ORIGINAL ARTICLE

Serum levels and genotype distribution of α₁-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 W Russi,6 Nicole M Probst-Hensch,2,3 Maurizio Luisetti1

ABSTRACT

Rationale α₁-Antitrypsin (AAT) deficiency is one of the commonest rare respiratory disorders worldwide. Diagnosis, assessment of risk for developing chronic obstructive pulmonary disease (COPD), and management of replacement therapy require the availability of precise and updated ranges for protein serum levels.

Objective This paper aims to provide ranges of serum AAT according to the main genotype classes in the general population.

Methods The authors correlated mean AAT serum levels with the main SERPINa7 variants (M1A/M1Val (rs6647), M3 (rs1303), M2/M4 (rs709932), S (rs17580) and Z (rs2892947)) in 6057 individuals enrolled in the Swiss Cohort Study on Air Pollution and Lung Diseases in Adults (SAPALDIA) cohort.

Results The following ranges (5th—95th percentile) of AAT were found in the serum (g/litre): 1.050—1.640 for PI*M*M, 0.880—1.369 for PI*M*MS, 0.730—1.060 for PI*S*S, 0.660—0.997 for PI*MZ and 0.490—0.660 for PI*SZ. There was very little overlap in AAT serum levels between genotype classes generally not believed to confer an enhanced health risk (MM and MS) and those associated with an intermediate AAT deficiency and a potentially mildly enhanced health risk (SZ, MZ).

Conclusion This work resulted in three important findings: technologically updated and narrower serum ranges for AAT according to PI genotype; a suggestion for a population-based ‘protective threshold’ of AAT serum level, used in decision-making for replacement therapy; and more precise ranges framing the intermediate AAT deficiency area, a potential target for future primary prevention.

INTRODUCTION

One of the few unambiguously ascertained individual risk factors for chronic obstructive pulmonary disease (COPD) is the serum level of α₁-antitrypsin (AAT), which in turn is strongly determined by the AAT genotype variant system, classically named PI type. A large body of evidence suggests that the degree of risk for COPD is inversely related to the serum AAT level according to the hierarchy PI NullNull > PI ZZ > PI SZ > PI MZ.1–4 Therefore, accuracy in AAT serum level determination is a relevant factor in COPD risk assessment. Above the area termed ‘severe’ AAT deficiency (AATD), bounded by the AAT protective threshold level of 11 μM5 and at high risk for developing COPD, lies the area of ‘intermediate’ AATD, whose threshold has not been determined but is currently used as a proxy for the PI*MZ genotype.

Notably, the currently used standard reference values for AAT in serum6 show a broad and overlapping range of values for the PI MM, PI MZ, PI MS and PI SS classes and do not represent data from the general population. In the absence of such data, only AAT serum values below 11 μM are of use for the assessment of severe AATD and for COPD risk prediction, whereas meaningful reference values to classify intermediate AATD associated with different AATD genotypes are lacking. Careful evaluation of serum AAT concentration is the initial diagnostic test in patients with suspected AATD.7 This measurement can be routinely performed in any clinical chemistry laboratory, and it is the determining factor that justifies further analysis such as genotyping and sequencing, which are performed in dedicated laboratories.8 Thus, the need for updated reference intervals for AAT according to the different PI types is of clinical relevance. This is especially true for reference values related to PI MZ, which may also require clinical attention in the form of smoking counselling in light of evidence for an increased risk of developing airflow obstruction.4,9

Key messages

What is the key question?

What are the ranges of serum α₁-antitrypsin (AAT) level in the general population?

What is the bottom line?

State-of-art methodologies allowed identification of AAT ranges according to the major genotypes, narrower than those previously available. Moreover, the authors defined the intermediate AAT deficiency area (0.92–0.49 g/ litre) of particular interest being a possible target for future interventional options and clarified the longstanding controversy in the conversion from μM to g/litre of the ‘protective threshold’ of AAT serum level, used in decision-making for replacement therapy.

Why read on?

