



Research Paper

Whole Exome Sequencing Reveals Severe Thrombophilia in Acute Unprovoked Idiopathic Fatal Pulmonary Embolism



Matt Halvorsen^a, Ying Lin^c, Barbara A. Sampson^c, Dawei Wang^c, Bo Zhou^c, Lucy S. Eng^c, Sung Yon Um^c, Orrin Devinsky^b, David B. Goldstein^a, Yingying Tang^{c,*}

^a Institute for Genomic Medicine, Columbia University Medical Center, United States

^b Department of Neurology, NYU Langone Medical Center, United States

^c Molecular Genetics Laboratory, New York City Office of Chief Medical Examiner, United States

ARTICLE INFO

Article history:

Received 22 December 2016

Received in revised form 19 January 2017

Accepted 26 January 2017

Available online 31 January 2017

Keywords:

Idiopathic fatal pulmonary embolism

Whole exome sequencing

Severe thrombophilia

Natural anticoagulant deficiency

ABSTRACT

Background: Acute unprovoked idiopathic fatal pulmonary embolism (IFPE) causes sudden death without an identifiable thrombogenic risk. We aimed to investigate the underlying genomic risks of IFPE through whole exome sequencing (WES).

Methods: We reviewed 14 years of consecutive out-of-hospital fatal pulmonary embolism records ($n = 1478$) from the ethnically diverse population of New York City. We selected 68 qualifying IFPE cases for WES. We compared the WES data of IFPE cases to those of 9332 controls to determine if there is an excess of rare damaging variants in the genome using ethnicity-matched controls in collapsing analyses.

Findings: We found nine of the 68 decedents (13.2%) who died of IFPE had at least one pathogenic or likely pathogenic variant in one of the three anti-coagulant genes: *SERPINC1* (Antithrombin III), *PROC*, and *PROS1*. The odds ratio of developing IFPE as a variant carrier for *SERPINC1* is 144.2 (95% CI, 26.3–779.4; $P = 1.7 \times 10^{-7}$), for *PROC* is 85.6 (95% CI, 13.0–448.9; $P = 2.0 \times 10^{-5}$), and for *PROS1* is 56.4 (95% CI, 5.3–351.1; $P = 0.001$). The average age-at-death of anti-coagulant gene variant carriers is significantly younger than that of non-carriers (28.56 years versus 38.02 years; $P = 0.01$).

Interpretation: This study showed the important role of severe thrombophilia due to natural anti-coagulant deficiency in IFPE. Evaluating severe thrombophilia in out-of-hospital fatal PE beyond IFPE is warranted.

© 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Pulmonary embolism (PE), a complication of deep venous thrombosis (DVT), is a serious and growing public health problem, causing considerable morbidity and mortality worldwide. It is estimated that PE causes >500,000 deaths in the European Union (population 500 million) (Cohen et al., 2007) and a similar number of deaths in the United States (population 300 million) every year (Raskob GE, 2014). The annual incidence of PE and DVT increases with age (>50 years), is higher in people of African descent and lower in Asians, compared to white individuals, and is higher in men than in women when estrogen and pregnancy is not considered (Di Nisio et al., 2016).

Fatal PE events are investigated either in-hospital and emergency rooms, or out-of-hospital in the offices of Chief Medical Examiner. In most parts of the United States, Medical Examiners (ME) are responsible for determining the underlying cause of death due to acute fatal PE

in the out-of-hospital setting, when unattended by a physician. After autopsy and review of clinical history, ME aims to identify one or more thrombogenic risk factors in Virchow's Triad (Wolberg et al., 2015): vascular damages (e.g. surgery or trauma), venous stasis (e.g. immobilization), and hypercoagulability (e.g. cancer). Frequently, ME finds no strong thrombogenic risk to explain a fatal PE in a young decedent, which prompts them to pursue a thrombophilia test. Heritable factors commonly assessed clinically confer mild thrombogenic risks: factor V Leiden (4–10 odds ratio (OR)) and prothrombin G20210A variants (2–4 OR). Mild thrombophilia plays little role in fatal PE as we reported previously (Tang et al., 2011). Severe thrombophilia due to deficiency of natural anticoagulants (antithrombin III, protein S, or protein C), are strong thrombogenic risk factors; however, the status of severe thrombophilia is largely unknown, particularly in the decedents who died of acute fatal PE events.

Acute unprovoked idiopathic fatal pulmonary embolism (IFPE) is pulmonary embolism leading to sudden death without identifiable thrombogenic risks and previous history. In this study, we aimed to identify the underlying genomic risk factors of IFPE through whole exome sequencing (WES). We reviewed 14 years of consecutive

* Corresponding author at: Molecular Genetics Laboratory, Office of Chief Medical Examiner, 421 East 26th Street, New York, NY 10016, United States.
E-mail address: ytang@ocme.nyc.gov (Y. Tang).

out-of-hospital fatal PE records ($n = 1478$ cases) in the New York City Office of Chief Medical Examiner and identified IFPE decedents who were under the age of 50 years without a known thrombogenic risk and negative for the factor V Leiden and prothrombin G20210A variants. The WES testing results are reported here.

