ONLINE DATA SUPPLEMENT

Serum levels and genotype distribution of alpha$_1$-antitrypsin in the general population

Ilaria Ferrarotti[1], Gian Andri Thun[2,3], Michele Zorzetto[1], Stefania Ottaviani [1], Medea Imboden [2,3], Christian Schindler[2,3], Arnold von Eckardstein[4], Lucia Rohrer[4], Thierry Rochat[5] , Erich Russi[6], Nicole M Probst-Hensch [2,3], Maurizio Luisetti [1]

SUPPLEMENTARY APPENDIX

Methods

SNP analysis

Results

Epidemiology of SERPINA1 Genotypes in Swiss language groups

ROC analysis for accuracy of predicting genotype classes

Analysis of normal M variant subtypes

Discussion

Frequency of SERPINA1 Gene variants in Switzerland

Effect of SERPINA1 M variants on AAT serum concentration

Methods

SNP analysis

The PCR conditions were identical for all applications: 0.125 μl of 20X working stock of SNPGenotypingAssay, 2.5 μl LightCycler 480 Probes Master (Roche Diagnostics), and 20 ng DNA sample, in a total volume of 5 μl. PCR cycling conditions were also identical for all assays: initial denaturation step of 95°C for 10 minutes, followed by 40 cycles of denaturation at 95°C for 15 seconds, and annealing at 60°C for 1 minute. After the amplification, melting curves were generated by slowly (ramp rate 2.2°C/second) cooling the sample to 40°C. The sequences of primers and probes are available on request. Subjects with low AAT serum levels who had been previously genotyped for S (rs17580) and Z (rs28929474) deficient variants by PCR-RFLP[E1], were re-analzyed by Taq-Man probes and had a good reproducibility score with an agreement of 99.7% and 99.6% for S and Z SNPs, respectively.

Results

Frequency of Z and S alleles in the Swiss language groups

We analyzed the frequency of S and Z alleles in the three language groups: German (n=3,288), French (n=1,938) and Italian (n=831) (Table E1). Interestingly, while the PI*Z allele was

homogeneously distributed among the three language groups, the frequency of the PI*S allele was significantly higher in the French subgroup (0.053) than in the German (0.036) and in the Italian (0.025) ($p<0.001$ for both comparisons).

ROC analysis for accuracy of predicting genotype classes

As we had previously reported differences in circulating AAT concentrations by sex, smoking status and CRP levels[E2], we tried to assess whether prediction of PI*SS or MZ genotypes could be improved by considering the influence of these factors on AAT blood levels. We found a marginal enhancement of prediction quality when adjusting for sex and current smoking status, but not for CRP (AUC=0.9927, p=0.05). We subsequently recalculated values for normal (PI*MM and PI*MS) vs. intermediate deficient genotype classes (PI*MZ and PI*SS) by sex and current smoking (Table E3). Despite slight differences in the stratified specific means, the impact of the genotype played a much bigger role than that of sex and current smoking.

Analysis of normal M variant subtypes

The analysis of the three SNPs for normal variants (M1Ala/M1Val - rs6647; M3 - rs1303; M2/M4 - rs709932) by haplotype reconstruction revealed 14 normal genotypic classes in the PI*MM group, 5 classes in the PI*MS group and 5 classes in the PI*MZ group (Table E4). Among these, the most common were PI*M1(Val)M1(Val), PI*M1(Ala)M1(Val), and M1(Val)M2 (frequencies of 0.28, 0.21, and 0.15, respectively). The reference intervals (5^{th}-95^{th} percentiles) for AAT serum concentration have been calculated in each group (Figure E3). Comparison of means within each group revealed only a significant difference between PI*M1(Val)Z and PI*M1(Ala)Z (0.819 vs. 0.754 g/L, p=0.003).

Discussion

Frequency of SERPINA1 Gene variants in Switzerland.

The estimated mean gene frequencies for PiS and PiZ in Switzerland are 0.0384 and 0.0073, respectively[E3] according to data based on the analysis of three Swiss cohorts[E4,E5] previously phenotyped for PI. The estimate is similar to ours for the S allele, but we obtained a slightly higher frequency for the Z allele. When we divided the population into the three language groups, we found evidence of a significantly higher frequency for the S allele in the French subgroup ($p<0.001$ for both comparisons, Table E1). This is in agreement with the hypothesis that the S mutation, which arose in the Portuguese population[E6], moved eastbound to the rest of Europe as a consequence of the late-glacial resettlement of Europe[E7], resulting in decreasing frequencies from southwest to the north and east.

