## SUPPLEMENTARY MATERIAL

## Table of contents

## Supplementary Tables

Supplementary Table 1 .....  2
Supplementary Table 2 .....  3
Supplementary Table 3 .....  5
Supplementary Table 4 .....  7
Supplementary Figures
Supplementary Figure 1. ..... 8
Supplementary Figure 2. .....  9
Supplementary Figure 3. ..... 10

Supplementary Table 1. Common SERPINC1 haplotype in unrelated carriers of c .1332-1336delAAGAG ( $\mathrm{n}=3$ )

| dbsnp | MAF |
| :--- | :--- |
| $r s 2227612$ | 0.135 |
| $r s 677$ | 0.110 |
| rs148783362 | 0.013 |
| rs61827936 | 0.133 |
| rs1799876 | 0.470 |
| rs5878 synonymous | 0.464 |
| rs5877 synonymous | 0.450 |
| rs2227597 | 0.237 |
| rs2227596 | 0.349 |
| $r s 2227595$ | 0.126 |
| $r s 2227594$ | 0.237 |
| $r s 2227593$ | 0.238 |
| $r s 2227590$ | 0.243 |

MAF: Minor allele frequency.

Supplementary Table 2. Summary of SERPINC1 exon 7 deletions

| ID | cDNA <br> -(deleted nucleotides) | Protein (reading frame) ${ }^{\dagger}$ | Amino acid sequence encoded by cDNA (exon 7) $\ddagger$ |
| :---: | :---: | :---: | :---: |
| 1 | $\begin{gathered} \text { c.1237-1239del } \\ \text {-GAA } \end{gathered}$ | p.Glu413del (inframe) | 407 VNEEGSAAASTAVVIAGRSLNPNRVTFKANRPFLVFIREVPLNTIIFMGRVANPCVK* |
| 2 | $\begin{gathered} \text { c.1264del } \\ -A \end{gathered}$ | $\begin{aligned} & \text { p.lle422Leufs*5 } \\ & (+2) \end{aligned}$ | 407 VNEEGSEAAASTAVVLLAVR * TPTG* ${ }^{\text {a }}$ SRPTGLSWFL*EKFL*TLLSSWAE*PTLVLSK |
| 3 | $\begin{gathered} \text { c.1272-1274del } \\ \text {-CCG } \end{gathered}$ | p.Arg425del (inframe) | 407 VNEEGSEAAASTAVVIAGSLNPNRVTFKANRPFLVFIREVPLNTIIFMGRVANPCVK* |
| 4 | $\begin{gathered} \text { c. } 1312 \mathrm{del} \\ -A \end{gathered}$ | $\begin{gathered} \text { p.Arg438Glyfs*6 } \\ (+2) \end{gathered}$ | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANGLSWFL*EKFL* $T L L S S W A E * P T L V L S K$ |
| 5 | $\begin{gathered} \text { c.1319del } \\ -T \end{gathered}$ | p.Phe440Serfs*4 (+2) | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPSWFL*EKFL*TLLSSWAE*PTLVLSK |
| 6 | $\begin{gathered} \text { c. } 1326 \mathrm{del} \\ -T \end{gathered}$ | p.Phe443Leufs*1 (+2) | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVL*EKFL* $T L L S S W A E * P T L V L S K$ |
| 7 | $\begin{gathered} c .1332 \_1333 \mathrm{del} \\ \text {-AA } \end{gathered}$ | $\begin{aligned} & \text { p.Ile444Mfs*19 } \\ & (+1) \end{aligned}$ | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFMRSSSEHYYLHGQSSQPLC* |
| 8 | $\begin{gathered} c .1332 \_1336 \mathrm{del} \\ \text {-AAGAG } \end{gathered}$ | $\begin{aligned} & \text { p.Arg445Serfs*17 } \\ & (+1) \end{aligned}$ | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFISSSEHYYLHGQSSQPLC* |
| 9 | $\begin{gathered} \text { c. } 1347 \mathrm{del} \\ -G \end{gathered}$ | p.Asn450Thrfs*8 (+2) | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFIREVPLTLLSSWAE*PTLVLSK |
| 10 | $\begin{gathered} \text { c. } 1366 \mathrm{del} \\ -\mathrm{G} \end{gathered}$ | p.Gly456Alafs*2 (+2) | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFIREVPLNTIIFMAE*PTLVLSK |
| 11 | c.1373_1384del <br> -TAGCCAACCCTT | p.Val458Glyfs*3 (+2) | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFIREVPLNTIIFMGRGVK* |
| 12 | c.1375-1383del -GCCAACCCT | p.Ala460-Pro462del (inframe) | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFIREVPLNTIIFMGRVCVK* |
| 13 | c.1390-1393del -AAGT | p.Stop465Metfs*13 (+2) | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFIREVPLNTIIFMGRVANPCVKMFLFFAPLPIFGL* |