It is important to clearly identify the protective threshold for AAT deficiency and, in turn, the serum level of AAT characterising patients with severe AAT deficiency. It is also important to correctly diagnose patients with intermediate AAT deficiency (mostly with the PI*MZ genotype).
The aim of this paper was to correlate serum AAT levels with the main PI variants, using current standards of measurement and diagnosis, including the molecular characterisation of the SERPINA1 gene encoding AAT. To the best of our knowledge, this information has not been published for a large general population sample.

We took advantage of the first follow-up examination of the (Swiss Cohort Study on Air Pollution and Lung Diseases in Adults) SAPALDIA cohort, which included 8047 people randomly selected from eight population registries representing the three major Swiss language regions, including both urban and rural areas. The SAPALDIA biobank, which includes blood and DNA samples for more than 6000 people, was used to perform a previous study on the SERPINA1 molecular characterisation of 1399 samples displaying reduced serum AAT levels. In this study all 6057 samples from subjects who gave consent for genetic analyses, including the 1399 mentioned above, were investigated for normal M and deficient Z and S SERPINA1 variants. The data presented were used to define the prevalence of those variants and the levels of serum AAT according to the main genotypes in the general population.

MATERIALS AND METHODS

Subjects
The SAPALDIA cohort has been previously described. At the baseline in 1991 the subjects, who were 18–60 years old and predominantly Caucasian of Swiss nationality, were randomly selected from eight population registries. The current cross-sectional investigation of serum AAT is restricted to follow-up data collected in 2002–2003 when the biobank was established and includes 6057 subjects who donated blood and consented to genetic analysis. The study was approved by the Central Ethics Committee of the Swiss Academy of Medical Science and Cantonal Ethics Committees for each of the eight examination areas.

Serum analysis
AAT (g/litre) and C-reactive protein (CRP, mg/litre) concentrations were determined by latex-enhanced immunoturbidimetric assay (COBAS Integra analyzer, Roche Diagnostics, Indianapolis, Indiana, USA), a robust assay with principles that are perfectly comparable to those of nephelometry. The interassay coefficient of variation (CV) was 3.6–4.6%; lower detection thresholds for the AAT and CRP assays were 0.21 g/litre and 1 mg/litre, respectively, and reference values were 0.9–2.0 g/litre and <3 mg/litre, respectively. Each new batch of antiserum was analysed with previous batches for value recovery and proportionality in actual assays. A clarified, delipidated, commercially available serum calibrant (Calibrator f.a.s. Proteins, Roche Diagnostics) was used during the study; the same calibration batch, buffers and other reagents were used throughout the entire study.

Single nucleotide polymorphism analysis
All subjects were typed for give SNPs: S (rs17580), Z (rs28929474), M1Ala/M1Val (rs6647), M5 (rs1305), M2/M4 (rs709932). Typing was performed by PCR with fluorescently labelled TaqMan probes (Vic or Fam labels) on a LightCycler480 (Roche Diagnostics). All single nucleotide polymorphisms (SNPs) were in Hardy–Weinberg equilibrium. Further details on SNP analysis are available in the online data supplement.

Detection of rare deficient variants
The presence of rare deficient mutations was determined by sequencing the coding region of the SERPINA1 gene, as previously described, on selected samples as reported by Zorzetto and coworkers.

Statistical analysis
AAT concentrations were normally distributed and analysis of variance (ANOVA) was applied to compare means in different subgroups. Reference values covered the range from the 5th to the 95th percentile of AAT serum values. Linear and quantile regression was used to calculate adjusted means and percentiles. Covariates in the regression models were selected according to a former publication and they were all significantly associated with AAT concentrations. The receiver operating characteristic (ROC) curve was used to estimate the predictive accuracy of serum AAT, and maximization of the Youden index (ie, the sum of sensitivity and specificity minus 1) defined the optimal threshold for discrimination of genotype classes. Bootstrapping procedures were used to estimate the 95% CIs of the optimal thresholds. Statistical analysis was performed with MedCalc 9.4.2.0 (MedCalc Software, Mariakerke, Belgium), Stata V.10.1 IC and SAS V9.2.

