U.PORTO

MESTRADO EM ONCOLOGIA

ESPECIALIZAÇÃO EM ONCOLOGIA LABORATORIAL

GWAS-reported SNPs related to venous thromboembolism and their impact on the clinical outcome of ovarian cancer patients

Valéria Tavares

M

2019

U.PORTO

Valéria Tavares

GWAS-reported SNPs related to venous thromboembolism and their impact on the clinical outcome of ovarian cancer patients

GWAS-reported SNPs related to venous thromboembolism and their impact on the clinical outcome of ovarian cancer patients

Valéria Delgado Tavares

M.ICBAS 2019

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

Valéria Delgado Tavares

GWAS-reported SNPs related to venous thromboembolism and their impact on the clinical outcome of ovarian cancer patients

Dissertação de Candidatura ao grau de Mestre em Oncologia - Especialização em Oncologia Laboratorial, submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador - Professor Doutor Rui Manuel de Medeiros Melo Silva

Categoria - Professor Associado com Agregação

Afiliação - Coordenador do Grupo de Oncologia Molecular e Patologia Viral do Centro de Investigação do Instituto Português de Oncologia do Porto

Co-orientador – Doutora Maria Deolinda Paulino Pereira de Sousa Pereira

Categoria - Médica Oncológica

Afiliação - Diretora do serviço de Oncologia médica do Instituto Português de Oncologia do Porto

Acknowledgements

As Clarice Lispector once said, happiness comes to those who recognize the importance of people who have been in their lives. During my master degree in Oncology, especially during my master project, I was lucky enough to have special people in my life, who gave me all the support that I needed to overcome many obstacles that I myself faced. For that, I would like to leave here a special thank you for all those people, although words are not enough to express my gratitude.

First of all, I would like to thank the Oncology Master's Coordinating Committee, currently led by Professor Doctor Carmen Jerónimo, for the opportunity to enter this master's degree program, which I feel gave me the tools to pursue a career in cancer research.

Furthermore, I would like to show my deepest gratitude to Professor Doctor Rui Medeiros, my advisor, for accepting me in the Molecular Oncology & Viral Pathology Group, allowing me to perform my master project. Thank you for guiding me during this journey, for reviewing and correcting my reports and for all the tremendous support and orientation you gave me.

I am also deeply grateful for Doctor Deolinda Pereira, my co-advisor, for the support and guidance, especially regarding the clinical component of this project.

I would like also to thank researcher MSc Ricardo Pinto and PhD student Joana Assis, who have guided me since the beginning, helping me in all the steps of the project. Thank you for everything you taught me, for all the constructive critics, for your support and patience, and for always believing in me and showing availability to help, particularly during the writing of my manuscripts.

Moreover, I would like to leave here a special thank you to each and every member of the Molecular Oncology & Viral Pathology Group, who were always remarkably kind, and helped me go through this period, making my learning process even more pleasant. Thank you for your good suggestions, for your patience and support and for always showing me concern and caring.

Furthermore, I would also like to thank my parents, who always believed and supported me in all ways possible, contributing to my never fading of enthusiasm and

helping me to accomplish success. Thank you, Antónia Tavares and Carlos Tavares, I could not have done it without you.

Just like someone once said, friends are the family that we chose. I would like to express my gratitude to all my friends, particularly to Javidson Miranda, for all the support, patience, kindness and help throughout this journey. You were and you continue to be one of my biggest inspirations, encouraging me always to give more of myself. Thank you for believing in me even when I do not believe in myself.

Last but not least, I am also grateful to all my master colleagues, in particular to Márcia Ferreira and Ana Rita Pinto, who gave me the support when I needed it the most, especially during this project, encouraging me to always keep on going.

Thank you to all of you for helping me during this important journey of my life and for giving me the support I needed to achieve success!

Abstract

***Introduction**:* The pathogenesis of venous thromboembolism (VTE) involves both acquired and inherited factors. Over the years, several susceptibility single-nucleotide polymorphisms (SNPs) in candidate genes have been identified. More recently, genome-wide association studies (GWAS) have reported novel VTE-associated SNPs. Conversely, VTE and cancer have a two-way association, with many haemostatic components found to be implicated in processes that sustain cancer growth and progression. Therefore, VTE-associated SNPs constitute potential cancer-related biomarkers currently needed, in particular, for ovarian cancer (OC) patients. The OC represents the most lethal gynaecological neoplasia, and thus, additional therapeutic management strategies are required. Conversely, among solid tumours, OC is one of the most frequently associated with VTE, and haemostatic components might promote OC progression. Hence, the clinical implications of VTE-associated SNPs should be assessed in OC patients, as they could provide directions for personalized cancer treatment to achieve a better clinical outcome.

With this knowledge in mind, this study aimed to analyse the implications of VTE genetic markers reported by GWAS in the clinical outcome of OC patients.

***Methods**:* The VTE GWAS-reported variants to be evaluated in this study were selected based on specific criteria. A retrospective hospital-based cohort study was conducted with 336 epithelial ovarian cancer (EOC) patients admitted for first-line treatment. The genotyping was performed using the TaqMan® Allelic Discrimination methodology. Overall survival (OS) and disease-free survival (DFS) were the two clinical outcomes evaluated in this study. The risk of recurrence and death were assessed through the Cox regression method. The predictive ability of different prognostic factors was performed using the concordance (c) index. For all statistical tests, a 5% level of significance was considered.

***Results**:* we selected the impact of the genetic variants *Zinc finger protein, FOG family member 2* (*ZFPM2*) rs4734879, *Solute Carrier Family 19 Member 2* (*SLC19A2*) rs2038024, *Contactin 6* (*CNTN6*) rs6764623, *OTU Deubiquitinase 7A* (*OTUD7A*) rs7164569, *Coagulation factor 11* (*F11*) rs4253417 and *Protein C receptor* (*PROCR*) rs10747514 in the clinical outcome of EOC patients. Patients carrying the *ZFPM2* rs4734879 G allele presented a significantly higher 5-year and 10-year OS and a higher DFS compared to AA

Abstract

genotype patients in the FIGO I/II stage subgroup (*P*=0.009, *P*=0.001 and *P*=0.003 respectively). Likewise, patients with rs4734879 AA genotype had a threefold increase in the risk of recurrence and a sixfold increase in the 10-year risk of death compared to patients with G allele (*P*=0.027 and *P*=0.004, respectively). Regarding *SLC19A2* rs2038024 polymorphism, patients carrying the CC genotype presented a significantly lower 5-year and 10-year OS and a lower DFS compared to A allele patients in the FIGO I/II stage subgroup (*P*<0.001, *P*=0.004 and *P*=0.005 respectively). Likewise, patients with rs2038024 CC genotype had a 14-fold increase in the risk of recurrence and a ninefold increase in the 10-year risk of death compared to patients with A allele (*P*=0.001 and *P*=0.005, respectively). As for *CNTN6* rs6764623 polymorphism, patients with the CC genotype presented a significantly lower 5-year OS compared to A allele patients in the FIGO I/II stage subgroup (*P*=0.015). Likewise, patients with CC genotype had a fivefold increase in the risk of recurrence and a ninefold increase in the 5-year risk of death compared to patients with A allele genotypes (*P*=0.013 and *P*=0.010, respectively). Regarding *OTUD7A* rs7164569 polymorphism, patients with the GG genotype presented a prolonged DFS compared to A allele patients (*P*=0.025). As for *F11* rs4253417 and *PROCR* rs10747514, no impact on EOC patients' survival was observed. In terms of predictive ability, the predictive model including the genetic information concerning *SLC19A2* rs2038024 and *ZFPM2* rs4734879 polymorphisms, age, histologic subtype, surgical resection and hormonal status had the highest predictive ability (c=0.768).

Conclusions: Given the two-way association between VTE and OC cancer, VTE-associated SNPs identified by GWAS might be potential prognostic and/or predictive factors currently need for better therapeutic management of OC patients. In our study, rs4734879, rs2038024, rs6764623 and rs7164569 polymorphisms exert a significant impact on the clinical outcome of EOC patients. However, future studies are required to validate these results and to uncover the biological mechanisms underlying our results.

Keywords: Venous thromboembolism, GWAS, SNPs, ovarian cancer, clinical outcome

Resumo

Introdução: Ao longo dos anos, vários *single-nucleotide polymorphisms* (SNPs) de risco para o tromboembolismo venoso (TEV) foram identificados em genes candidatos. Mais recentemente, os *genome-wide association studies* (GWAS) têm contribuído para a identificação de novas variantes genéticas de risco. Visto que TEV e cancro têm uma relação bilateral, estes SNPs constituem potenciais biomarcadores preditivos e/ou de prognóstico atualmente necessários, principalmente para doentes com cancro de ovário (CO). Esta neoplasia está associada à uma taxa de sobrevivência relativa a cinco anos inferior a 40%, e é frequentemente associado ao TEV, sendo que os componentes hemostáticos parecem influenciar a agressividade tumoral. Portanto, é necessário avaliar o impacto clínico destes SNPs em doentes com CO, sobretudo porque podem ser a chave para um tratamento oncológico mais personalizado e uma melhor evolução clínica destas doentes.

Com as evidências atuais em mente, o presente estudo teve como objetivo estudar o impacto de marcadores genéticos de TEV reportados por GWAS na evolução clínica de doentes.

Métodos: Para selecionar as variantes genéticas a serem estudadas, critérios específicos foram aplicados. Foi realizado um estudo retrospetivo do tipo coorte de base hospitalar com 336 doentes com cancro epitelial do ovário (CEO) admitidas para tratamento de primeira linha. A genotipagem das variantes genéticas foi realizada utilizando a metodologia de descriminação alélica TaqMan®. A sobrevivência global (SG) e a sobrevivência livre de doença (SLD) foram as duas medidas de evolução clínica avaliadas neste estudo. O risco de recorrência e o risco de morte foram também avaliados usando o método de regressão de Cox. A análise da capacidade preditiva de diferentes fatores de prognóstico foi realizada atráves do índice de concordância (índice C). Para todos os testes estatísticos, foi considerado um nível de significância de 5%.

Resultados: Foi estudado o impacto das variantes *FOG family member 2* (*ZFPM2*) rs4734879, *Solute Carrier Family 19 Member 2* (*SLC19A2*) rs2038024, *Contactin 6* (*CNTN6*) rs6764623, *OTU Deubiquitinase 7A* (*OTUD7A*) rs7164569, *Coagulation factor 11* (*F11*) rs4253417 e *Protein C receptor* (*PROCR*) rs10747514 na evolução clínica das

pacientes. As doentes com o alelo G do *ZFPM2* rs4734879 apresentaram um SG aos 5 e 10 anos e uma SLD significativamente superiores em comparação as doentes com genótipo AA no subgrupo FIGO I/II (*P* = 0,009, *P* = 0,001 e *P* = 0,003, respetivamente). Da mesma forma, doentes com genótipo AA tiveram um aumento de três vezes no risco de recorrência e um aumento de seis vezes no risco de morte aos 10 anos em comparação com doentes com alelo G (*P* = 0,027 e *P* = 0,004, respetivamente). Em relação a variante *SLC19A2* rs2038024, as doentes com o genótipo CC apresentaram uma SG aos 5 e 10 anos e uma SLD significativamente inferiores em comparação as doentes com alelo A no subgrupo FIGO I / II (*P* <0,001, *P* = 0,004 e *P* = 0,005, respetivamente). Da mesma forma, as doentes com genótipo CC tinham 14 vezes o risco de recorrência e nove vezes o risco de morte aos 10 anos comparativamente as doentes com alelo A (*P* = 0,001 e *P* = 0,005, respetivamente). Quanto ao polimorfismo *CNTN6* rs6764623, as doentes com o genótipo CC apresentaram uma SG aos 5 anos significativamente menor comparativamente as doentes com alelo A no subgrupo FIGO I / II (*P* = 0,015). As doentes com genótipo CC tiveram um aumento de cinco vezes no risco de recorrência e um aumento de nove vezes no risco de morte aos 5 anos em comparação as doentes com alelo A (*P* = 0,013 e *P* = 0,010, respetivamente). Em relação ao polimorfismo *OTUD7A* rs7164569, as doentes com o genótipo GG apresentaram uma SLD mais prolongada em comparação as doentes com alelo A (*P* = 0,025). Quanto aos polimorfismos *F11* rs4253417 e *PROCR* rs10747514, não foi observado um impacto significativo na sobrevida das doentes com CEO. O modelo incluindo idade, subtipo histológico, extensão cirúrgica, status hormonal e os polimorfismos *SLC19A2* rs2038024 e *ZFPM2* rs4734879 apresentou a maior capacidade preditiva (c = 0,768).

***Conclusões*:** Os SNPs associados à suscetibilidade ao TEV, identificados por GWAS, podem ser potenciais fatores preditivos e/ou prognósticos atualmente necessários para um melhor tratamento das doentes com CO. Neste estudo, verificou-se um impacto significativo dos polimorfismos rs4734879, rs2038024, rs6764623 e rs7164569 na evolução clínica das doentes com CEO. No entanto, estudos adicionais são necessários.

***Palavras-chave*:** Tromboembolismo venoso, GWAS, *single-nucleotide polymorphisms*, cancro de ovário, evolução clínica

Index

Index

Figure index

Figure index

Table index

Abbreviations

A

A	Adenine
ABO	ABO, alpha 1-3-N-acetylgalactosaminyltransferase and alpha 1-3-galactosyltransferase
aHR	Adjusted hazard ratio
AKt	Protein kinase B
APC	Activated Protein C
AT	Antithrombin

B

B3GAT2	Beta-1,3-glucuronyltransferase 2

C

C	Cytosine
CCDC181	Coiled-coil domain containing 181
CD/CV	Common disease/common variant
CI	Confidence interval
CNTN6	Contactin 6
COX7A2L	Cytochrome c oxidase subunit 7A2 like
CP	Cancer Pro-coagulant
CYP4V2	Cytochrome P450 family 4 subfamily V member 2
C4BPA	Complement component 4 binding protein alpha
C4BPB	Complement component 4 binding protein beta

D

DD	D-Dimer
DFS	Disease free survival
DNA	Deoxyribonucleic Acid
DVT	Deep Vein Thrombosis

E

EDTA	Ethylenediamine tetraacetic acid
EGFR	Epidermal Growth Factor Receptor
EOC	Epithelial ovarian cancer
EPCR	Endothelial Protein C Receptor
EPHA3	EPH receptor A3

F

FGA	Fibrinogen alpha chain
FGB	Fibrinogen beta chain
FGG	Fibrinogen gamma chain
FIGO	International Federation of Gynaecology and Obstetrics
FUNDC2	FUN14 domain containing 2
FVa	Coagulation factor V activated
FVII	Coagulation factor VII
FVIII	Coagulation factor VIII
FVIIIa	Coagulation factor VIII activated
FVL	Factor V Leiden
FX	Coagulation factor X
FXa	Coagulation factor X activated
FXI	Coagulation factor XI
F2	Coagulation factor II
F5	Coagulation factor V
F8	Coagulation factor VIII
F11	Coagulation factor XI

G

G	Guanine
GP6	Glycoprotein VI platelet
GWAS	Genome-wide association studies

H

HR	Hazard ratio

L

LD	Linkage disequilibrium
LEMD3	LEM domain containing 3
LY86	Lymphocyte antigen 86

M

MAF	Minor allele frequency
MMP	Metalloproteinase
MPs	Microparticles
mTOR	Mammalian target of rapamycin

N

NME7	NME/NM23 family member 7
NK-kB	Nuclear factor kappa B

O

OC	Ovarian cancer
OR	Odds ratio
OS	Overall survival
OTUD7A	OTU Deubiquitinase 7A

P

PAI-1	Plasminogen activator inhibitor-1
PAR	Protease-activated receptors
PC	Protein C
PCR	Polymerase Chain Reaction
PE	Pulmonary embolism
PI3K	Phosphatidylinositol 3-kinase
PROC	Protein C
PROCR	Protein C receptor
PROS1	Protein S
PS	Protein S
PTEN	Phosphatase and tensin homolog

R

RIMS1	Regulating synaptic membrane exocytosis 1
RNA	Ribonucleic acid
RR	Relative risk

S

sEPCR	Soluble Endothelial Protein C Receptor
SERPINC1	Serpin family C member 1
SLC19A2	Solute carrier family 19 member 2
SLC44A2	Solute carrier family 44 member 2
SMAP1	Small ArfGAP 1
SNP	Single-nucleotide polymorphism
SUSD1	Sushi domain containing 1
SV2C	Synaptic vesicle glycoprotein 2C

T

T	Thymine
TF	Tissue factor
THTR1	Thiamine transporter 1
THBD	Thrombomodulin
TIMP	Tissue inhibitor of metalloproteinases
TKTL1	Transketolase-like-1
TMEM170B	Transmembrane protein 170B
TP53	Tumour protein p53
TSPAN15	Tetraspanin 15

V

VEGF	Vascular Endothelial Growth Factor
VTE	Venous thromboembolism
vWF	von Willebrand factor

Z

ZFPM2	Zinc finger protein, FOG family member 2

1. Introduction

1.1. Genome-wide association studies: uncovering the genetic architecture of human diseases

Common diseases have a unique underlying genetic architecture that characterizes the distribution of effect sizes for disease-causal variants. Over the last few decades, in an attempt to uncover the genetic contribution to disease development, the genome research associated with human diseases has been completely revolutionised, which had led to a better understanding of human genetic variability [1, 2].

The human DNA sequence is characterized by a remarkable amount of variability, which is best exemplified by genetic polymorphisms, its most common form [3, 4]. This type of genetic variation is distinguished from mutations by a requirement for the minor allele frequency (MAF) to be, at least, one per cent in a particular population [5, 6]. The simplest form of polymorphism, known as single nucleotide polymorphism (SNPs), corresponds to a single DNA base substitution estimated to occur every 100 to 300 base-pairs across the human genome, accounting for almost 90% of human DNA sequence variations [7, 8]. Apart from the genetic variability, SNPs constitute also the main source of human phenotypic variation [9]. Spread all across the genome, these genetic variants may fall within coding and non-coding regions of the genome or even intergenic regions, with their functionality depending on their location [10]. Although most of them are likely to be functionally neutral (i.e. silent), to some SNPs is attributed biological effects in terms of gene expression and protein function or structure modification [11]. Functional SNPs are thought to underlie differences in human susceptibility to a wide range of disorders, particularly common and multifactorial diseases [11].

According to the common disease/common variant (CD/CV) hypothesis, common diseases are most likely influenced by genetic variants that are also common within a given population [12]. This hypothesis implies that the effect size of common genetic variants, such as SNPs, must be small (low penetrance) while comparing to mutations associated with monogenic diseases (Figure 1) [12, 13]. Thus, given the low penetrance of common genetic variants, which contrasts with the substantial heritability of common pathologies, disease susceptibility is probably influenced by several genetic factors spread across the genome, which interact with environmental factors to produce phenotypic alterations [13].

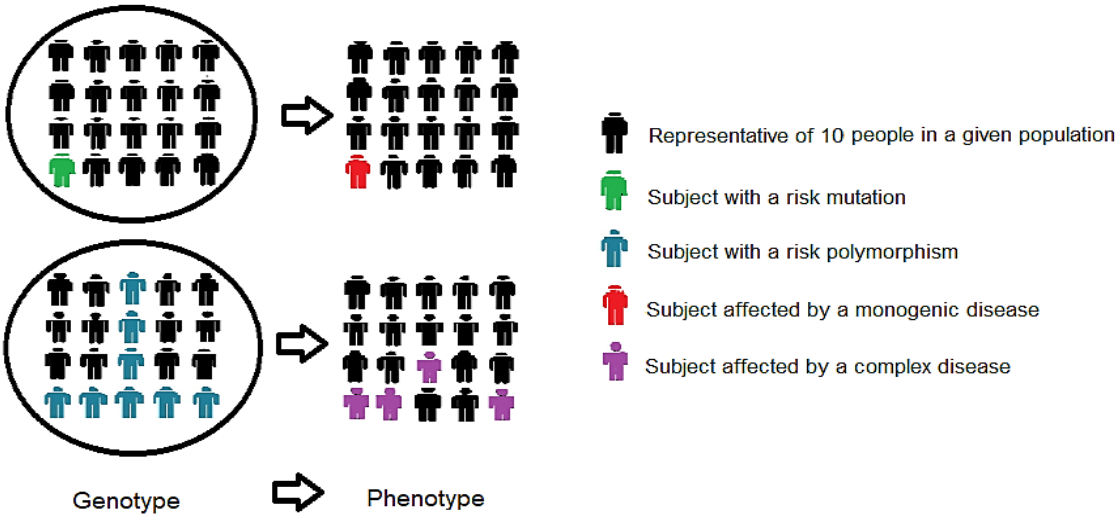

Figure 1 - Differences between mutations and polymorphisms regarding their frequency and penetrance

Given the impact on human health, SNPs constitute an attractive field to assess diseases susceptibility and prognosis based on individual's SNP genotype [14]. However, genotype all SNPs in the human genome to search for disease-causal variants would be extremely expensive [15]. Nonetheless, relying on the principle of linkage disequilibrium (LD), it is possible to scan the entire genome in a cost-effective manner [15]. LD defines the degree to which an allele of one SNP is correlated with an allele of another SNP within a given population [13]. This property is often measured in terms of r^2, which indicates the ratio of observations in which two alleles of different SNPs occur concurrently. Therefore, high r^2 values indicate that the existence of a certain allele of one SNP is strongly predictive of the presence of the allele of the other SNP, meaning that the SNPs are in strong LD [16]. Thus, this property is specific of each population, given that a substantial number of recombination events may disrupt haplotype blocks containing linked alleles of different SNPs until they are inherited independently (linkage equilibrium) [13, 17].

The improvements in our knowledge of human genetic variability, greatly driven by the achievement of the Human Genome and HapMap projects, which have allowed to catalogue all common SNPs and the extent of their LD patterns, respectively, as well as the availability of improved methods for genome research, such as DNA chips for high-throughput genotyping, have enabled the gradual shifting from linkage and candidate-gene association studies to an era of genome-wide associations studies (GWAS) [3, 16, 18, 19]. This genetic tool conglomerate epidemiological study designs and genetic research

methods to test in thousands of samples, simultaneously and in a cost-effective manner, the association of hundreds of thousands of SNPs with a particular phenotype [19, 20]. In opposition to candidate gene studies that test a small number of genetic variants based on preselected loci, GWAS investigate the entire genome, relying on the principle of LD, to identify common SNPs underlying susceptibility to common and complex diseases [12].

Although the cost-efficient assessment of the common genetic variants is possible, given that certain SNPs, known as tag SNPs, can be representative of a set of genetic variants within a haplotype block, in a GWAS it is not possible to assume that the identified SNPs are the actual disease-causal variants (Figure 2) [21, 22]. According to the LD principle, there are two possibilities for GWAS results: (1) the SNP statistically associated with the disorder is, in fact, the functional SNP and thus, modulates disease risk (Figure 2A), or (2) the identified SNP might just be in strong LD with the functional one, requiring additional studies to locate the functional SNP (Figure 2B) [13, 16, 18].

Since the accomplishment of the first GWAS in 2005, some criticism has been made in light of the inability to identify several of the once recognised disease-associated variants and the perceived ‘missing heritability, leading the CD/CV hypothesis to be questioned [23-25]. The heritability proportion explained by a set of genetic variants is defined as the ratio of the heritability due to these variants estimated through their observed effects, to the total heritability of the trait inferred indirectly from the population data [26]. data [26]. Despite identifying several SNPs underlying susceptibility to a variety of common and presumed heritable diseases, GWAS appear to explain only a minority of disease heritability [26-28]. One possible explanation for the missing heritability is the existence of disease-causal variants poorly tagged by the individual GWAS markers (Figure 2C), and/or the existence of several independent variants in the same locus tagged by the most-significant GWAS marker (Figure 2D) [27, 29, 30]. Apart from the challenges while choosing proper tag-SNPs, four main hypotheses have been addressed as possible explanations for the missing heritability: (1) rare genetic variants with a frequency of less than one percent (i.e., mutations) might have a substantial impact on disease risk, (2) existence of common variants with minor effect sizes that are not captured by the current genetic tools, (3) by not excluding shared environmental effects, the heritability obtained in previous family studies could be overestimated, or even (4) the missing heritability could be due to the fact that GWAS are currently blind to non-genetic (epigenetic) information that is also inherited and greatly contributes to human phenotype variation [28, 31].

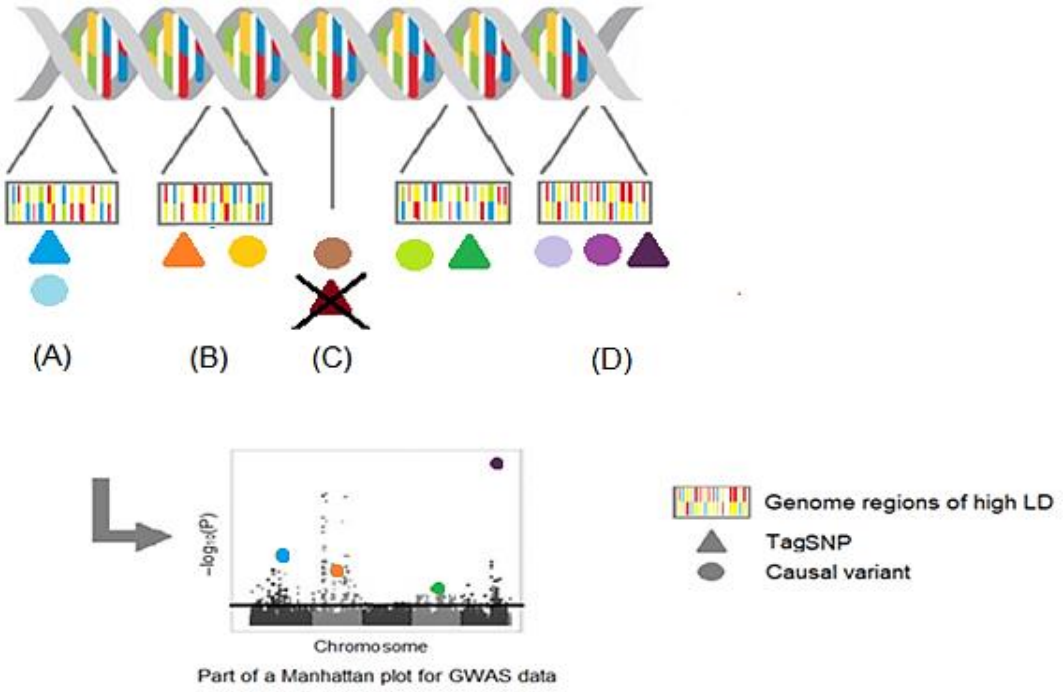

Figure 2 – Assessment of disease causal-variants using tag SNPs and its challenges. (A): the tag SNP, which was found to be statistically associated with the disorder, is the causal variant but it is in strong LD with the causal one. (B): the tag SNP, which was found to be statistically associated with the disorder, is not the causal variant. (C): The causal variant is not tagged by any individual GWAS marker (tag SNP), and therefore, the GWAS cannot detect this causal variant. (D): existence of independent causal variants in the same haplotype block tagged by the most-significant tag SNP.

Apart from the missing heritability, one main challenge in GWAS is attributed to data analysis, namely in the discrimination between the true associations from the false-positive ones [20, 32]. Thereby, researchers make use of several post-GWAS studies, including replication and validation studies, fine mapping and functional analysis, which can confirm or refute the findings. Furthermore, these studies may provide additional information regarding the functional SNP and target gene, as well as the molecular mechanism by which the genetic variant may introduce modifications at a phenotypic level, consequently narrowing down the knowledge gap existent between DNA sequence and functional consequence [33-36]. For instance, validation studies are an important complement to the GWAS' discovery phase, as it allows to validate or refute the results, and also access the accurate effect size of the identified variants, which, due to the experimental design of GWAS, are often overestimated in the discovery phase. Unlike replication studies, validation studies are performed using an independent sample set (i.e., a sample drawn

from a different population than the one used in the GWAS, known as confirmation sample), or using different methods [36, 37].

Despite facing numerous technical challenges, GWAS constitute a powerful tool to unravel the genetic basis of human diseases, and thus, might have potential applicability in the clinical practice both in disease prevention, treatment and prognosis setting [13]. On the prevention scope, although still debatable, susceptibility SNPs could be used to personalize the current prophylaxis measures for a wide sort of diseases, based on individual risk [38, 39]. Regarding treatment, in the era of personalized medicine, GWAS can be useful by enlightening the involvement of specific biological pathways in disease pathogenesis, and thus, revealing new therapy targets [16]. Furthermore, given that disease pathogenesis and drug metabolism are both influenced by individual unique genetic signature, GWAS findings might give insights to predict therapy response and define more accurate biomarkers [40, 41].

Over the years, GWAS have confirmed and revealed many genetic variants (Figure 3), including susceptibility SNPs for a large set of common diseases characterized by a complex multifactorial aetiology, including venous thromboembolism (VTE) [42, 43].

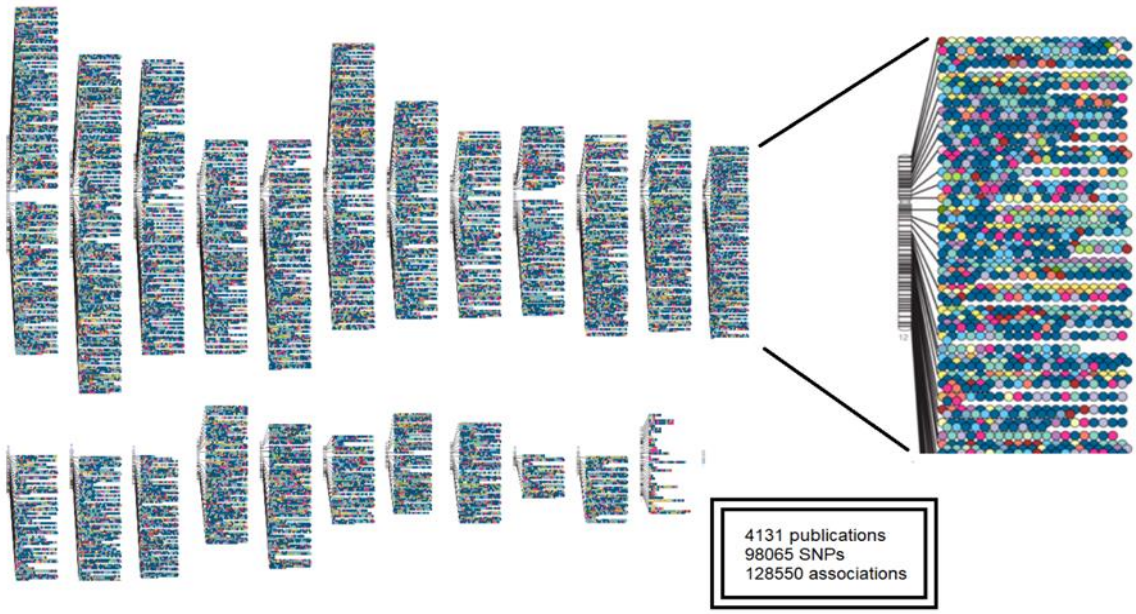

Figure 3 – Chromosomal localization of 98065 GWAS-identified variants (P<5.00 × 10^{-8}) [42]

1.2. Venous thromboembolism

As the third most common cardiovascular disease, after ischemic heart disease and stroke, VTE represents a serious health problem, affecting approximately 1 to 3 in 1000 people in developed countries [43-46]. Besides being a common disease, VTE also represents a prominent cause of morbidity and mortality [47]. The mortality rate is about 10%, which is mainly due to the pulmonary embolism (PE) events, one of the VTE forms of manifestation [44, 45].

In fact, VTE is a collective term for deep vein thrombosis (DVT) and PE [48]. The former represents the most common form of VTE manifestation, which occurs when an excessive blood clot, known as thrombus, is presented in veins located deeply inside the body, mostly legs veins, causing impairment of the blood flow (i.e., embolism) [48, 49]. As for PE, the leading cause of mortality related to VTE, is typically a DVT complication that arises when a thrombus creates an embolus that further lodges and obstructs pulmonary arteries, compromising the blood flow towards the lung, and consequently the gas exchange, which can be fatal [48, 50, 51].

The pathophysiology of VTE was first summarised in 1856 by the German pathologist Rudolf Virchow, who described venous stasis, blood hypercoagulability and vascular injury as the three major factors that contribute to disease development [52, 53]. These factors, called the arms of Virchow's triad, directly affect the haemostatic system, which involves a set of highly regulated physiological mechanisms that are trigger following endothelial damage to both reduce blood loss and guarantee vascular integrity [52, 54].

In the event of endothelial damage, a series of mechanisms are activated to initiate the haemostatic cascade with its underlying pro-inflammatory state (Figure 4) [55]. Initially, the blood loss is minimized due to the formation of a provisional plug by the platelets, a process known as primary haemostasis [56, 57]. Simultaneously, the blood coagulation system is activated, with a group of plasma proteins prompting a complex set of enzyme activation reactions that integrate the coagulation cascade, leading to the conversion of soluble fibrinogen to insoluble fibrin that deposits at the vascular injury site [58, 59]. Consequently, a fibrin clot is formed, which consolidate the previously established platelet plug (secondary haemostasis) [58-60]. Once the haemostatic plug is formed and an immune barrier against pathogens had been established, a two-step process of vascular repair is started [61]. In the first phase, known as the proliferative phase, inflammatory mediators trigger fibroblast migration to the wound site to synthesize extracellular matrix components, whereas the

angiogenic factors produced during haemostasis are responsible to promote the repair and the establishment of new blood vasculature [62, 63]. As for the second phase, the remodelling phase, it is mainly characterized by changes in the extracellular matrix components previously formed, which culminates in the formation of final scar tissue and new epithelium [64]. At this second stage, it is crucial the level of matrix metalloproteinases (MMPs) and endogenous tissue inhibitors of MMP's (TIMPs) since it allows to maintain the balance between synthesis and degradation of extracellular matrix proteins, therefore allowing a normal wound healing [65].