2. Methods

2.1. Case Selection

All out-of-hospital fatal PE deaths in the City of New York are investigated in the Office of Chief Medical Examiner (OCME). The detailed forensic investigation procedures include complete autopsy (gross and microscopic examination), toxicology, scene investigation, and review of clinical history (when possible). Postmortem specimens are routinely collected at autopsy and submitted by a ME for the molecular testing of two mild thrombophilia variants, factor V Leiden (FVL) and prothrombin (FII) G20210A, to the College of American Pathologists (CAP) accredited Molecular Genetics Laboratory within OCME, where the cases are accessioned and the specimens are tested. For this WES-based study, we reviewed 1478 consecutive cases with PE as immediate cause of death between 2001 and 2014 and selected 72 that met the selection criteria (see below). Four cases were removed due to insufficient DNA, and a total of 68 cases were selected for WES study at Institute for Genome Medicine (IGM) at Columbia University Medical Center (CUMC).

Inclusion criteria for the IFPE cases are: age under 50 years; BMI ≤ 30 kg/m²; negative FVL and FII G20210A; no thrombogenic risks (i.e. various level of immobility, recent surgeries or trauma, arteriosclerotic and hypertensive cardiovascular disease, cancer et al.). A history of oral contraceptive use was *not* an exclusionary criterion for this study because oral (progestin-only) contraceptives and transdermal estradiol reportedly carry minimal or no thrombotic risk (Trenor et al., 2011), and the thrombotic risk is affected by estrogen dose, type of progestin, mechanism of delivery, and length of therapy, for which we often do not have the full information. Early pregnancy alone without the presence of other thrombogenic risk factors was also not an exclusionary criterion.

This study is *not* regulated by 45 CFR Part 46 because only cadaver specimens were used. OCME approved this study for diagnosis of the underlying cause of PE.

2.2. WES and Data Processing

Postmortem tissue samples (spleen, liver, or heart) preserved in RNAlater® (Qiagen, Valencia, CA) or dry bloodstain card samples were used for DNA extraction and subsequent WES. At least 1.5 μ g of genomic DNA was used for WES. Sequencing was performed at IGM at CUMC on an Illumina Highseq 2500, using a Roche Nimblegen EZCap V3 capture kit. WES data were analyzed using the DNA sequence data alignment and variant calling pipeline utilized at IGM (see Supplemental Methods). The pipeline used in processing the cases matches that of controls, eliminating the possibility of case/control differences being confounded by differential processing.

2.3. Control Selection

We used samples that have been sequenced previously at the IGM as controls. Samples were required to be formally approved for control use, and to pass stringent bioinformatics QC at the IGM. Samples were only included if their self-declared ethnicity matched a group represented in the case cohort (African, Hispanic, White or South Asian ancestry). We additionally required controls to not be listed as having a broad phenotype characterized by cardiovascular or pulmonary disease. After controlling for sufficient coverage and a lack of cryptic relatedness (see details in Supplemental Methods), we were left with 68 cases and

9332 controls that passed QC and could be included in our case/control study. The control set included 1424 African Americans, 1917 Hispanics, 10 South Asians and 5981 Caucasians. A breakdown of the broad phenotypes represented across the control cohort was shown in Table S1, and the statistical summary of sequencing data for both controls and cases was shown in Table S2. We unfortunately lack easily accessible detailed clinical data for controls (including age); however, this should not affect our power to detect a strong genetic signal in cases relative to controls, if one is present.

2.4. Cases Versus Controls - Collapsing Analyses

For all collapsing analyses, we tested for the differential burden of rare nonsynonymous genetic variation within a protein coding *genetic unit* defined via particular criteria: for testing of known pathogenic burden, we defined the *genetic unit* as a set of 12 genes listed by OMIM as being associated with thrombophilia; for testing of general rare nonsynonymous variant burden, we defined the *genetic unit* as a single gene. All collapsing analyses were run under a dominant model, where samples were discretized into those with or without a single qualifying rare variant. For a single genetic unit we compared the portion of cases with a qualifying variant to the proportion of controls. For gene-level collapsing analyses done across the exome we require a P-value below the bonferroni-adjusted threshold to consider the result to be exome-wide significant (see details in Supplemental Methods). We additionally took all male cases and controls and performed gene-based collapsing analyses on X and Y chromosome genes only to identify evidence for any single gene where a haploid damaging nonsynonymous variant conferred PE risk.

For the variants that contributed to the collapsing analysis signal in the anticoagulant gene-set, we classified with regard to their clinical significance. We estimated the minor allele frequency (MAF) of a variant using population databases, e.g. ESP6500 data from NHLBI Grand Opportunity Exome Sequencing Project (ESP) (Exome Variant Server) and Exome Aggregation Consortium (ExAC) (Lek et al., 2016) and determined if they were previously reported by searching the variants in HGMD (Stenson et al., 2012) and ClinVar (Landrum et al., 2016a). If a variant is previously reported with strong evidence (family or function studies), the variant is classified as pathogenic variant. If a variant is not reported, but the same amino acid residue change has been previously reported with supporting evidence from family or function studies, the variant is classified as likely pathogenic variant. If a variant was not reported previously or there was insufficient information to classify it as benign or pathogenic, the variant was classified as a VUS (variant of uncertain significance).

2.5. Statistical Analysis

Calculations of odds ratios and P-values for collapsing analyses were done using a two-sided fisher's exact test (R statistical analysis software version 3.2.3). The significance of the age-at-death in carriers versus non-carriers of the anticoagulant gene variants was evaluated by Wilcoxon signed-rank test. The significant association of *SERPINC1* carriers and females was evaluated by binomial exact test (Graphpad prism 7).