RESULTS
As a first step we identified the number of subjects belonging to different SERPINA1 genotype classes and determined their frequency in the general population (table 1). The PI*MM genotype accounted for 5598 individuals (89.12% of the overall population), whereas PI*MZ was the second genotype in order of frequency (7.48%), followed by PI*MZ (2.36%). Only one subject carrying the PI*ZZ genotype (0.02%) was identified. The two classes defined in table 1 as rare variants and novel variants, accounting for 42 subjects (0.69%), were very heterogeneous groups of variants and were therefore excluded from further analyses on the relationship between AAT serum levels and SERPINA1 genotypes. Nevertheless, means and ranges of AAT serum levels did not notably change if rare and novel variants were not excluded (data not shown). The frequencies of S and Z alleles in the three main Swiss language groups (German, French, Italian) are shown in online table E1 and are further described in the online data supplement.

Table 1 Frequency of the SERPINA1 genotype classes detected in 6057 subjects from the Swiss Cohort Study on Air Pollution and Lung Diseases in Adults (SAPALDIA) cohort

<table>
<thead>
<tr>
<th>Genotype</th>
<th>PI*MM</th>
<th>PI*MS</th>
<th>PI*SS</th>
<th>PI*MZ</th>
<th>PI*SZ</th>
<th>PI*ZZ</th>
<th>Rare variants</th>
<th>Novel variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of subjects</td>
<td>5398</td>
<td>453</td>
<td>10</td>
<td>143</td>
<td>10</td>
<td>34</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Frequency (%)</td>
<td>89.12</td>
<td>7.48</td>
<td>0.16</td>
<td>2.36</td>
<td>0.17</td>
<td>0.56</td>
<td>0.02</td>
<td>0.13</td>
</tr>
</tbody>
</table>

† The PI*MM genotype class encompasses different combinations of normal M variants (M1Ala/Val, M3, M2/M4). Details are reported in the online data supplement.
‡ This class includes subjects heterozygous for rare deficient variants, such as I, Pcomo, Mabisko, Mbolzano, etc. These data were previously analysed in more detail.
§ This class includes novel putative deficient variants detected during SERPINA1 gene sequencing. These data were previously analysed in more detail.
Unadjusted and adjusted means and reference intervals for AAT serum concentration in the six main SERPINA1 genotype classes are presented in table 2. Adjusting for age, sex, study area, alcohol intake (yes/no), systolic blood pressure, body mass index (BMI), smoking status (never, former, current) and C-reactive protein levels.

Since AAT is a acute phase protein, we recalculated the reference intervals of AAT serum concentrations in subgroups of subjects according to systemic inflammatory status. A CRP value of 8 mg/litre, which is the upper normal limit for this protein as suggested by the equipment used in the present study, was used as the cutoff to stratify subjects as being without (<8 mg/litre) or with (≥8 mg/litre) systemic inflammation (table E2). Comparison of AAT means between the two CRP strata revealed higher values in the systemic inflammation stratum for all genotype classes. This difference was statistically significant in the PI*MM subgroup (p<0.001).

As a next step we assessed the accuracy of predicting genotype classes which are not believed to represent a risk for developing emphysema (PI*MM and PI*MS) and those associated with intermediate AATD and arguably a slightly increased risk for developing emphysema (PI*SS and PI*MZ) from AAT concentrations using ROC statistics. For this analysis we included the rare variant carriers of the respective groups to get a representative sample for the general population. The findings were highly accurate for the area under the curve (AUC = 0.9907) (figure E2). The optimal threshold according to the Youden index provided a cutoff at 1.00 g/litre AAT level (95% CI 0.97 to 1.06), which presents a sensitivity of 95.8% and a specificity of 94.8%. For discrimination between PI*MM and any other genotype carrying at least one S or Z allele an optimal cutoff at 1.10 g/litre was determined (75.4% sensibility, 88.5% specificity). The impact of sex, smoking status and CRP levels on these genotype discriminations are described in the online data supplement (table E3).

Finally we analysed the influence of the different PI*M subtypes on AAT serum level. This result is reported in the online data supplement (table E4 and figure E3).