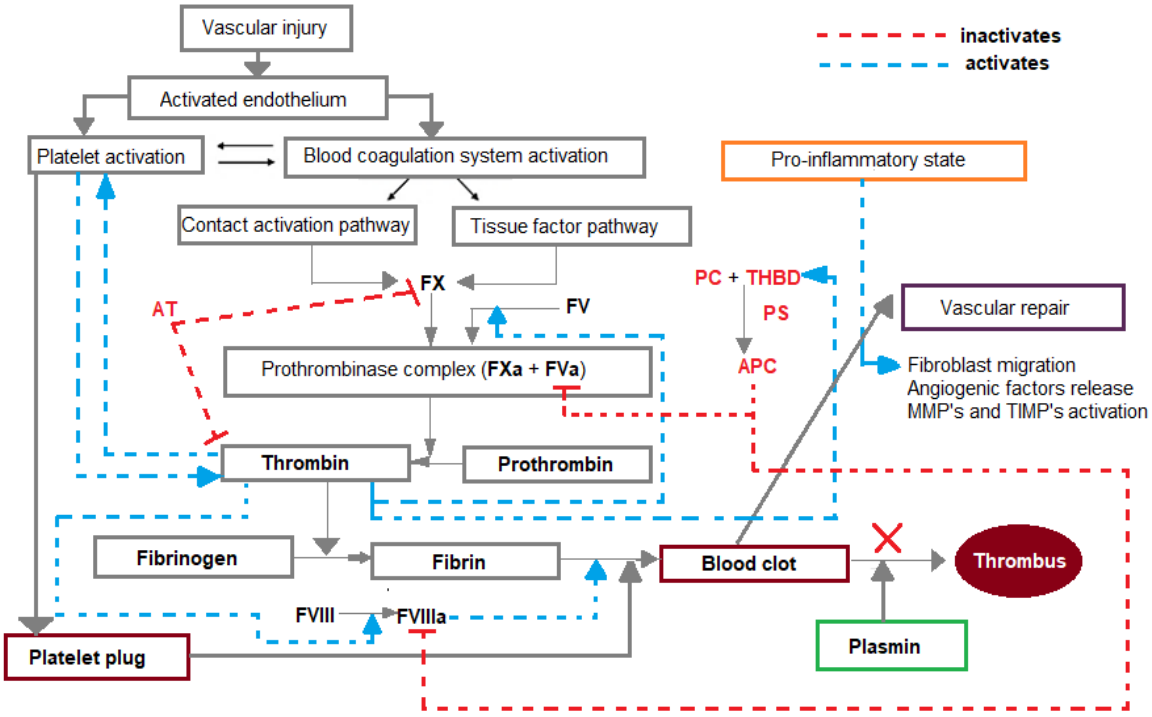

Figure 4 – Schematic representation of the haemostatic system in a simplistic perspective. Platelets are responsible for primary haemostasis; Coagulation factors promote thrombin generation and fibrin deposition (secondary haemostasis); Anticoagulant proteins highlighted in red, including antithrombin (AT), protein C(PC), protein S (PS), thrombomodulin (THBD), APC (activated protein C), are responsible to inactivate the coagulation factors preventing fibrin formation; Fibrinolytic proteins (e.g., plasmin) mediate the lysis of fibrin.

The activity of the haemostatic system is highly regulated (Figure 4) and it depends on the blood flow, vascular integrity and blood composition [66]. Under physiological conditions, a delicate balance exists between the pro-clotting processes that minimize the

bleeding following vascular injury and the anti-clotting processes that avoid thrombus formation (i.e., thrombosis) [67]. The maintenance of this balance is essential as it allows proper hemostasis and prevents the pathological states of blood hypercoagulability (thrombosis) and blood hypocoagulability (haemorrhage) [68, 69]. However, vascular endothelium injuries, defects in the blood flow (i.e., venous stasis) and changes in the blood composition, i.e., the arms of the Virchow's triad, can slope the haemostatic balance toward blood clot formation leading to a state of blood hypercoagulability. Particularly in terms of blood composition, a set of events may lead to this pathological state, including increased number or activity of the pro-clotting elements (platelets, red blood cells and coagulation factors) or reduced amount or activity of anti-clotting elements including anticoagulant proteins that limit the formation of fibrin and fibrinolytic proteins that are required to dissolve the fibrin clots already formed [69, 70].

Fibrinolysis (i.e., the degradation of fibrin) is mediated by plasmin, which is the active form of plasminogen, a plasma protein [71]. Although hyperfibrinolysis is more common, leading to severe haemorrhage, impairment of this system can also occur, promoting thrombogenesis [72, 73]. In opposition, an imbalance between the anti-coagulant and pro-coagulant proteins is a frequent cause of thrombosis and further thromboembolic events, including VTE [74-76].

1.2.1. Genetic susceptibility for VTE

In terms of pathophysiology, VTE is well known as a multifactorial disease that results from the complex interaction between acquired (i.e., environmental) and inherited factors, which explains the variation in disease susceptibility [77, 78]. Among acquired risk factors, the more relevant are advanced age, high body mass index and obesity, autoimmune diseases, and cancer [79, 80]. In addition to the nonmodifiable, trauma, immobilization, hospitalization, surgery, use of central venous catheters and access devices, oral contraception, hormone replacement, pregnancy and puerperium are some transient conditions with a reported pro-thrombotic effect [43, 79].

Although an acquired risk factor is commonly recognized, almost half of VTE cases are assumed to be idiopathic, without any acquired factor contributing to disease development [81]. This may be explained by genetic factors given that the disease familiar history is associated with almost threefold increase of VTE risk [82, 83]. Furthermore, it has

long been recognized that VTE is associated with a significant genetic burden given that 50% to 60% of the variability in the disease susceptibility can be attributed to genetic susceptibility factors. However, the inherited risk factors are not as well understood as the acquired ones [84-87]. Nevertheless, substantial advances in the studies of human disease genetics have improved our understanding of the importance of genetic determinants in VTE pathophysiology in the general population [88, 89].

During many years, the genetic research in VTE susceptibility scope was focused on a limited group of candidate genes known to play roles in the haemostatic system and, thereby, potentially leading to a blood hypercoagulability state. As a result of linkage and candidate gene studies conducted at the time, mostly in European populations, numerous genetic determinants were associated with VTE risk (Table 1) [90].

Table 1 - Overview of genetic factors reported to modulate VTE risk

Discover*	Gene	Genetic alteration	Pro-thrombotic effect	OR/RR$^{\gamma}$	Reference
1965	*SERPINC1*	Multiple mutations	Partial antithrombin deficiency	~10	[91, 92]
1969	*ABO*	[O, A2] vs. [A1, B]	Increased vWF and FVIII levels	~2.00	[93-95]
1981	*PROC*	Multiple mutations	Protein C deficiency	~10	[92, 96]
1984	*PROS1*	Multiple mutations	Protein S deficiency	~10	[92, 97]
1994	*F5*	rs6025	Resistance to activated protein C	~3.00	[92, 98-100]
1996	*F2*	rs1799963	Higher levels of prothrombin	2.80	[101]
2004	*PROCR*	rs867186	Reduced activation of anticoagulation pathway associated with protein C	1.80	[102, 103]
2005	*FGB/FGA/FGG*	rs2066865	Reduced levels of γ' chains	2.40	[104-106]
2007	*F11*	rs2289252	Higher levels of factor XI	1.30	[107]
2008	*GP6*	rs1613662	Increased platelet activation and aggregation	1.15	[108, 109]

*: Year in which the genetic alteration was associated with VTE risk. γ: Associated with the risk allele
OR: odds ratio; RR: relative risk; vWF: von Willebrand factor.

The genetic contribution to VTE risk was first recognized by establishing an association between hereditary deficiency of antithrombin (AT), a natural anticoagulant, and disease susceptibility [91]. Since then, several loss-of-function mutations linked to

deficiencies in AT but also other anticoagulants proteins including protein C (PC) and protein S (PS) were identified [110-112]. Heterozygous carriers of these inherited factors have an increased VTE risk of approximately 10-fold [88].

Apart from high-risk mutations, VTE susceptibility was found to be also attributed to other genetic factors, particularly SNPs, which given their frequency, and in light of the CD/CV hypothesis, might explain better the variation of VTE risk in the general population. One gene that is well-known to harbour genetic variants with impact in VTE pathogenesis is *ABO* [93]. Non-O blood groups are associated with almost twofold increase in VTE risk, with many *ABO* SNPs influencing disease risk most likely by modulating the plasma levels of factor VIII and von Willebrand factor, two known blood-clotting proteins [94, 113-116].

Another gene also linked with VTE pathogenesis is *coagulation factor V* (*F5*), which harbours the most common risk SNP among VTE patients, known as Factor V Leiden (FVL) (rs6025) [117]. Factor V, which is coded by *F5*, is a protein that once activated (FVa) acts as a cofactor required to convert prothrombin in thrombin, which further promotes both platelet activation and fibrin formation [118]. To inhibit thrombin generation, activated PC (APC) degrades FVa. However, carriers of FVL polymorphism have a partial resistance of FVa to cleavage by APC, resulting in an inadequate anticoagulant response, which explains why the heterozygotes with rs6025-A allele have a fourfold increase in VTE risk (OR,4.10; 95% CI, 3.23–5.21) [108, 119]. The role of *F5* in VTE pathogenesis is strengthened by the observation that other SNPs in this gene appear to modulate disease risk [108]. Another major genetic determinant for disease risk is rs1799963, a variant in the *coagulation factor II* (*F2*), a gene that codes for prothrombin [90]. The rs1799963-A allele is associated with higher levels of prothrombin, which increases thrombin generation and, consequently, VTE risk in approximately threefold (OR, 2.80; 95% CI, 1.40-5.60) [90, 101].

More recently, novel genes have been recognised to harbour SNPs with weak to moderate effect on disease susceptibility [92]. One of them is *PROCR*, which encodes for endothelial protein C receptor (EPCR) at the endothelial membrane, whose function is to enhance the activation of the anticoagulation pathway related to the PC, restricting thrombus formation (Table 1) [102]. In concordance, *PROCR* rs867186-G allele has been associated with an increased shedding of EPCR from the endothelial membrane resulting into increased levels of soluble EPCR (sEPCR), which trap activated PC (APC) impairing the anticoagulation pathway and marginally increasing the risk of developing VTE (OR=1.80; 95% CI, 1.20–2.60) [102, 103]. Also known to modulate VTE risk is *FGA/FGB/FGG* gene cluster, which encodes respectively for α, β and γ chains, three

subunits that compose the blood-clotting protein named fibrinogen (Table 1) [120, 121]. The *FGG* rs2066865 variant, the most likely functional SNP in *FGB/FGA/FGG* gene cluster, was found to be associated with plasma levels of fibrinogen γ' chains (obtained by alternative splicing of γ chains) and the γ'/γ ratio, meaning that carriers of rs2066865-T allele present reduced plasma levels of fibrinogen γ' chain and an increased VTE risk (OR, 2.40; 95% CI, 1.50–3.90) [92, 104]. An additional susceptibility gene is *coagulation factor XI* (*F11*), which encodes for factor XI (FXI), a known procoagulant protein whose levels have been associated with VTE risk [122]. Apart from rs2289252, which was found to be associated with increased plasmatic levels of FXI and an increased VTE risk (OR,1.30; 95% CI, 1.05–1.60), many *F11* SNPs with thrombotic effects have been revealed through the candidate gene approach [107, 108]. Additionally, in 2008, a large-scale genome association study restricted to 19682 SNPs, mainly non-synonymous variants, led to the identification of two novel genes specifically associated with DVT development, namely *GP6* and *CYP4V2* [108]. The former encodes for a platelet membrane glycoprotein implicated in collagen-induced activation and platelets aggregation [123]. In concordance, it has been shown that *GP6* rs1613662-A allele increases platelet activation and aggregation, with its carriers having an increased DVT risk of 1.15-fold (95% CI, 1.01-1.30) [108, 109, 113]. As for *CYP4V2*, no apparent role in the haemostatic system is known. However, evidence indicates that this gene is associated with FXI levels [108]. Therefore, the association between *CYP4V2* rs13146272 and VTE risk is probably due to the high LD observed between *CYP4V2* and *F11* genes polymorphisms [124].

1.2.2. VTE GWAS: what was left to be discovered?

Since 2009, 12 VTE GWAS have been conducted, which have allowed to confirm previously known associations but also to identify novel risk loci, including some not directly related to coagulation/fibrinolysis system (Supplementary Table 1) [42]. To be noted, most of the genetic variants reported to be associated with VTE (DVT, PE, or both) risk reached the genome-wide significance level ($P<5.00 \times 10^{-8}$). However, in some GWAS, not all SNPs that reached the genome-wide threshold were further analysed in replication studies, and/or the authors did not show the combined data when both discovery and replication phases were performed (Supplementary Table 1) [106, 125-130].

1. Introduction

The first GWAS that addressed VTE susceptibility was conducted by Trégouët and colleagues in 2009 [106]. The authors analysed 291 872 SNPs in 1228 controls and 419 VTE cases of European ancestry, reporting, in the discovery phase, five SNPs significantly associated with VTE risk ($P<3.50\times10^{-7}$) (Supplementary Table 1). Specifically, *CCDC181* rs1208134 (OR, 2.29; 95% CI, 1.58–3.32; $P=3.47\times10^{-7}$) and *F5* rs2420371 (OR, 2.27; 95% CI, 1.62–3.18; $P=8.08\times10^{-10}$), both at 1q24.2, as well as three *ABO* SNPs, namely, rs657152, rs630014 and rs505922, with the latter being more significantly associated with disease risk (OR,1.91; 95% CI, 1.53–2.39; $P=1.48\times10^{-14}$). Following discovery phase, the authors have performed two replication studies, which confirmed the GWAS findings in terms of the role of rs2420371 and rs1208134 in disease pathogenesis, although rs1208134 did not maintain the statistical significance in one of the reported replication studies (Supplementary Table 1) [106]. Nevertheless, after adjusting for *F5* rs6025, the associations regarding rs2420371 and rs1208134 were no longer observed, which might be explained by their moderate LD with FVL ($r^2=0.54$ and $r^2=0.50$, respectively). Concordantly, *F5* rs6025 was associated with a considerably higher disease risk, compared to rs2420371 and rs1208134 in both replication sets (OR, 2.01; 95% CI, 1.63–2.48; $P=9.91\times10^{-11}$ and OR, 2.46; 95% CI, 1.55–3.93; $P=1.50\times10^{-4}$). Likewise, the identified and further replicated *ABO* SNPs did not remain significantly associated with VTE risk after adjusting for rs8176719 and rs8176750, which tag the ABO blood O and A2, respectively. To be noted that rs8176719 was associated with a lower VTE risk compared to the other *ABO* SNPs in the replication studies (OR, 0.33; 95% CI, 0.26–0.42; $P=1.70\times10^{-18}$ and OR, 0.53; 95% CI, 0.41–0.69; $P=2.21\times10^{-6}$). This GWAS confirmed the previously reported role of both *ABO* and *F5* in the VTE outset, however, it failed in terms of identifying novel genes associated with disease susceptibility at genome-wide significance. Following the first publication, more 11 GWAS were conducted, including two in Afro-Americans and nine in European ancestry populations (inclusively one concerning paediatric VTE and other to search for pairwise SNP interactions associated with disease risk) [125-135].

Particularly in terms of VTE GWAS conducted among adults of European ancestry, they have allowed to confirm many susceptibility genes, namely, *ABO*, *F5*, *F2*, *F11*, *FGB/FGA/FGG* and *PROCR*, but also have reported novel genes, including *C4BPB/C4BPA*, *CCDC181, SLC19A2, NME7, CNTN6* (nearby gene to rs6764623)*, SV2C, OTUD7A, SUSD1, ZFPM2, SLC44A2, TSPAN15, FUNDC2, COX7A2L, EPHA3 (*nearby gene to rs60942712) and *TMEM170B (*nearby gene to rs113092656). Generally, considering only the ones that remain significantly associated with disease risk in replication

studies, 46 risk SNPs were reported (Figure 5). Among them, 37 SNPs were exclusively identified by GWAS (i.e., novel SNPs), including 28 variants at genome-wide significance. In this set of SNPs, the four most significant ones were *ABO* rs529565 (P=4.23×10^{-75}), *ABO* rs687621 (P=1.55×10^{-52}), *F5* rs6427196 (P=4.47×10^{-51}) and *ABO* rs505922 (P=1.39×10^{-34}) (Supplementary Table 1) [130-132]. Regarding the effect sizes, excluding the high-risk SNPs previously reported by candidate gene studies, most of the GWAS-identified variants presented, as expected, low-to-moderate effects (OR<2.00). As for the accomplishment of validation studies, only 14 novel SNPs were further analysed in validation studies, with 11 being successfully validated (Figure 5). Though, it should be noted that only validation studies performed in general population (i.e., meaning with no strong risk determinants) with incident VTE were considered (Supplementary Table 1) [136-138]. Finally, in this set of GWAS, a highlight should be given to the study performed by Germain *et al*. (2015), which was the first to identify susceptibility SNPs in novel genes reaching the genome-wide significance threshold, namely, *ZFPM2* rs4602861, *TSPAN15* rs78707713 and *SLC44A2* rs2288904 [130]. Remarkably, except for *ZFPM2*, which is potentially implicated in platelet formation, none of the novel genes have a clear role in the haemostatic system, which suggests that other VTE pathologic mechanisms remain unknown [139].

The fact that VTE constitutes essentially an age-related disease highlights the difficulty to conduct studies on the genetic factors that modulate susceptibility to paediatric VTE. Therefore, aiming to determine the genetic factors of VTE in children, Rühle *et al*. (2017) conducted a GWAS on paediatric VTE in families of European origin [128]. The SNPs that were more significantly associated disease risk, namely, rs2748331 (OR, 0.49; 95% Cl, 0.36–0.67; P=1.80×10^{-5}), rs9293858 (OR, 0.48; 95% Cl, 0.34–0.67; P=8.00×10^{-6}) and rs1304029 (OR, 0.48; 95% Cl, 0.36–0.65; P=2.00×10^{-6}) are located at 6q13 chromosome region, which includes *SMAP1*, *RIMS1* and *B3GAT2* genes. No clear role in VTE development is known for *RIMS1* and *B3GAT2,* whereas *SMAP1* is presumed to be implicated in platelet spreading and also clot retraction [128, 140, 141]. Remarkably, rs1304029 and rs2748331 were also found to be associated with VTE susceptibility in adults of European origin (OR,1.18; 95% Cl, 1.02–1.36; P=0.03 and OR,1.20; 95% Cl, 1.02–1.40; P=0.02 and, respectively). Contrariwise, SNPs in genes previously reported to modulate VTE risk in adults (*F11*, *ABO*, and *CYP4V2*) were also associated with risk of VTE in children in this study (Supplementary Table 1) [128].

Apart from single SNPs associated with VTE in European populations, Greliche *et al*. (2013) reported a genome-wide search of pairwise SNP interactions associated with

disease risk [134]. Although the combined data allowed the identification of 37 common SNP-SNP interactions significantly associated with VTE risk (Supplementary Table 2), none of the interactions reached the Bonferroni significance level. Hence, the authors concluded that it is unlikely that common SNPs exert robust interactive effects on disease risk. Nevertheless, the interaction between rs9804128 and rs4784379, both intergenic variants, showed substantial interactive effect (P=4.83×10^{-5}) on the plasmatic FVIII levels, which is a quantitative VTE biomarker [142].

Although VTE GWAS conducted in European ancestry populations have identified and confirmed many susceptibility SNPs, these genetic variants, particularly rs1799963 and rs6025, are almost absent in African ancestry populations, suggesting that different biological mechanisms might be involved in the disease pathogenesis in these populations [135]. African ancestry individuals have an incidence of VTE that is 30%-74% higher compared to other ethnicities [143, 144]. Thus, to provide insight into this topic, two GWAS with African-Americans were conducted. Briefly, in the first GWAS, performed by Hernandez and his colleagues, the two identified and replicated pro-thrombotic SNPs, rs1998081 (OR,1.94; 95% CI, 1.10–3.50; P=0.02) and rs2144940 (OR,1.89; 95% CI, 1.10–3.30; P=0.02), located near *THBD*, were found to be associated with a decrease in the expression of this gene, which encodes for thrombomodulin, a known natural anticoagulant [135, 145]. As for the second GWAS, conducted by Heit and colleagues, three pro-thrombotic variants were identified and replicated, including rs142143628 (OR, 4.97; 95% CI, 2.80–8.83; P=4.35×10^{-8}), rs3804476 (OR,1.83; 95% CI, 1.48–2.26; P=1.97×10^{-8}) and rs138916004 (OR, 3.17; 95% CI, 2.13–4.72; P=1.27×10^{-8}), located in *LOC100130298, LY86*, and *LEMD3*. The former has no connection to any known phenotype, whereas *LY86* and *LEMD3* are likely implicated in the innate immunity, which was previously associated with VTE pathogenesis [129, 146]. In both GWAS, the identified and replicated novel variants were found to be almost exclusive to African ancestry populations, although some previously reported risk SNPs in individuals of European origin were also associated with VTE risk in this population set, particularly *ABO* SNPs (Supplementary Table 1) [129, 135].

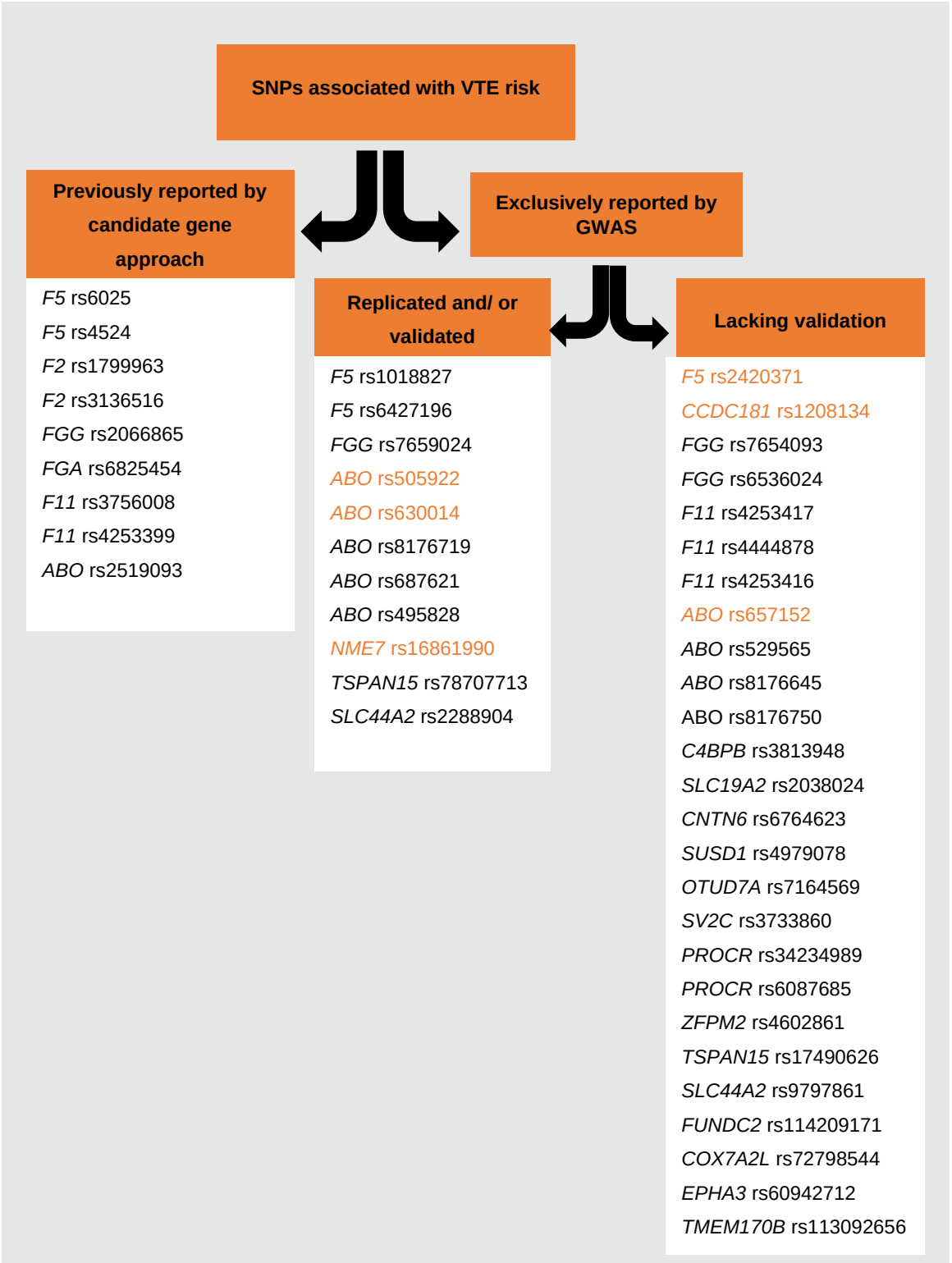

Figure 5 - GWAS-reported SNPs that are associated with VTE risk in adults of European ancestry, according to the data present in Supplementary Table 3. The SNPs highlighted in orange were reported to be no longer significantly associated with VTE after adjusting for other genetic variants. SNPs: Single-nucleotide polymorphisms; VTE: Venous thromboembolism; GWAS: Genome-wide association studies

Over the years, several VTE genetic determinants have been identified, mostly through candidate gene studies although GWAS also contribute to the identification of novel and unsuspected susceptibility genes. However, there is still a major proportion of disease heritability to be clarified. Thereby, and given that VTE is strongly heritable, other genetic factors are likely left to be revealed [130, 132, 147]. As for the genetic variants already identified, there is an increasing interest in studying the effect of these genetic determinants in cancer individuals, which constitute a significant proportion of VTE patients [148].

1.3. VTE and cancer: a two-way association

Cancer represents the most frequent cause of death worldwide, with more than 9.6 million deaths being estimated in 2018 [149]. In the last decades, in an attempt to improve patient prognosis, many efforts have been addressed to identify biological mechanisms that underpin cancer progression and, consequently, affect therapeutic response and clinical outcome [150, 151].

It is known that VTE is a common complication in patients with active cancer, and that depending on several factors, the estimated risk of developing VTE is four- to eightfold higher for cancer patients compared to cancer-free subjects [47, 152-155]. Inclusively, malignant disease is an independent risk factor for VTE outset, and even in its absence, the majority of cancer patients present several irregularities on blood coagulation tests, revealing a blood hypercoagulability state [155-157]. In accordance, a tight relationship exists between both diseases, which was firstly described in 1865 by Armand Trousseau, and it is currently been studied with particular interest given the increase in cancer-related VTE incidence in the last decades [47, 158, 159]. Remarkably, more than 15% of all cancer patients are diagnosed with VTE and up to 20% of patients with VTE events have active cancer [160]. In fact, this cardiovascular disease is recurrently the first manifestation of occult cancer, and up to 10% of the patients have a diagnose of cancer within one year [161]. Furthermore, VTE appears to have a negative impact on the prognosis of cancer patients, being the second most common cancer-related mortality [162-166]. This negative effect on prognosis can be due to the occurrence of VTE itself, however, it is becoming more evident that the increased mortality cannot be only attributed to the thromboembolic event but also to the biological mechanisms that underlie its pathogenesis (coagulation,

fibrinolysis, inflammation, angiogenesis and the extracellular matrix remodelling by MMPs) [167, 168].

The current evidence advocate that both clinical and biological determinants are implicated in the pathophysiology of cancer-related VTE [169]. The clinical determinants, briefly indicated in Table 2, are characterised as cancer-related, treatment-related and patient-related, with cancer treatment representing the most impacting risk factor [153, 158]. In terms of biological determinants, several studies have demonstrated that the haemostatic system and cancer biology are intricately connected [170]. On the one hand, tumour cells are known to be able to activate the blood coagulation system in many different ways, specifically by releasing pro-inflammatory cytokines, fibrinolytic regulators, proangiogenic factors, and pro-clotting proteins, including factor VII, cancer pro-coagulant (CP), as well as tissue factor (TF) alone but also within microparticles (MPs). Furthermore, they can trigger a blood hypercoagulability state also by inducing the release of pro-clotting components by host vascular cells, and by producing pro-aggregating components to stimulate leucocytes, endothelial cells and platelets, which induce a further pro-thrombotic cascade [170-173]. Contrariwise, the generated pro-thrombotic cascade, with the underlying pro-inflammatory state, promotes thrombogenesis, which not only can lead to VTE development but also promotes cancer progression, further enhancing thrombogenesis, which complicates the therapeutic management of the patients [174]. So, the relationship between cancer and VTE is bidirectional, with one stimulating the other [175].

Table 2 - Clinical determinants for cancer-related VTE

Patient-related	Cancer-related	Treatment-related
Older age (≥65 years old)	Histology	Surgery
Personal history of DVT	Stage	Chemotherapy
Performance status	Primary site	Radiotherapy
Obesity	Size	Anti-angiogenic therapy
Comorbidities (pulmonary disease, diabetes, heart failure, pulmonary disease, hypertension, and chronic renal disease)	Time since cancer diagnosis	Use of central vein catheters
		Supportive care agents
		Hormonal therapy
		Hospitalisation
Race		
Gender		
Pregnancy		
Immobility		

DVT: Deep vein thrombosis.
Data obtained from [153, 158, 176-179].

Over the years, cellular and circulating haemostatic components have been associated with cancer aggressiveness, namely platelets, fibrinolytic and blood-clotting proteins, such as thrombin, TF, and fibrin itself, which acts as a matrix for angiogenesis [180, 181]. Similarly, many VTE GWAS-identified genes, including some not directly involved in the haemostatic system, appear to have putative roles in several cancer hallmarks (Figure 6). Thereby, studies assessing the impact of VTE-associated SNPs on the clinical outcome of patients with cancer are warranted.

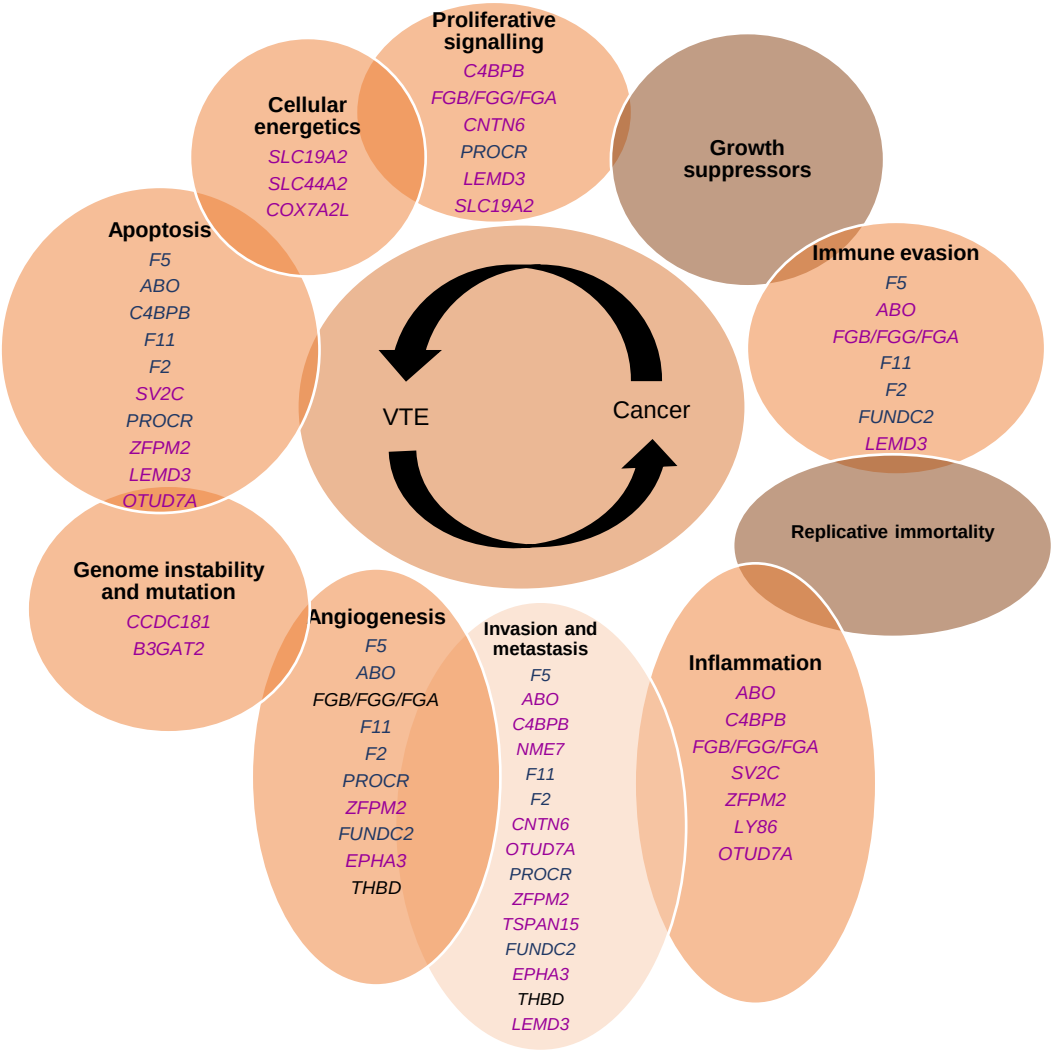

Figure 6 - VTE-related genes reported by GWAS and their putative roles in cancer pathways, according to the data present in Supplementary Table 4. The genes highlighted in blue are associated with the respective cancer hallmark through coagulation-related mechanisms in which previously mentioned haemostatic components are involved (i.e., thrombin, fibrin and platelets). The genes highlighted in purple are associated with the respective cancer hallmark through coagulation-independent mechanisms. The genes highlighted in black are associated

with individual cancer hallmark through both mechanisms. Besides, for some VTE-related genes (*SUSD1* and *LOC100130298*), no apparent role in cancer hallmarks is known.

Furthermore, although out of the scope of this dissertation, emerging evidence suggests that these genetic variants might also play a role in VTE prevention among cancer patients. The growing incidence of VTE, with its underlying impact on survival among these patients, have encouraged the implementation of prophylaxis approaches. Though, this measure has been hampered mainly due to the increased risk of bleeding complications associated with anticoagulation therapy [160, 182-187]. Hence, predictive models, such as Khorana risk score that incorporates laboratory and clinical variables to assess VTE risk in cancer patients, are essential as they allow to identify those with a positive benefit-to-harm ratio for prophylactic intervention before cancer therapy [177, 188-190]. Remarkably, VTE susceptibility SNPs might also have a role in cancer-related VTE, and thus, they might be used as predictive biomarkers of VTE among cancer patients, paving the way for better targeting thromboprophylaxis approaches based on individual risk [148]. However, the predictive capability of these SNPs, particularly those identified by GWAS, have been scarcely studied in cancer patients [163].

1.4. Impact of VTE and putative roles of haemostatic components in ovarian cancer

Ovarian cancer (OC) is the eighth leading cancer and the eighth-most frequent cause of cancer death in women worldwide, with reports of more than 295 414 OC new cases and 184 799 related deaths in 2018 [191]. Despite the relatively low incidence, OC represents the most lethal gynaecological neoplasia in industrialised countries, which is mainly attributed to the late diagnosis of the disease [181, 191]. Over 70% of the patients are diagnosed at advanced disease stages (FIGO III and IV) due to the absence of specific symptoms and the lack of methods with appropriate sensitivity and specificity for its early diagnosis [192-195].

As a heterogeneous disease, OC is subdivided into three major histological types: epithelial, sex cord and ovarian stroma, and germ cell tumours. Nevertheless, epithelial ovarian cancer (EOC) accounts for 90% of OC cases, being subdivided into high-grade

serous (70%), endometrioid (10%), clear cell (10%), mucinous (3%) and low-grade serous tumours (<5%), each one with specific genetic profiles, cellular origins, molecular alterations, pathogenesis, as well as prognosis [196-199].