3. Results

3.1. Characteristics of the Cases

We reviewed 14 years (between 2001 and 2014) of consecutive out-of-hospital fatal pulmonary embolism records (1478 cases) in the New York City. Consistent with our previous report (Tang et al., 2011), African Americans (non-Hispanic Blacks) represented a majority of deaths due to fatal PE (58%), followed by non-Hispanic Whites (25%), Hispanics (15%), and Asians (2%). This ethnic breakdown is in drastic contrast to the

general population of NYC (8.5 million people) in which non-Hispanic Whites represent 32% of the population, followed by Hispanics (29%), African Americans (23%), and Asians (15%) (Li W and G, 2016).

Most cases had coexisting major and minor thrombogenic risks. The breakdown of the leading thrombogenic risks in the 1478 fatal PE cases are: 50% had various level of acute or chronic immobility due to trauma and/or surgical procedures (non-pregnancy related), 23% obesity, 7% cancer, 3% hypertensive cardiovascular or cardiomyopathy, 2% abdominal tumors, 2% mild thrombophilia (factor V Leiden and/or Prothrombin G20210A variants), 2% phlebitis, 2% pregnancy status post C-section, and 3% various chronic disorders (See Supplemental Fig. S1).

Seventy-two instances of acute unprovoked idiopathic fatal PE (IFPE) occurred in individuals who were ≤ 50 years old (see Methods for case selection criteria), representing 5% of the total cases. Sixty-eight IFPE cases were tested via WES (four of the 72 IFPE cases were removed due to insufficient DNA). The average age-at-death, mean BMI, and gender ratio of the 68 cases were summarized in Table 1.

3.2. WES and Collapsing Analysis for Cases Versus Controls

We compared exome sequence data for the 68 PE cases against 9332 internal controls (see Methods for control sample selection). We hypothesized that PE cases would have a high burden of damaging coding variants in genes previously implicated in thrombophilia. To test this, we performed a single collapsing analysis on all 12 genes listed in OMIM that are associated with a dominant form of thrombophilia, where the variants were required to be almost entirely absent from the population controls (ExAC and ESP), as well as previously reported as pathogenic or within 2 nucleotides of a previously reported pathogenic variant. This screen showed a very high portion of cases carried a pathogenic variant in one of these genes (8/68 cases vs. 63/9332 controls, OR = 19.58, two-sided fisher's exact test, $P = 3.64 \times 10^{-8}$). We repeated this test in a self-declared African ancestry subset of the case/control population and observed the same strong enrichment statistics (5/52 cases vs. 8/1424 controls, OR = 18.69, two-sided fisher's exact test, $P = 4.6 \times 10^{-5}$).

We further investigated the variants from the eight cases and noted that they affected three anticoagulant genes, *SERPINC1* (Antithrombin III), *PROC* and *PROS1*. We next considered the possibility that there may be a hidden layer of variants in these three genes that are pathogenic but have yet to be classified as such. We performed a single collapsing analysis on those PE cases and controls without thrombophilia pathogenic variants to see if an elevated portion of PE cases have a predicted damaging coding variant in one of these three genes compared to controls. The results showed 1/60 PE cases vs. 57/9269 controls had a predicted damaging coding variant. Although the results did not reach statistical significance (OR = 2.74, $P = 0.31$), the variant in the one qualifying case was in fact reportedly pathogenic upon further investigation (Kim et al., 2014). Taken together, nine out of the 68 fatal PE cases (13.2%) had a rare damaging coding variant in the three anticoagulant genes.

We estimated the odds ratio conferred by the variants in each of the three anticoagulant genes by comparing burden per gene in cases

versus controls. We went through the variants that were proximal to a previously reported pathogenic variant, and kept them only if they truly affected the same amino acid as a previously reported pathogenic variant. Additionally, we required that the variants were of annotation content with the proximal variant (ex: missense near a previously reported pathogenic missense). Doing this left us with a total of 9/68 PE cases versus 14/9332 controls with a qualifying variant. Based on these case vs. control statistics, the odds ratio of developing IFPE as a variant carrier for the *SERPINC1* gene is 144.2 (95% CI, 26.3–779.4; two-sided fisher's exact test $P = 1.7 \times 10^{-7}$), for the *PROC* gene is 85.6 (95% CI, 13.0–448.9; two-sided fisher's exact test $P = 2.0 \times 10^{-5}$), and for the *PROS1* gene is 56.4 (95% CI, 5.3–351.1; two-sided fisher's exact test $P = 0.001$). The gene-level statistics support a stronger risk conferred via damage variants in the *SERPINC1* and *PROC* genes than in the *PROS1* gene (Fig. 1).

A series of careful gene-based collapsing analyses on remaining samples were carried out to determine if there was additional undetected signal in our cohort. We took a set of 40 cases and 970 controls of African American ancestry that cluster well according to principal component axes produced by EIGENSTRAT. All samples utilized were negative for known and potentially deleterious *SERPINC1*, *PROC* and *PROS1* coding variants. We then subjected this cohort to further gene-based collapsing analyses to see if any single genes had exome-wide significant ($P < 2.75 \times 10^{-6}$) differences in rare nonsynonymous variant burden between cases and controls. We did not observe any single genes that exceeded this threshold (Fig. S2). We performed a small set of collapsing analyses on X and Y chromosome genes focused on the male-only subset of this cohort, and failed to identify any additional exome-wide significant signal.