### DISCUSSION

This study ideally represents the most valid setting to date to derive reference values for serum AAT by genotype group in the general population. We applied state-of-the-art technology for the assessment of serum AAT and SERPINA1 genotypes in the Swiss population, which is a combination of three language groups that adequately represent the genetic structure of the European population. To the best of our knowledge, only a few studies have been performed in the general population that measure circulating AAT protein or SERPINA1 gene variants or both. The most comparable study is the Copenhagen City Heart study, a longitudinal survey of 7963 subjects from whom only a small sample of AAT concentrations in blood were measured (n=592). In the study by Sveger, blood from 200 000 infants was drawn for simultaneous AAT determination by semi-quantitative electroimmunoassay and Pi typing with isoelectric focusing. However, the analytical methods used in this study were out of date and therefore these data can no longer be used as a reference. The study by Silverman et al. applied an automated immunoassay to measure AAT in plasma samples from 20 000 blood donors from the St Louis area. Plasma samples that met criteria of <50% of plasma pool reactivity were examined by isoelectric focusing to determine PI type. The reported St Louis Z allele frequency was 0.0116, but no data on the reference values for the concentration of AAT in plasma were extrapolated. In summary, current AAT serum level—genotype relationships seem obsolete as updated diagnostic standards for AATD have never been applied to a general population sample.

Our work generated a number of outputs. First, a more precise allele and genotype frequency was identified for SERPINA1 variants in Switzerland’s general population. We calculated an updated PI*S gene frequency of 0.0401 whereas that of PI*Z was 0.0130. Compared with previous estimates for this population,
PI*S frequency was similar whereas PI*Z frequency was slightly higher. Further discussion of this topic is provided in the online data supplement.

A second major outcome of the study was the analysis of AAT serum concentration and genotypes of the SERPINA1 gene in a large cohort representative of the general population. This enabled the extrapolation of updated protein ranges according to the main normal and deficient genotype classes and evaluation of whether reference values need to be specific for sex, smoking status and CRP level.

To date, no epidemiological data on AAT serum concentration according to the SERPINA1 genotype have been reported for the general population. Some studies identified mean values or reference intervals of AAT serum concentration but they were limited to target groups, mostly COPD cases or populations in which the frequency of AATD is low, such as Japanese, and Korean populations, or in patients with a peculiar clinical phenotype, such as Peyronie’s disease, or characteristic cohorts, such as paediatric subjects and pregnant women. The use of advanced technologies makes this study innovative. These technologies include SNP detection for genotyping and the exclusion of potentially confounding genetic factors, that is, deficient variants other than S and Z alleles, from the final analysis. The presence of other variables that could affect the AAT concentration in serum were considered and adjusted for in an additional analysis. This resulted in much narrower serum AAT ranges than those presented in the ATS/ERS consensus document, with a drastic reduction in the overlap among genotypes (table 3). Moreover, the mean values of AAT concentration according to SERPINA1 genotypical classes (table 2), were lower than those reported so far, likely due to a smaller upper dispersion of measurements. Although 70–80% of the variation in total AAT serum concentration is explained by the Pi type (after age and sex adjustment), other factors can influence variation. Since AAT is an acute-phase reactant, inflammatory status may increase the serum level of AAT. As shown in table E2, when SAPALDIA subjects were stratified according to the presence or absence of an inflammatory condition, most SERPINA1 genotypes showed even narrower ranges, particularly those with CRP <8 mg/litre. However, ranges for individuals with elevated CRP were generally higher. In the real world of routine AAT serum measurement, our data do not justify systematic measurement of the inflammatory status and the stratified ranges can be used only in reference laboratories.

Other factors that impact AAT concentration variability include active smoking and age. An in-depth analysis of the same SAPALDIA population has shown that an inter-relationship among circulating AAT, smoke exposure, gender and systemic inflammatory status exists. However, in this study with narrower diagnostic purposes, we demonstrated that inclusion of sex and current smoking status to predict intermediate deficiency genotypes was not necessary. We also addressed a further putative factor for AAT serum concentration variability, that is, the intrinsic effect of the different PI*M subtypes (figure E3). The effect of this variable seems to be negligible because only two PI*MZ haplotypes displayed significant changes in AAT serum concentration.

