)	6 (40.0)	1 (50.0)		

BMI (kg/m^2)	26.9 ± 4.8	27.2 ± 4.9	26.8 ± 3.1	23.0 ± 0.7	0.523	0.868
Post-bronchodilator FEV1, L	2.6 ± 0.8	2.7 ± 1.0	2.9 ± 0.9	2.6 ± 0.6	0.752	0.310
Post-bronchodilator FEV1 % predicted	96.9 ± 18.2	94.5 ± 18.7	104.0 ± 15.0	107.8 ± 30.5	0.262	0.467
Post-FEV1/FVC	72.3 ± 8.9	72.4 ± 8.4	74.2 ± 7.8	67.1 ± 14.8	0.901	0.698
FEV1 reversibility, %	4.7 ± 7.4	6.0 ± 7.9	3.2 ± 4.9	1.5 ± 2.1	0.282	0.526
FEF25-75	2.0 ± 1.0	2.0 ± 1.0	2.2 ± 0.9	1.6 ± 0.5	0.622	0.603
DLCO, % predicted	111.0 ± 22.9 [14][a]	119.6 ± 29.5 [4][a]	108.2 ± 22.1 [2]	64.9 ± 29.8[b]	**0.003**	**0.003**
RV, % predicted	113.9 ± 31.2 [19]	118.0 ± 30.1 [3]	102.1 ± 15.7	127.5 ± 13.4	0.208	0.461
FRC, % predicted	108.0 ± 24.8 [18]	112.0 ± 25.0 [3]	101.9 ± 14.8	128.5 ± 1.8	0.233	0.486
TLC, % predicted	116.9 ± 17.5 [17]	120.3 ± 22.2 [3]	114.3 ± 10.5	130.1 ± 2.8	0.409	0.459
LAA-950 (%)	3.6 ± 3.5 [56][a]	3.7 ± 4.3 [10][a]	3.1 ± 1.7[a]	25.1 ± 17.9[b]	**<0.001**	**<0.001**
PD15	-915.6 ± 20.9 [56][a]	-915.2 ± 21.4 [10][a]	-919.0 ± 10.9	-960.9 ± 20.3[b]	**0.020**	**0.032**
Pi10_all	3.9 ± 0.1 [56]	3.9 ± 0.1 [10]	4.0 ± 0.2	4.1 ± 0.2	0.419	0.272
Pi10_leq20	3.9 ± 0.1 [56]	3.8 ± 0.1 [10]	3.9 ± 0.1	4.0 ± 0.1	0.379	0.519
MRC scales (Level 1-5)	1.3 ± 0.6 [11]	1.3 ± 0.6 [3]	1.2 ± 0.4	1.5 ± 0.7	0.619	0.745
MRC scales ≥ 3, n (%)	11 (2.8) [11]	2 (3.8) [3]	0 (0.0)	0 (0.0)	0.791	0.973
CAT score	5.5 ± 4.7 [10]	6.4 ± 5.0 [1]	2.8 ± 3.3	5.0 ± 4.2	0.072	0.08
SGRQ-Total score	10.2 ± 12.5 [218]	10.4 ± 10.9 [23]	2.3 ± 1.5 [10]	25.0 ± . [1]	0.155	0.251

See footnotes of Table 3.

Table S3. Clinical characteristics by *SERPINA1* genotyping groups with MZ heterozygotes-only in group 3.