Despite the late diagnosis, most OC patients initially respond to the standard treatment, which consists of cytoreductive surgery followed by systemic treatment with platinum-taxane chemotherapy [200, 201]. However, chemoresistance and tumour recurrence frequently occur, which, in combination with late diagnosis, translates into a 5-year survival rate of 20-40% [201, 202]. Given the poor prognosis, new therapeutic management strategies are required, bringing up the need for the identification of suitable biomarkers of OC progression, which so far have proven to be a challenge [181, 203-205]. Nevertheless, since there is a bilateral relationship between VTE and cancer, VTE-associated SNPs might constitute potential cancer-related biomarkers that could be used in the therapeutic management of OC patients. However, so far, few studies have assessed the impact of this cardiovascular disease and the haemostatic components in their clinical setting [181]. Therefore, clarification of OC-related VTE pathogenesis is warranted.

Among solid tumours, OC is one of the most commonly associated with VTE, as nearby 5%-20% of OC patients have a VTE event as a result of an activated coagulation system [164, 181, 206, 207]. Furthermore, studies have reported that VTE also has a detrimental effect on the prognosis of these patients, although the results are still very controversial since VTE is not an independent predictor factor of clinical outcome in all OC stages [181, 208, 209]. Nevertheless, the increased manifestation and severity of thromboembolic events in these patients appears to reflect the presence of a more aggressive malignancy, which ultimately leads to an earlier patient's death [164].

From a clinical perspective, the aetiology of OC-related VTE depends particularly on the period between cancer diagnosis and VTE diagnosis, with early events being mainly associated with the effect of cancer therapy, whereas later events are mostly associated with cancer and patient-related features [206]. Among them, the initial events are more frequent, with over 75% of all VTE events occurring after major surgery or even during chemotherapy, two medical interventions responsible for direct vascular injury and venous stasis, both arms of Virchow's triad [206, 207, 210]. Besides the standard treatment scheme, the use of Bevacizumab as a therapeutic approach in OC has also been associated with VTE risk [179, 211]. By blocking VEGF, bevacizumab appears to constrain the inhibitory effect of VEGF on plasminogen activator inhibitor-1 (PAI-1) expression, consequently increasing their levels [212]. As a crucial regulator of the fibrinolytic system,

PAI-1 prevents the activation of plasminogen, and hence, by increasing *PAI-1* expression, bevacizumab impairs the fibrinolytic pathway allowing thrombus development, which increases VTE risk [213-216].

Apart from cancer treatment and general patient-related factors, some critical cancer-related factors might also modulate VTE-risk in OC patients. Those factors encompass high histological grade and advanced cancer stage at the time of diagnosis, both underlying the presence of an aggressive malignancy, and thus, associated with an apparent more thrombogenic phenotype [207, 217]. Additionally, large size tumours can cause venous obstruction as a result of vascular compression, thus leading to venous stasis and an increased risk of thrombogenesis [218, 219]. Another cancer-related factor worthy of mention is the histological subtype, with clear cell carcinomas being more frequently associated with VTE risk than the others [207, 217].

Contrary to clinical factors, the biological mechanisms underlying the pathogenesis of OC-related VTE are not entirely understood [181]. However, TF overexpression along with the liberation of TF+ and TF-FVII+ MPs, in the presence of D-dimer (DD) elevated concentration, are thought to be significant stimulators of the pro-thrombotic state in OC patients [181, 220]. Namely, both are more pronounced in clear cell carcinomas, which could be the reason for an increased VTE risk particularly with this histological subtype [220-222].

Regarding TF, an initiator of thrombin formation (tissue factor pathway), it is reported to be constitutively expressed in many tumour cell lines, including OC cells [223-226]. Although varying among different cell lines, its overexpression appears to be promoted by oncogenic events, including the loss of *tumour protein p53 (TP53)* and *phosphatase and tensin homolog (PTEN),* activation of *KRAS proto-oncogene (KRAS)* and amplification of *epidermal growth factor receptor* (*EGFR*), as well as pro-inflammatory factors, being further enhanced under hypoxia [226-231]. Conversely, it is suggested that TF can influence multiple tumour processes either through signalling events or through thrombin generation [168, 232-234]. As a signalling protein, the role of TF in OC progression is scarcely studied [226]. Nevertheless, in other tumour cell lines, it has been reported that TF can promote the expression of factors that are not only crucial for the vascular repair process following haemostasis, under physiological conditions, but can also stimulate tumour invasion and metastasis (through MMP-2 and MMP-9), as well as promoting tumour neoangiogenesis (mediated by VEGF and interleukin-8) [235, 236]. Despite TF role as a signalling protein, it is considered that it is the TF-mediated local generation of thrombin the primary determinant

for cancer progression [168]. Apart from leading to fibrin formation, thrombin can promote OC progression in many other ways, such as modulating the expression of epithelial-mesenchymal transition related-proteins (as interleukin-6, vimentin and MMPs), which stimulate invasion and metastasis and play roles in angiogenesis, chemoresistance and apoptosis in OC cells [237-241]. Likewise, this clotting protein is also known to recruit and activate platelets, which further release several growth mediators, support tumour dissemination and immune evasion [242, 243]. Moreover, TF is also a major modulator of inflammation, which in return enhances the coagulation cascade and simultaneously promotes cancer progression [244-246]. Concordantly, TF tumoral expression has been reported to be associated with poor prognosis in OC patients [247]. Summing up, through signalling events and thrombin generation, TF favours cancer growth and metastatic dissemination, which worsens patients' prognosis.

Apart from TF, whose role both in cancer and VTE is already well established, DD is a specific product of cross-linked fibrin degradation, reflecting the concentration of fibrin deposited [218]. Concordantly, the plasma levels of DD are frequently increased in the presence of VTE, thus constituting a circulating marker of blood hypercoagulability [248, 249]. Remarkably, regardless of VTE, increased levels of preoperative DD have been associated with advanced cancer stage, chemoresistance and poor survival in OC patients, which implicates DD as a potential prognostic and predictive biomarker of OC [250-252].

Taken together, all these findings related to VTE and OC show that haemostatic factors, particularly coagulation proteins, not only can promote VTE development, but also support tumour growth, neoangiogenesis, immune evasion, migration, invasion and metastatic dissemination. Concordantly, the role of coagulation proteins in cancer progression is strengthened by the observation that, in the absence of VTE, the use of anticoagulants improves survival rates of OC patients [181]. Therefore, by reflecting the hyperactivity of haemostatic components, which could potentially be implicated in cancer progression, VTE-associated SNPs constitute potential OC-related biomarkers that are currently needed.

Although these genetic variants have been scarcely studied in OC patients, it is clear that VTE-associated genes are potentially involved in many OC-related processes. For example, Ducros *et al*. (2012) have explored the role of *PROCR* in tumour cells, including OC cells [253]. In the study, the authors observed increased levels of sEPCR, which was found to be correlated with increased cancer antigen 125 (CA-125) levels. Moreover, an enhanced cell invasion and survival, as well as immune down-regulation, was observed due to the cytoprotective effect attributed to APC [254-259]. Therefore, OC cells overexpressing

PROCR might benefit from these APC effects enhanced by EPCR. Furthermore, in the study, the increased levels of sEPCR/EPCR in OC patients were associated with expression of the *PROCR* A3 haplotype, which is tagged by rs867186-G allele, one known VTE risk factor [253, 260].

In general, the VTE susceptibility genes, and putatively the respective SNPs, are implicated in many processes that may promote OC progression, and thus, could pinpoint the molecular mechanisms associated with chemoresistance in OC therapy. As a final point, given the clinic relevance as potential OC-related biomarkers, VTE-associated SNPs might give insights needed to improve the OC patient survival, and thereby, their clinical and biological implications in these patients should be explored.

2. Aims

2.1. Main aim

The main objective of this study was to assess VTE GWAS-reported genetic variants as putative predictive biomarkers in a cohort of EOC patients from the North region of Portugal.

2.2. Specific aims

- Perform a literature review regarding VTE susceptibility GWAS;
- Select the genetic variants associated with VTE susceptibility;
- Assess the influence of the selected GWAS-associated variants in the clinical outcome of a cohort of EOC patients.

3. Materials and Methods

3.1. Study population description

It was conducted a retrospective hospital-based cohort study on histologically diagnosed EOC patients with European ancestry, admitted for first-line treatment, from January 1996 to December 2012, in the department of gynaecology and oncology of the Portuguese Institute of Oncology, Porto, Portugal (IPO-Porto). From the initial cohort of patients, were excluded those that: (1) under 18 years of age, (2) that were only submitted to specific treatment techniques, (3) only admitted for a second opinion or (4) those whose follow up was made in other institutions. After the exclusion, a cohort of 336 EOC patients from the North region of Portugal, for whom biological material was available was enrolled.

The EOC cases were all staged in accordance with the International Federation of Gynaecology and Obstetrics (FIGO) staging system and the evaluation of tumour response to chemotherapy was performed according to the Rustin criteria [261, 262]. The clinicopathological and follow-up data were obtained from patients' medical data files. The mean age of the enrolled patients, which also corresponds to the median age, was 55 years (minimum = 21 years; maximum = 80 years), from which 61.6% were postmenopausal. Most of the patients were diagnosed at advanced cancer stage (59.8% FIGO III/IV). Particularly in terms of histological subtype, 56.8% were diagnosed with serous tumours, 12.5% with clear cell, 10.1% with endometrioid, 9.8% with mucinous, and the other 10.8% with less common subtypes. Regarding the therapeutic management, most patients were submitted to the standard treatment (87.8%), with cytoreductive surgery followed by chemotherapy with a combination of paclitaxel and carboplatin (51.5%) or cisplatin (41.7%). Other options of first-line treatment also used were neoadjuvant chemotherapy (5.7%), chemotherapy alone (2.7%) or only surgery (1.2%). Considering the level of residual disease, optimal surgical resection was accomplished for 46.7% of the cases, while 19.9% and 2.7% presented residual disease ≥ 1 cm and < 1 cm, respectively. Regarding therapeutic response, 79.8% and 78.9% of the patients had a complete response in terms of imaging criteria and CA-125 levels, respectively. The mean follow-up in the study was 144.4 months (median = 148.0 months; minimum = 132 months; maximum = 163 months).

This study received approval by the ethics committee at IPO-Porto (CES IPO:286/2014). Furthermore, from each patient, it was obtained a written consent according to the principles of the Helsinki Declaration, prior to their enrolment in this study.

3.2. Laboratory procedures

3.2.1. Sample collection and genomic DNA extraction

Peripheral venous blood samples of the patients were obtained using a standard technique and further collected in EDTA-containing tubes.

From the blood samples, genomic DNA was extracted using the extraction kit Qiagen®, QIAmp DNA Blood Mini Kit (Qiagen® 51106), as indicated by the manufacturer's procedure.

3.2.2. SNP selection

To select the polymorphisms to be genotyped, all genetic variants statistically associated with VTE susceptibility were primarily gathered by screening the GWAS catalogue database [42]. From the initial list of variants, only the ones associated with disease risk in adults of European ancestry were considered, which were then submitted to the GWAS4D, an intuitive web server that analyses GWAS findings and efficiently prioritize disease-causal regulatory polymorphisms based on functional genomic and epigenomic data [263]. Based on (1) the priority rank returned by GWAS4D, (2) the MAF in the Iberian population (>14%), (3) the putative relevance in VTE pathways and the formerly accomplishment of validation studies and (4) the presumed role in ovarian cancer pathways, were selected: the *Zinc finger protein, FOG family member 2* (*ZFPM2*) rs4602861, *Solute Carrier Family 19 Member 2* (*SLC19A2*) rs2038024, *Contactin 6* (*CNTN6*) rs6764623, *OTU Deubiquitinase 7A* (*OTUD7A*) rs7164569, *Coagulation factor 11* (*F11*) rs4253417 and *Protein C receptor* (*PROCR*) rs34234989.

3.2.3. SNP genotyping

Genotyping for *ZFPM2* rs4602861, *SLC19A2* rs2038024, *CNTN6* rs6764623, *OTUD7A* rs7164569, *F11* rs4253417 and *PROCR* rs34234989 was conducted using TaqMan® Allelic Discrimination methodology (Figure 7), through Real Time Polymerase

Chain Reaction (Real time-PCR). The information regarding the validated assays used in the study are indicated in Table 3. Since predesigned TaqMan® assay were not available for *PROCR* rs34234989 and *ZFPM2* rs4602861, assays corresponding to SNPs in strong LD in the Iberian population were used, namely *PROCR* rs10747514 (r^2=1) and *ZFPM2* rs4734879 (r^2=0.94), respectively.

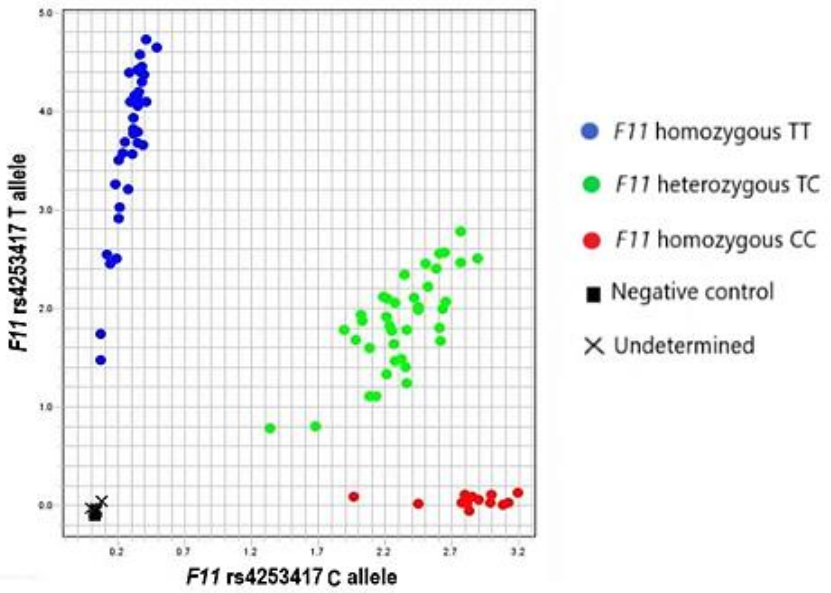

Figure 7 - Example of an allelic discrimination plot for *F11* rs4253417 polymorphism

The real time-PCR reactions were performed using 6.0 µL volumes with the following components: 2.5 µL of TaqPath™ ProAmp™ Master Mix (1x), 2.375 µL of sterile water, 0.125 µL of TaqMan® Genotyping Assay Mix and 1.0 µL of genomic DNA.

DNA amplification was conducted according to the following thermal conditions: (1) 95°C for 10 minutes to activated the Taq DNA Polymerase, (2) 45 cycles of 95°C for 15 seconds for DNA chain denaturation, and (3) 60°C for 1 minute to allow primers pairing and extension. The DNA amplification was detected and the obtained data was further analysed through the StepOne Plus Real-Time PCR system and StepOne Software (version 2.3 Applied Biosystems).

To guarantee the quality of SNP genotyping, two negative controls were included in each amplification reaction, to prevent false positives results, and double sampling was conducted in, at least, 10% of the samples randomly chosen, with an accuracy above 99%.

The genotyping results were individually evaluated by two researchers, both with no previous knowledge regarding the patient' clinicopathological data.

Table 3 - Genetic variants and respective assays used in the study

SNP	Alteration	Assay	VIC/ FAM	Flanking Sequence
rs4734879	A>G	C___1315535_10	A/G	TAGTTAGGCCAAGTTTGAACTTTTA**[A/G]**AA AGAATTATTTTTACAAGATTATT
rs2038024	C>A	C__11975194_10	A/C	TGATGGATTTCTTCATTTAAAAACA**[A/C]**AA AAGAGGTCTTGGCATAGTTTTCA
rs6764623	A>C	C__26850683_10	A/C	CAGCAACCCAAATAGTGACAAGTTA**[A/C]**AA CTTTACTCAGTAAGGACAATAAA
rs7164569	A>G	C___1698935_20	A/G	GCTGGTCTCTCTGTTCCATGGACAC**[A/G]**AG GGCAGAGAAATGGGCTTGATCAT
rs4253417	T>C	C__32291269_10	C/T	TCTTGCTCTGTCACTCAGATTTGGT**[C/T]**GC ACTGGGTGTGATCTCAGCCCACT
rs10747514	G>A	C___1825060_10	A/G	AAGGTACTACTTGGAGGGATTCTCT**[A/G]**AA AATTTTTGCCCATGTCCAAAAGC

3.2.4. In *silico* analyses

Given the lack of knowledge concerning the functionality of the selected genetic variants, *in silico* analyses were conducted. The MotifMap online tool was used to identify putative transcription factor binding sites (TFBS) generated by the different alleles of each polymorphism. Furthermore, Human Splicing Finder 3.1 was used to predict putative effects on gene pre-mRNA splicing. Additionally, the GTEx portal was used to conducted eQTL (expression quantitative trait loci) and sQTL (splicing quantitative trait loci) analysis. Particularly for *OTUD7A* rs7164569, which is a synonymous SNP, the ESEfinder 3.0 was used to identify putative exonic splicing enhancers (ESEs) motifs for the binding of the human SR proteins involved in RNA splicing (SF2/ASF, SRp40, SC35 and SRp55) and to predict whether this SNP may influence such elements.

3.2.5. Statistical analysis

Data analysis was performed using the computer software IBM® SPSS® Statistics for Windows™ (version 25.0, SPSS Inc, 2016).

Assessment of the associations between genetic variants and patients' clinicopathological data was conducted using chi-square test (x^2) and student´s t-test, for categorical and continuous variables (age), respectively.

The survival curves were obtained using the Kaplan-Meier method and probabilities of survival were analysed using the log-rank test.

In terms of clinical outcome, it was measured OS (overall survival), defined as the time period from patient diagnosis until either patient death by EOC (i.e., EOC specific survival) or last clinical evaluation, and DFS (disease-free survival), defined as the interval of time between patient diagnosis and either date of first recurrence or patients' last clinical evaluation for those with complete response to first-line treatment. Endpoint definition was based on RECIST (Response Evaluation Criteria in Solid Tumours) criteria updated in 2009 (RECIST 1.1) [264].

The risk of death and recurrence were calculated using Cox proportional-hazards model, which was adjusted for the following potential confounders: histologic subtype (serous vs. others), surgical resection (complete vs. others), age (<60. vs. ≥60 years), and hormonal status (pre vs. post-menopausal). The cause of death of each patient was determined by their medical records. Assessment of the predictive ability of different prognostic factors was performed using the concordance (c) index, with C > 0.5 being considered as good prediction ability.

All tests conducted were two-sided and it was stablished a 5% level of significance.

4. Results

4.1. Descriptive statistics of the GWAS-identified genetic variants

The genotype distribution in our study cohort for each of the GWAS-identified variants are summarized in Table 4**.** In general, the data did not show significant statistical differences between the different genotypes of each genetic variant and patients' clinicopathological characteristics, namely regarding FIGO stage (I and II vs. III and IV), age (<60 years vs. ≥60 years), histological subtype (serous vs. others), tumour grade (well-differentiated vs. others), surgical extension (complete vs. others) and hormonal status (pre- vs. post-menopausal) (*P*>0.05). However, an exception was observed for *ZFPM2* rs4734879 polymorphism and age, as there was a higher frequency of GG genotype patients among older patients (18.1% compared to 8.7% in younger patients; *P*=0.015).

For the overall cohort, the mean DFS and OS were 115.17 and 138.24 months, respectively.

Table 4 - Genotype distribution of VTE-associated SNPs in a cohort of 336 EOC patients

SNP	MAF_{IP} (MA)	MAF_{ST} (MA)	Genotype distribution						Genotyping failure
				n (%)		n (%)		n (%)	n (%)
ZFPM2 rs4734879	0.34 (G)	0.31 (G)	**AA**	161 (47.9)	**AG**	123 (36.6)	**GG**	39 (11.6)	13 (3.9)
SLC19A2 rs2038024	0.18 (C)	0.13 (C)	**AA**	244 (72.6)	**AC**	75 (22.3)	**CC**	7 (2.1)	10 (3.0)
CNTN6 rs6764623	0.25 (C)	0.23 (C)	**AA**	187 (55.7)	**AC**	109 (32.4)	**CC**	19 (5.7)	21 (6.3)
F11 rs4253417	0.42 (C)	0.43 (C)	**TT**	109 (32.4)	**TC**	142 (42.3)	**CC**	68 (20.2)	17 (5.1)
OTUD7A rs7164569	0.34 (G)	0.35 (G)	**AA**	150 (44.6)	**AG**	127 (37.8)	**GG**	49 (14.9)	10 (3.0)
PROCR rs34234989	0.22 (A)	0.35 (A)	**GG**	146 (43.5)	**AG**	127 (37.8)	**AA**	48 (14.2)	15 (4.5)

MAF_{IP} – Minor allele frequency in the Iberian population; MAF_{ST} – Minor allele frequency in our study cohort; MA – Minor allele

4.2. Impact of *ZFPM2* rs4734879 genotypes on the clinical outcome of EOC patients

Regarding the survival curves and probabilities obtained using Kaplan-Meier method and log-rank test, respectively, no significant impact of *ZFPM2* rs4734879 genotypes on patients' 5-year OS was observed, considering the log-addictive genetic model (*P*=0.691). However, when stratifying the analysis according to FIGO stage at diagnosis (I/II vs. III/IV), a significant impact of rs4734879 genotypes on 5-year OS was observed for early disease stage patients (FIGO I/II; *P*=0.031). This association is even more evident when considering a dominant genetic model (AG/GG vs. AA; *P*=0.009) (Figure 8). Namely, the patients with the AA genotype presented a mean 5-year OS of 56.29 months, whereas the patients with G allele exhibited 59.31 months. The protective effect of rs4734879-G allele was further corroborated when considering 10-year OS (*P*=0.001) (Figure 8). However, among advanced disease stage patients, no significant association regarding survival was observed, regardless of the genetic model used (*P*>0.05).

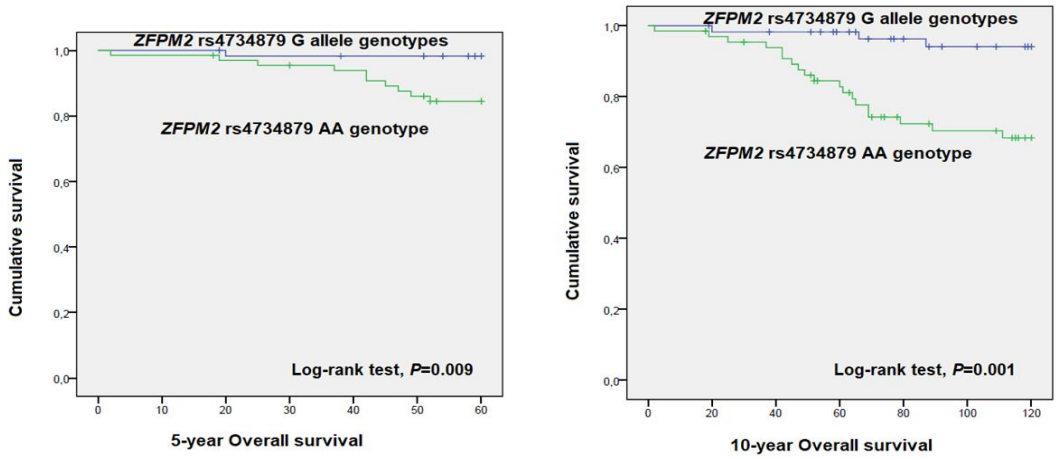

Figure 8 – Overall survival by Kaplan-Meier and long-rank test for patients with early disease stage at diagnosis, according to *ZFPM2* rs4734879 genotypes (dominant genetic model). Patients with G allele genotypes had higher 5-year and 10-year OS compared to patients with AA genotypes (*P*=0.009 e *P*=0.001, respectively).

Concordantly, by considering a dominant genetic model, a significant impact of rs4734879 genotypes on overall DFS was also observed for early disease stage patients (*P*=0.003) (Figure 9). Namely, G allele genotype patients exhibited a prolonged time to disease recurrence when compared to AA genotype patients (mean DFS of 215.85 and

164.11 months, respectively). The same was not verified for advanced disease stage patients (*P*=0.488).

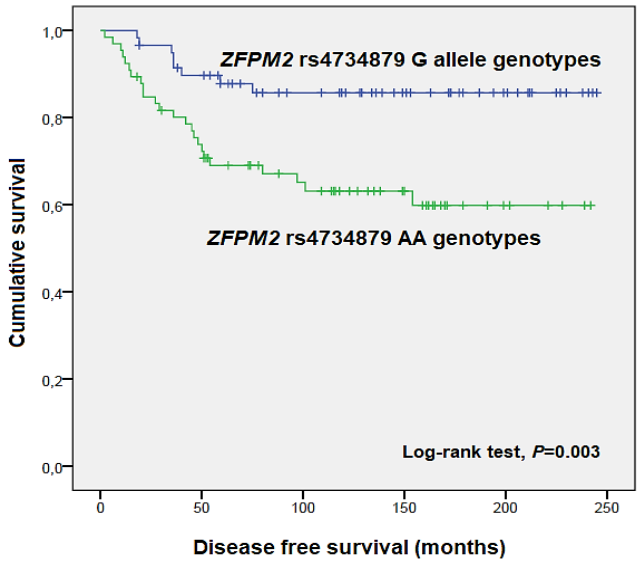

Figure 9 - Disease free survival by Kaplan-Meier and log-rank test for patients with early disease stage at diagnosis, according to *ZFPM2* rs4734879 genotypes (dominant genetic model). The patients with G allele had significantly higher DFS compared to patients with AA genotypes (*P*=0.003).

Moreover, a multivariate analysis for the risk of recurrence and risk of death of EOC patients was performed, adjusting for histological subtype, surgical extension, age and hormonal status. This analysis showed, under the dominant genetic model, a predictive impact of *ZFPM2* rs4734879 genotypes concerning the risk of recurrence and death (Table 5). Namely, the data showed that patients with AA genotype had a threefold increase in the risk of recurrence and a sixfold increase in the 10-year risk of death compared to patients with G allele (aHR, 2.67; 95% CI, 1.12-6.39; *P*=0.027 and aHR, 6.11; 95% CI, 1.78-20.39; *P*=0.004, respectively) (Table 5).

Table 5 - Multivariate analysis, using Cox regression, on the risk of recurrence and death in FIGO I/II patients considering clinicopathological variables and *ZFPM2* rs4734879 genotypes (dominant genetic model).

	Risk of Recurrence			10-year risk of death		
Variable	**HR**	**95% CI**	***P*-value**	**HR**	**95% CI**	***P*-value**
ZFPM2 rs4734879 (AG/GG vs. AA)	2.67	**1.12-6.39**	**0.027**	6.11	**1.78-20.96**	**0.004**
Histological subtype (Serous vs. others)	0.67	0.31-1.47	0.320	0.75	0.31-1.80	0.516
Surgical extension (complete vs. others)	2.90	**1.04-8.12**	**0.043**	1.67	0.47-5.98	0.430
Age (<60 vs. >=60 years)	0.93	0.37-2.30	0.866	1.63	0.65-4.09	0.301
Hormonal status (pre- vs. post-menopause)	2.27	0.84-6.17	0.108	-	-	-
Bold values were considered as statistically significant						

4.3. Impact of *SLC19A2* rs2038024 genotypes on the clinical outcome of EOC patients

Regarding *SLC19A2* rs2038024 variant and patients' 5-year OS, no statistically significant associations were noticed considering the log-addictive model (P=0.290). Conversely, by stratifying the analysis by FIGO stage, a significant impact of rs2038024 genotypes on 5-year OS was observed for patients at early cancer stages (P<0.001). The same conclusion was reached when considering a recessive genetic model (AC/AA vs. CC; P<0.001) (Figure 10). Namely, CC genotype patients presented a mean 5-year OS of 43.67 months, while the ones with A allele genotypes exhibited a mean 5-year OS of 58.09 months. Likewise, within this FIGO stage subgroup, it was also observed a significant impact of rs2038024 genotypes on 10-year OS (P=0.004), which was even more evident under the recessive genetic model (P=0.001) (Figure 10). In opposition, no significant association regarding patient survival was observed for advanced cancer stage patients, regardless of the genetic model used (P>0.05).

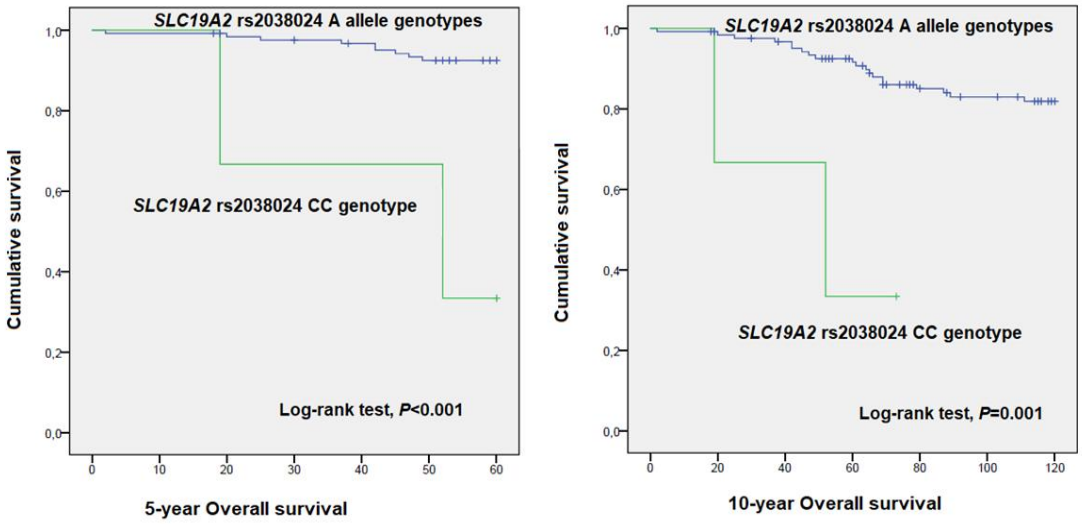

Figure 10 – Overall survival by Kaplan-Meier and long-rank test for patients with early disease stage at diagnosis, according to *SLC19A2* rs2038024 genotypes (recessive genetic model). Patients with CC genotypes had lower 5-year and 10-year OS compared to patients with A allele genotypes (P<0.001 e P=0.001, respectively).

Concordantly, by considering a recessive genetic model, a significant impact of rs2038024 genotypes on overall DFS was also observed for patients at early cancer stages (P=0.005) (Figure 11). Namely, patients with CC genotype exhibited a mean DFS of 34.00 months, whereas, for the same time survival period, the ones with A allele genotypes presented 192.32 months. The same impact was not observed for advanced disease stage patients (P=0.975).

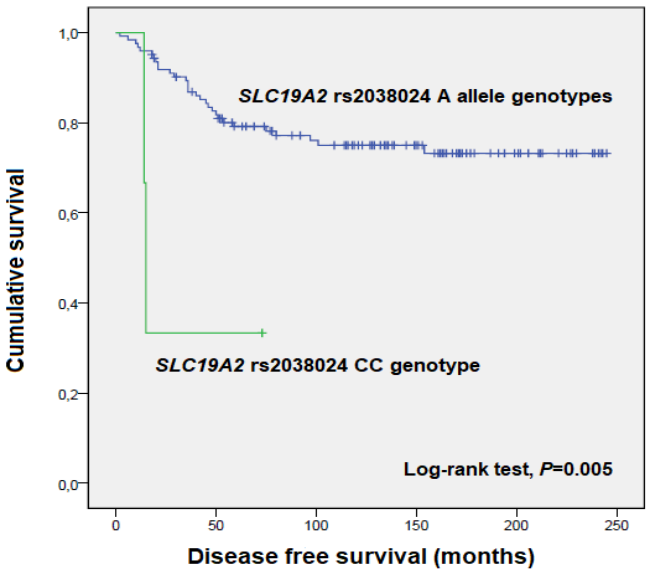

Figure 11 - Disease free survival by Kaplan-Meier and log-rank test for patients with early disease stage at diagnosis, according to *SLC19A2* genotypes (recessive genetic model). The patients with CC genotype had significantly lower DFS compared to patients with A allele genotypes (*P*=0.005).

Additionally, a multivariate analysis for the risk of recurrence and risk of death of EOC patients was performed, adjusting for histological subtype, surgical extension, age and hormonal status. Considering the recessive genetic model, this analysis showed a predictive impact of *SLC19A2* rs2038024 genotypes concerning the risk of recurrence and death of patients at FIGO I/II stages (Table 6). Namely, patients with CC genotype had a 14-fold increase in the risk of recurrence and a ninefold increase in the 10-year risk of death compared to patients with A allele (aHR, 13.78; 95% CI, 2.89-65.71; *P*=0.001 and aHR, 8.77; 95% CI, 1.93-39.79; *P*=0.005, respectively).

Although a significant association was observed between *SLC19A2* rs2038024 genotypes and EOC patients' survival, the results should be analysed carefully given the reduced number of patients carrying the CC genotype in the entire cohort (n=7) and in FIGO I/II subgroup (n=3).

Table 6 - Multivariate analysis, using Cox regression, on the risk of recurrence and death in FIGO I/II patients considering clinicopathological variables and *SLC19A2* rs2038024 genotypes (recessive genetic model).

	Risk of Recurrence			10-year risk of death		
Variable	**HR**	**95% CI**	**P-value**	**HR**	**95% CI**	**P-value**
SLC19A2 rs2038024 (AC/AA vs. CC)	13.78	**2.89-65.71**	**0.001**	8.77	**1.93-39.79**	**0.005**
Histological subtype (Serous vs others)	0.60	0.27-1.31	0.198	0.64	0.26-1.55	0.322
Surgery (complete vs others)	3.59	**1.32-9.72**	**0.012**	2.14	0.62-7.38	0.228
Age (<60 vs >=60 years)	0.76	0.31-1.88	0.550	1.31	0.52-3.28	0.566
Hormonal status (pre- vs post-menopause)	3.00	**1.08-8.32**	**0.035**	-	-	-
Bold values were considered as statistically significant						

4.4. Impact of *CNTN6 r*s6764623 genotypes on the clinical outcome of EOC patients

Regarding *CNTN6* rs6764623 variant and patients' 5-year OS, statistically significant associations were only noticed after stratifying the analysis by FIGO stage, with rs6764623 genotypes (CC vs AC vs AA) presenting a significant impact on 5-year OS of FIGO I/II patients (*P*=0.028). This association was even more evident after considering a recessive genetic model (AC/AA vs. CC; *P*=0.015) (Figure 12). Namely, in terms of mean 5-year OS, the patients with the CC genotype presented 47.50 months, whereas the ones with A allele genotypes exhibited 58.24 months. Nonetheless, the results should be analysed carefully given that only six patients with FIGO I/II exhibited rs6764623 CC genotype. Furthermore, no significant associations were observed between *CNTN6* rs6764623 genotypes and 10-year OS for FIGO I/II patients, regards of the genetic model used (data not showed). Nonetheless, there is a trend for the patients with CC genotype continuing to exhibit lower survival time when compared to patients with A allele genotypes. For advanced cancer stage patients, no significant impact of rs6764623 on patient OS was observed, regardless of the genetic model used (*P*>0.05).