Effect of SERPINA1 M variants on AAT serum concentration.

The normal variants of AAT, usually called M, are characterized by point mutations that neither change the phenotype nor the serum concentration of the AAT protein. The most common are M1Ala/M1Val (213Ala/213Val)[E8], M2 (213Val, 376Asp, and 101His)[E9], M3 (213Val and 376Asp)[E10], M4 (213Val and 101His)[E11]. It is well known that individuals bearing these mutations have normal AAT serum levels and that the protein functions normally as an inhibitor of neutrophil elastase.

Evaluation of the crystallographic structure of AAT[E12] revealed that the substitution at position 101 occurs in helix D and the 376 substitution occurs in sheet 4B of the molecule, which are areas where changes are likely to cause minor conformational changes in the molecule.

Nevertheless, epidemiological studies of these normal variants are limited to their geographic distribution or prevalence data, and comparison among different genotypes, in terms of serum concentration of AAT, has never been performed. A side aim of this paper, was to confirm the absence of quantitative differences among PI*MM subtypes; therefore, an analysis of all possible combinations of M1, M2, M3, and M4 alleles in the cluster PI*MM, was performed. The effect of the M1(Ala) allele, in combination with the Z allele, in reducing the mean concentration of AAT in serum (Figure E3), should be considered with caution, because these data are not supported by biochemical explanations[E13].

Supplement References

E1. **Zorzetto M**, Russi EW, Senn O, *et al*. SERPINA1 gene variants in subjects from the general population with reduced alpha1-antitrypsin level. *Clin Chem* 2008; **54**:1331-8.

E2. **Senn O**, Russi EW, Schindler C, *et al*. Circulating alpha1-antitrypsin in the general population: determinants and association with lung function.Circulating alpha1-antitrypsin in the general population : determinants and association with lung function. *Respir Res* 2008;**9**:35

E3. **de Serres FJ**. Worldwide racial and ethnic distribution of alpha1-antitrypsin deficiency: summary of an analysis of published genetic epidemiologic surveys. *Chest* 2002;**122:**1818-29.

E4. **Scheffrahn W**, Ziggiotti E. Electrophoretic alpha 1-antitrypsin variation in the Swiss population. *Anthropol Anz* 1982;**40**:137-43.

E5. **Bär W**, Kratzer A. Polymorphism of alpha-1-antitrypsin (Pi) in the Swiss population determined by isoelectric focusing with an immobilized pH gradient. *Hum Hered* 1988;**38**:106-10.

E6. **Seixas S**, Garcia O, Trovoada MJ, *et al*. Patterns of haplotype diversity within the serpin gene cluster at 14q32.1: insights into the natural history of the alpha1-antitrypsin polymorphism. *Hum Genet* 2001;**108**:20-30.

E7. **Pereira L**, Richards M, Goios A, *et al*. High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. *Genome Res* 2005;**15**:19-24.

E8. **NukiwaT**, Satoh K, Brantly ML, *et al*. Identification of a second mutation in the protein-coding sequence of the Z type alpha 1-antitrypsin gene. *J Biol Chem* 1986; **261**:15989-94.

E9. **Nukiwa T**, Brantly ML, Ogushi F, *et al*. Characterization of the gene and protein of the common alpha 1-antitrypsin normal M2 allele. *Am J Hum Genet* 1988;**43**:322-30.

E10. **Graham A**, Hayes K, Weidinger S, *et al*. Characterisation of the alpha-1-antitrypsin M3 gene, a normal variant. *Hum Genet* 1990;**85**:381-2.

E11. **Okayama H**, Holmes MD, Brantly ML, *et al*. Characterization of the coding sequence of the normal M4 alpha 1-antitrypsin gene. *Biochem Biophys Res Commun* 1989;**162**:1560-70.

E12. **Loebermann H**, Tokuoka R, Deisenhofer J, *et al*. Human alha1-proteinase inhibitor. Crystal structure analysis oft wo crystal modifications, molecular model and preliminary analysis oft he implications for function. *J Mol Biol*. 1984;**177**:531-57.