3.3. Pathogenic and Likely Pathogenic Variants in IFPE Cases

Following the WES collapsing analysis of cases versus controls, we focused on the detailed variant analyses of the nine cases that had a pathogenic or likely pathogenic variant in the three anticoagulant genes (*SERPINC1*, *PROC* and *PROS1*) (Table 2).

Four cases had *SERPINC1* gene variants, all of which are absent from ExAC and ESP databases (Table 2). CaseID FBMG06-229 had a previously reported pathogenic variant p.L302P that was found in an individual who had type I anti-thrombin 3 deficiency (AT3D) and a positive family history (Tait et al., 1994). CaseID FB03-094 had a novel variant (p.C279R) - a previously reported pathogenic variant affecting the same amino acid residue (p.C279W) showed type I AT3D (Luxembourg et al., 2011). CaseID FBMG077083 had a novel variant (p.S426W) - a previously reported pathogenic variant affecting the same amino acid position (p.S426L) showed type II (Denver/Milano-2 type) AT3D, deprived of inhibitory activity (Olds et al., 1989; Miyata et al., 2009), and recurrent miscarriage (Neki et al., 2014) in the literature and in ClinVar (Landrum et al., 2016b). Case FBMG06-365 had two variants: a previously reported pathogenic codon deletion variant (p.F155del) in ClinVar (Landrum et al., 2016c) (the importance of the amino acid residue F155 was supported by a thrombosis-associated missense variant p.F155C (Ding et al., 2013)), and a novel variant (p.D232N) of uncertain significance.

Three cases had *PROC* gene variants (Table 2). CasesID FBMG08-325 had a previously reported variant p.R51H that was found in a DVT patient who had a type II protein C deficiency (PCD) (Faioni et al., 2000). This variant is absent from ExAC and ESP databases. The importance of the amino acid residue R51 was further supported by a reported PCD-associated missense variant p.R51C (Reitsma et al., 1995). CaseID FBMG08-338 had a previously reported variant p.V367M that was identified in a patient with PCD (Alhenc-Gelas et al., 2000). This variant was found in 1/4300 European American alleles tested in ESP and 1/121,032 alleles tested in ExAC. The importance of the amino acid residue V367 was supported by another previously reported PCD-associated missense variant (p.V367A) (Witt et al., 1994). CasesID FBMG12-322 had a

Table 1
Characteristics of the 68 idiopathic fatal pe cases tested by WES.

Variables	African Americans	Whites	Hispanics	Asians	Total
Case, n	52	6	9	1	68
Average Age (years old)	36.3	41.3	37.6	23	36.7
Gender(male; female)	29;23	4;2	7;2	0;1	40;28
BMI, mean, kg/m ²	26	29	28	24	27
<i>SERPINC1</i> gene variants carriers, n	3	0	0	1	4
<i>PROC</i> gene variants carriers, n	2	0	1	0	3
<i>PROS1</i> gene variants carriers, n	1	1	0	0	2

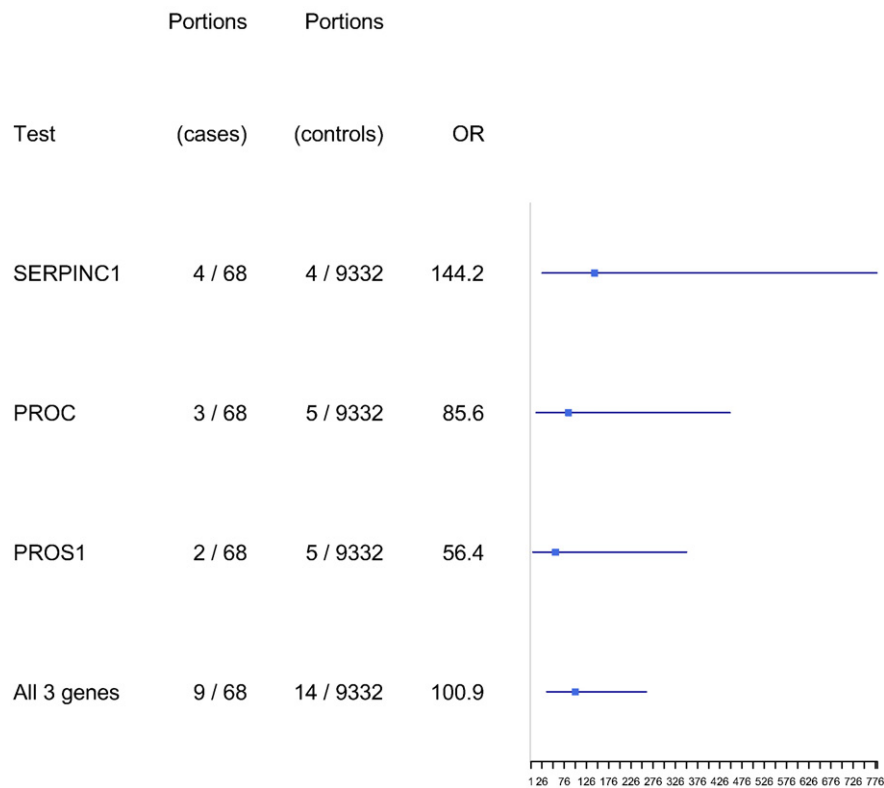


Fig. 1. Odds ratios at the single gene level for pathogenic variant burden between cases and controls. For each gene, a depiction of the 95% confidence interval for the OR has been included. *SERPINC1* appears to confer the highest PE risk amongst the three genes assessed.

novel variant p.R264K that is absent in both ESP and ExAC databases - a previously reported pathogenic variant affecting the same amino acid position (p.R264G) was found in two patients with <50% PC activity (Kim et al., 2014).