[^0]Supplementary Table 2. Summary of SERPINC1 exon 7 deletions (cont.)

| ID | $\begin{gathered} \text { MW } \\ \text { (KDa)t } \end{gathered}$ | IP† | N (overall) $\ddagger$ | N (our cohort) | Age of $1^{\text {st }}$ thromb. event (y.o.) ${ }^{\S}$ | AT activity (\%)ๆ | HGMD ${ }^{\text {® }}$ | Ref. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 52.47 | 6.290 | 1 | 0 | 37 | 48 | CD994760 | $\begin{aligned} & \text { Tsuda } \\ & \text { (1999) } \end{aligned}$ |
| 2 | 48.3 | 5.422 | 1 | 0 | - | . | CD982444 | Jochmans (1998) |
| 3 | 52.45 | 5.890 | 2 | 1 | 30 | 55 | CD036071 | $\begin{gathered} \text { Raja } \\ \text { (2003) } \end{gathered}$ |
| 4 | 50.19 | 5.625 | 1 | 0 | - | - | CD117608 | Celinska (2011) |
| 5 | 50.28 | 5.749 | 1 | 1 | 19 | 34 | New | . |
| 6 | 50.22 | 5.749 | 1 | 0 | - | 52 | CD930909 | Chowdhury (1993) |
| 7 | 52.44 | 5.860 | 2 | 1 | - | 50 | CD941603 | Emmerich (1994) |
| 8 | 52.27 | 5.737 | 6 | 5 | 28 (16-45) | 39 (33-51) | CD930910 | Millar (1993) |
| 9 | 51.85 | 5.638 | 1 | 0 | - | - | CD061469 | Schleithoff (2006) |
| 10 | 51.88 | 5.638 | 1 | 0 | - | - | CD024464 | $\begin{gathered} \mathrm{Fu} \\ (2002) \end{gathered}$ |
| 11 | 52.17 | 6.080 | 1 | 1 | - | 75 | New | - |
| 12 | 52.32 | 6.074 | 1 | 0 | - | - | CD961791 | $\begin{gathered} \text { Perry } \\ (1996) \end{gathered}$ |
| 13 | 108.72 | 6.070 | 1 | 0 | - | - | CD125926 | Castaldo (2012) |