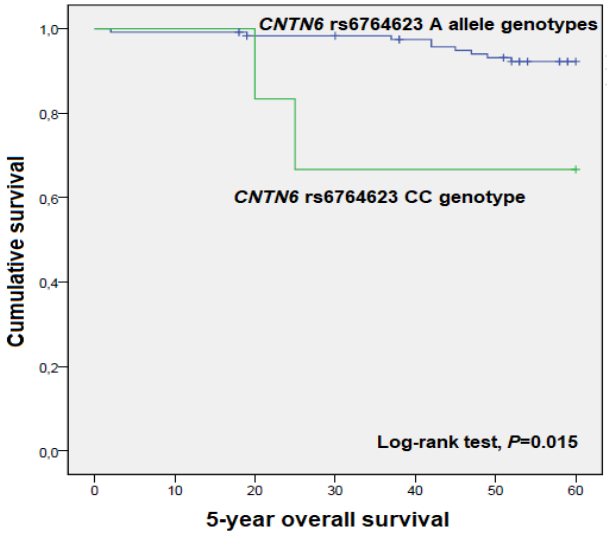

Figure 12 – Overall survival by Kaplan-Meier and long-rank test for patients with early disease stage at diagnosis, according to *CNTN6* rs6764623 genotypes (recessive genetic model). Patients with CC genotypes had lower 5-year OS compared to patients with A allele genotypes (*P*=0.015).

Considering DFS as the outcome, no significant impact of rs6764623 genotypes on patients' DFS was observed, regardless of the genetic model used and FIGO stage considered (*P*>0.05). Although not statistically significant, early cancer stage patients with CC genotype presented a reduced time to disease recurrence when compared to patients with A allele genotypes (mean DFS of 123.33 and 192.24, respectively; *P*=0.072).

Additionally, a multivariate analysis for the risk of recurrence and risk of death of EOC patients was performed, adjusting for histological subtype, surgical extension, age and hormonal status. Considering the recessive genetic model (AC/AA vs. CC), this analysis showed a predictive impact of *CNTN6* rs6764623 genotypes concerning the risk of recurrence and death of patients at FIGO I/II stages (Table 7). Namely, patients with CC genotype had a fivefold increase in the risk of recurrence and a ninefold increase in the 5-year risk of death compared to patients with A allele genotypes (aHR, 5.05; 95% CI, 1.41-18.18; *P*=0.013 and aHR, 9.41; 95% CI, 1.71-51.69; *P*=0.010, respectively).

Table 7 - Multivariate analysis, using Cox regression, on the risk of recurrence and death in FIGO I/II patients considering clinicopathological variables and *CNTN6* rs6764623 genotypes (recessive genetic model).

	Risk of Recurrence			5-year risk of death		
Variable	**HR**	**95% CI**	**P-value**	**HR**	**95% CI**	**P-value**
CNTN6 rs6764623 (AC/AA vs. CC)	5.05	**1.41-18.18**	**0.013**	9.41	**1.71-51.69**	**0.010**
Histological subtype (Serous vs others)	0.65	0.30-1.41	0.276	0.92	0.27-3.09	0.890
Surgery (complete vs others)	3.38	**1.25-9.14**	**0.016**	4.494	**1.18-17.15**	**0.028**
Age (<60 vs >=60 years)	0.91	0.37-2.23	0.830	3.71	1.001-13.712	0.050
Hormonal status (pre- vs post-menopause)	2.86	**1.04-7.91**	**0.043**	-	-	-

Bold values were considered as statistically significant

4.5. Impact of *OTUD7A* rs7164569 genotypes on the clinical outcome of EOC patients

Considering 5-year OS as the outcome, no significant impact of *OTUD7A* rs7164569 genotypes on EOC patients' survival was observed, regardless of the genetic model used and FIGO stage considered (P>0.05). Furthermore, the multivariate analysis adjusted for patient' clinicopathological factors did not reveal a predictive impact of rs7164569 genotypes regarding the risk of death of EOC patients (data not showed).

Likewise, considering the entire cohort, no significant associations between rs7164569 genotypes and patients' DFS was found (P>0.05). However, for early cancer stage patients, rs7164569 polymorphism presented a significant impact on this outcome (Figure 13), as patients with GG genotype exhibited a prolonged time to disease recurrence when compared to patients with AG e AA genotypes (mean DFS of 216.11, 164.78 and 195.85 months, respectively; P=0.025). However, the multivariate analyses for risk for recurrence adjusted for patient' clinicopathological factors did not reveal a significant impact of rs7164569 genotypes on the risk of recurrence of EOC patients (data not showed).

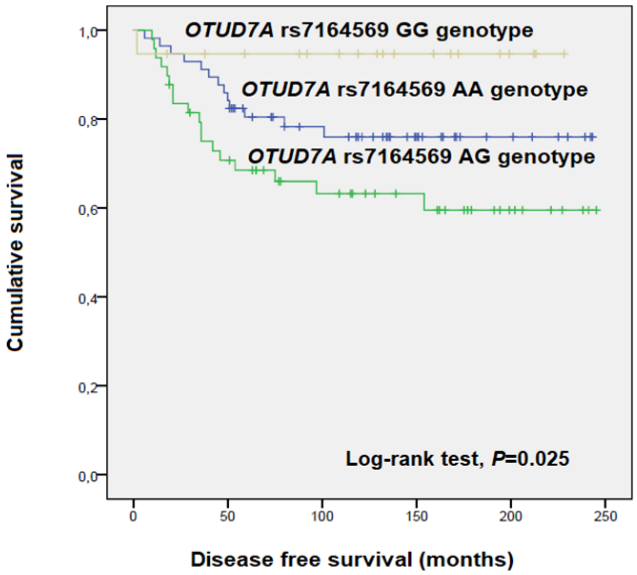

Figure 13 - Disease free survival by Kaplan-Meier and log-rank test for patients with early disease stage at diagnosis, according to *OTUD7A* genotypes (log-addictive genetic model). The patients with GG genotype had significantly higher DFS compared to patients with the other genotypes (*P*=0.025).

4.6. Impact of *F11* rs4253417 and *PROCR* rs10747514 polymorphisms on the clinical outcome of EOC patients

In general, for *F11* rs4253417 and *PROCR* rs10747514 polymorphisms, no statistically significant impact on patients' 5-year OS was observed independently of FIGO stage or genetic model considered in the analysis (*P*>0.05). However, for early cancer stage patients is possible to observe a certain tendency regarding the polymorphisms' genotypes and patient survival (Table 8). In concordance, the multivariate analysis adjusted for patient' clinicopathological factors did not reveal a significant predictive impact of these polymorphisms regarding the risk of death (data not showed).

Likewise, considering patients' DFS as the outcome, no statistically significant associations were observed for *F11* rs4253417 and *PROCR* rs10747514 polymorphisms, considering the entire cohort (*P*>0.05). Likewise, by stratifying the analysis according to FIGO stages, no significant associations between patients' DFS and *F11* rs4253417 and *PROCR* rs10747514 genotypes were noticed, respectively, neither considering FIGO I/II (Table 8) or FIGO III/IV stages (*P*>0.05). In concordance, the multivariate analysis adjusted

for patient' clinicopathological factors did not reveal a significant predictive impact of these polymorphisms regarding the risk of recurrence (data not showed).

Table 8 – Overall survival and disease-free survival by Kaplan-Meier and log-rank test for patients with FIGO I/II, according to *F11* rs4253417 and *PROCR* rs10747514 polymorphisms

Polymorphism	Minor allele	Genotypes	Mean 5-year OS (months)	*P*-value	Mean DFS (months)	*P*-value
F11 rs4253417	C	TT_TC	57.19	0.264	192.76	0.453
		CC	59.52		173.46	
PROCR rs10747514	A	GG	57.72	0.837	185.70	0.841
		AG	58.17		194.02	
		AA	56.23		174.96	

4.7. Combined analysis of *ZFPM2* rs4734879 and *SLC19A2* rs2038024 polymorphisms regarding the clinical outcome of EOC patients

For both *ZFPM2* rs4734879 and *SLC19A2* rs2038024 polymorphisms, a significant impact on 5-year and 10-year OS of patients with FIGO I/II stages at diagnosis was observed, which has led us to conduct a combined analysis of both polymorphisms restricted to early cancer stage patients.

A significant impact of *ZFPM2* rs4734879 and *SLC19A2* rs2038024 polymorphisms on patients´10-year OS was observed, as it was expected (rs2038024 CC and rs4734879 AA genotypes vs. others, *P*=0.001; rs2038024 CC and/or rs4734879 AA genotype vs. others, *P*<0.001) (Figure 14). In terms of total OS, the patients with both risk genotypes had a mean OS of 48.00 months compared to 206.49 months exhibited by the ones with other combination of rs2038024 and rs4734879 genotypes (*P*=0.001). Furthermore, the patients with at least one of the risk genotypes presented a mean OS of 175.04 months, whereas the ones with other combination of rs2038024 and rs4734879 genotypes exhibited 234.35 months (*P*<0.001).

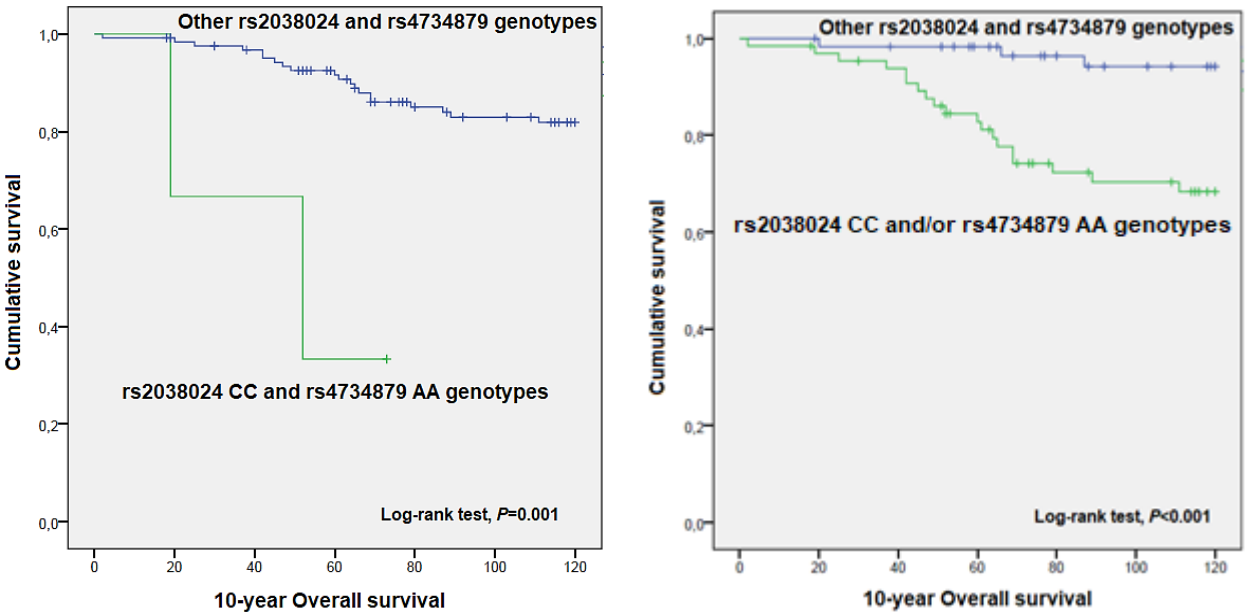

Figure 14 - OS by Kaplan-Meier and long-rank test for patients with early disease stage, according to according to *SLC19A2* rs2038024 and *ZFPM2* rs4734879 genotypes. The patients with both risk genotypes or at least one risk genotype had significantly lower 10-year OS compared to patients with other combination of *SLC19A2 rs2038024* and *ZFPM2* rs4734879 genotypes (P=0.001 and P<0.001, respectively).

Concordantly, these genotypes presented also a significant impact on patients' DFS (Figure 15). Namely, the patients with both risk genotypes had a mean DFS of 34.00 months compared to 192.32 months exhibited by the ones with other combination of rs2038024 and rs4734879 genotypes (P=0.005). Moreover, the patients with at least one the risk genotypes presented a mean DFS of 164.11 months, whereas the ones with other combination of rs2038024 and rs4734879 genotypes exhibited 216.37 months (P=0.003).

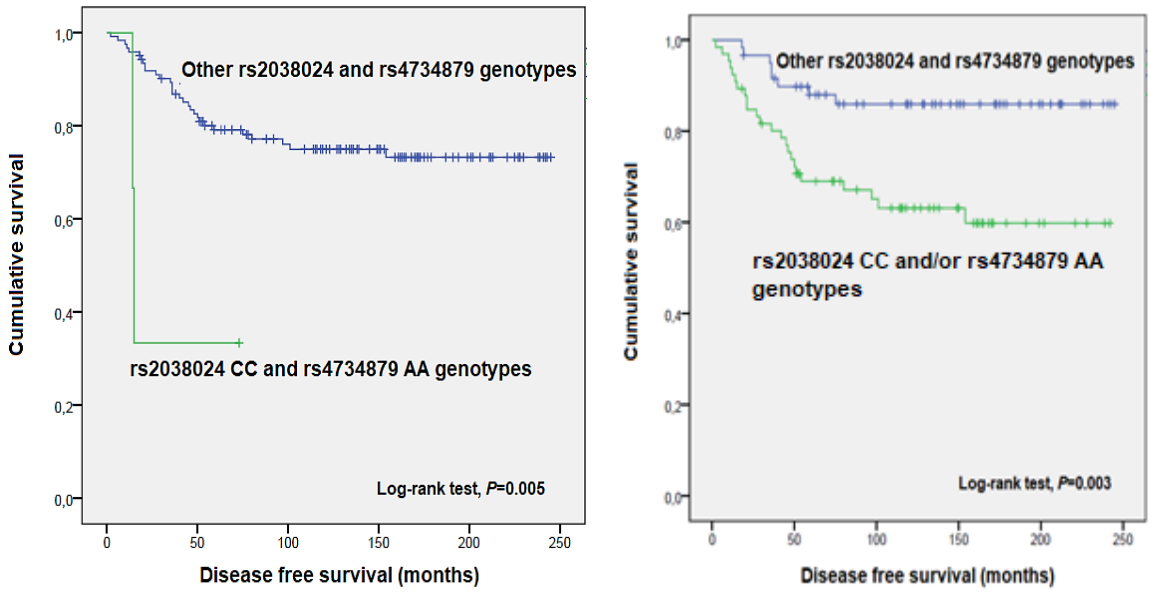

Figure 15 - Disease free survival by Kaplan-Meier and log-rank test for patients with early disease stage at diagnosis, according to *SLC19A2* rs2038024 and *ZFPM2* rs4734879 genotypes. The patients with both risk genotypes or at least one risk genotype had significantly lower DFS compared to patients with other combination of *SLC19A2* rs2038024 and *ZFPM2* rs4734879 genotypes (*P*=0.005 and *P*=0.003, respectively).

Furthermore, a multivariate analysis for the risk of recurrence and risk of death of EOC patients was performed, adjusting for histological subtype, surgical extension, age and hormonal status (Table 9). In this analysis, two models were considered. Namely, model A considering the presence of both risk genotypes of each polymorphism (others vs. rs2038024 CC and rs4734879 AA genotypes) and model B considering the presence of at least one risk genotype of each polymorphism (others vs. rs2038024 CC and/or rs4734879 AA genotype). In model A, patients with both rs2038024 CC and rs4734879 AA genotypes had a 14-fold increase in the risk of recurrence and a ninefold increase in the risk of death compared to patients with other combination of rs2038024 and rs4734879 genotypes (aHR, 13.78; 95% CI, 2.89-65.71; *P*=0.001 and aHR, 8.77; 95% CI, 1.93-39.79; *P*=0.005, respectively). As for model B, patients with rs2038024 CC genotype and/or rs4734879 AA genotype had a threefold increase in the risk of recurrence and a fivefold increase in the 10-year risk of death compared to patients with other rs2038024 and rs4734879 genotypes (aHR, 2.71; 95% CI, 1.13-6.49; *P*=0.025 and aHR, 5.35; 95% CI, 1.53-18.70; *P*=0.009, respectively).

Table 9 - Multivariate analysis, using Cox regression, on the risk of recurrence and death in FIGO I/II patients considering clinicopathological variables, according to model A and model B.

	Risk of Recurrence			10-year risk of death		
Model A						
Others vs. rs2038024 CC and rs4734879 AA genotypes	13.78	**2.89-65.71**	**0.001**	8.77	**1.93-39.79**	**0.005**
Histological subtype (Serous vs. others)	0.60	0.27-1.31	0.198	0.64	0.26-1.55	0.322
Surgery (complete vs. others)	3.59	**1.32-9.72**	**0.012**	2.14	0.62-7.38	0.228
Age (<60 vs. >=60 years	0.76	0.31-1.88	0.552	1.31	0.52-3.28	0.566
Hormonal status (pre- vs. post-menopausal)	3.00	**1.08-8.32**	**0.035**	-	-	-
Model B						
Others vs. rs2038024 CC and/or rs4734879 AA genotypes	2.71	**1.13-6.49**	**0.025**	5.35	**1.53-18.70**	**0.009**
Histological subtype (Serous vs. others)	0.68	0.32-1.48	0.33	0.82	0.32-2.08	0.674
Surgery (complete vs. others)	2.90	**1.04-8.14**	**0.043**	1.77	0.49-6.42	0.387
Age (<60 vs. >=60 years)	0.93	0.37-2.30	0.870	1.30	0.44-3.84	0.634
Hormonal status (pre- vs. post-menopausal)	2.31	0.85-6.27	0.101	1.93	0.58-6.46	0.285

Bold values were considered as statistically significant

4.8. Predictive ability of prognostic factors in EOC patients

The predictive ability of different prognostic factors regarding the 10-year risk of death among EOC patients with FIGO I/II at diagnosis was assessed according to different predictive models (Table 10). The most complex model, which includes patients' age, histologic subtype, surgical resection, hormonal status and genetic information regarding rs2038024 and rs4734879 genotypes, had the highest predictive ability (model 4; c=0.768).

Table 10 - Predictive ability of different models related to 10-year risk of death for FIGO I/II patients

	HR	95% CI	*P*-value	c-index
Model 1				
Histological subtype (Serous vs. others)	0.70	0.29-1.68	0.424	0.664
Surgery (complete vs. others)	2.82	0.94-8.45	0.065	
Age (<60 vs. >=60 years)	1.08	0.38-3.04	0.887	
Hormonal status (pre- vs. post-menopausal)	1.97	0.63-6.15	0.242	
Model 2				
Others vs. rs2038024 CC genotype and/or rs4734879 AA genotype	6.75	**1.98-23.00**	**0.002**	0.751
Histological subtype (Serous vs. others)	0.67	0.29-1.58	0.362	
Age (<60 vs. >=60 years)	1.66	0.69-4.02	0.261	
Model 3				
Others vs. rs2038024 CC genotype and/or rs4734879 AA genotype	6.24	**1.82-21.36**	**0.004**	0.757
Histological subtype (Serous vs. others)	0.75	0.31-1.81	0.527	
Age (<60 vs. >=60 years)	1.64	0.65-4.12	0.295	
Surgery (complete vs. others)	1.67	0.47-6.00	0.429	
Model 4				
Others vs. rs2038024 CC genotype and/or rs4734879 AA genotype	5.35	**1.53-18.70**	**0.009**	**0.768**
Histological subtype (Serous vs. others)	0.82	0.32-2.08	0.674	
Age (<60 vs. >=60 years)	1.30	0.44-3.84	0.634	
Surgery (complete vs. others)	1.77	0.49-6.42	0.387	
Hormonal status (pre- vs. post-menopausal)	1.93	0.58-6.46	0.285	

5. Discussion

5. Discussion

The VTE is a common cardiovascular disease that arises from the imbalance between proclotting and anticlotting mechanisms towards a blood hypercoagulation state. Although the disease pathogenesis is not fully understood, it is known that VTE results from a complex interplay between acquired and inherited factors that affect the haemostatic system [77-79]. Over the years, many susceptibility SNPs have been reported, particularly by GWAS. Since 2009, 12 VTE GWAS have been conducted, leading to the identification and confirmation of several genetic variants directly associated with disease risk, most of them located within or near genes involved in the haemostatic system, as it was expected [42]. Conversely, haemostatic components and tumour biology are intricately connected, with cancer and VTE sharing biological processes that prompt mutual interaction, particularly, inflammation, angiogenesis, and extracellular matrix remodelling [170, 243, 265, 266].

Among solid tumours, OC is one of the most frequently associated with VTE, which adversely affects patients' prognosis [207]. In general, OC patients have a poor prognosis and the manifestation and severity of thromboembolic events seems to reflect the tumour aggressiveness, and thus, VTE itself constitutes a predictor factor of the clinical outcome of these patients [181, 201, 202]. Moreover, even in the absence of VTE, a blood hypercoagulation state is often present, which indicates the presence of an activated coagulation system, with many haemostatic components playing key roles in OC progression [181].

The VTE-associated SNPs might reflect the presence of a more active coagulation system, which is characteristic of OC patients, making them not only more predisposed to experience VTE events, but also contributing to cancer aggressiveness. Thereby, these genetic variants might constitute attractive prognostic and treatment prediction biomarkers currently needed for better therapeutic management of OC patients. Thus, with the existent literature in mind, and given the potential clinical applicability, we designed the present study to assess the impact of the VTE-associated SNPs reported by GWAS on the clinical outcome of OC patients from the North region of Portugal.

5.1. Impact of *ZFPM2* rs4734879 genotypes on the clinical outcome of EOC patients

The polymorphism rs4734879 is classified as an intronic variant which leads to the substitution of an adenine (A) for a guanine (G) in the *ZFPM2* gene (also known as *FOG-2*) [267]. This gene codes for a zinc finger protein, a member of the FOG family of transcription factors, which is known to modulate the activity of transcription factors of the GATA family [268]. The GATA-binding proteins regulate, in a tissue-specific manner, the cell proliferation and differentiation, and the expression of several genes involved in cardiac and gonadal development, which they do so by binding to a consensus GATA motif, (A/T)GATA(A/G) in enhancers and promoters of these genes [269, 270]. The transcription factors of the GATA family can be divided into two groups according to the pattern of tissue-specific expression. Namely, GATA-1/-2/-3, which are implicated in the hematopoietic system development, and GATA-4/-5/-6, which are expressed in several endoderm- and mesoderm-derived tissues, including gonads, where they regulate gene expression [271, 272]. Abnormal expression of GATA-4 and GATA-6 is frequently reported in ovarian cancer cell lines, suggesting that these proteins may have key roles in ovarian carcinogenesis [273, 274]. In fact, given that GATA-binding proteins maintain the cell differentiation states, abnormalities in their expression or their regulatory pathways, which includes ZFPM2, can lead to dedifferentiation of epithelial cells, which ultimately can contribute to ovarian carcinogenesis [272, 273].

The zinc finger protein ZFPM2 is able to interact with all GATA-binding proteins by binding to the N-terminal zinc finger of each protein, leading to a stimulatory or repressive modulation of their transcriptional activity depending on the promoter and specific cell type [275]. Consequently, this modulation can activate or down-regulate the expression of GATA-target genes [276-278]. For instance, GATA-4 is able to regulate the expression of the *vascular endothelial growth factor* (*VEGF*), which is an important pro-angiogenic factor, by binding to its promoter and further enhancing transcription [279]. Therefore, given that ZFPM2 is a modulator of GATA-4, abnormalities in this gene can lead to the activation or downregulation of *VEGF* expression, with further impact on angiogenesis, one of the recognized cancer hallmarks [280]. Moreover, it was reported an increased GATA-4 expression in MYCN-amplified neuroblastomas, while tumours with favourable prognosis presented higher levels of ZFPM2, suggesting that it may play a tumour suppressed role [281].

5. Discussion

Although this protein is thought to be only a nuclear transcriptional coregulator, a significant body of evidence suggests that ZFPM2 can also be found at the cytoplasm, where it may play additional roles [282, 283]. Inclusively, it can bind to the regulatory subunits (dp60/p85α) of the *phosphatidylinositol 3-kinase* (*PI3K*), inhibiting its activation [283, 284]. The phosphatidylinositol 3-kinase-protein kinase B (PI3K-Akt),along with the mammalian target of rapamycin (mTOR) signalling pathways, is crucially involved in many processes, including cell growth, adhesion, metabolism, proliferation, invasion as well as cell survival, both in physiological and pathological conditions [285, 286]. These pathways are frequently activated in cancer, including OC, promoting its onset and progression, which also suggests that ZFPM2 might have a tumour suppressor role, given that activation of the PI3K-Akt pathway depends on the activation of PI3K [283, 287, 288]. Furthermore, also at the cytoplasm, ZFPM2 may also indirectly down-regulate tumour neoangiogenesis given that PI3K/AKT/mTOR pathway is also known to increase *VEGF* levels [289].

Corroborating the impact on *VEGF* expression, many genetic variants in *ZFPM2* have been associated with VEGF protein levels, including rs4734879 and its tightly linked SNP rs6993770 (r^2=1 in Europeans) [290, 291]. The genetic variant rs4734879 was found to be significantly associated with *VEFG* levels, with carriers of the G allele presenting lower levels of VEGF (P=7.00×10^{-30}) [291]. Therefore, this allele may be associated with a less effective tumour neoangiogenesis, and consequently, to a more favourable patient clinical outcome. Apart from *VEGF*, rs4734879-G allele was found to be associated with lower levels of other ZFPM2 protein targets that are implicated both in thrombosis and angiogenesis, including, PAI-1, platelet-derived growth factors, and Angiopoietin 1, which also suggests a protective effect of the G allele in both VTE susceptibility and cancer progression. In accordance, the results of our study suggest that the rs4734879 G allele might have indeed a protective effect, as the EOC patients carrying this allele exhibited prolonged OS and DFS.

Although the effect rs4734879 polymorphism in the expression of ZFPM2 target genes has been addressed, little is known about its functional consequence. Indeed, polymorphisms in intron regions may have different possible functional impacts, namely by disrupting a TFBS or splicing site, or even by regulating the expression (eQTL) or splicing (sQTL) of near or distant genes [292, 293]. However, according to MotifMap online tool and Human Splicing Finder 3.1 analysis, it is unlikely that rs4734879 is in a TFBS or a splicing site. Furthermore, this polymorphism is unlikely to be an eQTL or sQTL, according to the GTEx portal database. Nevertheless, recently it was found that rs4734879 G allele was also

associated with higher levels of two microRNAs known as miR-145-5p and miR-143-3p, which shared the protein targets of ZFPM2, as individuals with higher levels of either miRNA exhibited lower circulating levels of the ZFMP2 proteins targets. Thereby, the authors concluded that rs4734879 G allele may exert its effect on ZFMP2 proteins targets through miR-145-5p and miR-143-3p, which suggests that this SNP is likely to be a regulatory variant, although the exact mechanism is not described [294]. In parallel, miR-145-5p and miR-143-3p are downregulated in ovarian cancer and high grade serous ovarian cancer when compared to normal ovarian tissues, also suggesting a tumoral suppressor role for these miRNAs [295, 296].

In our study, although significant associations were not observed for the entire cohort, for EOC patients with FIGO I/II at diagnosis, our results show that rs4734879 polymorphism had a significant impact on both 5-year and 10-year OS, particularly under the dominant genetic model (AG/GG vs. AA; *P*=0.009 and *P*=0.001, respectively). In this patient subgroup, the carriers of AA genotype exhibited lower survival time compared to carriers of G allele (175.04 and 234.14 months, respectively; *P*<0.001). Concordantly, in the same subgroup, patients with AA genotype exhibited a reduced time to disease recurrence when compared to patients with G allele genotypes (mean DFS of 164.11 and 215.85 months, respectively; *P*=0.003). Moreover, in the multivariate Cox regression analysis considering only patients with FIGO I/II at diagnosis, the ones carrying AA genotype had a threefold increase in risk of recurrence and a sixfold increase in risk of death compared to patients with G allele (aHR, 2.67; 95% CI, 1.12-6.39; *P*=0.027 and aHR, 6.11; 95% CI, 1.78-20.96; *P*=0.004, respectively). The multivariate analysis showed that rs4734879 is a better predictive biomarker of risk of recurrence than the risk of death, as the association of this polymorphism with risk of recurrence remains statistically significant even after adjusting for hormonal status.

Significant associations between rs4734879 polymorphism and EOC patient survival were only observed among patients with early cancer stage at diagnosis (FIGO I/II). In general, the emerging evidence supports a tumoral suppressor role for *ZFPM2*, and given its interaction with GATA-binding proteins and the PI3K/AKT/mTOR pathway, this protein appears to regulate gene expression, cell growth, differentiation, proliferation, adhesion, invasion, survival, glucose metabolism and angiogenesis [275, 283, 284]. Despite the irrefutable role of all these processes in OC initiation and progression, and consequently, treatment response, the impact exerted by rs4734879 on *ZFPM2* and/or its gene targets might not be determinant in advance cancer stages [280].

5. Discussion

The OC cells are thought to have a unique biology, as they are only superficially invasive and primarily disseminate inside the peritoneal cavity. Nonetheless, this neoplasia is a highly deadly disease, particularly in advance cancer stages, as it is characterized by rapid proliferation, compressing visceral organs, and it becomes chemo-resistant in a short time period [297]. According to the passive dissemination model, during metastasis OC cells endure an epithelial-to-mesenchymal transition, which is promoted by up-regulation of proteolytic pathways and a change in the expression of adhesion molecules, being further carried by the peritoneal fluid, avoiding anoikis and attaching on the abdominal peritoneum or omentum, where they regain the epithelial phenotype [298]. Although this passive dissemination model is widely accepted to explain OC metastatic process, recent findings support also the existence of a hematogenous metastatic process, instigating the rethinking of the relevance of the "seed-and-soil" hypothesis for OC metastasis [299]. Conversely, the molecular mechanisms underlying OC metastasis are not entirely comprehended, which is in part attributed to OC molecular heterogenicity and the complexity of the molecular pathways that are implicated in the metastatic process [196-199]. Nonetheless, considering the emergence evidence, we hypothesized that *ZFPM2* rs4734879 might have a more relevant impact on early cancer stages, particularly in terms of angiogenesis, cell differentiation and survival, compared to advance cancer stages in which the metastatic process might be under the influence of other more impacting molecular pathways.

Despite the body of evidence supporting the tumoral suppressor role of *ZFPM2*, inclusively in ovarian carcinogenesis, and the protective effect of rs4734879 G allele, to the best of our knowledge, the impact of rs4734879 polymorphism on this neoplasia was not previously studied. Therefore, and given our study results, further analyses are warranted, particularly concerning the regulatory network that this polymorphism might be implicated, as *ZFPM2* and GATA-binding proteins seem to regulate gene expression in a tissue-specific manner.

Taking into account the previous assumptions and considering our results, rs4734879 polymorphism might constitute an attractive prognostic biomarker in OC patients with early cancer stages.

5.2. Impact of *SLC19A2* rs2038024 genotypes on the clinical outcome of EOC patients

The polymorphism rs2038024 is classified as an upstream genetic variant which leads to the substitution of an A for a cytosine (C) in the *SLC19A2* gene [267]. This gene codes for thiamine transporter 1 (THTR1), which is responsible for the cellular intake of thiamine [300]. Also known as vitamin B1, thiamine plays a crucial role in the intracellular glucose metabolism, by acting as a co-enzyme (thiamine pyrophosphate) for pyruvate dehydrogenase, transketolase (TK) and α-ketoglutarate dehydrogenase [301]. Furthermore, thiamine can regulate the expression of genes that code for these thiamine-dependent enzymes [302, 303]. In parallel, increased levels of TK and transketolase-like-1 (TKTL1) has been reported in several cancer types, including OC, suggesting that these thiamine-dependent enzymes may have a key role in carcinogenesis and in the non-oxidative pentose-phosphate pathway, which is also significantly up-regulated in cancer patients (Warburg effect) [304-306]. Concordantly, increase in TKTL1 expression was found to be correlated with increased tumour progression and patient' poor prognosis [306].

Furthermore, a substantial body of evidence suggests that thiamine might have a cytoprotective effect, particularly under hypoxic stress, as it may inhibit both p53 and PARP-1 activity, preventing hypoxia-induced apoptosis [307-309]. Therefore, given that cancer cells may exploit thiamine itself and thiamine-dependent enzymes for proliferative, survival and metabolic purposes, it is plausible that *SLC19A2* expression might be up-regulated in cancer cells, ultimately contributing to cancer progression. In fact, THTR1 as well as other thiamine-associated proteins, was found to be significantly up-regulated in both tissues and cell lines of breast cancer. Likewise, increased intracellular thiamine levels were also observed [310].

Although *SLC19A2* rs2038024 C allele was previously associated with APC resistance and increased VTE risk (OR, 1.53; 95% CI, 1.32–1.78; P=1.12×10^{-8}) (Supplementary Table 1), little is known about its functional consequence [126, 311]. According to *Ensembl* database, this upstream gene variant might have a regulatory impact given that it overlaps a promoter (ENSR00000015472) region [267]. Therefore, it is plausible that this polymorphism may influence *SLC19A2* expression.

Polymorphisms in promoter regions can affect gene expression in different ways, namely by influencing promoter activity, altering a TFBS, DNA methylation and even histone modifications, which can further lead to distant transcriptional effects [312-315]. Since it is

unlikely that rs2038024 might be in a TFBS, according to MotifMap online tool, and it is also unlikely that this polymorphism may induce histone modifications and alter DNA methylation, given that there is no CpG island in the rs2038024 flanking region (Table 3), we hypothesized that the most plausible functional impact is the alteration of the promoter activity modifying *SLC19A2* expression.

Although no significant associations were found for the entire cohort, for early cancer stage patients the rs2038024 CC genotype appeared to exert a negative impact in EOC patients' 5-year and 10-year OS. For instance, CC genotype patients presented a mean 5-year OS of 43.67 months, while the ones with A allele genotypes exhibited a mean 5-year OS of 58.09 months (P<0.001). Furthermore, patients with CC genotype exhibited a mean DFS of 34.00 months, whereas, for the same time survival period, the ones with A allele genotypes presented 192.32 months (P=0.005). Moreover, in the multivariate Cox regression analysis considering only patients with FIGO I/II at diagnosis, the ones carrying the CC genotype had a 13-fold increase in risk of recurrence and a ninefold increase in risk of death compared to patients with A allele genotypes (aHR, 13.78; 95% CI, 2.89-65.71; P=0.001 and aHR, 8.77; 95% CI, 1.93-39.79; P=0.005, respectively). The multivariate analysis showed that rs2038024 is a better predictive biomarker of risk of recurrence than the risk of death, as the association of this polymorphism with risk of recurrence remains statistically significant even after adjusting for hormonal status.