Characteristics	Group 1: no deficiency n=1,149	Group 2: mild deficiency n=147	Group 3: MZ n=40	Group 4: severe deficiency n=4	P value*	P value adj. age, sex, and study sites
Age (years; mean ± SD)	66.8 ± 9.7	64.9 ± 9.6	69.9 ± 8.0	68.3 ± 10.6	**0.028**	-
AAT levels (g/L)	1.31 ± 0.17 [12][a]	1.10 ± 0.14 [1][b]	0.88 ± 0.13 [2][c]	0.50 ± 0.21[d]	**<0.001**	**<0.001**
Sex – no. of patients (%)					0.307	-
Male	647 (56.3)	85 (57.8)	27 (67.5)	1 (25.0)		
Female	502 (43.7)	62 (42.2)	13 (32.5)	3 (75.0)		
BMI (kg/m^2)	27.6 ± 5.0 [1]	27.9 ± 5.3	27.2 ± 4.1	25.7 ± 4.8	0.735	0.917
Smoking status – no. of patients (%)					0.523	0.725
Never	401 (34.9)	55 (37.4)	9 (22.5)	2 (50.0)		
Former	561 (48.8)	70 (47.6)	26 (65.0)	2 (50.0)		
Current	187 (16.3)	22 (15.0)	5 (12.5)	0 (0.0)		
Pack-years (ever-smokers)	16.7 ± 22.5 [19]	16.9 ± 21.9 [1]	16.3 ± 19.7	11.7 ± 22.2	0.956	0.983
Post-bronchodilator FEV1, L	2.55 ± 0.81	2.62 ± 0.88	2.6 ± 0.9	2.18 ± 0.62	0.654	0.976
Post-bronchodilator FEV1 % predicted	91.5 ± 20.4	90.8 ± 19.7	89.5 ± 21.5	90.1 ± 27.8	0.948	0.852
Post-FEV1/FVC	69.5 ± 10.4	69.4 ± 9.8	66.3 ± 12.3	59.7 ± 13.0	0.166	0.107
FEV1 reversibility, %	5.4 ± 7.9	6.7 ± 8.9	6.6 ± 12.6	7.9 ± 12.9	0.496	0.282
FEF25-75	1.77 ± 1.04	1.79 ± 1.04	1.6 ± 0.8	1.05 ± 0.68	0.421	0.482
DLCO, % predicted	106.4 ± 24.3 [56][a]	109.9 ± 26.5 [11][a]	95.5 ± 21.2 [4][b]	69.4 ± 19.1[b]	**<0.001**	**<0.001**
RV, % predicted	120.1 ± 34.8 [61]	125.1 ± 34.9 [10]	112.9 ± 34.5 [2]	145.9 ± 33.1	**0.036**	0.089
FRC, % predicted	110.3 ± 25.6 [54]	113.6 ± 25.2 [10][a]	104.4 ± 26.0 [2][b]	136.9 ± 19.7	**0.014**	**0.032**
TLC, % predicted	118.1 ± 17.4 [52]	120.2 ± 18.3 [11]	116.1 ± 15.2 [2]	133.8 ± 11.9	0.116	0.117
LAA-950 (%)	4.30 ± 4.52 [135][a]	4.11 ± 4.23 [25][a]	5.00 ± 5.68 [1][a]	19.89 ± 12.18[b]	**0.008**	**<0.001**
PD15	-917.2 ± 21.5 [135][a]	-916.5 ± 19.8 [25][a]	-922.3 ± 18.2 [1][a]	-954.5 ± 15.0[b]	**0.007**	**0.002**
Pi10_all	3.96 ± 0.16 [135]	3.97 ± 0.14 [25]	4.01 ± 0.16	4.00 ± 0.14	0.16	0.204
Pi10_leq20	3.87 ± 0.11 [135]	3.88 ± 0.11 [25]	3.91 ± 0.12 [1]	3.90 ± 0.08	0.174	0.215
MRC scales (Level 1-5)	1.44 ± 0.65 [51]	1.40 ± 0.67 [11]	1.5 ± 0.9 [2]	1.75 ± 0.50	0.333	0.77
MRC scales ≥ 3, n (%)	58 (5.3) [51]	8 (5.9) [11]	5 (13.2) [2]	0 (0.0)	0.199	0.284
CAT score	6.8 ± 5.9 [25]	7.9 ± 6.7 [3]	6.3 ± 8.4	5.8 ± 2.9	**0.039**	0.122
SGRQ-Total score	13.3 ± 14.4 [255]	15.5 ± 17.0 [26]	14.6 ± 18.5 [10]	20.0 ± 4.4 [1]	0.336	0.324
Self-reported comorbidities, n (%)						
Hypertension	402 (35.0)	51 (34.7)	14 (35.0)	2 (50.0)	0.929	0.778
CVD (any CVD excluding Hypertension)	325 (28.3)	40 (27.2)	15 (37.5)	1 (25.0)	0.581	0.893
Diabetes	123 (10.7)	13 (8.8)	5 (12.5)	0 (0.0)	0.808	0.96