In this study, we also analysed the limits of the so-called ‘protective threshold’ and we tried to address the controversy and confusion about the expression of serum AAT concentration. The term ‘protective threshold’ derives from evidence that subjects with AATD and an AAT serum level above the threshold are at reduced risk of developing emphysema. This is not merely a theoretical cutoff because it is considered the decisional cutoff below which subjects with AATD are eligible for AAT replacement therapy. Therefore, it is a very important concept in AATD patient management. An excellent discussion on this topic by Tonelli and Brantly has recently been published. The concept was originally developed by Hutchinson et al and Stockley, based on evidence that subjects displaying the PI SZ phenotype had a reduced risk of developing emphysema compared with those displaying the PI ZZ phenotype and are therefore seldom suitable for replacement therapy. The threshold was fixed at the serum AAT level corresponding to 0.8 g/litre, measured by radial immunodiffusion. A few years later, to resolve the lack of standardisation among laboratories that caused so much confusion in the definition of the AAT measurements, a highly purified AAT standard, expressed as μM, was introduced. In the same report, the protective threshold using the highly purified AAT standard determined by nephelometry was fixed at 11 μM; that is, the 10th percentile of the AAT serum range for subjects with PI SZ, which is considered adequate to protect the lungs from proteolytic attack. Since then, in countries where the AAT concentration was expressed as g/litre, the 0.8 threshold was often considered equivalent to 11 μM. However, radial immunodiffusion cannot be considered equivalent to nephelometry because the former, obsolete method overestimates the real AAT concentration by about

**Figure 1** Suggested areas corresponding to severe α1-antitrypsin (AAT) deficiency (below the protective threshold) and intermediate AAT deficiency (above the protective threshold and below the 10th percentile of the AAT range for subjects carrying the PI*MS genotype). Bars represent 5th/95th percentiles of AAT serum levels.
50%.30 The value corresponding to the 10th percentile of AAT serum concentration for the PI*SZ group in SAPALDIA, which is suitable to derive threshold values in the general population, is 0.49 g/litre.

Intermediate deficiency is a term usually referred to as synonymous with the PI*MZ genotype, which may represent a slightly increased risk of developing COPD.4 We believe that correct diagnosis of subjects carrying the PI*MZ genotype is a critical issue for a number of reasons. First, having been identified as a group at risk of developing COPD, they are subjects particularly suitable for an effective prevention and smoking cessation campaign, as suggested by an increased rate of attempting to quit smoking following genetic testing.31

Intermediate AATD. We believe that these findings will be helpful in the future for the investigation of AATD-related risk for COPD and for a more precise definition of when to implement AAT replacement therapy. Finally, the reported data show that the gene-environmental analysis is critical in the ongoing SAPALDIA longitudinal assessment of the impact of SERPINA1 (alpha-1-antitrypsin) on pulmonary health.