E13. **Nukiwa T**, Brantly M, Ogushi F, *et al.* Characterization of the M1(Ala213) type of alpha 1-antitrypsin, a newly recognized, common "normal" alpha 1-antitrypsin haplotype. *Biochemistry* 1987;**26**:5259-67.

Legends to Supplemental Figures

Figure E1 – Intervals (5^{th}-95^{th} percentiles) for unadjusted serum AAT levels in the main SERPINA1 genotypic classes;, 1^{st} -99^{th} percentiles are represented with lines, where possible. In this figure, the g/L values of our analysis were mathematically converted to μM, based on a molecular weight of 52kDa.

Figure E2: ROC curve for predicting MM/MS vs SS/MZ genotype classes from unadjusted AAT blood level (rare variants included).

Figure E3 – Intervals (5^{th}- 95^{th} percentiles) for AAT serum concentration in individuals stratified according to the haplotype reconstruction resulting in 24 genotypic classes.

Table E1 – Frequencies of Z and S alleles in the cohort and the three language groups

	Z allele (%)	S allele (%)
General population (n=6,057)	1.30	4.01
German subgroup (n=3,288)	1.22	3.65
French subgroup (n=1,938)	1.44	5.26
Italian subgroup (n=831)	1.26	2.53
p-value (chi-square-test)	0.61	<0.001

Table E2 – Intervals (5^{th}- 95^{th} percentiles) of AAT serum concentration (g/L) in individuals stratified according to genotype and CRP serum concentration. In the case of low number of samples, intervals were not calculated, and the lowest and highest values were reported.

	PI*MM		PI*MS		PI*SS		PI*MZ		PI*SZ	
	CRP<8	CRP≥8	CRP<8	CRP≥8	CRP<8	CRP≥8	CRP<8	CRP≥8	CRP<8	CRP≥8
N	5,092	274	426	25	9	1	136	7	10	0
5^{th} perc.	0.960	1.230	0.870	1.107			0.660		0.490	
95^{th} perc.	1.590	2.086	1.320	1.545			0.990		0.660	
Lowest value					0.730	1.060		0.820		
Highest value					0.910	1.060		1.060		

Table E3 - Influence of sex and current smoking status on unadjusted AAT (g/L) reference values

	n	median	5^{th}/95^{th} perc.
MM/MS all	5,848	1.26	1.01/1.63
MM/MS, male-smoker	791	1.29	1.01/1.59
MM/MS, female-smoker	663	1.36	1.07/1.74
MM/MS, male-nonsmoker	2,127	1.19	0.98/1.48
MM/MS, female-	2,267	1.28	1.03/1.69

nonsmoker			
SS/MZ, all	155	0.79	0.66/1.01
SS/MZ, male-smoker	16	0.79	0.69/1.02
SS/MZ, female-smoker	14	0.88	0.76/1.07
SS/MZ, male-nonsmoker	55	0.75	0.62/0.91
SS/MZ, female-nonsmoker	70	0.81	0.66/1.04

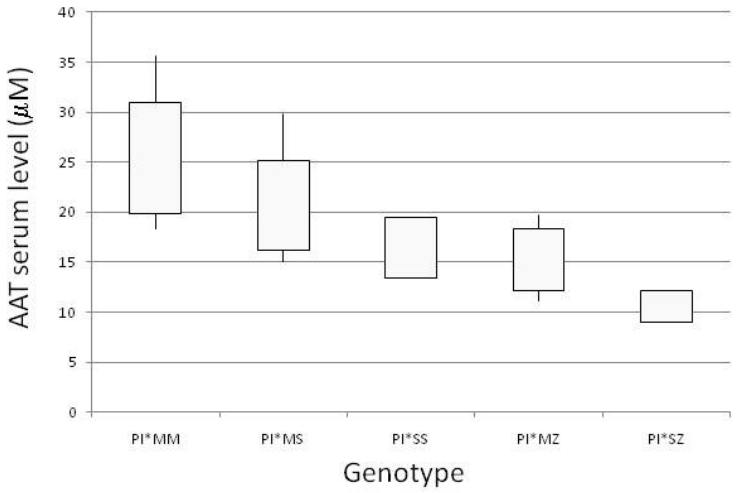
Table E4 – Frequencies of the genotypic classes

PI Group	**PI Genotype**	**Frequencies**
PI*MM		
	M1(Val) M1(Val)	0.2512
	M1(Ala) M1(Val)	0.1850
	M1(Val) M2 or M3 M4	0.1620
	M1(Val) M3	0.0908
	M1(Ala) M2	0.0559