Asthma	267 (23.2)	32 (21.8)	13 (32.5)	1 (25.0)	0.478	0.29
Respiratory medications, n (%)						
SAMA/SABA	35 (3.0)	6 (4.1)	1 (2.5)	0 (0.0)	0.778	0.931
LABA± SAMA/SABA	3 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	1	1
LAMA± SAMA/SABA	10 (0.9)	3 (2.0)	1 (2.5)	0 (0.0)	0.19	0.478
LAMA+LABA± SAMA/SABA	71 (6.2)	11 (7.5)	2 (5.0)	0 (0.0)	0.844	0.962
ICS± SAMA/SABA	2 (0.2)	0 (0.0)	0 (0.0)	0 (0.0)	1	1
LABA+ICS± SAMA/SABA	99 (8.6)	10 (6.8)	1 (2.5)	1 (25.0)	0.233	0.383
LAMA+ICS± SAMA/SABA	3 (0.3)	1 (0.7)	0 (0.0)	0 (0.0)	0.46	0.819
LAMA+LABA+ICS± SAMA/SABA	32 (2.8)	5 (3.4)	3 (7.5)	0 (0.0)	0.264	0.544
Any above medications	255 (22.2)	36 (24.5)	8 (20.0)	1 (25.0)	0.87	0.944

See footnotes of Table 3.

Table S4. Primers to sequence the *SERPINA1* gene

	Forward	**Reverse**	**Amplicon Size (bp)**
Exon 2	5'- AATGCATTGCCAAGGAGAGTT C -3'	**5'- CAAGTACTTGGCACAGGC** TG **-3'**	866
Exon 3	5'- CCTCAGTCCCAACATGGCTAA G -3'	**5'- AGGGATGTGTGTCGTCAA** GG **-3'**	460
Exon 4	5'- CTTCCTCAGCCTCAGGAC AG -3'	**5'- AGTTCCCATCTTAGTGTGGGT** G **-3'**	376
Exon 5	**5'- GGATTTACAGTCACATGCAGG** C **-3'**	5'- GGGAGTGAGCGCTTC CTG -3'	387

Sequencing primers are shown in bold.

Table S5. Primers for allele-specific PCR

AS-PCR specificity	Allele-Specific Forward Primer sequence	Common Reverse Primer sequence	Size of the amplicon	PCR conditions*
rs1303 Glu376Asp	5'- CATGAAGAGGGGAGACTTGGTATTTTGG -3'	5'- CTGAGTTCGCCTTCAGCCTATACC-3'	4598 bp	**A**
	5'- CATGAAGAGGGGAGACTTGGTATTTTGT -3			
rs1303 Glu376Asp	5'- CATGAAGAGGGGAGACTTGGTATTTTGG -3'	5'- CAAGTACTTGGCACAGGCTGG -3'	4888 bp	**A**
	5'- CATGAAGAGGGGAGACTTGGTATTTTGT -3'			
rs1303 Glu376Asp	5'- CATGAAGAGGGGAGACTTGGTATTTTGG -3'	5'- AGGGATGTGTGTCGTCAAGG -3'	2789 bp	**A**
	5'- CATGAAGAGGGGAGACTTGGTATTTTGT -3'			
rs1379209512 Leu-2Gln	5'- CCTGGTCCCTGTCTCCCA -3'	5'- CCTCAGTCCCAACATGGCTAAG -3'	2382 bp	**B**
	5'- CCTGGTCCCTGTCTCCCT -3'			
rs759736224 Pro255Thr	5'- GTGCTGTAGTTTCCCCTCATCAGT -3'	5'- CTGGCTGAGTTCGCCTTCAG -3'	2149 bp	**C**
	5'- GTGCTGTAGTTTCCCCTCATCAGG -3'			
rs766025736 Arg101Cys	5'- TGAAGGCTTCCAGGAACTCCTCT -3	5'- CCTCAGTCCCAACATGGCTAAG -3	2082 bp	**D**
	5'- TGAAGGCTTCCAGGAACTCCTCC -3			
rs28929474 Glu342Lys	5'- CCAGCAGCTTCAGTCCCTTTCTT -3'	5'- AGGGATGTGTGTCGTCAAGG -3'	2680 bp	**E**
	5'- CCAGCAGCTTCAGTCCCTTTCTC -3'			
rs745463238 Ser14Phe	5'- GCTGCCCAGAAGACAGATACATT -3'	5'- CCTCAGTCCCAACATGGCTAAG -3'	2342 bp	**F**
	5'- GCTGCCCAGAAGACAGATACATC -3'			
rs9630 Ala332Ala	5'- GTCAGCACAGCCTTATGCACA -3'	5'- AGGGATGTGTGTCGTCAAGG -3'	2650 bp	**D**
	5'- GTCAGCACAGCCTTATGCACG -3'			
rs1049800 Asp256Asp	5'- AGGTGCTGTAGTTTCCCCTCG -3'	5'- CTGAGTTCGCCTTCAGCCTATACC -3'	2148 bp	**A**
	5'- CAGGTGCTGTAGTTTCCCCTCA -3'			
rs200414579 Gly167Gly	5'- GCTCCTTGACCAAATCCACAATTTTT -3'	5'- CTGAGTTCGCCTTCAGCCTATACC -3'	435 bp	**G**
	5'- GCTCCTTGACCAAATCCACAATTTTC -3'			
rs147283849 Leu-10Leu	5'- ATCCTCCTGCTGGCAGGCT -3'	5'- CCTCAGTCCCAACATGGCTAAG -3'	2408 bp	**A**
	5'- ATCCTCCTGCTGGCAGGCC -3'			

*see Table S6.