Two cases had *PROS1* gene variants (Table 2). CaseID FS03-082 had a previously reported variant p.R90C found in a patient with autosomal dominant protein S deficiency (PSD) (Rezende et al., 2002). This variant is absent in both ExAC and ESP databases. The importance of the amino acid residue R90 was further supported by another reported PSD-associated missense variant p.R90H (Gandrille et al., 1995). CaseID FBMG12-058 had a previously reported variant p.R316C found in DVT patient with PSD (Okada et al., 2006), possibly through an aberrant splicing effect (Xiong et al., 2015). This variant was found in 1/2200 African American alleles tested in ESP and 1/120,422 alleles tested for this variant in ExAC. The importance of the amino acid residue R316 was

further supported by another PSD associated missense variant p.R316H (Kim et al., 2014).

All variants identified in the IFPE cases are heterozygous, which is consistent with the established autosomal-dominant inheritance of natural anticoagulant deficiency.

3.4. Characteristics of IFPE Cases with Anti-coagulant Gene Variants

We compared the average age-at-death for the nine anti-coagulant gene variant carriers and 59 non-carriers, and found that the carriers died at significantly younger ages than that of non-carriers (28.56 years versus 38.02 years; $P = 0.01$, Wilcoxon test, Fig. 2). This difference is not due to the age-at-death for females versus males in the cohort (37.26 years versus 36.32 years). The age-at-death is much younger for the compound heterozygote carrier of the two

Table 2
Pathogenic and likely pathogenic anticoagulant gene variants.

Case ID	Age (years)	Sex	Ethnicity	BMI (kg/m ²)	Brief history	Variant ID	Gene name	Preferred gene transcript (AA change)
FBMG06-229	42	Female	African American	21	Taking Oral contraceptives	1-173878938-A-G	<i>SERPINC1</i>	NM_000488.3(p.L302P)
FB03-094	23	Female	Indian	24	Taking oral contraceptives	1-173879008-A-G	<i>SERPINC1</i>	NM_000488.3(p.C279R)
FBMG07-083	20	Female	African American	30	First trimester pregnancy	1-173873145-G-C	<i>SERPINC1</i>	NM_000488.3(p.S426W)
FBMG06-365	13	Female	African American	25	Family history of DVT	1-173879960-C-T 1-173881096-AAAG-A	<i>SERPINC1</i> <i>SERPINC1</i>	NM_000488.3(p.D232N) NM_000488.3(p.F155del) CODON_CHANGE_PLUS_CODON_DELETION
FBMG08-325	33	Male	African American	29	No significant PMH	2-128178940-G-A	<i>PROC</i>	NM_000312.3(p.R51H)
FBMG08-338	17	Female	African American	24	On a contraceptive patch	2-128186235-G-A (rs140582220)	<i>PROC</i>	NM_000312.3(p.V367M)
FBMG12-322	39	Female	Hispanic	27	No significant PMH	2-128184793-G-A	<i>PROC</i>	NM_000312.3(p.R264K)
FS03-082	29	Male	White	28	No significant PMH	3-93629541-G-A	<i>PROS1</i>	NM_000313.3(p.R90C)
FBMG12-058	41	Male	African American	23	Familial DVT unknown type (father died at 35y, grandma and uncle)	3-93615439-G-A	<i>PROS1</i>	NM_000313.3(p.R316C)

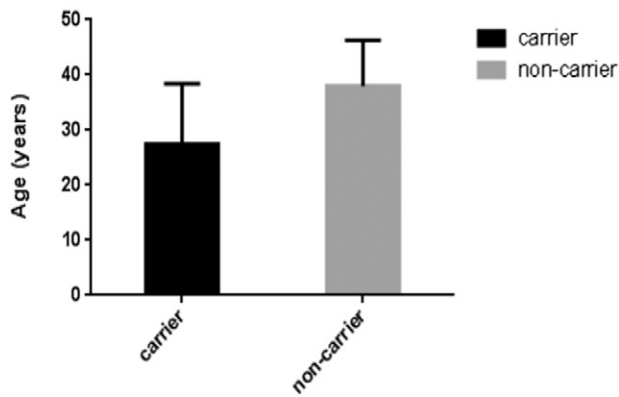


Fig. 2. Age-at-death for anticoagulant gene variant carriers versus non-carriers (a standard deviation has been included).

SERPINC1 gene variants (13 years old, caseID FBMG06-365) compared to those individuals who were heterozygote carriers of a single *SERPINC1* gene variant (all older than 20 years old) (Table 2).