Acknowledgements Current SAPALDIA team: study directorate: T Rochat (p), JM Gasco (p), N Kandú (e/exp), LIS Lu (exp), NM Probst Hensch (e/g), C Schindler (s). Scientific team: JC Barthélémy (c), W Berger (g), R Bettschart (a), B Bircher (a), G Bolognini (p), D Brändli (p), C Brombach (n), M Brutsche (p), L Burdet (p), M Frey (p), U Frey (p), MW Gerbase (p), D Gold (e/c/p), E de Groot (c), W Karrer (p), R Keller (p), B Knöpfli (p), B Martin (pa), D Miedinger (o), U Neu (exp), L Nicod (p), M Pons (p), F Roché (c), T Rotte (p), E Russi (p), P Schmid-Grendelmeier (a), A Schmidt-Trucksäss (pa), A Turk (p), J Schwartz (e), D Stolz (p), P Straehl (exp), JM Tschopp (p), A von Eckardt (cc), E Zemp Stutz (e). Scientific team at coordinating centres: M Adam (e/g), E Boes (p), GD Briveaux (p), D Carballo (c), E Conradi (e), J Curciarj (e), J Dratva (e), A Di Pasquale (s), L Grize (s), D Keidel (s), S Kneimler (pa), A Kumar (g), M Imboden (g), N Maitre (s), A Mehta (e), F Meier (e), H Phuleria (exp), E Schaffner (s), GA Thun (g) A Aineck (exp), M Ragetti (e), M Ritter (exp), T Schikowski (e), G Stem (p/d), M Tarantino (s), M Tasi (e), M Wanner (pa), (a) allergology, (c) cardiology, (cc) clinical chemistry, (e) epidemiology, (exp) exposure, (g) genetic and molecular biology, (m) meteorology, (n) nutrition, (o) occupational health, (p) pneumology, (pa) physical activity, (pd) pediatrics, (s) statistics. The study could not have been done without the help of the study participants, technical and administrative support and the medical teams and field workers at the local study sites. Local fieldworkers: Aarau: S Brun, G Giger, M Spersen, M Stahel, Basel: C Bärli, C Dahler, N Oertli, I Harrehr, F Karrer, G Novic, N Wytenbachen; Davos: A Savar, P Senn, R Winzeler; Geneva: G Bonfils, B Blizchart, C Gandolt, J Rochat; Lugano: S Boccia, E Gehrig, MT Manda, G Solar, B Viscardi; Montana: AP Bieri, C Danioy, M Maire; Payrins: F D’don D’anelis, A Vonnez; Wald: D Bodmer, E Hochstrasser, R Kunz, C Meier, J Räük, U Schafroth, A Walder. Administrative staff: C Gabriel, R Gutknecht.

Contributors Conception and design of the study: ML, NPH, IF, MZ; data acquisition and analysis: IF, GAT, MZ, SO, CS, AV, LR; drafting the manuscript for important intellectual content: ML, IF, NPH, GAT, MI, TR.

Funding Unrestricted grant to IF and G-A T and eALTA 2006 assigned to IF, both provided by Talecris Biotherapeutics Inc.; funds from the Fondazione IRCCS Policlinico San Matteo—Ricerca Corrente (RC345) and from the Fondazione Cariplo. Support to SAPALDIA: The Swiss National Science Foundation (grants 33CSC0.134276/1, 33CSCD-108796/1, 3274BO-104283, 3274BO-104284, 3274BO-104285, 3247F-105698, 3105-00320, 3200-005270, 3200-004532, 4026-020909, 3233-054986, PNP3-123171), the Federal Office for Forest, Environment and Landscape, the Federal Office of Public Health, the Federal Office of Roads and Transport, the canton’s government of Aargau, Basel-Stadt, Basel-Land, Geneva, Luzern, Ticino, Zurich, the Swiss Lung League, the canton’s Lung League of Basel Stadt/Basel Landschaft, Geneva, Ticino and Zurich, SUVA, Freiwillige Akademische Gesellschaft, UBS Wealth Foundation, Talecris Biotherapeutics GmbH, Abbott Diagnostics, European Commission 019896 (GABRIEL), Wellcome Trust WT 084703MA.

Competing interests IF has received educational and consultation fees, a research grant (eALTA Award) and an unrestricted grant, and travel support from Talecris Biotherapeutics GmbH and Kedron SpA. Part of the salary costs of GAT are covered by an unrestricted research grant from Talecris Biotherapeutics GmbH. TR has received fees for consulting once in 2011 by Talecris Biotherapeutics GmbH. SO has received travel support from Grifols International SA and consultation fees from Kedron SpA. ML has received travel costs to ERS and ATS congresses from Talecris Biotherapeutics GmbH. He has performed paid lectures for Kedron SpA, and has obtained research funds from Talecris Biotherapeutics GmbH, as well as funds for staff members. NPH has received an unrestricted research grant from Talecris Biotherapeutics GmbH. The grant money was applied to covering part of the salary costs for GAT. The company was not involved in defining specific aims, conduct of data analysis or data interpretation. All other authors declare no conflict of interest.

Patient consent Obtained.