Table S6. AS-PCR cycling conditions

PCR conditions	Taq polymerase	Initial Denaturation	Denaturation	Annealing Temperature	Elongation
A	NEB Q5	98°C, 30 sec	98°C, 10 sec	68°C, 30 sec	72°C, 3 min
B	NEB Q5	98°C, 30 sec	98°C, 10 sec	69°C, 30 sec	72°C, 3 min
C	NEB Q5	98°C, 30 sec	98°C, 10 sec	70°C, 30 sec	72°C, 3 min
D	NEB Q5	98°C, 30 sec	98°C, 10 sec	72°C, 30 sec	72°C, 3 min
E	HotStarTaq	95°C, 15 min	94°C, 15 sec	61.5°C, 40 sec	72°C, 2.5 min
F	HotStarTaq	95°C, 15 min	94°C, 15 sec	63°C, 40 sec	72°C, 2.5 min
G	HotStarTaq	95°C, 15 min	94°C, 15 sec	57°C, 30 sec	72°C, 1 min

A, B, C and **D**: Final reaction volume of 50 µL, 1X PCR buffer, 1X Q5 GC enhancer buffer, 200 µM of each dNTP, 500 nM of each primer, 2.5 Units of Taq Polymerase and 35 PCR cycles of denaturation-annealing-elongation.

E and **F**: Final reaction volume of 50 µL, 1X PCR buffer, 1X Q-Solution, 160 µM of each dNTP, 400 nM of each primer, 2 Units of Taq Polymerase and 35 PCR cycles of denaturation-annealing-elongation.

G: Final reaction volume of 25 µL, 1X PCR buffer, 1X Q-Solution, 160 µM of each dNTP, 200 nM of each primer, 1 Units of Taq Polymerase and 35 PCR cycles of denaturation-annealing-elongation.

NEB Q5: Q5® High-Fidelity DNA Polymerase, New England Biolabs Ltd., Ontario, Canada

HotStarTaq: HotStarTaq DNA polymerase, Qiagen, Ontario, Canada

Table S7. Frequencies, allelic background, AAT serum levels, and lung phenotypes among carriers of rare genetic variants with no Pi typing.

Exon	rs #	Protein	Background	# of carriers	AAT levels	DLCO, % predicted	LAA-950
2	rs1343069141	Ser-19Ser	M3	1	1.22	NA	4.44
2	rs147283849	Leu-10Leu	M1 (Ala^{213})	3	1.06±0.03	133.24±33.09	3.59±1.48 [1]
2	rs1379209512	Leu-2Gln	M1 (Ala^{213})	2	1.27±0.06	91.34±35.24	2.32 [1]
2	rs150784949	Phe33Phe	M1	2	1.17±0.10	103.72±14.21	4.19 [1]
2	rs113817720	Ala58Thr	M1	1	1.12	129.44	2.39
2	rs758820515	Pro88Leu	M1 (Ala^{213})	1	1.00	136.63	2.43
2	rs766025736	Arg101Cys	S	1[a]	0.99	146.27	4.20
2	rs1344951022	Thr113Thr	M1	1	1.24	NA	2.75
2	rs200414579	Gly167Gly	M3	1	1.39	112.26	7.81
3	New1	Lys222Lys	M1	1	1.32	96.62	1.30
3	rs759736224	Pro255Thr	M3	1	1.38	72.48	1.39
3	rs1049800	Asp256Asp	M1 or M1 (Ala^{213})[b]	6	1.27±0.17	110.16±17.70	8.52±10.72
3	New2	Lys274Asn	M1 (Ala^{213})	1	1.24	101.40	0.55
4	rs141620200	Ala284Ser	M1 (Ala^{213})	11	1.20±0.21 [1]	96.11±17.97 [2]	3.32±4.26 [1]
4	rs139964603	Val302Ile	M1	1	1.16	104.97	12.98
4	New3	Leu318Phe	M1	1	1.50	97.17	3.29
5	rs9630	Ala332Ala	M1 or M1 (Ala^{213})[c]	2	1.25±0.16	117.53 [1]	2.84±1.17
5	rs201318727	Ile340Val	R	1	1.34	114.92	2.82

The number of missing values is indicated in square brackets

[a]MS

[b]Background M1 for individuals of Asian ancestry (n=5) and M1 (Ala^{213}) for European ancestry (n=1).