All *SERPINC1* variant carriers are females (Tables 1 and 2), a significant association considering that there were only 28 females amongst the 68 PE cases ($P = 0.029$, binomial exact test). The positive rate of identifying a pathogenic/likely pathogenic anticoagulant gene variant in the three anti-coagulant genes in females is three times that in the males (6/28, 21.43% vs. 3/40, 7.5%). All but two female anti-coagulant gene variant carriers were taking oral contraceptives or were in first-trimester pregnancy. We were unable to perform statistical analysis here because oral contraceptive history is not available for all females of the cohort.

4. Discussion

Our WES study showed that 13.2% of the decedents who died of acute idiopathic fatal PE had a pathogenic or likely pathogenic variant in one of the three anticoagulant genes, highlighting the important role of severe thrombophilia due to natural anticoagulant deficiency in causing fatal PE. Because biochemical-based blood testing for natural anticoagulants is not feasible in the postmortem setting, direct gene testing is the only tool to accurately diagnosis the underlying cause of PE.

We found that the odds ratio for developing IFPE as an anticoagulant gene variant carrier is the highest for *SERPINC1* (144.2), followed by *PROC* (85.6), and *PROS1* (56.4). All three genes encode established anticoagulants in balancing the effects of the procoagulants in clotting and bleeding. The *SERPINC1* gene product, antithrombin-III, is the most important serine protease inhibitor in plasma, regulating the coagulation cascade by inhibiting serine proteases of the intrinsic pathway (e.g. thrombin, factors IXa, Xa and XIa). Protein C is a vitamin K-dependent serine protease that regulates blood coagulation by inactivating factors Va and VIIIa in the presence of calcium ions and phospholipids (Kovacs et al., 2015); it also protects the endothelial cell barrier function. Protein S is a cofactor to activated protein C in the degradation of coagulation factors Va and VIIIa; it inhibits coagulation and stimulates fibrinolysis. It has been estimated that amongst the three anticoagulants, antithrombin III deficiency is the least common (~1/2000–5000 people), followed by protein C deficiency (1/200–500 people), and protein S deficiency (1/500 people) (Lipe and Ornstein, 2011). The reversed relationship between the prevalence of the anticoagulant deficiency and the risk to DVT/PE was reported (Lipe and Ornstein, 2011). Collectively, natural anticoagulant deficiency represents the severe type of thrombophilia and confers high thromboembolic risks as previously reported (Wolberg et al., 2015), (Holzhauer et al., 2012) and showed in this study.

We also noted that 1) the carriers of anticoagulant gene variants are associated with significantly younger age-at-death compared to non-carriers (28.56 years vs. 38.02 years), and 2) a significant association of *SERPINC1* variant carriers being females who were oral contraceptive users. These findings further highlight the importance of testing the severe thrombophilia in the decedents who died of fatal PE. The knowledge gained from testing a deceased family member will guide the physicians for birth control options, such as levonorgestrel-secreting intrauterine device (Mirena) or progestin-only pills, which are not associated with a risk for thrombosis (Trenor et al., 2011). Furthermore, the knowledge of a specific anticoagulant deficiency could be lifesaving in the event of a thrombosis management in at-risk family members. Without concurrent treatment with another anticoagulant, Warfarin alone can worsen the DVT and PE in a carrier of anticoagulant deficiency as it inhibits protein C (Lipe and Ornstein, 2011).

Our ability to identify the underlying causes of IFPE in 13.2% cases using a small case cohort of 68 with well-defined phenotype demonstrated the effectiveness and power of the collapsing analysis in WES-based study. We hypothesize the following reasons for a negative WES testing result in a PE case: 1) A regulatory region (non-coding) variant (Toderici et al., 2016) affecting the expression of an anticoagulant gene or vitamin response elements, or a deletion or gene rearrangement (Caspers et al., 2012) in the anticoagulant gene that is undetectable by the WES testing. Target gene testing that includes the regulatory regions and a testing method for copy number variants (CNV) may improve the positive testing yield. 2) Lack of power for identifying a new gene enriched with PE associated rare variants due to small-sized cohort in case versus control study. WES collapsing analysis for cases versus controls usually requires larger sized cohort. It is quite remarkable that we obtained statistical significance for the anticoagulant genes from only 68 PE cases. New gene identification would benefit from either a large-sized cohort, or family studies with multiple affected individuals. 3) Post-transcriptional or translational modifications of the coagulation cascade beyond genome level changes that alter the balance between coagulation and anticoagulation. 4) Other PE associated factors, such as platelets, leukocytes and microvesicles, endothelial dysfunction, as well as environmental factors beyond those already established and screened in our study.

This is an original study in which WES was used to identify cause of death in acute unprovoked IFPE in the out-of-hospital setting. We identified severe thrombophilia due to natural anticoagulant deficiency in 13.2% of the IFPE cases. We also showed high odds ratio for developing IFPE as an anticoagulant gene variant carrier. Our data support the need of additional study in understanding the role of natural anticoagulant deficiency in a broader fatal PE cases beyond IFPE.

Funding Sources

The funding sources include New York City Office of Chief Medical Examiner and Columbia University Medical Center. The funding sources have no role in study design, data collection, data analysis, interpretation, or writing of the report.

Conflict of Interest

All authors have no conflict of interest.

Author Contributions

DBG and YT designed the study and provided the oversight of the project. MH, YL, DBG and YT participated in the data analysis. MH, YL, OD, BAS, DBG and YT participated in manuscript writing. DW, BZ, LSE, and SYU participated in case review and specimen preparation. All authors reviewed and approved the manuscript.