Given the previous assumptions, we hypothesized that the CC genotype might lead to higher *SLC19A2* expression, increasing the intracellular availability of thiamine in OC cells, which is further used by the tumour cells for proliferative, survival and metabolic processes to promote cancer progression.

However, the negative impact on patient survival concerning rs2038024 CC genotype was not observed for advanced disease stage patients, which let us to hypothesized that, similarly to *ZFPM2* rs4734879 polymorphism, rs2038024 might have a more relevant impact on the early cancer stages, particularly in terms of cancer cell metabolism, proliferation and survival, compared to advance cancer stages in which the metastatic process might be under the influence of other more impacting molecular pathways [316-319].

Nonetheless, future studies with a larger sample size analysing the impact of *SLC19A2* rs2038024 are required, given that in our study there was a reduced number of EOC patients carrying the CC genotype in the entire cohort (n=7) and in FIGO I/II subgroup (n=3), and therefore, the results should be taken carefully.

5.3. Impact of *CNTN6* rs6764623 genotypes on the clinical outcome of EOC patients

The polymorphism rs6764623 is classified as an intergenic variant located 100 kb from the near protein-coding gene, *CNTN6*, leading to the substitution of an A for a C [267]. Although the functional impact of this polymorphism is not described, given that it is within an intergenic region, its most reasonable functional effect is the regulation of gene transcription through cis and trans effects [320]. Therefore, assuming that rs6764623 indeed modulates VTE susceptibility, it is plausible that it could affect *CNTN6* transcription, given that its encoded protein, though the Notch pathway, has been previously associated with inflammatory responses implicated in the pathogenesis of cardiovascular disorders [321].

The *CNTN6* gene codes for a neural adhesion molecule that facilitates cell surface interactions during the development of the nervous system [267, 322]. Furthermore, CNTN6 is thought to trigger the Notch pathway, in particular, one of the four Notch receptors known as Notch 1, which was reported to increase OC proliferation [323-325]. The role of the Notch signalling pathway in the carcinogenic process can be either as oncogenic or tumour suppressor depending on the specific cellular context [326]. In OC cells, this pathway was shown to regulate cancer stem cells and tumour chemoresistance to platinum, as tumour overexpression of Notch 3 seems to result in the expansion of cancer stem cells and increased platinum chemoresistance [327]. Additionally, Notch pathway might induce the epithelial-mesenchymal transition of OC cells, a feature that not only promotes tumour invasiveness and metastasis but also increases chemoresistance in OC [328-330]. Furthermore, this pathway is thought to increase OC cell growth, proliferation, survival and angiogenesis, particularly in serous OC [331, 332].

The rs6764623 C allele was associated with an increased VTE risk (OR, 1.80; 95% CI, 1.10–1.26; P=1.57×10^{-6}) (Supplementary Table 1) [132]. In this study, although no significant associations were found for the entire cohort, for early cancer stage patients the rs6764623 CC genotype appeared to exert a negative impact on patient's survival. Namely, CC genotype patients presented a mean 5-year OS of 47.50 months, while the ones with A allele genotypes exhibited a mean 5-year OS of 58.24months (P=0.015). Inclusively, in the multivariate analysis adjusted for histological subtype, surgical extension and age, rs6764623 was independently associated with 5-year risk of death (aHR, 9.41; 95% CI, 1.71-51.69; P=0.010). In the same patient subgroup, no significant associations were

observed between rs6764623 genotypes and 10-year OS for FIGO I/II patients, regards of the genetic model used. Nonetheless, there is a trend for the patients with CC genotype continuing to exhibit lower survival time when compared to patients with A allele genotypes. As for patient DFS, no significant impact of this polymorphism was observed, regardless of the genetic model used and FIGO stage considered, which suggests that additional clinicopathological factors might be more relevant for EOC recurrence. As for the multivariate analysis for the risk of recurrence adjusted for histological subtype, surgical extension, age and hormonal status, patients with CC genotype had a fivefold increase in the risk of recurrence compared to patients with A allele genotypes (aHR, 5.05; 95% CI, 1.41-18.18; *P*=0.013).

The negative impact of rs6764623 CC genotype on patient survival was not observed for advanced disease stage patients, suggesting a preponderant role of additional factors in the disease metastatic dissemination. In fact, the several possible outcomes of Notch signalling pathway, either promoting or inhibiting cancer progression, appears to depend on the tumour microenvironment and the Notch pathway crosslinks with other signalling pathways that are also implicated in cancer progression, namely Wnt, Hedgehog, transforming growth factor-β, EGFR/HER2 receptor tyrosine kinase family, PI3K/AKT/mTOR signalling pathways among others [333].

According to these results and considering the previously assumptions, we hypothesized that rs6764623 CC genotype might lead to a higher expression of *CNTN6*, promoting the activation of the Notch signalling pathway, which could favour OC progression by promoting tumour cell growth, survival, proliferation, invasiveness, and angiogenesis and metastasis, although the latter seems to be under the influence of additional factors with more preponderant role.

However, similarly to *SLC19A2* rs2038024, the results should be taken carefully and additional well-defined studies with a larger sample size analysing the impact of *CNTN6* rs6764523 on EOC patients' survival is required, as only six FIGO I/II patients carried the CC genotype.

5.4. Impact of *OTUD7A* rs7164569 genotypes on the clinical outcome of EOC patients

The polymorphism rs7164569 is classified as a synonymous variant which leads to the substitution of an A for a G in the *OTUD7A* gene [267]. This gene codes for a member of a family of deubiquitinases (DUBs) known as ovarian tumour proteases (OTU), which specifically cleaves linkages of ubiquitin from proteins [334].

The current body of evidence suggests that this protein can inhibit the expression of *nuclear factor kappa B* (*NF-kB*), by cleaving the polyubiquitin from its positive regulator, the TNF receptor-associated factor 6 (TRAF6). Therefore, given that NF-kB is known to regulate the expression of genes implicated in inflammation, cell survival, proliferation, and migration, OTUD7A as its negative regulator may suppress these cellular processes conferring a protective effect in cancer progression [335, 336]. In concordance, *OTUD7A* was found to be downregulated in hepatocellular carcinoma cells, which was correlated with a poor prognosis [336].

Regarding *OTUD7A* rs7164569 polymorphism, little is known about its functional consequence. According to MotifMap online tool, Human Splicing Finder 3.1 analysis, it is unlikely that this polymorphism is in a TFBS or a splicing site. Furthermore, according to the GTEx portal database, rs7164569 is also unlikely to be an eQTL or sQTL. Nonetheless, according to ESEfinder 3.0, this polymorphism appears to influence ESEs motifs for the binding of human SR proteins involved in RNA splicing (SF2ASF1, SF2ASF2, SRp40 and SC35). In fact, according to UniProt Knowledgebase, two *OTUD7A* mRNA isoforms are reported, suggesting the occurrence of alternative splicing events. Therefore, it is plausible that rs7164569 may affect the alternative splicing of *OTUD7A*, consequently affecting the OTUD7A activity.

The rs7164569 G allele was also previously found to be protective regarding VTE development (OR, 0.87; 95% CI, 0.81–0.92; P=3.27×10^{-6}) (Supplementary Table 1), which is plausible given that NF-kB was reported to be also a positive regulator of platelet responses, and thereby, by inhibiting NF-kB, OTUD7A may impair platelet activation and aggregation, and thus, decrease VTE risk [132, 335-338]. In our study, although there was no significant impact of rs7164569 in EOC patient's OS, regardless of the FIGO stage and genetic model considered, this polymorphism genotypes were significantly associated with EOC patient's DFS, considering FIGO I/II stages (P=0.025). Namely, the patients with GG genotype exhibited a prolonged time to disease recurrence when compared to patients with

AG e AA genotypes (mean DFS of 216.11, 164.78 and 195.85 months, respectively) (Table 8). This suggests that rs7164569 variant, despite probably not changing EOC patients' OS, may delay tumour recurrence. Contrariwise, the heterozygous patients were the ones exhibiting lower DFS, suggesting that in these patients, additional factors might favour cancer progression in a more relevant way.

Although the putative impact of this polymorphism in EOC patient's DFS, the multivariate analyses for risk for recurrence adjusted for histological subtype, surgery extension, age and hormonal status, did not reveal a significant impact of rs7164569 genotypes on the risk of recurrence, which also suggests that additional clinicopathologic factors may have a more preponderant role in EOC recurrence.

Nevertheless, the positive impact of rs7164569 GG genotype on EOC patient' DFS suggest that this polymorphism genotype may lead to the expression of a more active deubiquitinase, which could more efficiently downregulate NF-kB, inhibiting its oncogenic roles (cell survival, proliferation, and migration). Furthermore, this impact seems to more relevant in early cancer stages, as no significant association between this polymorphism and EOC patients' DFS was found in FIGO III/IV stages, suggesting that additional and more relevant factors might be implicated in OC metastatic dissemination.

5.5. Impact of *F11* rs4253417 genotypes on the clinical outcome of EOC patients

The polymorphism rs4253417 is classified as an intronic variant which leads to the substitution of an thiamine (T) for a C in the *F11* gene [267]. As it was mentioned before, *F11* codes for FXI, a coagulation factor of the contact activation pathway (intrinsic coagulation pathway), which culminates with thrombin generation [122, 339].

In terms of functional impact, the polymorphism rs4253417 does not seem to be in a TFBS or a splicing site and it is unlikely to be a sQTL, according to MotifMap online tool, Human Splicing Finder 3.1 and GTEX portal database, respectively. Nonetheless, rs4253417 was previously associated with FXI plasmatic levels, suggesting a putative role of this polymorphism in *F11* expression, which is corroborated by eQTL analysis in the GTEx portal [340, 341]. Furthermore, rs4253417 C allele was associated with an increase in VTE risk (OR, 1.27; 95% CI, 1.22–1.34; $P=1.21\times10^{-23}$) (Supplementary Table 1),

suggesting that this polymorphism could indeed modulate FXI plasmatic levels, which is a known VTE biomarker [130].

Conversely, increased levels of FXI and FXII transcripts were reported in EOC patients' peritoneum, suggesting that the contact activation pathway might contribute to EOC dissemination in the peritoneal cavity [342]. Although a direct role of this protein in cancer progression has not yet been established, FXI may have an indirect impact on carcinogenesis by favouring downstream factors of the coagulation pathway, including factor Xa, which is thought to inhibit apoptosis and anoikis favouring the metastatic dissemination, as well as thrombin, which has established roles in the apoptosis, immune evasion, angiogenesis and metastasis [343-352]. Therefore, given these putative roles of FXI in OC progression, we initially wonder whether *F11* rs4253417 polymorphism could have an impact on ovarian carcinogenesis. The results of our study, however, do not support a significant impact of this polymorphism on the clinical outcome of EOC patients. Therefore, as a direct role in carcinogenesis has not yet been established, and considering the indirect impact of this protein, we hypothesized that additional elements, inclusively in the coagulation cascade, could exert more impacting roles [353]. Nevertheless, given that this study is the first to analyse the impact of this polymorphism in OC progression, further analyses are required.

5.6. Impact of *PROCR* rs10747514 genotypes on the clinical outcome of EOC patients

The polymorphism rs10747514 is classified as an intronic variant which leads to the substitution of a G for an A in the *PROCR* gene [267]. This gene codes for EPCR, which, as mentioned before, is the endothelial receptor for PC [102].

Following a vascular injury, circulating PC binds to EPCR, which facilitates PC interaction with the complex formed by thrombin and thrombomodulin at the endothelial cell surface. Thrombin further cleaves PC generating APC, which once released from EPCR, inactivates the coagulation factors FVa and FVIIIa in the presence of PS, resulting in the downregulation of thrombin generation, and thus preventing thrombus formation [354]. Due to its anticoagulant activity, EPCR modulates the risk of developing thrombosis and further thromboembolic events, including VTE [69, 70]. However, in addition to thromboembolic events, EPCR has been also implicated in processes that promote cancer progression [354,

355]. In fact, *in vivo* and *in vitro* studies have demonstrated that through the anticoagulant pathway associated related to the PC, EPCR may play key roles in inflammation, angiogenesis, apoptosis, proliferation, migration and invasion by regulating several pathways independent of the haemostatic system [354].

In terms of inflammation, EPCR by binding APC was shown to reduce immune cells migration [356]. Likewise, this receptor was also reported to increase angiogenesis via the eNOS pathway and MMP activation [256, 357]. Furthermore, EPCR seems to increase cell proliferation by activating eNOS, PI3K and MAPK pathways [256]. Additionally, through the interaction with APC and protease-activated receptors (PAR), this receptor might inhibit apoptosis by downregulating p53, Bax and caspases 3,8 and 9 and upregulating Bcl-2 [358-361]. In the same conditions (i.e., interacting with APC and PAR), EPCR might also influence the activation of several signalling pathways that increase cell migration and invasiveness, a feature that was also reported in OC cells [257, 362]. Therefore, the emergent evidence strongly suggests that, through interaction with APC, EPCR might contribute to cancer progression. In fact, the roles of EPCR in carcinogenesis have been extensively addressed in breast, gastric, colorectal and ovarian cancer [354, 363-365].

As for *PROCR* rs10747514 polymorphism, little is known about its functional impact. According to MotifMap online tool and Human Splicing Finder 3.1 analysis, it is unlikely that this polymorphism is in a TFBS or a splicing site. However, according to GTEx portal database, this genetic variant is a putative eQTL and sQTL, exerting changes in the expression and splicing process of *PROCR* and surrounding genes. Therefore, it is plausible that this polymorphism might affect EPCR levels or activity. Furthermore, rs10747514 is in strong LD with two intronic SNPs that were previously associated with VTE susceptibility, namely rs34234989 and rs2069945 (r^2=1 in Iberian populations for both variants) [127, 267, 366]. The variant rs34234989 as a VTE risk SNP was identified in a GWAS and its effect is not described, whereas rs2069945 is part of *PROCR* haplotype 1 to which is attributed a protective effect regarding VTE development [127, 366]. Therefore, given that rs10747514 is in strong LD with rs34234989, a VTE GWAS-identified variant, it is possible that rs10747514 could be the causal variant, and consequently, could modulate EPCR levels or its activity with further implications in cancer progression.

However, the results of our study do not support a significant impact of rs10747514 polymorphism on the clinical outcome of EOC patients, which brings two possibilities: 1) this polymorphism may not be a causal variant and neither it is in strong LD with the causal variants modulating EPCR levels or activity, or 2) assuming that this polymorphism might

modulate ECPR levels or activity, or at least be in strong LD with the causal variant, additional and more preponderant factors might overshadow the impact of this polymorphism in cancer progression. Nevertheless, similarly to *F11* rs4253417 polymorphism, given that this study is the first to analyse the impact of rs10747514 polymorphism in OC progression, further analyses are required.

5.7. Combined impact of *ZFPM2* rs4734879 and *SLC19A2* rs2038024 genotypes on the clinical outcome of EOC patients

As it was expected, a significant impact of rs4734879 and rs2038024 polymorphisms on EOC patients' 10-year OS was observed, considering the presence of both risk genotypes of each polymorphism (others vs. rs2038024 CC and rs4734879 AA genotypes; *P*=0.001) or at least one of the risk genotype of each polymorphism (others vs. rs2038024 CC and/or rs4734879 AA genotypes; *P*<0.001). Likewise, a significant impact of these polymorphisms on EOC patients' DFS was also observed, considering the presence of both risk genotypes (*P*=0.005) or at least one of the risk genotypes of each polymorphism (*P*=0.003). Furthermore, the multivariate analysis for the risk of recurrence and risk of death of EOC patients, adjusted for patients' clinicopathological factors, revealed a significant predictive impact of *ZFPM2* rs4734879 and *SLC19A2* rs2038024 polymorphisms concerning the risk of recurrence and the risk of death, either considering the presence of both risk genotypes (aHR, 13.78; 95% CI, 2.89-65.71; *P*=0.001 and aHR, 8.77; 95% CI, 1.93-39.79; *P*=0.005, respectively) or the presence of at least one risk genotype (aHR, 2.71; 95% CI, 1.13-6.49; *P*=0.025 and aHR, 5.35; 95% CI, 1.53-18.70; *P*=0.009, respectively). In all these analyses, EOC patients with both rs4734879 AA and rs2038024 CC genotypes or at least one of the risk genotypes presented a worse clinical outcome, as it was expected given the putative and independent roles of both polymorphisms in cancer progression.

However, given the limited number of EOC patients with FIGO I/II stages at diagnosis carrying *SLC19A2* rs2038024 CC (n=3) and consequently the limited number of *SLC19A2* rs2038024 CC and *ZFPM2* rs4734879 AA genotypes carriers, future studies are required in order to proper study the predictive ability of the presence of both rs2038024 CC and rs4734879 AA genotypes in model A. As for model B, we observed a substantial decrease in the range of 95% CIs in comparison with the same analysis but considering rs2038024 and rs4734879 polymorphisms separately. Therefore, by considering this model, there is a

lower variance in the concluded outcomes, which suggests that this model is better suitable to predict the risk of recurrence and the 10-year risk of death in EOC patients with FIGO I/II stages, in comparison with the models considering rs2038024 and rs4734879 polymorphisms separately.

5.8. Predictive ability of prognostic factors in EOC patients

Using four different predictive models, we have analysed the predictive ability (c-index) of some OC-related prognostic factors concerning the 10-year risk of death among EOC patients with FIGO I/II at diagnosis (Table 10). Additionally, we analysed the impact on the predictive ability by including the genetic information regarding *SLC19A2* rs2038024 and *ZFPM2* rs4734879 polymorphisms.

The model 1, which includes the clinical characteristics histological subtype, surgical extension, age and hormonal status, revealed a c-index of 0.664. However, according to Cox regression analysis, none of these factors in model 1 had a significant independent impact on the 10-year risk of death (P>0.05). In fact, one of the biggest challenges in OC research is to identify strong prognostic and/or predictive biomarkers in early cancer stages [367]. Nonetheless, by including the genetic data concerning *SLC19A2* rs2038024 and *ZFPM2* rs4734879 polymorphisms (rs2038024 CC genotype or rs4734879 AA genotype vs. others) there was an increase in the ability to predict the 10-year risk of death (model 2, 3 and 4; C=0.751, C=0.757 and C=0.768, respectively). In all the models encompassing rs4734879 and rs2038024 genotypes, the genetic information concerning these polymorphisms emerged as the more impacting prognostic factor (model 2, 3 and 4; P=0.002, P =0.004 and P =0.009, respectively). Furthermore, the most complex model (model 4), which includes rs2038024 and *rs4734879* genotypes information, histological subtype, age, surgical resection and hormonal status had the highest predictive ability (c=0.768). Therefore, the results suggest that a predictive profile including both *SLC19A2* rs2038024 and *ZFPM2 rs4734879* polymorphisms' information could be used as a clinical tool to better predict the outcome of EOC patients with early disease stages, and therefore, enabling better therapeutic management strategies

However, as this is the first study to report the impact of rs2038024 and *rs4734879* polymorphisms in combination with clinical prognostic factors in EOC patients' survival,

future studies are required to replicate our findings in another study cohort with FIGO I/II at diagnosis.

6. Conclusion and future perspectives

6. Conclusion and futurue perspectives

Over the years, through candidate gene studies, the research for SNPs associated with trait phenotypes has been focused on nonsynonymous SNPs, meaning those in genome coding-regions affecting protein' amino acid sequence, and consequently, the protein structure and function [368]. More recently, substantial advances towards a better understanding of complex and common diseases have been made with the development of high-throughput genotyping technologies allowing the entry in the era of GWAS [3, 16, 18, 19]. In opposition to candidate gene studies, GWAS investigate the entire genome, relying on the principle of LD, to identify common SNPs underlying susceptibility to common and multifactorial diseases [12]. As a result, this genetic tool has identified several low-penetrance variants associated with a wide range of diseases [42]. Remarkably, most of these genetic variants although lying in noncoding regions and being supposedly irrelevant to disease pathogenesis, seem to exert an indirect effect on the expression and activity of near or distant genes by regulating complex genetic networks [369].

Particularly in terms of VTE, 12 GWAS have identified and confirmed several genetic variants underlying disease susceptibility, most of them with a putative regulatory role, which given the bilateral relationship between VTE and cancer, constitute potential prognostic and predictive biomarkers currently needed for better management of cancer patients, in particular, OC patients [42, 181]. Even in the absence of VTE, OC patients present a blood hypercoagulability state, which suggests that haemostatic components may play key roles in ovarian carcinogenesis [181]. To corroborate this theory, VTE susceptibility genes, and the respective SNPs, seem to be implicated in many processes including coagulation, inflammation, angiogenesis, metabolism and cellular signalling pathways that may ultimately promote OC progression, providing new directions for personalized cancer treatment. This is extremely important because, despite the improvements in OC treatment, patients' overall prognosis remains poor with a 5-year survival rate of 20-40% [201, 202]. Furthermore, identifying OC-related biomarkers with significant predictive and/or prognostic value has been proved to be a major challenge [205].

In this perspective, this study was first designed to study the association of six VTE-associated SNPs previously reported by GWAS with the clinic outcome of OC patients. To our best knowledge, this is the first study that attempts to analyse the impact of these polymorphisms on the clinical outcome of OC patients. Summarizing, we found significant associations between EOC patients' clinic outcome and the polymorphisms *ZFPM2* rs4734879, *SLC19A2* rs2038024, *CNTN6* rs6764623 and *OTUD7A* rs7164569, all

putatively relevant in FIGO I/II stages. Despite promising, these associations should be further analysed in larger cohorts of EOC patients, particularly concerning *SLC19A2* rs2038024 and *CNTN6* rs6764623 given the under-power in early cancer stages patients with rs2038024 CC genotype (n=3) and *CNTN6* rs6764623 CC genotype (n=6), respectively. Additionally, the moderate range of 95% CIs calculated in the multivariate analysis could imply the existence of moderate variance in the outcomes, which also bring up the need for additional studies to validate our findings. Likewise, future analysis including functional studies, fine mapping, eQTL and sQTL analyses are mandatory to better characterize the functional consequence of these polymorphisms, and consequently, better understand the biological mechanisms underlying the results [370-372].

As complex diseases often share biological pathways, VTE GWAS findings might bring a new understanding of the genetic architecture that influences OC initiation and progression and the acquisition of cellular chemoresistant behaviour, which could help to personalize the treatment approaches aiming a better clinical outcome [373]. Furthermore, these findings could bring new directions in pharmacogenomics, as it could also help to develop effective anticoagulants with an additional benefit for OC patients, given that the proteins associated with the genetic factors for VTE development might constitute potential therapeutic targets for both cancer and VTE treatment [181, 250, 374]. Besides, more research in the field is also important given that the knowledge may be further applied to other types of cancer, considering that the incidence of cancer-related VTE has been increasing over the last decades [158, 375, 376].

7. References

7. References

1. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-922.
2. Venter, J.C., et al., *The sequence of the human genome.* science, 2001. **291**(5507): p. 1304-1351.
3. Naidoo, N., et al., *Human genetics and genomics a decade after the release of the draft sequence of the human genome.* Human genomics, 2011. **5**(6): p. 577.
4. Cantsilieris, S. and S.J. White, *Correlating multiallelic copy number polymorphisms with disease susceptibility.* Human mutation, 2013. **34**(1): p. 1-13.
5. Brookes, A.J., *The essence of SNPs.* Gene, 1999. **234**(2): p. 177-186.
6. Karki, R., et al., *Defining "mutation" and "polymorphism" in the era of personal genomics.* BMC medical genomics, 2015. **8**(1): p. 37.
7. Gupta, P., J. Roy, and M. Prasad, *Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants.* Curr Sci, 2001. **80**(4): p. 524-535.
8. Collins, F.S., L.D. Brooks, and A. Chakravarti, *A DNA polymorphism discovery resource for research on human genetic variation.* Genome research, 1998. **8**(12): p. 1229-1231.
9. Feuk, L., A.R. Carson, and S.W. Scherer, *Structural variation in the human genome.* Nature Reviews Genetics, 2006. **7**(2): p. 85.
10. Shastry, B.S., *SNPs: impact on gene function and phenotype*, in *Single Nucleotide Polymorphisms*. 2009, Springer. p. 3-22.
11. Wang, X., et al., *Single nucleotide polymorphism in transcriptional regulatory regions and expression of environmentally responsive genes.* Toxicology and applied pharmacology, 2005. **207**(2): p. 84-90.
12. Manolio, T.A., *Genomewide association studies and assessment of the risk of disease.* New England journal of medicine, 2010. **363**(2): p. 166-176.
13. Bush, W.S. and J.H. Moore, *Genome-wide association studies.* PLoS computational biology, 2012. **8**(12): p. e1002822.
14. Wang, Z. and J. Moult, *SNPs, protein structure, and disease.* Human mutation, 2001. **17**(4): p. 263-270.
15. Wall, J.D. and J.K. Pritchard, *Haplotype blocks and linkage disequilibrium in the human genome.* Nature Reviews Genetics, 2003. **4**(8): p. 587.
16. Manolio, T.A., L.D. Brooks, and F.S. Collins, *A HapMap harvest of insights into the genetics of common disease.* The Journal of clinical investigation, 2008. **118**(5): p. 1590-1605.
17. Pritchard, J.K. and M. Przeworski, *Linkage disequilibrium in humans: models and data.* The American Journal of Human Genetics, 2001. **69**(1): p. 1-14.
18. Hirschhorn, J.N. and M.J. Daly, *Genome-wide association studies for common diseases and complex traits.* Nature Reviews Genetics, 2005. **6**(2): p. 95.
19. Hardy, J. and A. Singleton, *Genomewide association studies and human disease.* New England Journal of Medicine, 2009. **360**(17): p. 1759-1768.
20. Pearson, T.A. and T.A. Manolio, *How to interpret a genome-wide association study.* Jama, 2008. **299**(11): p. 1335-1344.
21. Gabriel, S.B., et al., *The structure of haplotype blocks in the human genome.* Science, 2002. **296**(5576): p. 2225-2229.
22. Visscher, P.M., et al., *Five years of GWAS discovery.* The American Journal of Human Genetics, 2012. **90**(1): p. 7-24.
23. Klein, R.J., et al., *Complement factor H polymorphism in age-related macular degeneration.* Science, 2005. **308**(5720): p. 385-389.
24. Lander, E.S., *Initial impact of the sequencing of the human genome.* Nature, 2011. **470**(7333): p. 187.
25. Clarke, A.J. and D.N. Cooper, *GWAS: heritability missing in action?* European Journal of Human Genetics, 2010. **18**(8): p. 859.

26. Zuk, O., et al., *The mystery of missing heritability: Genetic interactions create phantom heritability.* Proceedings of the National Academy of Sciences, 2012. **109**(4): p. 1193-1198.
27. Gusev, A., et al., *Quantifying missing heritability at known GWAS loci.* PLoS genetics, 2013. **9**(12): p. e1003993.
28. Bourrat, P., Q. Lu, and E. Jablonka, *Why the missing heritability might not be in the DNA.* BioEssays, 2017. **39**(7): p. 1700067.
29. Spencer, C., et al., *Quantifying the underestimation of relative risks from genome-wide association studies.* PLoS genetics, 2011. **7**(3): p. e1001337.
30. Plenge, R.M., et al., *Two independent alleles at 6q23 associated with risk of rheumatoid arthritis.* Nature genetics, 2007. **39**(12): p. 1477.
31. Yang, J., et al., *Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index.* Nature genetics, 2015. **47**(10): p. 1114.
32. Kraft, P., E. Zeggini, and J.P. Ioannidis, *Replication in genome-wide association studies.* Statistical Science: A review journal of the Institute of Mathematical Statistics, 2009. **24**(4): p. 561.
33. Visscher, P.M., et al., *10 years of GWAS discovery: biology, function, and translation.* The American Journal of Human Genetics, 2017. **101**(1): p. 5-22.
34. Low, S.-K., et al., *Genome-wide association study: a useful tool to identify common genetic variants associated with drug toxicity and efficacy in cancer pharmacogenomics*. 2014, AACR.
35. Curtin, S.J., et al., *Validating Genome-Wide Association candidates through quantitative variation in nodulation.* Plant physiology, 2017: p. pp. 01923.2016.
36. Igl, B.-W., I.R. König, and A. Ziegler, *What do we mean by 'replication'and 'validation'in genome-wide association studies?* Human heredity, 2009. **67**(1): p. 66-68.
37. Barnett, G.C., et al., *A genome wide association study (GWAS) providing evidence of an association between common genetic variants and late radiotherapy toxicity.* Radiotherapy and Oncology, 2014. **111**(2): p. 178-185.
38. Chamberlain, M., et al., *Unraveling a complex genetic disease: age-related macular degeneration.* Survey of ophthalmology, 2006. **51**(6): p. 576-586.
39. Roukos, D.H., *Genome-wide association studies: how predictable is a person's cancer risk?* Expert review of anticancer therapy, 2009. **9**(4): p. 389-392.
40. Baker, M., *Genome studies: genetics by numbers.* Nature News, 2008. **451**(7178): p. 516-518.
41. Salloum, N.C., et al., *Towards the clinical implementation of pharmacogenetics in bipolar disorder.* BMC medicine, 2014. **12**(1): p. 90.
42. MacArthur, J., et al., *The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog).* Nucleic acids research, 2016. **45**(D1): p. D896-D901.
43. Heit, J.A., F.A. Spencer, and R.H. White, *The epidemiology of venous thromboembolism.* Journal of thrombosis and thrombolysis, 2016. **41**(1): p. 3-14.
44. Raskob, G.E., et al., *Thrombosis: a major contributor to global disease burden.* Arteriosclerosis, thrombosis, and vascular biology, 2014. **34**(11): p. 2363-2371.
45. White, R.H., *The epidemiology of venous thromboembolism.* Circulation, 2003. **107**(23 suppl 1): p. I-4-I-8.
46. Alotaibi, G.S., et al., *Secular trends in incidence and mortality of acute venous thromboembolism: the AB-VTE population-based study.* The American journal of medicine, 2016. **129**(8): p. 879. e19-879. e25.
47. Walker, A.J., et al., *Incidence of venous thromboembolism in patients with cancer–a cohort study using linked United Kingdom databases.* European journal of cancer, 2013. **49**(6): p. 1404-1413.
48. Beckman, M.G., et al., *Venous thromboembolism: a public health concern.* American journal of preventive medicine, 2010. **38**(4): p. S495-S501.

49. Paydar, S., et al., *Management of deep vein thrombosis (DVT) prophylaxis in trauma patients.* Bulletin of Emergency & Trauma, 2016. **4**(1): p. 1.
50. Van Gent, J.-M., et al., *Pulmonary embolism without deep venous thrombosis: de novo or missed deep venous thrombosis?* Journal of Trauma and Acute Care Surgery, 2014. **76**(5): p. 1270-1274.
51. Rali, P., M. Rali, and M. Sockrider, *Pulmonary Embolism Part 1.* American journal of respiratory and critical care medicine, 2018. **197**(9): p. P15-P16.
52. Virchow, I., *Thrombose and Embolie. Gefassentzündung und septische Infection.* Gesammalte Abhandlungen zur Wissentschaftlichen Medizin Meidinger Sohn, Berlin, 1856: p. 220-227.
53. Bagot, C.N. and R. Arya, *Virchow and his triad: a question of attribution.* British journal of haematology, 2008. **143**(2): p. 180-190.
54. Versteeg, H.H., et al., *New fundamentals in hemostasis.* Physiological reviews, 2013. **93**(1): p. 327-358.
55. Jobling, L. and L. Eyre, *Haemostasis, blood platelets and coagulation.* Anaesthesia & Intensive Care Medicine, 2013. **14**(2): p. 51-53.
56. Golebiewska, E.M. and A.W. Poole, *Platelet secretion: From haemostasis to wound healing and beyond.* Blood reviews, 2015. **29**(3): p. 153-162.
57. Davì, G. and C. Patrono, *Platelet activation and atherothrombosis.* New England Journal of Medicine, 2007. **357**(24): p. 2482-2494.
58. Eyre, L. and F. Gamlin, *Haemostasis, blood platelets and coagulation.* Anaesthesia & Intensive Care Medicine, 2010. **11**(6): p. 244-246.
59. Smith, S.A., R.J. Travers, and J.H. Morrissey, *How it all starts: initiation of the clotting cascade.* Critical reviews in biochemistry and molecular biology, 2015. **50**(4): p. 326-336.
60. Davie, E.W., K. Fujikawa, and W. Kisiel, *The coagulation cascade: initiation, maintenance, and regulation.* Biochemistry, 1991. **30**(43): p. 10363-10370.
61. Velnar, T., T. Bailey, and V. Smrkolj, *The wound healing process: an overview of the cellular and molecular mechanisms.* Journal of International Medical Research, 2009. **37**(5): p. 1528-1542.
62. Enoch, S. and D.J. Leaper, *Basic science of wound healing.* Surgery (Oxford), 2008. **26**(2): p. 31-37.
63. George Broughton, I., J.E. Janis, and C.E. Attinger, *The basic science of wound healing.* Plastic and reconstructive surgery, 2006. **117**(7S): p. 12S-34S.
64. Witte, M.B. and A. Barbul, *General principles of wound healing.* Surgical Clinics of North America, 1997. **77**(3): p. 509-528.
65. Inoue, T., et al., *Vascular inflammation and repair: implications for re-endothelialization, restenosis, and stent thrombosis.* JACC: Cardiovascular Interventions, 2011. **4**(10): p. 1057-1066.
66. Rasche, H., *Haemostasis and thrombosis: an overview.* European Heart Journal Supplements, 2001. **3**(suppl_Q): p. Q3-Q7.
67. Kremers, R., et al., *The balance of pro-and anticoagulant processes underlying thrombin generation.* Journal of Thrombosis and Haemostasis, 2015. **13**(3): p. 437-447.
68. Gale, A.J., *Continuing education course# 2: current understanding of hemostasis.* Toxicologic pathology, 2011. **39**(1): p. 273-280.
69. Minors, D.S., *Haemostasis, blood platelets and coagulation.* Anaesthesia & intensive care medicine, 2007. **8**(5): p. 214-216.
70. Reitsma, P.H., H.H. Versteeg, and S. Middeldorp, *Mechanistic view of risk factors for venous thromboembolism.* Arteriosclerosis, thrombosis, and vascular biology, 2012. **32**(3): p. 563-568.
71. Chapin, J.C. and K.A. Hajjar, *Fibrinolysis and the control of blood coagulation.* Blood reviews, 2015. **29**(1): p. 17-24.