[c]Background M1 for one individual of European ancestry and M1 (Ala^{213}) for African ancestry (n=1).

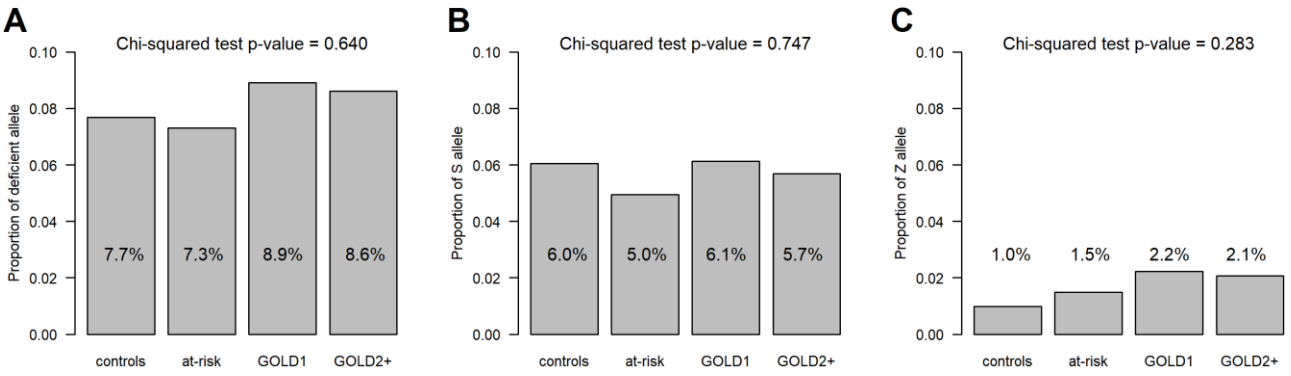

Figure S1. Proportion of all deficient alleles (A), S allele (B), and Z allele in CanCOLD subgroups including healthy control free of smoking history and airway obstruction (controls), ever-smokers free of airway obstruction (at-risk), individuals with mild COPD (GOLD1) and individuals with moderate-severe COPD (GOLD2+).

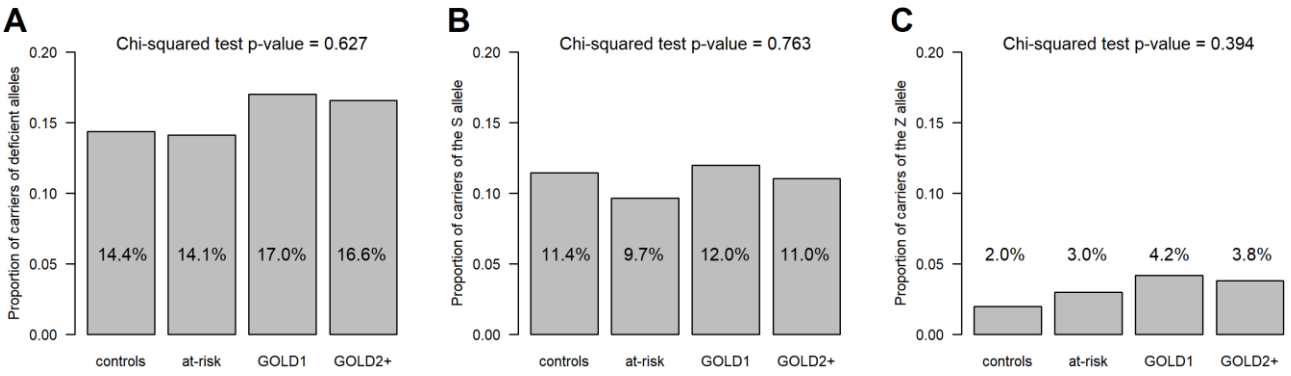

Figure S2. Proportion of carriers of deficient alleles (A), S allele (B), and Z allele in CanCOLD subgroups including healthy control free of smoking history and airway obstruction (controls), ever-smokers free of airway obstruction (at-risk), individuals with mild COPD (GOLD1) and individuals with moderate-severe COPD (GOLD2+).