[^1]
## Supplementary Table 3. Summary of SERPINC1 exon 7 insertions

| ID | cDNA | Protein (reading frame) ${ }^{\dagger}$ | Amino acid sequence encoded by cDNA (exon 7) $\ddagger$ |
| :---: | :---: | :---: | :---: |
| 14 | c.1233dupC | p.Ser412GInfs*1 (+1) | 407 VNEEGQ*SSCKYRCCDCWPFAKPQQGDFQGQQAFPGFYKRSSSEHYYLHGQSSQPLC* |
| 15 | c.1247dupC | $\begin{aligned} & \text { p.Ser417Lysfs*47 } \\ & (+1) \end{aligned}$ | 407 VNEEG SEAAAKYRCCDCWPFAKPQQGDFQGQQAFPGFYKRSSSEHYYLHGQSSQPLC* |
| 16 | $\begin{gathered} \text { c. } 1255-56 \text { ins } \\ \text { ACCG } \end{gathered}$ | $\begin{gathered} \text { p.Ala419Aspfs*46 } \\ (+1) \end{gathered}$ | 407 VNEEGSEAAASTDRCCDCWPFAKPQQGDFQGQQAFPGFYKRSSSEHYYLHGQSSQPLC* |
| 17 | c.1292dupG | $\begin{aligned} & \text { p.Val432Glyfs*32 } \\ & (+1) \end{aligned}$ | 407 VNEEG SEAAASTAVVIAGRSLNPNRGDFQGQQAFPGFYKRSSSEHYYLHGQSSQPLC* |
| 18 | c.1320-21insA | $\begin{aligned} & \text { p.Phe440Leufs*24 } \\ & (+1) \end{aligned}$ | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPLPGFYKRSSSEHYYLHGQSSQPLC* |
| 19 | c.1357dupA | $\begin{aligned} & \text { p.lle453Asnfs*11 } \\ & (+1) \end{aligned}$ | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFIREVPLNTINLHGQSSQPLC* |
| 20 | c.1366dupG | p.Arg457Glyfs*8 (+1) | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFIREVPLNTIIFMGQSSQPLC* |
| 21 | c.1390dupA | $\begin{gathered} \text { p.Stop465Valfs*18 } \\ (+2) \end{gathered}$ | 407 VNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFIREVPLNTIIFMGRVANPCVKVKCSYSLHLFLFLVCEQK* |

$\dagger$ The effect of the INDEL on the reading frame is indicated (inframe, frameshift +1 or frameshift +2 ).
$\ddagger$ Open reading frame is highlighted in bold, whereas the non-coding sequence is highlighted in italics. The aa sequence in which the reading frame changes is highlighted in red. cDNA: complementary DNA. *: Stop codon.

## Supplementary Table 3. Summary of SERPINC1 exon 7 insertions (cont.)

| ID | $\begin{aligned} & \text { MW } \\ & \text { (KDa) } \end{aligned}$ | IP $\dagger$ | N (overall) $\ddagger$ | N (our cohort) | Age of 1st thromb. event (y.o.) ${ }^{\S}$ | AT activity (\%)ๆ | HGMD ${ }^{\text {® }}$ | Ref. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 14 | 46.99 | 5.414 | 1 | 1 | 50 | 56 | - | . |
| 15 | 52.78 | 5.630 | 1 | 0 | 30 | 47 | CIO41962 | $\begin{gathered} \text { David } \\ \text { (2004) } \end{gathered}$ |
| 16 | 52.79 | 5.453 | 1 | 0 | 45 | 50 | CI941833 | Emmerich (1994) |
| 17 | 52.46 | 5.633 | 1 | 0 | 23 | 37 | CIO41963 | $\begin{gathered} \text { David } \\ \text { (2004) } \end{gathered}$ |
| 18 | 52.51 | 6.001 | 1 | 0 | 30 | - | CI941834 | Gandrille (1991) |
| 19 | 52.45 | 5.807 | 1 | 0 | - | 69 | CI941835 | Chowdhury (1993) |
| 20 | 52.48 | 5.752 | 1 | 0 | 9 | 52 | CI941836 | $\begin{aligned} & \text { Olds } \\ & \text { (1991) } \end{aligned}$ |
| 21 | 57.73 | 5.138 | 1 | 0 | - | 50 | Cl011219 | $\begin{gathered} \text { Lane } \\ \text { (1997) } \end{gathered}$ |

$\dagger$ Molecular weight (MW) and isoelectric point (IP) estimated by Protein isoelectric point calculator software (http://isoelectric.ovh.org/).
$\ddagger$ Number ( N ) of all cases, including the first reported in literature.
§ Age of first thrombotic event. If more than one case, median and interquartile range are shown.
If Antithrombin activity. If more than one case, median and interquartile range are shown.
AT: antithrombin. HGMD ${ }^{\circledR}$ : Human Genetic Mutation Database. MW: molecular weight. N: number of cases. Ref: reference