Acknowledgements

The authors sincerely thank the staff at the Genomic Analysis Facility in IGM at CUMC.

Appendix A. Supplementary Data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.ebiom.2017.01.037>.

References

- Alhenc-gelas, M., Gandrille, S., Aubry, M.L., Aiach, M., 2000. Thirty-three novel mutations in the protein C gene. French INSERM network on molecular abnormalities responsible for protein C and protein S. *Thromb. Haemost.* 83, 86–92.
- Caspers, M., Pavlova, A., Driesen, J., Harbrecht, U., Klamroth, R., Kadar, J., Fischer, R., Kemkes-matthes, B., Oldenburg, J., 2012. Deficiencies of antithrombin, protein C and protein S - practical experience in genetic analysis of a large patient cohort. *Thromb. Haemost.* 108, 247–257.
- Cohen, A.T., Agnelli, G., Anderson, F.A., Arcelus, J.L., Bergqvist, D., Brecht, J.G., Greer, I.A., Heit, J.A., Hutchinson, J.L., Kakkar, A.K., Mottier, D., Oger, E., Samama, M.M., Spannagl, M., 2007. Venous thromboembolism (VTE) in Europe. The number of VTE events and associated morbidity and mortality. *Thromb. Haemost.* 98, 756–764.
- Di Nisio, M., Van ES, N., Buller, H.R., 2016. Deep Vein Thrombosis and Pulmonary Embolism (*Lancet*).
- Ding, Q., Wang, M., Xu, G., Ye, X., Xi, X., Yu, T., Wang, X., Wang, H., 2013. Molecular basis and thrombotic manifestations of antithrombin deficiency in 15 unrelated Chinese patients. *Thromb. Res.* 132, 367–373.
- Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: <http://evs.gs.washington.edu/EVS/>) [December, 2016 accessed].
- Faioni, E.M., Hermida, J., Rovida, E., Razzari, C., Asti, D., Zeinali, S., Mannucci, P.M., 2000. Type II protein C deficiency: identification and molecular modelling of two natural mutants with low anticoagulant and normal amidolytic activity. *Br. J. Haematol.* 108, 265–271.
- Gandrille, S., Borgel, D., Eschwege-gufflet, V., Aillaud, M., Dreyfus, M., Matheron, C., Gaussem, P., Abgrall, J.F., Jude, B., Sie, P., et al., 1995. Identification of 15 different candidate causal point mutations and three polymorphisms in 19 patients with protein S deficiency using a scanning method for the analysis of the protein S active gene. *Blood* 85, 130–138.
- Holzhauser, S., Goldenberg, N.A., Junker, R., Heller, C., Stoll, M., Manner, D., Mesters, R., Krumpel, A., Stach, M., Nowak-gottl, U., 2012. Inherited thrombophilia in children with venous thromboembolism and the familial risk of thromboembolism: an observational study. *Blood* 120, 1510–1515.
- Kim, H.J., Seo, J.Y., Lee, K.O., Bang, S.H., Lee, S.T., Ki, C.S., Kim, J.W., Jung, C.W., Kim, D.K., Kim, S.H., 2014. Distinct frequencies and mutation spectrums of genetic thrombophilia in Korea in comparison with other Asian countries both in patients with thromboembolism and in the general population. *Haematologica* 99, 561–569.
- Kovacs, K.B., Pataki, I., Bardos, H., Fekete, A., Pfliegler, G., Haramura, G., Gindele, R., Komaromi, I., Balla, G., Adany, R., Muszbek, L., Berezky, Z., 2015. Molecular characterization of p.Asp77Gly and the novel p.Ala163Val and p.Ala163Glu mutations causing protein C deficiency. *Thromb. Res.* 135, 718–726.
- Landrum, M.J., Lee, J.M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Hoover, J., Jang, W., Katz, K., Ovetsky, M., Riley, G., Sethi, A., Tully, R., Villamarin-salomon, R., Rubinstein, W., Maglott, D.R., 2016a. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 44, D862–D868. <https://www.ncbi.nlm.nih.gov/clinvar/variation/18010/>.
- Landrum, M.J., Lee, J.M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Hoover, J., Jang, W., Katz, K., Ovetsky, M., Riley, G., Sethi, A., Tully, R., Villamarin-salomon, R., Rubinstein, W., Maglott, D.R., 2016b. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 44:D862–D868. <https://www.ncbi.nlm.nih.gov/clinvar/variation/188092/>.
- Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T., Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J., Pierce-hoffman, E., Berghout, J., Cooper, D.N., DeFlaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M., Poplin, R., Rivas, M.A., Ruano-rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K., Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-luna, M.T., Weisburd, B., Won, H.H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman, C.M., Kathiresan, S., Laakso, M., Mccarroll, S., Mccarthy, M.I., McGovern, D., McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M., Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G., Daly, M.J., MacArthur, D.G., 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291.