72. Ezihe-Ejiofor, J.A. and N. Hutchinson, *Anticlotting mechanisms 1: physiology and pathology.* Continuing Education in Anaesthesia, Critical Care & Pain, 2013. **13**(3): p. 87-92.
73. Booth, N.A., *Fibrinolysis and thrombosis.* Best Practice & Research Clinical Haematology, 1999. **12**(3): p. 423-433.
74. Gross, P.L. and W.C. Aird. *The endothelium and thrombosis*. in *Seminars in thrombosis and hemostasis*. 2000. Copyright© 2000 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New ….
75. Middeldorp, S. and A. van Hylckama Vlieg, *Does thrombophilia testing help in the clinical management of patients?* British journal of haematology, 2008. **143**(3): p. 321-335.
76. Crowther, M.A. and J.G. Kelton, *Congenital thrombophilic states associated with venous thrombosis: a qualitative overview and proposed classification system.* Annals of Internal Medicine, 2003. **138**(2): p. 128-134.
77. Rosendaal, F., *Venous thrombosis: a multicausal disease.* The Lancet, 1999. **353**(9159): p. 1167-1173.
78. Meltzer, M.E., et al., *Synergistic effects of hypofibrinolysis and genetic and acquired risk factors on the risk of a first venous thrombosis.* PLoS medicine, 2008. **5**(5): p. e97.
79. Crous-Bou, M., L.B. Harrington, and C. Kabrhel. *Environmental and genetic risk factors associated with venous thromboembolism*. in *Seminars in thrombosis and hemostasis*. 2016. Thieme Medical Publishers.
80. Zöller, B., et al., *Autoimmune diseases and venous thromboembolism: a review of the literature.* American journal of cardiovascular disease, 2012. **2**(3): p. 171.
81. Cushman, M., *Inherited risk factors for venous thrombosis.* ASH Education Program Book, 2005. **2005**(1): p. 452-457.
82. Zöller, B., et al., *Familial transmission of venous thromboembolism: a cohort study of 80,214 Swedish adoptees linked to their biological and adoptive parents.* Circulation: Genomic and Precision Medicine, 2014: p. CIRCGENETICS. 113.000341.
83. Sørensen, H.T., et al., *Familial risk of venous thromboembolism: a nationwide cohort study.* Journal of thrombosis and haemostasis, 2011. **9**(2): p. 320-324.
84. Souto, J.C., et al., *Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study.* The American Journal of Human Genetics, 2000. **67**(6): p. 1452-1459.
85. Ariëns, R.A., et al., *Activation markers of coagulation and fibrinolysis in twins: heritability of the prethrombotic state.* The Lancet, 2002. **359**(9307): p. 667-671.
86. Larsen, T.B., et al., *Major genetic susceptibility for venous thromboembolism in men: a study of Danish twins.* Epidemiology, 2003: p. 328-332.
87. Heit, J.A., et al., *Familial segregation of venous thromboembolism.* Journal of thrombosis and haemostasis, 2004. **2**(5): p. 731-736.
88. Morange, P.-E., P. Suchon, and D.-A. Trégouët, *Genetics of venous thrombosis: update in 2015.* Thrombosis and haemostasis, 2015. **114**(05): p. 910-919.
89. Trégouët, D., et al., *Is there still room for additional common susceptibility alleles for venous thromboembolism?* Journal of Thrombosis and Haemostasis, 2016. **14**(9): p. 1798-1802.
90. Morange, P. and D. Tregouet, *Lessons from genome-wide association studies in venous thrombosis.* Journal of Thrombosis and Haemostasis, 2011. **9**: p. 258-264.
91. Egeberg, O., *Inherited antithrombin deficiency causing thrombophilia.* Thrombosis and Haemostasis, 1965. **13**(01): p. 516-530.
92. Morange, P.E. and D.A. Tregouet, *Deciphering the molecular basis of venous thromboembolism: where are we and where should we go?* British journal of haematology, 2010. **148**(4): p. 495-506.
93. Jick, H., et al., *Venous thromboembolic disease and ABO blood type: a cooperative study.* The Lancet, 1969. **293**(7594): p. 539-542.

94. Tirado, I., et al., *The ABO blood group genotype and factor VIII levels as independent risk factors for venous thromboembolism.* Thrombosis and haemostasis, 2005. **93**(03): p. 468-474.
95. Wu, O., et al., *ABO (H) blood groups and vascular disease: a systematic review and meta-analysis.* Journal of thrombosis and haemostasis, 2008. **6**(1): p. 62-69.
96. Griffin, J.H., et al., *Deficiency of protein C in congenital thrombotic disease.* The Journal of clinical investigation, 1981. **68**(5): p. 1370-1373.
97. Schwarz, H.P., et al., *Plasma protein S deficiency in familial thrombotic disease.* Blood, 1984. **64**(6): p. 1297-1300.
98. Dahlbäck, B., M. Carlsson, and P.J. Svensson, *Familial thrombophilia due to a previously unrecognized mechanism characterized by poor anticoagulant response to activated protein C: prediction of a cofactor to activated protein C.* Proceedings of the National Academy of Sciences, 1993. **90**(3): p. 1004-1008.
99. Bertina, R.M., et al., *Mutation in blood coagulation factor V associated with resistance to activated protein C.* Nature, 1994. **369**(6475): p. 64.
100. Seligsohn, U. and A. Lubetsky, *Genetic susceptibility to venous thrombosis.* New England Journal of Medicine, 2001. **344**(16): p. 1222-1231.
101. Poort, S.R., et al., *A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis.* Blood, 1996. **88**(10): p. 3698-3703.
102. Saposnik, B., et al., *A haplotype of the EPCR gene is associated with increased plasma levels of sEPCR and is a candidate risk factor for thrombosis.* Blood, 2004. **103**(4): p. 1311-1318.
103. Qu, D., et al., *The Ser219–> Gly dimorphism of the endothelial protein C receptor contributes to the higher soluble protein levels observed in individuals with the A3 haplotype.* Journal of Thrombosis and Haemostasis, 2006. **4**(1): p. 229-235.
104. de Willige, S.U., et al., *Genetic variation in the fibrinogen gamma gene increases the risk for deep venous thrombosis by reducing plasma fibrinogen γ′ levels.* Blood, 2005. **106**(13): p. 4176-4183.
105. Grünbacher, G., et al., *The fibrinogen gamma (FGG) 10034C> T polymorphism is associated with venous thrombosis.* Thrombosis research, 2007. **121**(1): p. 33-36.
106. Trégouët, D.-A., et al., *Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach.* Blood, 2009. **113**(21): p. 5298-5303.
107. Smith, N.L., et al., *Association of genetic variations with nonfatal venous thrombosis in postmenopausal women.* Jama, 2007. **297**(5): p. 489-498.
108. Bezemer, I.D., et al., *Gene variants associated with deep vein thrombosis.* Jama, 2008. **299**(11): p. 1306-1314.
109. Snoep, J., et al., *The minor allele of GP6 T13254C is associated with decreased platelet activation and a reduced risk of recurrent cardiovascular events and mortality: results from the SMILE–Platelets project.* Journal of thrombosis and haemostasis, 2010. **8**(11): p. 2377-2384.
110. Tait, R., et al., *Prevalence of antithrombin deficiency in the healthy population.* British journal of haematology, 1994. **87**(1): p. 106-112.
111. Lane, D.A., et al., *Inherited Thrombophilia*: Part 2.* Thrombosis and haemostasis, 1996. **76**(06): p. 0824-0834.
112. Tait, R., et al., *Prevalence of protein C deficiency in the healthy population.* Thrombosis and haemostasis, 1995. **73**(01): p. 087-093.
113. Morange, P.E. and D.A. Trégouët, *Current knowledge on the genetics of incident venous thrombosis.* Journal of thrombosis and haemostasis, 2013. **11**: p. 111-121.
114. Koster, T., et al., *Role of clotting factor VIII in effect of von Willebrand factor on occurrence of deep-vein thrombosis.* The Lancet, 1995. **345**(8943): p. 152-155.

115. Jenkins, P.V. and J.S. O'Donnell, *ABO blood group determines plasma von Willebrand factor levels: a biologic function after all?* Transfusion, 2006. **46**(10): p. 1836-1844.
116. Morelli, V., et al., *ABO blood group genotypes and the risk of venous thrombosis: effect of factor V Leiden.* Journal of Thrombosis and Haemostasis, 2005. **3**(1): p. 183-185.
117. Kujovich, J.L., *Factor V Leiden Thrombophilia*, in *GeneReviews®[Internet]*. 2018, University of Washington, Seattle.
118. Nicolaes, G.A. and B.r. Dahlbäck, *Factor V and thrombotic disease: description of a janus-faced protein.* Arteriosclerosis, thrombosis, and vascular biology, 2002. **22**(4): p. 530-538.
119. Kujovich, J.L., *Factor v Leiden thrombophilia.* Genetics in Medicine, 2011. **13**(1): p. 1.
120. Koster, T., et al., *Factor VII and fibrinogen levels as risk factors for venous thrombosis.* Thrombosis and haemostasis, 1994. **72**(06): p. 719-722.
121. Weisel, J.W., *Fibrinogen and fibrin*, in *Advances in protein chemistry*. 2005, Elsevier. p. 247-299.
122. Meijers, J.C., et al., *High levels of coagulation factor XI as a risk factor for venous thrombosis.* New England Journal of Medicine, 2000. **342**(10): p. 696-701.
123. Massberg, S., et al., *A crucial role of glycoprotein VI for platelet recruitment to the injured arterial wall in vivo.* Journal of Experimental Medicine, 2003. **197**(1): p. 41-49.
124. Li, Y., et al., *Genetic variants associated with deep vein thrombosis: the F11 locus.* Journal of thrombosis and haemostasis, 2009. **7**(11): p. 1802-1808.
125. Buil, A., et al., *C4BPB/C4BPA is a new susceptibility locus for venous thrombosis with unknown protein S independent mechanism: results from genome-wide association and gene expression analyses followed by case-control studies.* Blood, 2010: p. blood-2010-01-263038.
126. Heit, J.A., et al., *A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24. 2 and 9q.* Journal of thrombosis and haemostasis, 2012. **10**(8): p. 1521-1531.
127. Hinds, D.A., et al., *Genome-wide association analysis of self-reported events in 6135 individuals and 252 827 controls identifies 8 loci associated with thrombosis.* Human molecular genetics, 2016. **25**(9): p. 1867-1874.
128. Rühle, F., et al., *Rare genetic variants in SMAP1, B3GAT2, and RIMS1 contribute to pediatric venous thromboembolism.* Blood, 2017. **129**(6): p. 783-790.
129. Heit, J.A., et al., *Identification of unique venous thromboembolism-susceptibility variants in African-Americans.* Thrombosis and haemostasis, 2017. **117**(4): p. 758.
130. Germain, M., et al., *Meta-analysis of 65,734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism.* The American Journal of Human Genetics, 2015. **96**(4): p. 532-542.
131. Germain, M., et al., *Genetics of venous thrombosis: insights from a new genome wide association study.* PloS one, 2011. **6**(9): p. e25581.
132. Tang, W., et al., *A genome-wide association study for venous thromboembolism: The extended cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium.* Genetic epidemiology, 2013. **37**(5): p. 512-521.
133. Klarin, D., et al., *Genetic analysis of venous thromboembolism in UK biobank identifies the ZFPM2 locus and implicates obesity as a causal risk factor.* Circulation: Cardiovascular Genetics, 2017. **10**(2): p. e001643.
134. Greliche, N., et al., *A genome-wide search for common SNP x SNP interactions on the risk of venous thrombosis.* BMC medical genetics, 2013. **14**(1): p. 36.
135. Hernandez, W., et al., *Novel genetic predictors of venous thromboembolism risk in African Americans.* Blood, 2016: p. blood-2015-09-668525.
136. Crous-Bou, M., et al., *Interactions of established risk factors and a GWAS-based genetic risk score on the risk of venous thromboembolism.* Thrombosis and haemostasis, 2016. **116**(10): p. 705-713.

137. Manco, L., et al., *Venous thromboembolism risk associated with ABO, F11 and FGG loci.* Blood Coagulation & Fibrinolysis, 2018. **29**(6): p. 528-532.
138. Heit, J.A., et al., *Genetic variation within the anticoagulant, procoagulant, fibrinolytic and innate immunity pathways as risk factors for venous thromboembolism.* Journal of thrombosis and haemostasis, 2011. **9**(6): p. 1133-1142.
139. Bunimov, N., N. Fuller, and C.P. Hayward. *Genetic loci associated with platelet traits and platelet disorders*. in *Seminars in thrombosis and hemostasis*. 2013. Thieme Medical Publishers.
140. Huang, Y., et al., *Arf6 controls platelet spreading and clot retraction via integrin α IIb β 3 trafficking.* Blood, 2016: p. blood-2015-05-648550.
141. Kon, S., et al., *Smap1 deficiency perturbs receptor trafficking and predisposes mice to myelodysplasia.* The Journal of clinical investigation, 2013. **123**(3): p. 1123-1137.
142. Jenkins, P.V., et al., *Elevated factor VIII levels and risk of venous thrombosis.* British journal of haematology, 2012. **157**(6): p. 653-663.
143. Zakai, N. and L. McClure, *Racial differences in venous thromboembolism.* Journal of Thrombosis and Haemostasis, 2011. **9**(10): p. 1877-1882.
144. White, R.H. and C.R. Keenan, *Effects of race and ethnicity on the incidence of venous thromboembolism.* Thrombosis research, 2009. **123**: p. S11-S17.
145. Esmon, C.T., *Thrombomodulin as a model of molecular mechanisms that modulate protease specificity and function at the vessel surface.* The FASEB Journal, 1995. **9**(10): p. 946-955.
146. Reitsma, P. and F. Rosendaal, *Activation of innate immunity in patients with venous thrombosis: the Leiden Thrombophilia Study.* Journal of thrombosis and haemostasis, 2004. **2**(4): p. 619-622.
147. Germain, M., et al., *Caution in interpreting results from imputation analysis when linkage disequilibrium extends over a large distance: a case study on venous thrombosis.* PLoS One, 2012. **7**(6): p. e38538.
148. Gran, O.V., S.K. Brækkan, and J.-B. Hansen, *Prothrombotic genotypes and risk of venous thromboembolism in cancer.* Thrombosis research, 2018. **164**: p. S12-S18.
149. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA: a cancer journal for clinicians, 2018. **68**(6): p. 394-424.
150. Urruticoechea, A., et al., *Recent advances in cancer therapy: an overview.* Current pharmaceutical design, 2010. **16**(1): p. 3-10.
151. Bor, G., I.D. Mat Azmi, and A. Yaghmur, *Nanomedicines for cancer therapy: Current status, challenges and future prospects.* Therapeutic delivery, 2019. **10**(2): p. 113-132.
152. Blom, J.W., et al., *Malignancies, prothrombotic mutations, and the risk of venous thrombosis.* Jama, 2005. **293**(6): p. 715-722.
153. Ay, C., I. Pabinger, and A.T. Cohen, *Cancer-associated venous thromboembolism: burden, mechanisms, and management.* Thrombosis and haemostasis, 2017. **117**(02): p. 219-230.
154. Cronin-Fenton, D., et al., *Hospitalisation for venous thromboembolism in cancer patients and the general population: a population-based cohort study in Denmark, 1997–2006.* British journal of cancer, 2010. **103**(7): p. 947.
155. Heit, J.A., et al., *Risk factors for deep vein thrombosis and pulmonary embolism: a population-based case-control study.* Archives of internal medicine, 2000. **160**(6): p. 809-815.
156. Goldhaber, S.Z., *Risk factors for venous thromboembolism.* Journal of the American College of Cardiology, 2010. **56**(1): p. 1-7.
157. Falanga, A. and L. Russo, *Epidemiology, risk and outcomes of venous thromboembolism in cancer.* Hämostaseologie, 2012. **32**(02): p. 115-125.
158. Timp, J.F., et al., *Epidemiology of cancer-associated venous thrombosis.* Blood, 2013. **122**(10): p. 1712-1723.

159. Dolens, T.A.P.A., *Lectures on clinical medicine.* London: The New Sydenham Society, 1868. **5**: p. 281-331.
160. Agnelli, G. and M. Verso, *Management of venous thromboembolism in patients with cancer.* Journal of thrombosis and haemostasis, 2011. **9**: p. 316-324.
161. Chao, C.-H., H.-Y. Wang, and C.-H. Kao, *Occult cancer and thromboembolism: current epidemiology and its practical implications.* Polish archives of internal medicine, 2018. **128**(9): p. 539-544.
162. SØrensen, H.T., et al., *Prognosis of cancers associated with venous thromboembolism.* New England Journal of Medicine, 2000. **343**(25): p. 1846-1850.
163. Chew, H.K., et al., *Incidence of venous thromboembolism and its effect on survival among patients with common cancers.* Archives of internal medicine, 2006. **166**(4): p. 458-464.
164. Rodriguez, A.O., et al., *Venous thromboembolism in ovarian cancer.* Gynecologic oncology, 2007. **105**(3): p. 784-790.
165. Lee, J.-C., et al., *Venous thromboembolism in patients with pancreatic cancer: Incidence and effect on survival in east Asian ethnic groups*. 2013, American Society of Clinical Oncology.
166. Donnellan, E. and A.A. Khorana, *Cancer and venous thromboembolic disease: a review.* The oncologist, 2017: p. theoncologist. 2016-0214.
167. Khorana, A.A., *Venous thromboembolism and prognosis in cancer.* Thrombosis research, 2010. **125**(6): p. 490-493.
168. Palumbo, J.S. *Mechanisms linking tumor cell–associated procoagulant function to tumor dissemination*. in *Seminars in thrombosis and hemostasis*. 2008. © Thieme Medical Publishers.
169. Falanga, A., L. Russo, and C. Verzeroli, *Mechanisms of thrombosis in cancer.* Thrombosis research, 2013. **131**: p. S59-S62.
170. Rickles, F.R. and A. Falanga, *Molecular basis for the relationship between thrombosis and cancer.* Thrombosis research, 2001. **102**(6): p. V215-V224.
171. Magnus, N., et al., *Oncogenes and the coagulation system–forces that modulate dormant and aggressive states in cancer.* Thrombosis research, 2014. **133**: p. S1-S9.
172. Regina, S., et al., *Increased tissue factor expression is associated with reduced survival in non–small cell lung cancer and with mutations of TP53 and PTEN.* Clinical chemistry, 2009. **55**(10): p. 1834-1842.
173. Prandoni, P., A. Piccioli, and A. Girolami, *Cancer and venous thromboembolism: an overview.* Haematologica, 1999. **84**(5): p. 437-445.
174. Rickles, F.R., S. Patierno, and P.M. Fernandez, *Tissue factor, thrombin, and cancer.* Chest, 2003. **124**(3): p. 58S-68S.
175. Baron, J.A., et al., *Venous thromboembolism and cancer.* The Lancet, 1998. **351**(9109): p. 1077-1080.
176. Verso, M. and G. Agnelli, *Venous thromboembolism associated with long-term use of central venous catheters in cancer patients.* Journal of Clinical Oncology, 2003. **21**(19): p. 3665-3675.
177. Khorana, A.A. and G.C. Connolly, *Assessing risk of venous thromboembolism in the patient with cancer.* Journal of Clinical Oncology, 2009. **27**(29): p. 4839.
178. Falanga, A., M. Marchetti, and L. Russo, *The mechanisms of cancer-associated thrombosis.* Thrombosis research, 2015. **135**: p. S8-S11.
179. Nalluri, S.R., et al., *Risk of venous thromboembolism with the angiogenesis inhibitor bevacizumab in cancer patients: a meta-analysis.* Jama, 2008. **300**(19): p. 2277-2285.
180. Buller, H., et al., *Cancer and thrombosis: from molecular mechanisms to clinical presentations.* Journal of Thrombosis and Haemostasis, 2007. **5**: p. 246-254.
181. Swier, N. and H.H. Versteeg, *Reciprocal links between venous thromboembolism, coagulation factors and ovarian cancer progression.* Thrombosis research, 2017. **150**: p. 8-18.

182. Chaukiyal, P., et al., *Thromboprophylaxis in cancer patients with central venous catheters.* Thrombosis and haemostasis, 2008. **99**(01): p. 38-43.
183. Agnelli, G., et al., *Semuloparin for thromboprophylaxis in patients receiving chemotherapy for cancer.* New England Journal of Medicine, 2012. **366**(7): p. 601-609.
184. Lyman, G.H., et al., *American Society of Clinical Oncology guideline: recommendations for venous thromboembolism prophylaxis and treatment in patients with cancer.* Journal of clinical oncology, 2007. **25**(34): p. 5490-5505.
185. Landefeld, C.S. and R.J. Beyth, *Anticoagulant-related bleeding: clinical epidemiology, prediction, and prevention.* The American journal of medicine, 1993. **95**(3): p. 315-328.
186. Kahn, S.R., et al., *Prevention of VTE in nonsurgical patients: antithrombotic therapy and prevention of thrombosis: American College of Chest Physicians evidence-based clinical practice guidelines.* Chest, 2012. **141**(2): p. e195S-e226S.
187. Di Nisio, M., et al., *Primary prophylaxis for venous thromboembolism in ambulatory cancer patients receiving chemotherapy.* The Cochrane Library, 2016.
188. Hachey, K.J., et al., *Caprini venous thromboembolism risk assessment permits selection for postdischarge prophylactic anticoagulation in patients with resectable lung cancer.* The Journal of thoracic and cardiovascular surgery, 2016. **151**(1): p. 37-44. e1.
189. Ferroni, P., et al., *Risk assessment for venous thromboembolism in chemotherapy-treated ambulatory cancer patients: a machine learning approach.* Medical Decision Making, 2017. **37**(2): p. 234-242.
190. Khorana, A.A., et al., *Development and validation of a predictive model for chemotherapy-associated thrombosis.* Blood, 2008. **111**(10): p. 4902-4907.
191. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA: a cancer journal for clinicians, 2018.
192. Ries, L., et al., *SEER Cancer Statistics Review, 1975–2001. 2004, Bethesda, MD: National Cancer Institute*. 2004, Tables XVI-1-9 Google Scholar.
193. Doubeni, C.A., et al., *Diagnosis and Management of Ovarian Cancer.* American family physician, 2016. **93**(11).
194. Mathieu, K.B., et al., *Screening for ovarian cancer: imaging challenges and opportunities for improvement.* Ultrasound in Obstetrics & Gynecology, 2018. **51**(3): p. 293-303.
195. Jayson, G.C., et al., *Ovarian cancer.* The Lancet, 2014. **384**(9951): p. 1376-1388.
196. McCluggage, W.G., *Morphological subtypes of ovarian carcinoma: a review with emphasis on new developments and pathogenesis.* Pathology, 2011. **43**(5): p. 420-432.
197. Prat, J., *Ovarian carcinomas: five distinct diseases with different origins, genetic alterations, and clinicopathological features.* Virchows Archiv, 2012. **460**(3): p. 237-249.
198. Chen, Y. and H. Du, *The promising PARP inhibitors in ovarian cancer therapy: From Olaparib to others.* Biomedicine & Pharmacotherapy, 2018. **99**: p. 552-560.
199. Network, C.G.A.R., *Integrated genomic analyses of ovarian carcinoma.* Nature, 2011. **474**(7353): p. 609.
200. Pepa, C., et al., *Ovarian cancer standard of care: are there real alternatives?* Chinese journal of cancer, 2015. **34**(1): p. 17.
201. Papa, A., et al., *Update on Poly-ADP-ribose polymerase inhibition for ovarian cancer treatment.* Journal of translational medicine, 2016. **14**(1): p. 267.
202. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2016.* CA: A Cancer Journal for Clinicians, 2016. **66**(1): p. 7-30.
203. Coyne, G.O.S., R. Piekarz, and A.P. Chen, *New Treatment Options for Ovarian Cancer*, in *Oncogenomics*. 2019, Elsevier. p. 533-540.
204. Pereira, D., et al., *Improvement of a predictive model in ovarian cancer patients submitted to platinum-based chemotherapy: implications of a GST activity profile.* European journal of clinical pharmacology, 2016. **72**(5): p. 545-553.

205. Pinto, R., et al., *Pharmacogenomics in epithelial ovarian cancer first-line treatment outcome: validation of GWAS-associated NRG3 rs1649942 and BRE rs7572644 variants in an independent cohort.* The pharmacogenomics journal, 2019. **19**(1): p. 25.
206. Tateo, S., et al., *Ovarian cancer and venous thromboembolic risk.* Gynecologic oncology, 2005. **99**(1): p. 119-125.
207. Saadeh, F.A., et al., *Venous thromboembolism in ovarian cancer: incidence, risk factors and impact on survival.* European Journal of Obstetrics & Gynecology and Reproductive Biology, 2013. **170**(1): p. 214-218.
208. Stålberg, K., et al., *The influence of comorbidity on mortality in ovarian cancer patients.* Gynecologic oncology, 2014. **133**(2): p. 298-303.
209. Metcalf, R., et al., *Thrombosis in ovarian cancer: a case control study.* British journal of cancer, 2014. **110**(5): p. 1118.
210. Peedicayil, A., et al., *Incidence and timing of venous thromboembolism after surgery for gynecological cancer.* Gynecologic oncology, 2011. **121**(1): p. 64-69.
211. Zangari, M., et al., *Thrombotic events in patients with cancer receiving antiangiogenesis agents.* Journal of Clinical Oncology, 2009. **27**(29): p. 4865-4873.
212. Chen, N., et al., *Bevacizumab promotes venous thromboembolism through the induction of PAI-1 in a mouse xenograft model of human lung carcinoma.* Molecular cancer, 2015. **14**(1): p. 140.
213. Huebner, B.R., et al., *Thrombin stimulates increased plasminogen activator inhibitor-1 release from liver compared to lung endothelium.* journal of surgical research, 2018. **225**: p. 1-5.
214. Collen, D. and H.R. Lijnen, *Basic and clinical aspects of fibrinolysis and thrombolysis.* Blood, 1991. **78**(12): p. 3114-3124.
215. Birkedal-Hansen, H., *Proteolytic remodeling of extracellular matrix.* Current opinion in cell biology, 1995. **7**(5): p. 728-735.
216. Bollen, L., et al., *Active PAI-1 as marker for venous thromboembolism: Case–control study using a comprehensive panel of PAI-1 and TAFI assays.* Thrombosis research, 2014. **134**(5): p. 1097-1102.
217. Pather, S. and M. Quinn, *Clear-cell cancer of the ovary—is it chemosensitive?* International journal of gynecological cancer, 2005. **15**(3): p. 432-437.
218. Satoh, T., et al., *High incidence of silent venous thromboembolism before treatment in ovarian cancer.* British journal of cancer, 2007. **97**(8): p. 1053.
219. Lee, A.Y., *Cancer and thromboembolic disease: pathogenic mechanisms.* Cancer treatment reviews, 2002. **28**(3): p. 137-140.
220. Uno, K., et al., *Tissue factor expression as a possible determinant of thromboembolism in ovarian cancer.* British journal of cancer, 2007. **96**(2): p. 290.
221. Yokota, N., et al., *Self-production of tissue factor-coagulation factor VII complex by ovarian cancer cells.* British journal of cancer, 2009. **101**(12): p. 2023.
222. Koizume, S., et al., *High-level secretion of tissue factor-rich extracellular vesicles from ovarian cancer cells mediated by filamin-A and protease-activated receptors.* Thrombosis and haemostasis, 2016. **116**(02): p. 299-310.
223. Koizume, S., et al., *Activation of cancer cell migration and invasion by ectopic synthesis of coagulation factor VII.* Cancer research, 2006. **66**(19): p. 9453-9460.
224. Rao, L.V.M., *Tissue factor as a tumor procoagulant.* Cancer and metastasis Reviews, 1992. **11**(3-4): p. 249-266.
225. Ueno, T., et al., *Tissue factor expression in breast cancer tissues: its correlation with prognosis and plasma concentration.* British journal of cancer, 2000. **83**(2): p. 164.
226. Koizume, S. and Y. Miyagi, *Tissue factor–factor VII complex as a key regulator of ovarian cancer phenotypes.* Biomarkers in cancer, 2015. **7**: p. BIC. S29318.
227. Grignani, G. and A. Maiolo, *Cytokines and hemostasis.* Haematologica, 2000. **85**(9): p. 967-972.

228. Solovey, A., et al., *Endothelial cell expression of tissue factor in sickle mice is augmented by hypoxia/reoxygenation and inhibited by lovastatin.* Blood, 2004. **104**(3): p. 840-846.
229. Joanne, L.Y., et al., *Oncogenic events regulate tissue factor expression in colorectal cancer cells: implications for tumor progression and angiogenesis.* Blood, 2005. **105**(4): p. 1734-1741.
230. Rong, Y., et al., *Epidermal growth factor receptor and PTEN modulate tissue factor expression in glioblastoma through JunD/activator protein-1 transcriptional activity.* Cancer research, 2009. **69**(6): p. 2540-2549.
231. Stojkovic, S., et al., *Tissue factor is induced by interleukin-33 in human endothelial cells: a new link between coagulation and inflammation.* Scientific reports, 2016. **6**: p. 25171.
232. van den Berg, Y.W., et al., *The relationship between tissue factor and cancer progression: insights from bench and bedside.* Blood, 2012. **119**(4): p. 924-932.
233. Kasthuri, R.S., M.B. Taubman, and N. Mackman, *Role of tissue factor in cancer.* Journal of Clinical Oncology, 2009. **27**(29): p. 4834.
234. Ruf, W., N. Yokota, and F. Schaffner, *Tissue factor in cancer progression and angiogenesis.* Thrombosis research, 2010. **125**: p. S36-S38.
235. Ünlü, B. and H.H. Versteeg, *Effects of tumor-expressed coagulation factors on cancer progression and venous thrombosis: is there a key factor?* Thrombosis research, 2014. **133**: p. S76-S84.
236. Belting, M., et al., *Regulation of angiogenesis by tissue factor cytoplasmic domain signaling.* Nature medicine, 2004. **10**(5): p. 502.
237. Zhong, Y.-C., et al., *Thrombin promotes epithelial ovarian cancer cell invasion by inducing epithelial-mesenchymal transition.* Journal of gynecologic oncology, 2013. **24**(3): p. 265-272.
238. Lamouille, S., J. Xu, and R. Derynck, *Molecular mechanisms of epithelial–mesenchymal transition.* Nature reviews Molecular cell biology, 2014. **15**(3): p. 178.
239. Collino, F., et al., *Epithelial–mesenchymal transition of ovarian tumor cells induces an angiogenic monocyte cell population.* Experimental cell research, 2009. **315**(17): p. 2982-2994.
240. Kajiyama, H., et al., *Chemoresistance to paclitaxel induces epithelial-mesenchymal transition and enhances metastatic potential for epithelial ovarian carcinoma cells.* International journal of oncology, 2007. **31**(2): p. 277-283.
241. Kurrey, N.K., et al., *Snail and slug mediate radioresistance and chemoresistance by antagonizing p53-mediated apoptosis and acquiring a stem-like phenotype in ovarian cancer cells.* Stem cells, 2009. **27**(9): p. 2059-2068.
242. Gay, L.J. and B. Felding-Habermann, *Contribution of platelets to tumour metastasis.* Nature Reviews Cancer, 2011. **11**(2): p. 123.
243. Lal, I., K. Dittus, and C.E. Holmes, *Platelets, coagulation and fibrinolysis in breast cancer progression.* Breast Cancer Research, 2013. **15**(4): p. 207.
244. Chu, A.J., *Tissue factor mediates inflammation.* Archives of biochemistry and biophysics, 2005. **440**(2): p. 123-132.
245. Levi, M., T. van der Poll, and H. ten Cate. *Tissue factor in infection and severe inflammation*. in *Seminars in thrombosis and hemostasis*. 2006. Copyright© 2006 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New ….
246. Grivennikov, S.I., F.R. Greten, and M. Karin, *Immunity, inflammation, and cancer.* Cell, 2010. **140**(6): p. 883-899.
247. Han, L.Y., et al., *Preoperative serum tissue factor levels are an independent prognostic factor in patients with ovarian carcinoma.* Journal of clinical oncology, 2006. **24**(5): p. 755-761.
248. Perrier, A., et al., *D-dimer testing for suspected pulmonary embolism in outpatients.* American journal of respiratory and critical care medicine, 1997. **156**(2): p. 492-496.
249. Mahmmod, K.A., *Evaluation of D-Dimer in the diagnosis of suspected deep vein thrombosis.* Al-Kindy College Medical Journal, 2014. **10**(2): p. 29-31.

250. Amirkhosravi, A., et al., *Blood clotting activation analysis for preoperative differentiation of benign versus malignant ovarian masses.* Blood Coagulation & Fibrinolysis, 2013. **24**(5): p. 510-517.
251. Sakurai, M., et al., *High pretreatment plasma D-dimer levels are associated with poor prognosis in patients with ovarian cancer independently of venous thromboembolism and tumor extension.* International Journal of Gynecological Cancer, 2015. **25**(4): p. 593.
252. Liu, P., et al., *Elevated preoperative plasma D-dimer level is a useful predictor of chemoresistance and poor disease outcome for serous ovarian cancer patients.* Cancer chemotherapy and pharmacology, 2015. **76**(6): p. 1163-1171.
253. Ducros, E., et al., *Endothelial protein C receptor expressed by ovarian cancer cells as a possible biomarker of cancer onset.* International journal of oncology, 2012. **41**(2): p. 433-440.
254. Mosnier, L.O., B.V. Zlokovic, and J.H. Griffin, *The cytoprotective protein C pathway.* Blood, 2007. **109**(8): p. 3161-3172.
255. Bouwens, E., F. Stavenuiter, and L. Mosnier, *Mechanisms of anticoagulant and cytoprotective actions of the protein C pathway.* Journal of Thrombosis and Haemostasis, 2013. **11**: p. 242-253.
256. Uchiba, M., et al., *Activated protein C induces endothelial cell proliferation by mitogen-activated protein kinase activation in vitro and angiogenesis in vivo.* Circulation Research, 2004. **95**(1): p. 34-41.
257. Kobayashi, H., et al., *Role of activated protein C in facilitating basement membrane invasion by tumor cells.* Cancer research, 1994. **54**(1): p. 261-267.
258. Suzuki, K. and T. Hayashi. *Protein C and its inhibitor in malignancy*. in *Seminars in thrombosis and hemostasis*. 2007. © Thieme Medical Publishers.
259. Azzazene, D., et al., *Plasma endothelial protein C receptor influences innate immune response in ovarian cancer by decreasing the population of natural killer and TH17 helper cells.* International journal of oncology, 2013. **43**(4): p. 1011-1018.
260. Kallel, C., et al., *Association of soluble endothelial protein C receptor plasma levels and PROCR rs867186 with cardiovascular risk factors and cardiovascular events in coronary artery disease patients: the Athero Gene study.* BMC medical genetics, 2012. **13**(1): p. 103.
261. Prat, J. and F.C.o.G. Oncology, *FIGO's staging classification for cancer of the ovary, fallopian tube, and peritoneum: abridged republication.* Journal of gynecologic oncology, 2015. **26**(2): p. 87-89.
262. Rustin, G.J.S., et al., *Definitions for response and progression in ovarian cancer clinical trials incorporating RECIST 1.1 and CA 125 agreed by the Gynecological Cancer Intergroup (GCIG).* International Journal of Gynecologic Cancer, 2011. **21**(2): p. 419-423.
263. Huang, D., et al., *GWAS4D: multidimensional analysis of context-specific regulatory variant for human complex diseases and traits.* Nucleic acids research, 2018. **46**(W1): p. W114-W120.
264. Eisenhauer, E.A., et al., *New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1).* European journal of cancer, 2009. **45**(2): p. 228-247.
265. Langer, F. and C. Bokemeyer, *Crosstalk between cancer and haemostasis.* Hämostaseologie, 2012. **32**(02): p. 95-104.
266. Schaffner, F., N. Yokota, and W. Ruf, *Tissue factor proangiogenic signaling in cancer progression.* Thrombosis research, 2012. **129**: p. S127-S131.
267. Zerbino, D.R., et al., *Ensembl 2018.* Nucleic acids research, 2017. **46**(D1): p. D754-D761.
268. Stelzer, G., et al., *The GeneCards suite: from gene data mining to disease genome sequence analyses.* Current protocols in bioinformatics, 2016. **54**(1): p. 1.30. 1-1.30. 33.
269. Orkin, S.H., *GATA-binding transcription factors in hematopoietic cells.* Blood, 1992. **80**(3): p. 575-581.
270. Evans, T., *Regulation of cardiac gene expression by GATA-4/5/6.* Trends in cardiovascular medicine, 1997. **7**(3): p. 75-83.