Supplementary Table 4. Molecular dynamics setup conditions

| Condition | Initial setup |
| :---: | :---: |
| Force field | Amber99sb |
| Number of sol/atoms/ <br> ions/molecules <br> Cut-off <br> Water model <br> Temperature <br> Non-bonded int | 0.9 |
| Ensemble | tip3p |
| Simulation time | 300 |
| Integration step | PME |
| Box type | NPT2 |
| P coupling | $100-680$ ns |
| T coupling | dodecahedron |

## Supplementary Figures



Supplementary Figure 1. Study design, patient flow chart and summary of SERPINC1 molecular defects identified. AT: antithrombin. C-term: C-terminal/C-terminus. CDG: congenital disorders of glycosylation.


Supplementary Figure 2. Native-urea gels of C-terminal variants associated to type II PE antithrombin deficiency due to an increase in latent AT. Western blot for AT in plasma samples of carriers of p.Phe434Leu, p.Arg438Gly, p.Pro439Thr, p.Pro439Thr, and p.Pro461Ser variants after electrophoresis in native conditions with 6 M urea. Samples from a healthy subject and a positive II PE control, corresponding to a type II PE variant that causes increased latent transformation, but is not located in SERPINC1 exon 7 (p.Met251Val), are also shown. Latent AT, that can be better appreciated in the lower panel, with higher exposition, is pointed with arrows. AT: antithrombin. II PE: II pleiotropic deficiency.


```
407 -V--N--E--E--G--S--E--A--A--A--S--T--A--V--V--I--A--G--R--S--L--N--P- 429
\begin{tabular}{llccc} 
& -T & \(-\hat{A}\) \\
\(+G\) & \(-\hat{A}\) & \(+\hat{A}\) & \(-T\) & \(-\hat{A M G A G}\)
\end{tabular}
1281 aacaGGGTGACTTTCAAGGCCaacaGGCCTttcetGGTTTTTATAAGAGAAGttcetCTGaacaCTATT 1356
```



```
        4 5 2
\begin{tabular}{rrrl} 
& -G & -GCCAACCCT & \(+\dot{A}\) \\
\(+\dot{A}\) & +G & -TAGCCAACCCTT & -A AGT
\end{tabular}
1357 ATCTTCATGGGCAGAGTAGCCAACCCTTGTGTTAAGTAAAATGTTCTTATTCTTTGCACCTCEtCCtTA 1417
4 5 3 ~ - I - - F - - M - - G - - R - - V - - A - - N - - P - - C - - V - - K - - ~ * - . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
1418 TTTTTGGTTTGTGaacaGAAGTAAAAATAAATACAAACTACTTCCATCTCA

Supplementary Figure 3. Cluster of INDEL in exon 7 of SERPINC1. Nucleotide and peptide sequences showing the 21 deletions/insertions collected from HGMD® and from our cohort. Deleted and duplicated nucleotides are preceded by the signs " - " and " + ", respectively. The repetitive sequences identified by Emmerich et al. between p.Phe440-Arg445, which we have also identified in the 3'UTR region, are shown in small letter.

HGMD \({ }^{\oplus}\) : Human Gene Mutation Database. 3’UTR: 3’ Untraslated región.```


[^0]:    † The effect of the INDEL on the reading frame is indicated (inframe, frameshift +1 or frameshift +2 ).
    $\ddagger$ Open reading frame is highlighted in bold, whereas the non-codifying sequence is highlighted in italics. The aa sequence in which the reading frame changes is highlighted in red. cDNA: complementary DNA. *: Stop codon.

[^1]:    $\dagger$ Molecular weight (MW) and isoelectric point (IP) estimated by Protein isoelectric point calculator software (http://isoelectric.ovh.org/).
    $\ddagger$ Number ( N ) of all cases, including the first reported in literature.
    $\S$ Age of first thrombotic event. If more than one case, median and interquartile range are shown.
    II Antithrombin activity. If more than one case, median and interquartile range are shown.
    AT: antithrombin. HGMD ${ }^{\circledR}$ : Human Genetic Mutation Database. MW: molecular weight. N: number of cases. Ref: reference.