Ethics approval Swiss Academy of Medical Science and Cantonal Ethics Committees for each of the eight examination areas.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES
Alpha-1-antitrypsin deficiency


ONLINE DATA SUPPLEMENT

Serum levels and genotype distribution of alpha1-antitrypsin in the general population

Ilaria Ferrarotti¹, Gian Andri Thun²,³, Michele Zorzetto¹, Stefania Ottaviani¹, Medea Imboden²,³, Christian Schindler²,³, Arnold von Eckardstein⁴, Lucia Rohrer⁴, Thierry Rochat⁵, Erich Russi⁶, Nicole M Probst-Hensch²,³, Maurizio Luisetti¹

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-analyzed 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\textsuperscript{12}, 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*M5) 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*M5 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\textsuperscript{th}-95\textsuperscript{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\(^3\) according to data based on the analysis of three Swiss cohorts\(^4,5\) 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\(^6\), moved eastbound to the rest of Europe as a consequence of the late-glacial resettlement of Europe\(^7\), 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)\(^8\), M2 (213Val, 376Asp, and 101His)\(^9\), M3 (213Val and 376Asp)\(^10\), M4 (213Val and 101His)\(^11\). 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$^{12}$ 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$^{13}$.

Supplement References


**Legends to Supplemental Figures**

Figure E1 – Intervals (5<sup>th</sup>-95<sup>th</sup> percentiles) for unadjusted serum AAT levels in the main SERPINA1 genotypic classes; 1<sup>st</sup>-99<sup>th</sup> 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<sup>th</sup>-95<sup>th</sup> 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

<table>
<thead>
<tr>
<th></th>
<th>Z allele (%)</th>
<th>S allele (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>General population (n=6,057)</td>
<td>1.30</td>
<td>4.01</td>
</tr>
<tr>
<td>German subgroup (n=3,288)</td>
<td>1.22</td>
<td>3.65</td>
</tr>
<tr>
<td>French subgroup (n=1,938)</td>
<td>1.44</td>
<td>5.26</td>
</tr>
<tr>
<td>Italian subgroup (n=831)</td>
<td>1.26</td>
<td>2.53</td>
</tr>
<tr>
<td>p-value (chi-square-test)</td>
<td>0.61</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Table E2 – Intervals (5th-95th 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.

<table>
<thead>
<tr>
<th></th>
<th>PI*MM</th>
<th>PI*MS</th>
<th>PI*SS</th>
<th>PI*MZ</th>
<th>PI*SZ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CRP&lt;8</td>
<td>CRP≥8</td>
<td>CRP&lt;8</td>
<td>CRP≥8</td>
<td>CRP&lt;8</td>
</tr>
<tr>
<td>N</td>
<td>5,092</td>
<td>274</td>
<td>426</td>
<td>25</td>
<td>9</td>
</tr>
<tr>
<td>5th perc.</td>
<td>0.960</td>
<td>1.230</td>
<td>0.870</td>
<td>1.107</td>
<td></td>
</tr>
<tr>
<td>95th perc.</td>
<td>1.590</td>
<td>2.086</td>
<td>1.320</td>
<td>1.545</td>
<td></td>
</tr>
<tr>
<td>Lowest value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.730</td>
</tr>
<tr>
<td>Highest value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.910</td>
</tr>
</tbody>
</table>
Table E3 - Influence of sex and current smoking status on unadjusted AAT (g/L) reference values

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>median</th>
<th>5&lt;sup&gt;th&lt;/sup&gt;/95&lt;sup&gt;th&lt;/sup&gt; perc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM/MS all</td>
<td>5,848</td>
<td>1.26</td>
<td>1.01/1.63</td>
</tr>
<tr>
<td>MM/MS, male-smoker</td>
<td>791</td>
<td>1.29</td>
<td>1.01/1.59</td>
</tr>
<tr>
<td>MM/MS, female-smoker</td>
<td>663</td>
<td>1.36</td>
<td>1.07/1.74</td>
</tr>
<tr>
<td>MM/MS, male-nonsmoker</td>
<td>2,127</td>
<td>1.19</td>
<td>0.98/1.48</td>
</tr>
<tr>
<td>MM/MS, female-nonsmoker</td>
<td>2,267</td>
<td>1.28</td>
<td>1.03/1.69</td>
</tr>
</tbody>
</table>
nonsmoker