- Li, W., H., M., Lee, E., Lasner-frater, L., Castro, A., Kelley, D., Kennedy, J., Maduro, G., Sebek, K., Sun, Y., Van, G. W., 2016. Summary of Vital Statistics, 2014. New York City Department of Health and Mental Hygiene, Office of Vital Statistics, 2016, New York, NY.
- Lipe, B., Ornstein, D.L., 2011. Deficiencies of natural anticoagulants, protein C, protein S, and antithrombin. *Circulation* 124, e365–e368.
- Luxembourg, B., Delev, D., Geisen, C., Spannagl, M., Krause, M., Miesbach, W., Heller, C., Bergmann, F., Schmeink, U., Grossmann, R., Lindhoff-last, E., Seifried, E., Oldenburg, J., Pavlova, A., 2011. Molecular basis of antithrombin deficiency. *Thromb. Haemost.* 105, 635–646.
- Miyata, T., Sato, Y., Ishikawa, J., Okada, H., Takeshita, S., Sakata, T., Kokame, K., Kimura, R., Honda, S., Kawasaki, T., Suehisa, E., Tsuji, H., Madoiwa, S., Sakata, Y., Kojima, T., Murata, M., Ikeda, Y., 2009. Prevalence of genetic mutations in protein S, protein C and antithrombin genes in Japanese patients with deep vein thrombosis. *Thromb. Res.* 124, 14–18.
- Neki, R., Miyata, T., Fujita, T., Kokame, K., Fujita, D., Isaka, S., Ikeda, T., Yoshimatsu, J., 2014. Nonsynonymous mutations in three anticoagulant genes in Japanese patients with adverse pregnancy outcomes. *Thromb. Res.* 133, 914–918.
- Okada, H., Yamazaki, T., Takagi, A., Murate, T., Yamamoto, K., Takamatsu, J., Matsushita, T., Naoe, T., Kunishima, S., Hamaguchi, M., Saito, H., Kojima, T., 2006. In vitro characterization of missense mutations associated with quantitative protein S deficiency. *J. Thromb. Haemost.* 4, 2003–2009.
- Olds, R.J., Lane, D., Caso, R., Tripodi, A., Mannucci, P.M., Thein, S.L., 1989. Antithrombin III Milano 2: a single base substitution in the thrombin binding domain detected with PCR and direct genomic sequencing. *Nucleic Acids Res.* 17, 10511.
- Raskob, G.E., A., P., AN, Blanco, Buller, H., Gallus, A., Hunt, B.J., Hylek, E.M., Kakkar, A., Konstantinides, S.V., Mccumber, M., Ozaki, Y., Wendelboe, A., Weitz, J.I., 2014. Thrombosis: a major contributor to the global disease burden. *J. Thromb. Haemost.* 12, 1580–1590.
- Reitsma, P.H., Bernardi, F., Doig, R.G., Gandrille, S., Greengard, J.S., Ireland, H., Krawczak, M., Lind, B., Long, G.L., Poort, S.R., et al., 1995. Protein C deficiency: a database of mutations, 1995 update. On behalf of the subcommittee on plasma coagulation inhibitors of the scientific and standardization committee of the ISTH. *Thromb. Haemost.* 73, 876–889.
- Rezende, S.M., Lane, D.A., Zoller, B., Mille-baker, B., Laffan, M., Dalhback, B., Simmonds, R.E., 2002. Genetic and phenotypic variability between families with hereditary protein S deficiency. *Thromb. Haemost.* 87, 258–265.
- Stenson, P.D., Ball, E.V., Mort, M., Phillips, A.D., Shaw, K., Cooper, D.N., 2012. The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. *Curr. Protoc. Bioinformatics* (Chapter 1, Unit1 13).
- Tait, R.C., Walker, I.D., Perry, D.J., Islam, S.J., Daly, M.E., McCall, F., Conkie, J.A., Carrell, R.W., 1994. Prevalence of antithrombin deficiency in the healthy population. *Br. J. Haematol.* 87, 106–112.
- Tang, Y., Sampson, B., Pack, S., Shah, K., Yon Um, S., Wang, D., Wang, T., Prinz, M., 2011. Ethnic differences in out-of-hospital fatal pulmonary embolism. *Circulation* 123, 2219–2225.
- Toderici, M., De la morena-barrio, M.E., Padilla, J., Minano, A., Anton, A.I., Iniesta, J.A., Herranz, M.T., Fernandez, N., Vicente, V., Corral, J., 2016. Identification of regulatory mutations in SERPINC1 affecting vitamin D response elements associated with antithrombin deficiency. *PLoS One* 11, e0152159.
- Trenor 3rd, C.C., Chung, R.J., Michelson, A.D., Neufeld, E.J., Gordon, C.M., Laufer, M.R., Emans, S.J., 2011. Hormonal contraception and thrombotic risk: a multidisciplinary approach. *Pediatrics* 127, 347–357.
- Witt, I., Beck, S., Seydewitz, H.H., Tasangil, C., Schenck, W., 1994. A novel homozygous missense mutation (Val 325- > Ala) in the protein C gene causing neonatal purpura fulminans. *Blood Coagul. Fibrinolysis* 5, 651–653.
- Wolberg, A.S., Rosendaal, F.R., Weitz, J.I., Jaffer, I.H., Agnelli, G., Baglin, T., Mackman, N., 2015. Venous thrombosis. *Nat Rev Dis Primers* 1, 15006.
- Xiong, H.Y., Alipanahi, B., Lee, L.J., Bretschneider, H., Merico, D., Yuen, R.K., Hua, Y., Guerousov, S., Najafabadi, H.S., Hughes, T.R., Morris, Q., Barash, Y., Krainer, A.R., Jovic, N., Scherer, S.W., Blencowe, B.J., Frey, B.J., 2015. RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science* 347, 1254806.