271. Zhang, Y., et al., *Hematopoietic hierarchy–an updated roadmap.* Trends in cell biology, 2018. **28**(12): p. 976-986.
272. Molkentin, J.D., *The zinc finger-containing transcription factors GATA-4,-5, and-6 ubiquitously expressed regulators of tissue-specific gene expression.* Journal of Biological Chemistry, 2000. **275**(50): p. 38949-38952.
273. Capo-Chichi, C.D., et al., *Anomalous expression of epithelial differentiation-determining GATA factors in ovarian tumorigenesis.* Cancer research, 2003. **63**(16): p. 4967-4977.
274. Cai, K.Q., et al., *Loss of GATA4 and GATA6 expression specifies ovarian cancer histological subtypes and precedes neoplastic transformation of ovarian surface epithelia.* PloS one, 2009. **4**(7): p. e6454.
275. Crispino, J.D., et al., *Proper coronary vascular development and heart morphogenesis depend on interaction of GATA-4 with FOG cofactors.* Genes & development, 2001. **15**(7): p. 839-844.
276. Lu, J.-r., et al., *FOG-2, a heart-and brain-enriched cofactor for GATA transcription factors.* Molecular and cellular biology, 1999. **19**(6): p. 4495-4502.
277. Tevosian, S.G., et al., *FOG-2: A novel GATA-family cofactor related to multitype zinc-finger proteins Friend of GATA-1 and U-shaped.* Proceedings of the National Academy of Sciences, 1999. **96**(3): p. 950-955.
278. Holmes, M., et al., *hFOG-2, a novel zinc finger protein, binds the co-repressor mCtBP2 and modulates GATA-mediated activation.* Journal of Biological Chemistry, 1999. **274**(33): p. 23491-23498.
279. Heineke, J., et al., *Cardiomyocyte GATA4 functions as a stress-responsive regulator of angiogenesis in the murine heart.* The Journal of clinical investigation, 2007. **117**(11): p. 3198-3210.
280. Fouad, Y.A. and C. Aanei, *Revisiting the hallmarks of cancer.* American journal of cancer research, 2017. **7**(5): p. 1016.
281. Hoene, V., et al., *GATA factors in human neuroblastoma: distinctive expression patterns in clinical subtypes.* British journal of cancer, 2009. **101**(8): p. 1481.
282. Clugston, R.D., W. Zhang, and J.J. Greer, *Gene expression in the developing diaphragm: significance for congenital diaphragmatic hernia.* American Journal of Physiology-Lung Cellular and Molecular Physiology, 2008. **294**(4): p. L665-L675.
283. Hyun, S., et al., *Conserved MicroRNA miR-8/miR-200 and its target USH/FOG2 control growth by regulating PI3K.* Cell, 2009. **139**(6): p. 1096-1108.
284. Manuylov, N., F. Smagulova, and S. Tevosian, *Fog2 excision in mice leads to premature mammary gland involution and reduced Esr1 gene expression.* Oncogene, 2007. **26**(36): p. 5204.
285. Bader, A.G., et al., *Oncogenic PI3K deregulates transcription and translation.* Nature Reviews Cancer, 2005. **5**(12): p. 921.
286. Porta, C., C. Paglino, and A. Mosca, *Targeting PI3K/Akt/mTOR signaling in cancer.* Frontiers in oncology, 2014. **4**: p. 64.
287. Mabuchi, S., et al., *The PI3K/AKT/mTOR pathway as a therapeutic target in ovarian cancer.* Gynecologic oncology, 2015. **137**(1): p. 173-179.
288. Klein, D., et al., *MicroRNA expression in alpha and beta cells of human pancreatic islets.* PloS one, 2013. **8**(1): p. e55064.
289. Karar, J. and A. Maity, *PI3K/AKT/mTOR pathway in angiogenesis.* Frontiers in molecular neuroscience, 2011. **4**: p. 51.
290. Debette, S., et al., *Identification of cis-and trans-acting genetic variants explaining up to half the variation in circulating vascular endothelial growth factor levels.* Circulation research, 2011. **109**(5): p. 554-563.
291. Sun, B.B., et al., *Genomic atlas of the human plasma proteome.* Nature, 2018. **558**(7708): p. 73.

292. Xiong, H.Y., et al., *The human splicing code reveals new insights into the genetic determinants of disease.* Science, 2015. **347**(6218): p. 1254806.
293. Pagenstecher, C., et al., *Aberrant splicing in MLH1 and MSH2 due to exonic and intronic variants.* Human genetics, 2006. **119**(1-2): p. 9-22.
294. Nikpay, M., et al., *Genome-wide identification of circulating-miRNA expression quantitative trait loci reveals the role of several miRNAs in the regulation of cardiometabolic phenotypes.* Cardiovascular research, 2019.
295. Miles, G.D., et al., *Identifying microRNA/mRNA dysregulations in ovarian cancer.* BMC research notes, 2012. **5**(1): p. 164.
296. Zhang, S., et al., *Clinically relevant microRNAs in ovarian cancer.* Molecular Cancer Research, 2015. **13**(3): p. 393-401.
297. Lengyel, E., *Ovarian cancer development and metastasis.* The American journal of pathology, 2010. **177**(3): p. 1053-1064.
298. Tania, M., M.A. Khan, and J. Fu, *Epithelial to mesenchymal transition inducing transcription factors and metastatic cancer.* Tumor Biology, 2014. **35**(8): p. 7335-7342.
299. Yeung, T.-L., et al., *Cellular and molecular processes in ovarian cancer metastasis. A review in the theme: cell and molecular processes in cancer metastasis.* American Journal of Physiology-Cell Physiology, 2015. **309**(7): p. C444-C456.
300. Amr, K., et al., *Whole exome sequencing identifies a new mutation in the SLC19A2 gene leading to thiamine-responsive megaloblastic anemia in an Egyptian family.* Molecular genetics & genomic medicine, 2019: p. e777.
301. Zastre, J.A., et al., *Linking vitamin B1 with cancer cell metabolism.* Cancer & metabolism, 2013. **1**(1): p. 16.
302. LU'O'NG, K.V.Q. and L.T.H. NGUYỄN, *The role of thiamine in cancer: possible genetic and cellular signaling mechanisms.* Cancer Genomics-Proteomics, 2013. **10**(4): p. 169-185.
303. Frank, R., F. Leeper, and B. Luisi, *Structure, mechanism and catalytic duality of thiamine-dependent enzymes.* Cellular and molecular life sciences, 2007. **64**(7-8): p. 892.
304. Krockenberger, M., et al., *Transketolase-like 1 expression correlates with subtypes of ovarian cancer and the presence of distant metastases.* International Journal of Gynecologic Cancer, 2007. **17**(1): p. 101-106.
305. Schmidt, M., et al., *Glucose metabolism and angiogenesis in granulosa cell tumors of the ovary: activation of Akt, expression of M2PK, TKTL1 and VEGF.* European Journal of Obstetrics & Gynecology and Reproductive Biology, 2008. **139**(1): p. 72-78.
306. Xu, X., et al., *Transketolase-like protein 1 (TKTL1) is required for rapid cell growth and full viability of human tumor cells.* International journal of cancer, 2009. **124**(6): p. 1330-1337.
307. McLure, K.G., M. Takagi, and M.B. Kastan, *NAD+ modulates p53 DNA binding specificity and function.* Molecular and cellular biology, 2004. **24**(22): p. 9958-9967.
308. Yang, Z., et al., *The expression of p53, MDM2 and Ref1 gene in cultured retina neurons of SD rats treated with vitamin B1 and/or elevated pressure.* Yan ke xue bao (2016), 2004. **20**(4): p. 259-263.
309. Shin, B.H., et al., *Thiamine attenuates hypoxia-induced cell death in cultured neonatal rat cardiomyocytes.* Mol Cells, 2004. **18**(2): p. 133-140.
310. Zastre, J.A., et al., *Up-regulation of vitamin B1 homeostasis genes in breast cancer.* The Journal of nutritional biochemistry, 2013. **24**(9): p. 1616-1624.
311. Gaunt, T.R., et al., *A gene-centric analysis of activated partial thromboplastin time and activated protein C resistance using the HumanCVD focused genotyping array.* European journal of human genetics, 2013. **21**(7): p. 779.
312. Wu, H., et al., *A novel functional TagSNP Rs7560488 in the DNMT3A1 promoter is associated with susceptibility to gastric cancer by modulating promoter activity.* PLoS One, 2014. **9**(3): p. e92911.
313. Schirmer, M.A., et al., *Relevance of Sp binding site polymorphism in WWOX for treatment outcome in pancreatic cancer.* JNCI: Journal of the National Cancer Institute, 2016. **108**(5).

314. Rintisch, C., et al., *Natural variation of histone modification and its impact on gene expression in the rat genome.* Genome research, 2014. **24**(6): p. 942-953.
315. Zhang, S., et al., *A variant in the CHEK2 promoter at a methylation site relieves transcriptional repression and confers reduced risk of lung cancer.* Carcinogenesis, 2010. **31**(7): p. 1251-1258.
316. Pavlova, N.N. and C.B. Thompson, *The emerging hallmarks of cancer metabolism.* Cell metabolism, 2016. **23**(1): p. 27-47.
317. Sotgia, F., et al., *Caveolin-1 and cancer metabolism in the tumor microenvironment: markers, models, and mechanisms.* Annual Review of Pathology: Mechanisms of Disease, 2012. **7**: p. 423-467.
318. Nakajima, E.C. and B. Van Houten, *Metabolic symbiosis in cancer: refocusing the Warburg lens.* Molecular carcinogenesis, 2013. **52**(5): p. 329-337.
319. Nieman, K.M., et al., *Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth.* Nature medicine, 2011. **17**(11): p. 1498.
320. He, H., et al., *Multiple functional variants in long-range enhancer elements contribute to the risk of SNP rs965513 in thyroid cancer.* Proceedings of the National Academy of Sciences, 2015. **112**(19): p. 6128-6133.
321. Quillard, T. and B. Charreau, *Impact of notch signaling on inflammatory responses in cardiovascular disorders.* International journal of molecular sciences, 2013. **14**(4): p. 6863-6888.
322. Mu, D., et al., *Cntn6 deficiency impairs allocentric navigation in mice.* Brain and behavior, 2018. **8**(6): p. e00969.
323. Cui, X.-Y., et al., *NB-3/Notch1 pathway via Deltex1 promotes neural progenitor cell differentiation into oligodendrocytes.* Journal of Biological Chemistry, 2004. **279**(24): p. 25858-25865.
324. Wong, C.-M., et al., *Adropin is a brain membrane-bound protein regulating physical activity via the NB-3/Notch signaling pathway in mice.* Journal of Biological Chemistry, 2014. **289**(37): p. 25976-25986.
325. Rose, S.L., et al., *Notch 1 signaling is active in ovarian cancer.* Gynecologic oncology, 2010. **117**(1): p. 130-133.
326. Rose, S.L., *Notch signaling pathway in ovarian cancer.* International Journal of Gynecologic Cancer, 2009. **19**(4): p. 564-566-564-566.
327. McAuliffe, S.M., et al., *Targeting Notch, a key pathway for ovarian cancer stem cells, sensitizes tumors to platinum therapy.* Proceedings of the National Academy of Sciences, 2012. **109**(43): p. E2939-E2948.
328. Espinoza, I. and L. Miele, *Deadly crosstalk: Notch signaling at the intersection of EMT and cancer stem cells.* Cancer letters, 2013. **341**(1): p. 41-45.
329. Gupta, N., et al., *Notch3 induces epithelial–mesenchymal transition and attenuates carboplatin-induced apoptosis in ovarian cancer cells.* Gynecologic oncology, 2013. **130**(1): p. 200-206.
330. Ahmed, N., et al., *Epithelial mesenchymal transition and cancer stem cell-like phenotypes facilitate chemoresistance in recurrent ovarian cancer.* Current cancer drug targets, 2010. **10**(3): p. 268-278.
331. Groeneweg, J.W., et al., *Notch signaling in serous ovarian cancer.* Journal of ovarian research, 2014. **7**(1): p. 95.
332. Chiaramonte, R., et al., *Notch pathway promotes ovarian cancer growth and migration via CXCR4/SDF1α chemokine system.* The international journal of biochemistry & cell biology, 2015. **66**: p. 134-140.
333. Ranganathan, P., K.L. Weaver, and A.J. Capobianco, *Notch signalling in solid tumours: a little bit of everything but not all the time.* Nature Reviews Cancer, 2011. **11**(5): p. 338.
334. Coordinators, N.R., *Database resources of the national center for biotechnology information.* Nucleic acids research, 2018. **46**(Database issue): p. D8.

335. Enesa, K., et al., *NF-κB suppression by the deubiquitinating enzyme cezanne a novel negative feedback loop in pro-inflammatory signaling.* Journal of Biological Chemistry, 2008. **283**(11): p. 7036-7045.
336. Xu, Z., et al., *Snail1-dependent transcriptional repression of Cezanne2 in hepatocellular carcinoma.* Oncogene, 2014. **33**(22): p. 2836.
337. EVANS, P.C., et al., *Isolation and characterization of two novel A20-like proteins.* Biochemical Journal, 2001. **357**(3): p. 617-623.
338. Malaver, E., et al., *NF-κB inhibitors impair platelet activation responses.* Journal of Thrombosis and Haemostasis, 2009. **7**(8): p. 1333-1343.
339. Seligsohn, U., *Factor XI in haemostasis and thrombosis: past, present and future.* Thrombosis and haemostasis, 2007. **98**(07): p. 84-89.
340. Handin, N., *Identification of new regulatory mechanisms that determine coagulation FXI plasma concentration*. 2015.
341. Sennblad, B., et al. *Novel mechanisms regulating Factor XI plasma levels*. in *Journal of Thrombosis and Haemostasis*. 2016. WILEY-BLACKWELL 111 RIVER ST, HOBOKEN 07030-5774, NJ USA.
342. Wang, E., et al., *Peritoneal and subperitoneal stroma may facilitate regional spread of ovarian cancer.* Clinical cancer research, 2005. **11**(1): p. 113-122.
343. Versteeg, H.H., et al., *Coagulation factors VIIa and Xa inhibit apoptosis and anoikis.* Oncogene, 2004. **23**(2): p. 410.
344. Von dem Borne, P., J. Meijers, and B. Bouma, *Feedback activation of factor XI by thrombin in plasma results in additional formation of thrombin that protects fibrin clots from fibrinolysis.* Blood, 1995. **86**(8): p. 3035-3042.
345. Emsley, J., P.A. McEwan, and D. Gailani, *Structure and function of factor XI.* Blood, 2010. **115**(13): p. 2569-2577.
346. Ahmad, R., et al., *Thrombin induces apoptosis in human tumor cells.* International journal of cancer, 2000. **87**(5): p. 707-715.
347. Schiller, H., et al., *Thrombin as a survival factor for cancer cells: thrombin activation in malignant effusions in vivo and inhibition of idarubicin-induced cell death in vitro.* International journal of clinical pharmacology and therapeutics, 2002. **40**(8): p. 329-335.
348. Brass, L.F., *Thrombin and platelet activation.* Chest, 2003. **124**(3): p. 18S-25S.
349. Tesfamariam, B., *Involvement of platelets in tumor cell metastasis.* Pharmacology & Therapeutics, 2016. **157**: p. 112-119.
350. Nierodzik, M.L. and S. Karpatkin, *Thrombin induces tumor growth, metastasis, and angiogenesis: Evidence for a thrombin-regulated dormant tumor phenotype.* Cancer cell, 2006. **10**(5): p. 355-362.
351. Van Hinsbergh, V.W., A. Collen, and P. Koolwijk, *Role of fibrin matrix in angiogenesis.* Annals of the New York Academy of Sciences, 2001. **936**(1): p. 426-437.
352. Hu, L., et al., *Role of endogenous thrombin in tumor implantation, seeding, and spontaneous metastasis.* Blood, 2004. **104**(9): p. 2746-2751.
353. Szlam, F., et al., *Elevated factor VIII enhances thrombin generation in the presence of factor VIII-deficiency, factor XI-deficiency or fondaparinux.* Thrombosis research, 2011. **127**(2): p. 135-140.
354. Beaulieu, L.M. and F.C. Church, *Activated protein C promotes breast cancer cell migration through interactions with EPCR and PAR-1.* Experimental cell research, 2007. **313**(4): p. 677-687.
355. Tsuneyoshi, N., et al., *Expression and anticoagulant function of the endothelial cell protein C receptor (EPCR) in cancer cell lines.* Thrombosis and haemostasis, 2001. **85**(02): p. 356-361.
356. Feistritzer, C., et al., *Endothelial protein C receptor-dependent inhibition of migration of human lymphocytes by protein C involves epidermal growth factor receptor.* The Journal of Immunology, 2006. **176**(2): p. 1019-1025.

357. Jackson, C.J., et al., *Activated protein C prevents inflammation yet stimulates angiogenesis to promote cutaneous wound healing.* Wound repair and regeneration, 2005. **13**(3): p. 284-294.
358. Cheng, T., et al., *Activated protein C blocks p53-mediated apoptosis in ischemic human brain endothelium and is neuroprotective.* Nature medicine, 2003. **9**(3): p. 338.
359. Guo, H., et al., *Activated protein C prevents neuronal apoptosis via protease activated receptors 1 and 3.* Neuron, 2004. **41**(4): p. 563-572.
360. Liu, D., et al., *Tissue plasminogen activator neurovascular toxicity is controlled by activated protein C.* Nature medicine, 2004. **10**(12): p. 1379.
361. Joyce, D.E., et al., *Gene expression profile of antithrombotic protein C defines new mechanisms modulating inflammation and apoptosis.* Journal of Biological Chemistry, 2001. **276**(14): p. 11199-11203.
362. Xue, M., et al., *Activated protein C stimulates proliferation, migration and wound closure, inhibits apoptosis and upregulates MMP-2 activity in cultured human keratinocytes.* Experimental cell research, 2004. **299**(1): p. 119-127.
363. Yan, Q., et al., *Prevalence of protein C receptor (PROCR) is associated with inferior clinical outcome in Breast invasive ductal carcinoma.* Pathology-Research and Practice, 2017. **213**(9): p. 1173-1179.
364. Wang, Q., et al., *EPCR promotes MGC803 human gastric cancer cell tumor angiogenesis in vitro through activating ERK1/2 and AKT in a PAR1-dependent manner.* Oncology letters, 2018. **16**(2): p. 1565-1570.
365. Wojtukiewicz, M., et al., *Endothelial protein C receptor (EPCR), protease activated receptor-1 (PAR-1) and their interplay in cancer growth and metastatic dissemination.* Cancers, 2019. **11**(1): p. 51.
366. Uitte de Willige, S., et al., *Haplotypes of the EPCR gene, plasma sEPCR levels and the risk of deep venous thrombosis.* Journal of Thrombosis and Haemostasis, 2004. **2**(8): p. 1305-1310.
367. SkIRNISDOTTIR, I., T. Seidal, and H. ÅkERUD, *The relationship of the angiogenesis regulators VEGF-A, VEGF-R1 and VEGF-R2 to p53 status and prognostic factors in epithelial ovarian carcinoma in FIGO-stages I-II.* International journal of oncology, 2016. **48**(3): p. 998-1006.
368. Preskill, C. and J.B. Weidhaas, *SNPs in microRNA binding sites as prognostic and predictive cancer biomarkers.* Critical Reviews™ in Oncogenesis, 2013. **18**(4).
369. Boyle, E.A., Y.I. Li, and J.K. Pritchard, *An expanded view of complex traits: from polygenic to omnigenic.* Cell, 2017. **169**(7): p. 1177-1186.
370. Zhu, Z., et al., *Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets.* Nature genetics, 2016. **48**(5): p. 481.
371. Machado, J., et al., *Alternative splicing regulation by GWAS risk loci for breast cancer.* bioRxiv, 2019: p. 766394.
372. Chen, R., et al., *Fine mapping the TAGAP risk locus in rheumatoid arthritis.* Genes and immunity, 2011. **12**(4): p. 314.
373. Cotsapas, C. and D.A. Hafler, *Immune-mediated disease genetics: the shared basis of pathogenesis.* Trends in immunology, 2013. **34**(1): p. 22-26.
374. Kim, J.-y., et al., *Antiangiogenic and anticancer effect of an orally active low molecular weight heparin conjugates and its application to lung cancer chemoprevention.* Journal of controlled release, 2015. **199**: p. 122-131.
375. Wang, X., et al., *Ovarian cancer, the coagulation pathway, and inflammation.* Journal of translational medicine, 2005. **3**(1): p. 25.
376. Jiang, M.-j., et al., *Aspirin in pancreatic cancer: chemopreventive effects and therapeutic potentials.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2016. **1866**(2): p. 163-176.
377. Rosendaal, F., et al., *High risk of thrombosis in patients homozygous for factor V Leiden (activated protein C resistance)[see comments].* Blood, 1995. **85**(6): p. 1504-1508.

378. Martinelli, I., et al., *Prothrombin A19911G polymorphism and the risk of venous thromboembolism.* Journal of Thrombosis and Haemostasis, 2006. **4**(12): p. 2582-2586.
379. Bruzelius, M., et al., *Predicting venous thrombosis in women using a combination of genetic markers and clinical risk factors.* Journal of Thrombosis and Haemostasis, 2015. **13**(2): p. 219-227.
380. McDaid, A., et al., *Risk prediction of developing venous thrombosis in combined oral contraceptive users.* PloS one, 2017. **12**(7): p. e0182041.
381. Monkovic, D.D. and P.B. Tracy, *Activation of human factor V by factor Xa and thrombin.* Biochemistry, 1990. **29**(5): p. 1118-1128.
382. Strand, S., T. Orntoft, and K. Sorensen, *Prognostic DNA methylation markers for prostate cancer.* International journal of molecular sciences, 2014. **15**(9): p. 16544-16576.
383. Haldrup, C., et al., *DNA methylation signatures for prediction of biochemical recurrence after radical prostatectomy of clinically localized prostate cancer.* Journal of Clinical Oncology, 2013. **31**(26): p. 3250-3258.
384. McKee, T.C. and J.V. Tricoli, *Epigenetics of prostate cancer*, in *Cancer Epigenetics*. 2015, Springer. p. 217-234.
385. Yang, R., et al., *An integrated model of clinical information and gene expression for prediction of survival in ovarian cancer patients.* Translational Research, 2016. **172**: p. 84-95. e11.
386. Gao, C., et al., *Exploration of methylation-driven genes for monitoring and prognosis of patients with lung adenocarcinoma.* Cancer cell international, 2018. **18**(1): p. 194.
387. Kiechl, S., et al., *Association of variation at the ABO locus with circulating levels of soluble intercellular adhesion molecule-1, soluble P-selectin, and soluble E-selectin: a meta-analysis.* Circulation: Cardiovascular Genetics, 2011. **4**(6): p. 681-686.
388. Franchini, M., G.M. Liumbruno, and G. Lippi, *The prognostic value of ABO blood group in cancer patients.* Blood Transfusion, 2016. **14**(5): p. 434.
389. Franchini, M., et al., *Relationship between ABO blood group and von Willebrand factor levels: from biology to clinical implications.* Thrombosis journal, 2007. **5**(1): p. 14.
390. Franchini, M., et al., *von Willebrand factor and cancer: a renewed interest.* Thrombosis Research, 2013. **131**(4): p. 290-292.
391. Rezende, S.M., R.E. Simmonds, and D.A. Lane, *Coagulation, inflammation, and apoptosis: different roles for protein S and the protein S–C4b binding protein complex.* Blood, 2004. **103**(4): p. 1192-1201.
392. Stitt, T.N., et al., *The anticoagulation factor protein S and its relative, Gas6, are ligands for the Tyro 3/Axl family of receptor tyrosine kinases.* Cell, 1995. **80**(4): p. 661-670.
393. Mudduluru, G., et al., *Regulation of Axl receptor tyrosine kinase expression by miR-34a and miR-199a/b in solid cancer.* Oncogene, 2011. **30**(25): p. 2888.
394. Linger, R.M., et al., *Mer or Axl receptor tyrosine kinase inhibition promotes apoptosis, blocks growth and enhances chemosensitivity of human non-small cell lung cancer.* Oncogene, 2013. **32**(29): p. 3420.
395. Paccez, J.D., et al., *The receptor tyrosine kinase Axl in cancer: biological functions and therapeutic implications.* International journal of cancer, 2014. **134**(5): p. 1024-1033.
396. Wang, C.H., et al., *A shRNA functional screen reveals Nme6 and Nme7 are crucial for embryonic stem cell renewal.* Stem Cells, 2012. **30**(10): p. 2199-2211.
397. Bilitou, A., et al., *The NM23 family in development.* Molecular and cellular biochemistry, 2009. **329**(1-2): p. 17-33.
398. Girmann, G., et al., *Immunosuppression by micromolecular fibrinogen degradation products in cancer.* Nature, 1976. **259**(5542): p. 399.
399. Davalos, D. and K. Akassoglou. *Fibrinogen as a key regulator of inflammation in disease*. in *Seminars in immunopathology*. 2012. Springer.

400. Sahni, A., et al., *Fibrinogen synthesized by cancer cells augments the proliferative effect of fibroblast growth factor-2 (FGF-2).* Journal of Thrombosis and Haemostasis, 2008. **6**(1): p. 176-183.
401. Groeneweg, J.W., et al., *Notch signaling in serous ovarian cancer.* Journal of ovarian research, 2014. **7**(1): p. 1.
402. Venkatesh, V., et al., *Targeting Notch signalling pathway of cancer stem cells.* Stem cell investigation, 2018. **5**.
403. Lobry, C., P. Oh, and I. Aifantis, *Oncogenic and tumor suppressor functions of Notch in cancer: it's NOTCH what you think.* Journal of Experimental Medicine, 2011. **208**(10): p. 1931-1935.
404. Dunn, A.R., et al., *Synaptic vesicle glycoprotein 2C (SV2C) modulates dopamine release and is disrupted in Parkinson disease.* Proceedings of the National Academy of Sciences, 2017. **114**(11): p. E2253-E2262.
405. Lan, Y.-L., et al., *Anti-cancer effects of dopamine in human glioma: involvement of mitochondrial apoptotic and anti-inflammatory pathways.* Oncotarget, 2017. **8**(51): p. 88488.
406. Wang, D., et al., *Protein C receptor stimulates multiple signaling pathways in breast cancer cells.* Journal of Biological Chemistry, 2018. **293**(4): p. 1413-1424.
407. Kumar, M.S., et al., *The GATA2 transcriptional network is requisite for RAS oncogene-driven non-small cell lung cancer.* Cell, 2012. **149**(3): p. 642-655.
408. Mantovani, A., I. Barajon, and C. Garlanda, *IL-1 and IL-1 regulatory pathways in cancer progression and therapy.* Immunological reviews, 2018. **281**(1): p. 57-61.
409. Choi, S.H., et al., *Six novel loci associated with circulating VEGF levels identified by a meta-analysis of genome-wide association studies.* PLoS genetics, 2016. **12**(2): p. e1005874.
410. Zhang, B., et al., *TSPAN15 interacts with BTRC to promote oesophageal squamous cell carcinoma metastasis via activating NF-κB signaling.* Nature communications, 2018. **9**(1): p. 1423.
411. Katz-Brull, R., et al., *Metabolic markers of breast cancer: enhanced choline metabolism and reduced choline-ether-phospholipid synthesis.* Cancer research, 2002. **62**(7): p. 1966-1970.
412. Ma, Q., et al., *Mitochondrial PIP3-binding protein FUNDC2 supports platelet survival via AKT signaling pathway.* Cell Death & Differentiation, 2019. **26**(2): p. 321.
413. Jain, S., J. Harris, and J. Ware, *Platelets: linking hemostasis and cancer.* Arteriosclerosis, thrombosis, and vascular biology, 2010. **30**(12): p. 2362-2367.
414. De Francesco, E.M., et al., *Vitamin C and Doxycycline: A synthetic lethal combination therapy targeting metabolic flexibility in cancer stem cells (CSCs).* Oncotarget, 2017. **8**(40): p. 67269.
415. Chen, X., et al., *EphA3 inhibits migration and invasion of esophageal cancer cells by activating the mesenchymal-epithelial transition process.* International journal of oncology, 2019. **54**(2): p. 722-732.
416. Lv, X.Y., et al., *EphA3 contributes to tumor growth and angiogenesis in human gastric cancer cells.* Oncology reports, 2018. **40**(4): p. 2408-2416.
417. Caivano, A., et al., *Epha3 acts as proangiogenic factor in multiple myeloma.* Oncotarget, 2017. **8**(21): p. 34298.
418. Li, R.-X., Z.-H. Chen, and Z.-K. Chen, *The role of EPH receptors in cancer-related epithelial-mesenchymal transition.* Chinese journal of cancer, 2014. **33**(5): p. 231.
419. Janes, P.W., et al., *EphA3 biology and cancer.* Growth Factors, 2014. **32**(6): p. 176-189.
420. Noda, M., et al., *Glycosyltransferase gene expression identifies a poor prognostic colorectal cancer subtype associated with mismatch repair deficiency and incomplete glycan synthesis.* Clinical Cancer Research, 2018. **24**(18): p. 4468-4481.
421. Xu, Y., et al., *Unique DNA methylome profiles in CpG island methylator phenotype colon cancers.* Genome research, 2012. **22**(2): p. 283-291.

422. Hinoue, T., et al., *Genome-scale analysis of aberrant DNA methylation in colorectal cancer.* Genome research, 2012. **22**(2): p. 271-282.
423. Yu, M., et al., *Methylated B3GAT2 and ZNF793 are potential detection biomarkers for Barrett's esophagus.* Cancer Epidemiology and Prevention Biomarkers, 2015. **24**(12): p. 1890-1897.
424. Xu, W., et al., *Genome-Wide Plasma Cell-Free DNA Methylation Profiling Identifies Potential Biomarkers for Lung Cancer.* Disease Markers, 2019. **2019**.
425. Hanly, A.M. and D.C. Winter. *The role of thrombomodulin in malignancy*. in *Seminars in thrombosis and hemostasis*. 2007. © Thieme Medical Publishers.
426. Zheng, N., et al., *Thrombomodulin reduces tumorigenic and metastatic potential of lung cancer cells by up-regulation of E-cadherin and down-regulation of N-cadherin expression.* Biochemical and biophysical research communications, 2016. **476**(4): p. 252-259.
427. Hanly, A., et al., *Thrombomodulin: tumour biology and prognostic implications.* European Journal of Surgical Oncology (EJSO), 2005. **31**(3): p. 217-220.
428. Liu, P.L., et al., *Decreased expression of thrombomodulin is correlated with tumor cell invasiveness and poor prognosis in nonsmall cell lung cancer.* Molecular carcinogenesis, 2010. **49**(10): p. 874-881.
429. Wu, H.-L., Y.-Y. Hsu, and G.-Y. Shi, *Thrombomodulin Promotes Cell Adhesion and Migration and Enhances Angiogenesis through Interaction with Fibronectin.* The FASEB Journal, 2017. **31**(1_supplement): p. 618.1-618.1.
430. Gold, L.I., *The role for transforming growth factor-beta (TGF-beta) in human cancer.* Critical reviews in oncogenesis, 1999. **10**(4): p. 303-360.
431. Buck, M.B. and C. Knabbe, *TGF-beta signaling in breast cancer.* Annals of the New York Academy of Sciences, 2006. **1089**(1): p. 119-126.
432. Kaminska, B., A. Wesolowska, and M. Danilkiewicz, *TGF beta signalling and its role in tumour pathogenesis.* ACTA BIOCHIMICA POLONICA-ENGLISH EDITION-, 2005. **52**(2): p. 329.
433. Lampropoulos, P., et al., *TGF-beta signalling in colon carcinogenesis.* Cancer letters, 2012. **314**(1): p. 1-7.
434. Gorelik, L. and R.A. Flavell, *Transforming growth factor-β in T-cell biology.* Nature Reviews Immunology, 2002. **2**(1): p. 46.
435. Fabregat, I., et al., *TGF-beta signaling in cancer treatment.* Current pharmaceutical design, 2014. **20**(17): p. 2934-2947.
436. Sasaki, S., et al., *Serum soluble MD-1 levels increase with disease progression in autoimmune prone MRLlpr/lpr mice.* Molecular immunology, 2012. **49**(4): p. 611-620.
437. Su, S., et al., *DNA methylation of the LY86 gene is associated with obesity, insulin resistance, and inflammation.* Twin Research and Human Genetics, 2014. **17**(3): p. 183-191.