<table>
<thead>
<tr>
<th>SS/MZ, all</th>
<th>155</th>
<th>0.79</th>
<th>0.66/1.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS/MZ, male-smoker</td>
<td>16</td>
<td>0.79</td>
<td>0.69/1.02</td>
</tr>
<tr>
<td>SS/MZ, female-smoker</td>
<td>14</td>
<td>0.88</td>
<td>0.76/1.07</td>
</tr>
<tr>
<td>SS/MZ, male-nonsmoker</td>
<td>55</td>
<td>0.75</td>
<td>0.62/0.91</td>
</tr>
<tr>
<td>SS/MZ, female-nonsmoker</td>
<td>70</td>
<td>0.81</td>
<td>0.66/1.04</td>
</tr>
</tbody>
</table>

Table E4 – Frequencies of the genotypic classes

<table>
<thead>
<tr>
<th>PI Group</th>
<th>PI Genotype</th>
<th>Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI*MM</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M1(Val) M1(Val)</td>
<td>0.2512</td>
</tr>
<tr>
<td></td>
<td>M1(Ala) M1(Val)</td>
<td>0.1850</td>
</tr>
<tr>
<td></td>
<td>M1(Val) M2 or M3 M4</td>
<td>0.1620</td>
</tr>
<tr>
<td></td>
<td>M1(Val) M3</td>
<td>0.0908</td>
</tr>
<tr>
<td></td>
<td>M1(Ala) M2</td>
<td>0.0559</td>
</tr>
<tr>
<td>Combination</td>
<td>Value</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>M1(Ala) M1(Ala)</td>
<td>0.0351</td>
<td></td>
</tr>
<tr>
<td>M1(Ala) M3</td>
<td>0.0328</td>
<td></td>
</tr>
<tr>
<td>M2 M3</td>
<td>0.0266</td>
<td></td>
</tr>
<tr>
<td>M2 M2</td>
<td>0.0253</td>
<td></td>
</tr>
<tr>
<td>M1(Val) M4</td>
<td>0.0113</td>
<td></td>
</tr>
<tr>
<td>M3 M3</td>
<td>0.0105</td>
<td></td>
</tr>
<tr>
<td>M1(Ala) M4</td>
<td>0.0060</td>
<td></td>
</tr>
<tr>
<td>M3 M4</td>
<td>0.0042</td>
<td></td>
</tr>
<tr>
<td>M4 M4</td>
<td>0.0007</td>
<td></td>
</tr>
<tr>
<td>PI*MS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1(Val) S</td>
<td>0.0379</td>
<td></td>
</tr>
<tr>
<td>M1(Ala) S</td>
<td>0.0173</td>
<td></td>
</tr>
<tr>
<td>M2 S</td>
<td>0.0110</td>
<td></td>
</tr>
<tr>
<td>M3 S</td>
<td>0.0078</td>
<td></td>
</tr>
<tr>
<td>M4 S</td>
<td>0.0013</td>
<td></td>
</tr>
<tr>
<td>PI*MZ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1(Val) Z</td>
<td>0.0133</td>
<td></td>
</tr>
<tr>
<td>M1(Ala) Z</td>
<td>0.0047</td>
<td></td>
</tr>
<tr>
<td>M3 Z</td>
<td>0.0028</td>
<td></td>
</tr>
<tr>
<td>M2 Z</td>
<td>0.0027</td>
<td></td>
</tr>
<tr>
<td>M4 Z</td>
<td>0.0003</td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Figure 1

![Box plot showing AAT serum level (µM) by Genotype. The plot displays five different genotypes: PIMM, PIIMS, PI^S, PI^MZ, and PI^SZ. Each genotype has a box plot indicating the distribution of AAT serum levels.]
Supplementary Figure 2

Area under ROC curve = 0.9907
Supplementary Figure 3

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