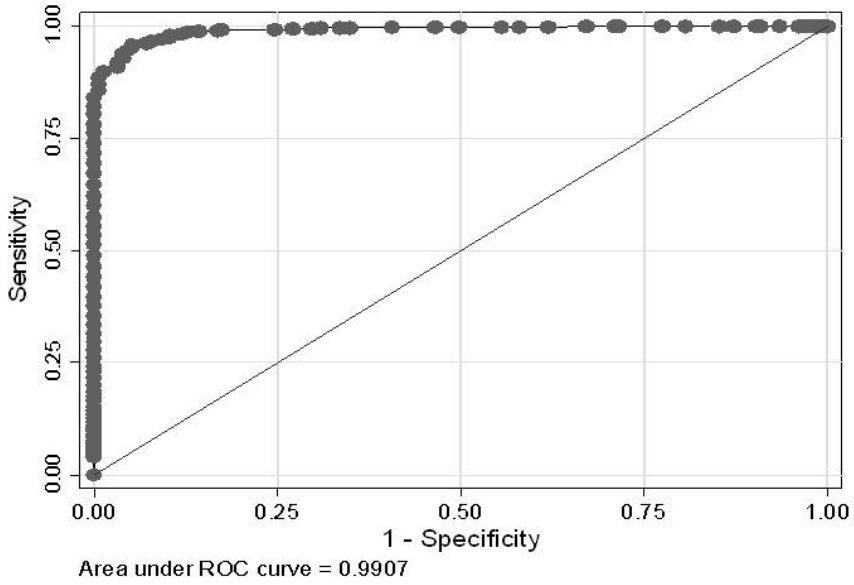
	M1(Ala) M1(Ala)	0.0351
	M1(Ala) M3	0.0328
	M2 M3	0.0266
	M2 M2	0.0253
	M1(Val) M4	0.0113
	M3 M3	0.0105
	M1(Ala) M4	0.0060
	M3 M4	0.0042
	M4 M4	0.0007
PI*MS		
	M1(Val) S	0.0379
	M1(Ala) S	0.0173
	M2 S	0.0110
	M3 S	0.0078
	M4 S	0.0013
PI*MZ		

	M1(Val) Z	0.0133
	M1(Ala) Z	0.0047
	M3 Z	0.0028
	M2 Z	0.0027
	M4 Z	0.0003

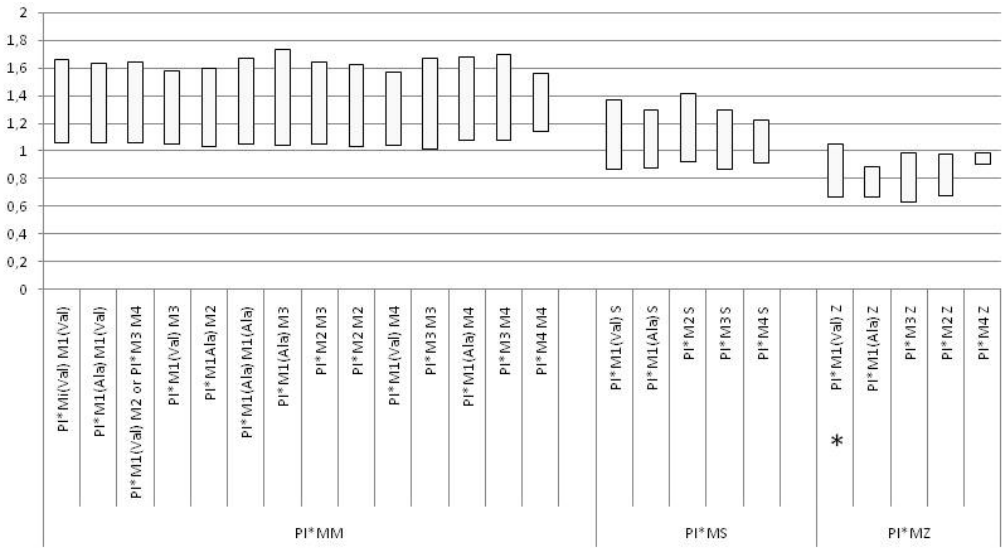
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



* p=0.003; PI*M1(Val) Z vs PI*M1(Ala) Z (0.819 and 0.754 g/L, respectively)