8. Appendix

Appendix 1

Supplementary Table 1 - SNP's identified by VTE susceptibility GWAS

Study	Associated SNP's	Population	No. cases/controls (combined)	MAF	Locus	Gene/ Variant	Overall risk	
							Allelic OR (95% CI)	*p-Value*
Trégouët et al. (2009)	rs2420371	European ancestry	419/1228 (Discovery phase)	0.15[a]	1q24.2	*F5*/intr	2.27 (1.62; 3.18) [c]	8.08×10^{-10}
	rs1208134			0.12[a]	1q24.2	*CCDC181*/ intr	2.29 (1.58; 3.32) [c]	3.47×10^{-7}
	rs657152			0.54[a]	chr9: 133263862 [b]	*ABO*/intr [b]	1.89 (1.51; 2.36) [c]	2.22×10^{-13}
	rs505922			0.52[a]	chr9: 133273813 [b]	*ABO*/intr [b]	1.91 (1.53; 2.39) [c]	1.48×10^{-14}
	rs630014			0.37[a]	9q34.2	*ABO*/intr	0.64 (0.51; 0.80) [c]	2.00×10^{-7}
	rs2420371¥	European ancestry	1150/801 (Replication phase I)	0.21[a]	1q24.2	*F5*/intr	1.39 (1.17;1.64) [c]	3.00×10^{-5}
	rs1208134¥			0.19[a]	1q24.2	*CCDC181*/ intr	1.57 (1.31; 1.88) [c]	2.89×10^{-7}
	rs6025			0.01	1q24.2	*F5*/mis	2.01 (1.63; 2.48) [c]	9.91×10^{-11}
	rs657152§			0.51[a]	chr9: 133263862 [b]	*ABO*/intr [b]	1.75 (1.51; 2.03) [c]	1.20×10^{-13}
	rs505922§			0.49[a]	chr9: 133273813 [b]	*ABO*/intr [b]	1.81 (1.56; 2.11) [c]	3.72×10^{-15}
	rs630014§			0.38[a]	9q34.2	*ABO*/intr	0.66 (0.57; 0.76) [c]	1.21×10^{-8}
	rs8176719			0.34	9q34.2	*ABO*/fra	0.33 (0.26; 0.42) [c]	1.70×10^{-18}
	rs8176750			0.05	9q34.2	*ABO*/fra	0.53 (0.38; 0.74) [c]	2.46×10^{-4}
	rs2420371¥	European ancestry	607/607 (Replication phase II)	0.10[a]	1q24.2	*F5*/intr	1.44 (1.07; 1.93) [c]	1.80×10^{-3}
	rs6025			0.01	1q24.2	*F5*/mis	2.46 (1.55; 3.93) [c]	1.50×10^{-4}
	rs657152§			0.47[a]	chr9: 133263862[b]	*ABO*/intr [b]	1.58 (1.34; 1.87) [c]	5.19×10^{-8}
	rs505922§			0.46[a]	chr9: 133273813[b]	*ABO*/intr [b]	1.65 (1.39; 1.95) [c]	7.25×10^{-9}
	rs630014§			0.38[a]	9q34.2	*ABO*/intr	0.63 (0.53; 0.74) [c]	5.01×10^{-8}
	rs8176719			0.34	9q34.2	*ABO*/fra	0.53 (0.41; 0.69) [c]	2.21×10^{-6}
Buil et al. (2010)	rs3813948	European ancestry	419/1228 (*in silico* GWAS)	0.09 [a]	1q32.1	*C4BPB*/nc	-	0.011
	rs3813948		1706/1379 (Replication phase)	0.09 [a]	1q32.1	*C4BPB*/nc	1.24 (1.00; 1.53)	0.046
Germain et al. (2011)	rs16861990	European ancestry	1542/1110 (Discovery phase)	0.13 [a]	1q24.2	*NME7*/intr	2.49 [c] -	2.75×10^{-15}
	rs1208134			0.13 [a]	1q24.2	*CCDC181*/ intr	2.53 [c] -	3.29×10^{-16}
	rs2420371			0.15 [a]	1q24.2	*F5*/intr	2.62 [c] -	8.44×10^{-19}
	rs2066865			0.28 [a]	4q32.1	*FGG*/inter	1.55 [c] -	1.17×10^{-10}
	rs6825454			0.30 [a]	4q31.3	*FGA*/inter	1.50 [c] -	1.32×10^{-9}

Supplementary Table 1 – SNP's identified by VTE susceptibility GWAS (cont.)

Study	Associated SNP's	Population	No. cases/controls (combined)	MAF	Locus	Gene/ Variant	Overall risk	
							Allelic OR (95% CI)	*p-Value*
Germain et al. (2011) (cont.)	rs10029715	European ancestry	1542/1110 (Discovery phase)	0.12 [a]	4q35.2	*F11-ASI*/intr	-	3.20 x 10^{-9}
	rs2073828			0.32 [a]	chr9: 133261737 [b]	*ABO*/intr [b]	-	3.57 x 10^{-9}
	rs657152			0.49 [a]	chr9: 133263862 [b]	*ABO*/intr [b]	1.70 [c]	1.10 x 10^{-18}
	rs500498			0.33 [a]	chr9: 133273232 [b]	*ABO*/intr [b]	-	1.03 x 10^{-12}
	rs505922			0.49 [a]	chr9: 133273813 [b]	*ABO*/intr [b]	1.85 [c]	1.06 x 10^{-23}
	rs630014			0.38 [a]	9q34.2	*ABO*/intr	0.63 [c]	4.40 x 10^{-14}
	rs495828			0.36 [a]	9q34.2	*ABO*/rr	1.64 [c]	1.78 x 10^{-14}
	rs1018827	European ancestry	1961/2338 (meta-analysis) [d]	0.07	1q24.2	*F5*/intr	2.52	2.41 x 10^{-26}
	rs7659024			0.30	4q31.3	*FGG*/inter	1.53	1.93 x 10^{-13}
	rs505922			0.35	chr9: 133273813 [b]	*ABO*/intr [b]	1.92	1.39 x 10^{-34}
	rs3756008			0.32	4q35.2	*F11*/inter	1.40	6.46 x 10^{-11}
Heit et al. (2012)	rs6025	98.64% European ancestry (USA)	1503/1459 (Discovery phase)	0.01	1q24.2	F5/mis	3.75 (2.76; 4.60)	1.68×10^{-22}
	rs8176719			0.34	9q34.2	*ABO*/fra	1.47 (1.32; 1.64)	5.68×10^{-12}
	rs2519093			0.14	chr9: 133266456 [b]	*ABO*/intr [b]	1.69 (1.48; 1.91)	8.08×10^{-16}
	rs495828			0.16	9q34.2	*ABO*/rr	1.65 (1.46; 1.86)	2.96×10^{-16}
	rs7538157[¥]			<0.01	1q24.2	*BLZF1*/intr	2.69 (2.09; 3.45)	1.04×10^{-14}
	rs16861990[¥]			0.06	1q24.2	*NME7*/intr	2.02 (1.66; 2.45)	1.69×10^{-12}
	rs2038024			0.13	1q24.2	*SLC19A2*/ nc	1.53 (1.32; 1.78)	1.12×10^{-8}
	rs1799963			<0.01	11p11.2	*F2*/utr	2.46 (1.70; 3.55)	1.69×10^{-6}
	rs6025	98.64% European ancestry (USA)	1407/1418 (Replication phase)	0.01	1q24.2	*F5*/mis	2.56 (1.97; 3.32)	1.40×10^{-12}
	rs8176719			0.34	9q34.2	*ABO*/fra	1.58 (1.40; 1.78)[e]	9.75×10^{-14} [e]
	rs2519093			0.14	chr9: 133266456 [b]	*ABO*/intr [b]	1.85 (1.61; 2.13)[e]	1.37×10^{-17} [e]
	rs495828			0.16	9q34.2	*ABO*/rr	1.76 (1.54; 2.01)[e]	3.60×10^{-17} [e]
	rs1799963			<0.01	11p11.2	*F2*/utr	1.71 (1.12; 2.63)[e]	0.01 [e]
	rs16861990			0.06	1q24.2	*NME7*/intr	1.79 (1.47; 2.18)	4.89×10^{-9}
							1.17 (0.89;1.54)[e]	0.25 [e]
	rs2038024			0.13	1q24.2	*SLC19A2*/ nc	0.77 (0.65;0.92)[e]	4.00×10^{-3} [e]

Supplementary Table 1 – SNP's identified by VTE susceptibility GWAS (cont.)

Study	Associated SNP's	Population	No. cases/controls (combined)	MAF	Locus	Gene/ Variant	Overall risk	
							Allelic OR (95% CI)	*p-Value*
Tang et al. (2013)	rs6427196	European ancestry	1618/44499 (Discovery phase)	0.09	1q24.2	*F5*/utr	1.82 (1.58; 2.10) [c]	1.97×10^{-16}
	rs687621			0.38	chr9: 133261662 [b]	*ABO*/intr [b]	1.37 (1.26;1.49) [c]	3.42×10^{-14}
	rs4253399			0.26	4q35.2	*F11*/intr	1.15 (1.06; 1.24) [c]	7.59×10^{-4}
	rs6536024			0.46	4q32.1	*FGG*/interg	0.79 (0.73; 0.87) [c]	4.04×10^{-7}
	rs6764623			0.35	3p26.3	*CNTN6*/interg	1.23 (1.11; 1.38) [c]	9.56×10^{-5}
	rs4979078			0.33	9q31.3	*SUSD1*/intr	1.31 (1.17; 1.47) [c]	2.46×10^{-6}
	rs7164569			0.33	15q13.3	*OTUD7A*/syn	0.84 (0.76; 0.92) [c]	3.54×10^{-4}
	rs3733860			0.17	5q13.3	*SV2C*/utr	1.22 (1.09; 1.37) [c]	6.27×10^{-4}
	rs6427196	European ancestry	3231/3536 (Replication phase)	0.09	1q24.2	*F5*/utr	2.31 (2.04; 2.62) [c]	2.56×10^{-38}
	rs687621			0.38	chr9: 133261662 [b]	*ABO*/intr [b]	1.75 (1.62; 1.89) [c]	1.20×10^{-44}
	rs4253399			0.26	4q35.2	*F11*/intr	1.32 (1.23; 1.43) [c]	2.07×10^{-13}
	rs6536024			0.46	4q32.1	*FGG*/interg	0.81 (0.75; 0.87) [c]	5.59×10^{-8}
	rs6764623			0.35	3p26.3	*CNTN6*/interg	1.14 (1.05; 1.24) [c]	2.00×10^{-3}
	rs4979078			0.33	9q31.3	*SUSD1*/intr	1.11 (1.00; 1.24) [c]	4.70×10^{-2}
	rs7164569			0.33	15q13.3	*OTUD7A*/syn	0.88 (0.82; 0.95) [c]	2.00×10^{-3}
	rs3733860			0.17	5q13.3	*SV2C*/utr	1.17 (1.05; 1.30) [c]	3.00×10^{-3}
	rs6427196	European ancestry	4849/48035 (Combined data of all nine studies)	0.09	1q24.2	*F5*/utr	2.07 (1.89; 2.28) [c]	4.47×10^{-51}
	rs687621			0.38	chr9: 133261662 [b]	*ABO*/intr [b]	1.55 (1.47; 1.64) [c]	1.55×10^{-52}
	rs4253399			0.26	4q35.2	*F11*/intr	1.24 (1.17; 1.31) [c]	2.78×10^{-14}
	rs6536024			0.46	4q32.1	*FGG*/interg	0.80 (0.76; 0.85) [c]	1.75×10^{-13}
	rs6764623			0.35	3p26.3	*CNTN6*/interg	1.18 (1.10; 1.26) [c]	1.57×10^{-6}
	rs4979078			0.33	9q31.3	*SUSD1*/intr	1.21 (1.11; 1.30) [c]	3.06×10^{-6}
	rs7164569			0.33	15q13.3	*OTUD7A*/syn	0.87 (0.81; 0.92) [c]	3.27×10^{-6}
	rs3733860			0.17	5q13.3	*SV2C*/utr	1.19 (1.10; 1.29) [c]	8.06×10^{-6}
Germain et al. (2015)	rs6025	European ancestry	7507/52632 (Discovery phase)	0.01	1q24.2	*F5*/mis	3.25 (2.91; 3.64)	1.10×10^{-96}
	rs4524			0.27	1q24.2	*F5*/mis	1.20 (1.14; 1.26)	2.65×10^{-11}
	rs2066865			0.30	4q32.1	*FGG*/ inter	1.24 (1.18; 1.31)	1.03×10^{-16}

Supplementary Table 1 – SNP's identified by VTE susceptibility GWAS (cont.)

Study	Associated SNP's	Population	No. cases/controls (combined)	MAF	Locus	Gene/ Variant	Overall risk	
							Allelic OR (95% CI)	*p-Value*
Germain et al. (2015) (cont.)	rs4253417	European ancestry	7507/52632 (Discovery phase)	0.30	4q35.2	*F11*/intr	1.27 (1.22; 1.34)	1.21×10^{-23}
	rs529565			0.37	chr9: 133274084[b]	*ABO*/intr [b]	1.55 (1.48; 1.63)	4.23×10^{-75}
	rs1799963			<0.01	11p11.2	*F2*/utr	2.29 (1.75; 2.99)	1.73×10^{-9}
	rs6087685			0.39	20q11.22	*PROCR*/intr	1.15 (1.10; 1.21)	1.65×10^{-8}
	rs4602861			0.39	8q23.1	*ZFPM2*/intr	1.20 (1.13; 1.27)	3.48×10^{-9}
	rs78707713			0.05	10q22.1	*TSPAN15*/intr	1.28 (1.19; 1.39)	5.74×10^{-11}
	rs2288904			0.18	19p13.2	*SLC44A2*/mis	1.19 (1.12; 1.26)	1.07×10^{-9}
	rs78707713	European ancestry	3009/2586 (Replication phase)	0.05	10q22.1	*TSPAN15*/intr	1.42 (1.24; 1.62)	2.21×10^{-7}
	rs2288904			0.18	19p13.2	*SLC44A2*/mis	1.28 (1.16; 1.40)	2.64×10^{-7}
	rs4602861	European ancestry	10516/55218 (combined data)	0.39	8q23.1	*ZFPM2*/intr	-	5.04×10^{-7}
	rs78707713			0.05	10q22.1	*TSPAN15*/intr	-	1.67×10^{-16}
	rs2288904			0.18	19p13.2	*SLC44A2*/mis	-	2.75×10^{-15}
Hernandez et al. (2016)	rs62322307[#]	West African Ancestry [f] (80%) European and Asian ancestry	146/432 (Discovery phase)	0.15 [a]	4q22.2	*ATOH1*/inter	2.79 (1.80; 4.30)	2.25×10^{-7}
	rs73692310			0.15 [a]	7p12.3	*IGFBP3*/inter	3.04 (2.00;4.70)	1.73×10^{-9}
	rs58952918[#]			0.17 [a]	18p11.32	*AP005230.1*/ intr	2.48 (1.70; 3.70)	1.07×10^{-8}
	rs28496996			0.17 [a]	18p11.32	*AP005230.1*/ intr	2.44 (1.60; 3.60)	1.13×10^{-8}
	rs2144940			0.31 [a]	20p11.21	*THBD*, CD93/inter	2.18 (1.60; 2.90)	3.52×10^{-7}
	rs2567617[#]			0.31 [a]	20p11.21	*THBD*, CD93/inter	2.17 (1.60; 2.90)	4.01×10^{-7}
	rs1998081			0.27 [a]	20p11.21	*THBD*, CD93/inter	2.28 (1.60; 3.10)	5.17×10^{-7}
	rs687621			0.38	chr9: 133261662 [b]	*ABO*/intr [b]	1.55 (1.20; 2.00)	2.00×10^{-3}
	rs505922			0.35	chr9: 133273813[b]	*ABO*/intr [b]	1.52 (1.20; 2.00)	2.00×10^{-3}
	rs657152			0.39	chr9: 133263862 [b]	*ABO*/intr [b]	1.39 (1.10; 1.80)	0.03
	rs73692310	West African Ancestry [f] (77%) European and Asian ancestry	94/65 (Replication phase)	0.09 [a]	7p12.3	*IGFBP3*/inter	1.27 (0.04; 2.70)	0.60
	rs28496996			0.13 [a]	18p11.32	*AP005230.1*/ intr	1.34 (0.60; 2.60)	0.45
	rs2144940			0.35 [a]	20p11.21	*THBD*, *CD93*/inter	1.89 (1.10; 3.30)	0.02
	rs1998081			0.30 [a]	20p11.21	*THBD*, *CD93*/inter	1.94 (1.10; 3.50)	0.02

Supplementary Table 1 – SNP's identified by VTE susceptibility GWAS (cont.)

Study	Associated SNP's	Population	No. cases/controls (combined)	MAF	Locus	Gene/ Variant	Overall risk	
							Allelic OR (95% CI)	*p-Value*
Hernandez et al. (2016) (cont.)	rs73692310	West African Ancestry [f] (79%) European and Asian ancestry	240/497 (Combined data)	0.02	7p12.3	*IGFBP3*/inter	-	2.48×10^{-8}
	rs28496996			0.03	18p11.32	*AP005230.1*/ intr	-	6.37×10^{-8}
	rs2144940			0.12	20p11.21	*THBD*, *CD93*/inter	-	1.88×10^{-8}
	rs1998081			0.11	20p11.21	*THBD*, *CD93*/inter	-	4.62×10^{-8}
Hinds et al. (2016)	rs6025	European ancestry	6135/252827 (Discovery phase)	0.01	1q24.2	*F5*/mis	2.93 (2.72; 3.15)	3.60×10^{-137}
	rs7654093			0.31	4q32.1	*FGG*/inter	1.22 (1.17; 1.27)	2.00×10^{-19}
	rs4444878			0.32	4q35.2	*F11-ASI*/intr	0.81 (0.78; 0.84)	7.00×10^{-28}
	rs1799963			<0.01	11p11.2	*F2*/utr	0.51 (0.46; 0.58)	1.30×10^{-24}
	rs34234989			0.39	20q11.22	*PROCR*/intr	0.89 (0.85; 0.92)	6.70×10^{-9}
	rs529565			0.37	chr9: 133274084 [b]	*ABO*/intr [b]	0.72 (0.70; 0.75)	7.10×10^{-63}
	rs9797861			0.21	19p13.2	*SLC44A2*/ intr	1.15 (1.09; 1.20)	6.10×10^{-9}
	rs114209171			0.24	Xq28	*FUNDC2*/nc	1.15 (1.11; 1.20)	7.00×10^{-13}
	rs72798544			0.01	2p21	*COX7A2L*/ intr	0.73 (0.65; 0.82)	1.90×10^{-7}
	rs17490626			0.04	10q22.1	*TSPAN15*/ intr	1.17 (1.10; 1.24)	2.90×10^{-7}
	rs113092656			0.01	6p24.1	*TMEM170B*/*ADTRP*/inter	0.73 (0.65; 0.82)	4.40×10^{-7}
	rs60942712			0.06	3p11.1	*EPHA3*/inter	1.21 (1.12; 1.31)	8.00×10^{-7}
	rs114209171	European ancestry	26112 participants (Replication phase)	0.24	Xq28	*FUNDC2*/nc	1.08 (1.02; 1.14)	0.01
Rühle et al. (2017)	rs1304029	European ancestry	212 children with VTE / 424 parents and siblings (Discovery phase)	0.48	6q13	*B3GAT2*/intr	0.48 (0.36; 0.65)	2.00×10^{-6} [h]
	rs9293858			0.26	6q13	*RIMS1*/intr	0.48 (0.34; 0.67)	8.00×10^{-6} [h]
	rs2748331			0.41	6q13	*B3GAT2*/rr	0.49 (0.36; 0.67)	1.80×10^{-5} [h]
	rs10498910			0.12	6q14.1	*LOC105377862*/ intr [b]	2.21 (1.47; 3.31)	6.89×10^{-5} [h]
	rs914958			0.23	1p22.1	*ABCA4*/intr	0.50 (0.36; 0.70)	1.80×10^{-5} [h]
	rs4529013			0.28	4q21.3	*MAPK10*/intr	0.53 (0.39; 0.72)	2.00×10^{-5} [h]
	rs9957519			0.27	18q23	-/inter	0.46 (0.32; 0.68)	2.10×10^{-5} [h]
	rs1865590			0.31	2q22.1	*THSD7B*/intr	1.97 (1.44; 2.68)	2.40×10^{-5} [h]

Supplementary Table 1 – SNP's identified by VTE susceptibility GWAS (cont.)

Study	Associated SNP's	Population	No. cases/controls (combined)	MAF	Locus	Gene/ Variant	Overall risk	
							Allelic OR (95% CI)	*p-Value*
Rühle et al. (2017) (cont.)	rs9606534	European ancestry European ancestry	212 children with VTE / 424 parents and siblings (Discovery phase)	0.17	chr22: 16916985 [b]	*IGKV2OR22-4*/rr	0.43 (0.29; 0.63)	3.30×10^{-5} [h]
	rs495828			0.16	9q34.2	*ABO*/rr	-	6.44×10^{-4}
	rs505922			0.35	chr9: 133273813[b]	*ABO*/intr [b]	-	4.03×10^{-4}
	rs657152			0.39	chr9: 133263862 [b]	*ABO*/intr [b]	1.77 (1.34; 2.32)	3.44×10^{-5}
	rs13146272			0.44	4q35.1	*CYP4V2*/miss	-	9.58×10^{-4}
	rs925451			0.29	4q35.2	*F11*/intr	-	2.76×10^{-3}
	rs11128790			0.06	3p24.3	*RFTN1*/intr	2.95 (1.78; 4.90)	3.40×10^{-5} [h]
	rs4792119			0.21	17p12	*SHISA6*/Intr	0.51 (0.37; 0.71)	3.50×10^{-5} [h]
	rs9399770			0.48	6q16.3	-/inter	0.55 (0.42; 0.74)	4.00×10^{-5} [h]
	rs17576372			0.27	1p22.1	*TGFBR3*/intr	1.84 (1.37; 2.47)	4.57×10^{-5} [h]
	rs10247053			0.25	7p15.2	-/inter	0.53 (0.39; 0.72)	5.35×10^{-5} [h]
	rs636434			0.34	6q12	*EYS*/intr	1.79 (1.34; 2.39)	5.35×10^{-5} [h]
	rs10190178			0.31	2q22.1	*THSD7B*/intr	1.91 (1.40; 2.62)	6.15×10^{-5} [h]
	rs5014872			0.12	2p16.3	*LOC730100*/ Intr [b]	0.46 (0.32; 0.68)	6.21×10^{-5} [h]
	rs3823606			0.04	7q11.21	*TPST1*/intr	-	6.27×10^{-5} [h]
	rs1565242			0.11	15q26.1	*LOC105370982*/intr [b]	0.44 (0.29; 0.67)	7.23×10^{-5} [h]
	rs1958059			0.31	14q13.1	*NPAS3*/intr	0.45 (0.31; 0.67)	7.28×10^{-5} [h]
	rs1521882			0.23	2q33.1	*KIAA2012*/intr	2.13 (1.46; 3.11)	7.48×10^{-5} [h]
	rs17781793			0.05	12q15	*MRPL40P1*/ inter	0.38 (0.23; 0.63)	7.81×10^{-5} [h]
	rs4775384			0.31	15q22.2	*AC104574.2*/ intr	0.41 (0.26; 0.65)	8.16×10^{-5} [h]
	rs1948650			0.33	15q14	*DPH6-DT*/intr	1.84 (1.34; 2.51)	8.71×10^{-5} [h]
	rs436985			0.34	5q12.1	*C5orf64*/intr	0.58 (0.44; 0.76)	9.13×10^{-5} [h]
	rs4926448			0.47	1q44	*SCCPDH*/intr	0.57 (0.43; 0.76)	9.38×10^{-5} [h]
	rs11153626			0.22	6q22.1	*FAM162B*/ inter	1.85 (1.34; 2.54)	9.49×10^{-5} [h]
	rs2214810			0.26	7p15.2	-/inter	0.54 (0.40; 0.74)	9.62×10^{-5} [h]

Supplementary Table 1 – SNP's identified by VTE susceptibility GWAS (cont.)

Study	Associated SNP's	Population	No. cases/controls (combined)	MAF	Locus	Gene/ Variant	Overall risk	
							Allelic OR (95% CI)	*p-Value*
Rühle et al. (2017) (cont.)	rs2748331	European ancestry	413 children/ 826 parents and siblings (combined data of discovery phase and replication phase I)	0.41	6q13	*B3GAT2*/rr	-	7.88×10^{-7}
	rs9446340			0.23	6q13	*B3GAT2*/Inter	-	1.48×10^{-3}
	rs10498910			0.12	6q14.1	*LOC105377862*/intr [b]	-	5.74×10^{-5}
	rs2748331	European ancestry	651 adults with VTE/ 1356 controls (Replication phase II)	0.41	6q13	*B3GAT2*/rr	1.20 (1.02; 1.40)	0.02 [g]
	rs1304029			0.48	6q13	*B3GAT2*/intr	1.18 (1.02; 1.36)	0.03 [g]
Heit et al. (2017)	rs138916004 [ж]	African ancestry (African-Americans)	393/4941 (Discovery phase)	< 0.01	12q14.3	*LEMD3*/intr	3.17 (2.13; 4.72) [j]	1.27×10^{-8} [j]
	rs3804476 [ж]			0.28	6p25.1	*LY86*/intr	1.83 (1.48; 2.26) [j]	1.97×10^{-8} [j]
	rs142143628 [ж]			< 0.01	8q12.2	*LOC100130298*/intr [b]	4.97 (2.80; 8.83) [j]	4.35×10^{-8} [j]
	rs6025			0.01	1q24.2	*F5*/mis	5.00 (2.02; 11.03) [j]	2.00×10^{-4} [j]
	rs8176746			0.15	9q34.2	*ABO*/mis	1.33 (1.09; 1.62) [j]	5.00×10^{-3} [j]
	rs8176719			0.34	9q34.2	*ABO*/fra	1.30 (1.11; 1.53) [j]	2.00×10^{-3} [j]
	rs77121243 [β]			0.03	11p15.4	*HBB*/miss	1.51 (1.11; 2.06)	9.00×10^{-3}
Klarin et al. (2017)	rs6025	European ancestry	3290/116868 (Discovery phase)	0.01	1q24.2	*F5*/mis	3.49 (2.96; 4.11)	7.10×10^{-50}
	rs2066865			0.30	4q32.1	*FGG*/inter	1.21 (1.15; 1.29)	3.10×10^{-11}
	rs4253416			0.41	4q35.2	*F11*/intr	1.18 (1.12; 1.24)	2.00×10^{-10}
	rs2519093			0.14	chr9: 133266456 [b]	*ABO*/intr [b]	1.41 (1.32; 1.50)	6.00×10^{-26}
	rs8176645			0.38	9q34.2	*ABO*/intr	1.28 (1.22; 1.35)	4.40×10^{-21}
	rs1799963			<0.01	11p11.2	*F2*/utr	2.63 (2.03; 3.40)	4.90×10^{-13}
	rs3136516			0.28	11p11.2	*F2*/intr	1.10 (1.04; 1.15) [k]	3.30×10^{-4} [k]
	rs4602861			0.39	8q23.1	*ZFPM2*/intr	1.08 (1.03; 1.15)	4.50×10^{-3}
	rs4602861	European ancestry	10516/55218 (Replication phase)	0.39	8q23.1	*ZFPM2*/intr	1.13 (1.08; 1.19)	5.04×10^{-7}
	rs3136516			0.28	11p11.2	*F2*/intr	1.10 (1.06; 1.15) [k]	5.65×10^{-6} [k]
	rs4602861	European ancestry	13806/ 172086 (combined data)	0.39	8q23.1	*ZFPM2*/intr	1.11 (1.07; 1.15)	4.88×10^{-10}
	rs3136516			0.28	11p11.2	*F2*/intr	1.10 (1.06; 1.13) [k]	7.60×10^{-9} [k]

8. Appendix

The data shown in the **Supplementary Table 1** concerning locus, type of genetic variant, as well as MAF values for all populations were obtained on the "Ensembl" database. For intergenic variants, the nearest gene was indicated.

MAF: minor allele frequency; **OR**: odds ratio; **Inter**: Intergenic variant, **Intr**: Intronic variant, **Mis**: missense variant, **Fra**: frameshift variant, **Nc**: non coding transcript exon variant, **Syn**: synonymous variant, **UTR**: 3 prime UTR variant, **RR:** regulatory region variant.

a: MAF values for cases in the study; **b**: Data obtained from “NCBI” database; **c**: OR/RR associated with the minor allele; **d**: 99 SNP’s reached genome-wide significant ($p < 2\times10^{-8}$), but only the hit SNPs of each locus (*F5*, *FGG*, *F11* and *ABO*) were included in the table; **e**: Data after adjusting for rs6025; **f**: SNP’s predominantly found in populations of African descent; **g**: After Bonferroni correction, the *P*-values became insignificant; **H**: p-values of permutation testing; **j**: Results after adjusting for sickle cell risk variant (*HBB* rs77121243-T allele) and other cofactors; **k**: Results after adjusting for rs1799963.

¥: SNP’s not significantly associated with VTE risk after adjusting for rs6025; **§:** SNP’s not significantly associated with VTE risk after adjusting for *ABO* blood group (rs8176719 and rs8176750); #: SNPs not tested in replication cohort due to high LD or due to failed assay; **Ж:** SNPs further replicated using parametric bootstrap, internal cross-validation and meta-analysis methods; **β**: SNP merged into rs334 according to “NCBI” database

Appendix 2

Supplementary Table 2 - Genome-wide search for VTE-associated pairwise SNP interactions

Study	Pairwise SNP interactions++	Population	No. cases/controls (combined)	MAF	Locus	Gene/Variant	Overall risk	
							OR	*p-Value*
Greliche et al. (2013)	rs493014 rs886090	European ancestry	1953/2338 (Combined data of two previous GWAS)	0.30 0.32	9q34.2 9q34.2	*SURF6*/Inter *SURF6*/mis	1.64	6.00×10^{-11}
	rs1336472 rs4715555			0.40 0.38	1p31.3 6p12.1	*AK4*/utr *HMGCLL1*/inter	1.54	4.24×10^{-10}
	rs380904 rs8086028			0.29 0.30	8q24.3 18p11.22	*ZC3H3*/intr *PIEZO2*/utr	1.67	4.51×10^{-10}
	rs6815916 rs6092326			0.09 0.47	4q34.3 20q13.31	*TENM3-AS1*/ inter *FAM209B*/inter	2.10	6.84×10^{-10}
	rs2282015 rs13050454			0.41 0.42	10q26.13 21q21.3	*AL160290.2*/intr *AP001595.1*/ inter	1.50	8.36×10^{-10}
	rs7648704 rs4868644			0.33 0.49	3p22.3 5q35.2	*TRIM71*/rr *RNF44*/inter	1.56	9.89×10^{-10}
	rs1985317 rs827637			0.41 0.46	9q33.1 10p14	*AL445644.1*/inter *AC044784.1*/inter	0.66	1.32×10^{-9}
	rs2321744 rs6497540			0.10 0.42	13q13.2 16p13.2	*RFC3*/inter *GRIN2A*/intr	0.49	1.38×10^{-9}
	rs315122 rs884483			0.30 0.12	12q15 15q23	*YEATS4*/intr *TLE3*/inter	2.05	1.42×10^{-9}
	rs1423386 rs6491679			0.20 0.29	5q12.1 13q33.1	*LRRC70*/inter *FGF14*/intr	1.73	1.63×10^{-9}
	rs7714670 rs12880735			0.44 0.35	5q13.2 14q12	*ARHGEF28*/miss *AL390334.1*/intr	1.52	1.75×10^{-9}
	rs9392653 rs7780976			0.28 0.19	6p25.1 7p21.2	*PPP1R3G*/inter *DGKB*/inter	1.74	1.83×10^{-9}
	rs9804128 rs4784379			0.26 0.24	1p36.13 16q12.2	*IGSF21*/inter *IRX3*/inter	1.71	1.90×10^{-9}
	rs1364505 rs1204660			0.32 0.16	7q32.3 20q11.22	*PLXNA4*/ intr *UQCC1*/intr	1.80	2.10×10^{-9}
	rs2288073 rs10771022			0.29 0.34	2q23.3 12p12.1	*FAM228A*/miss *SOX5*/intr	1.60	2.11×10^{-9}
	rs1367228 rs3905075			0.44 0.40	2p16.1 13q33.3	*EFEMP1*/intr *FAM155AIT1*/ intr	1.49	2.20×10^{-9}
	rs536477 rs1937920			0.43 0.27	1q43 10p15.1	*CHRM3*/intr *AKR1C2*/inter	0.63	2.93×10^{-9}
	rs2710201 rs3780293			0.06 0.35	7q36.2 9q21.2	*ACTR3B* /inter *GNA14*/intr	0.40	3.30×10^{-9}
	rs12541254 rs305009			0.34 0.23	8p22 15q23	*DLC1*/intr *TLE3*/inter	1.65	3.33×10^{-9}

Supplementary Table 2 – Genome-wide search for VTE-associated pairwise SNP interactions (cont.)

Study	Pairwise SNP interactions++	Population	No. cases/controls (combined)	MAF	Locus	Gene/ Variant	Overall risk	
							OR	*p-Value*
Greliche et al. (2013) (cont.)	rs4507975 rs9914518	European ancestry	1953/2338 (Meta-analysis of two previous GWAS)	0.29 0.47	1q25.2 17p13.1	*PAPPA2*/intr *GSG1L2*/intr	0.65	3.58×10^{-9}
	rs2771051 rs827637			0.37 0.46	9q33.1 10p14	-/inter -/inter	0.67	3.82×10^{-9}
	rs10516089 rs11072930			0.31 0.29	5q35.1 15q25.1	*SMIM23*/inter *ARNT2*/inter	0.63	3.86×10^{-9}
	rs10504130 rs2847351			0.14 0.31	8q11.22 18p11.22	*PCMTD1*/intr *APCDD1*/inter	1.88	4.46×10^{-9}
	rs318497 rs7019259			0.49 0.07	6p25.2 9q21.2	*AL133351.3*/nc *PSAT1*/inter	0.43	4.54×10^{-9}
	rs6695223 rs1763510			0.13 0.39	1p22.3 6q23.2	*WDR63*/intr *SGK1*/Intr	1.86	4.70×10^{-9}
	rs1336708 rs1423386			0.25 0.20	13q33.1 5q12.1	*FGF14-IT1*/intr *CKS1BP3*/inter	0.58	4.85×10^{-9}
	rs6771316 rs10986432			0.13 0.17	3p13 9q33.3	*LINC00877*/intr *OLFML2A*/intr	2.13	5.26×10^{-9}
	rs664910 rs877228			0.30 0.46	3q21.3 15q22.2	*MGLL*/intr *RORA*/intr	1.50	6.63×10^{-9}
	rs9945428 rs4823535			0.30 0.27	18q22.3 22q13.32	*FBXO15*/intr *FAM19A5*/inter	0.62	6.88×10^{-9}
	rs1910358 rs9981595			0.23 0.11	5q14.2 21q22.2	*C5orf17*/inter *BRWD1*/intr	2.03	7.14×10^{-9}
	rs6771725 rs10507246			0.27 0.09	3q26.31 12q24.21	*NAALADL2*/intr *TBX5*/intr	2.22	8.60×10^{-9}
	rs16865717 rs2009579			0.28 0.36	2p25.2 20q12	*RSAD2*/intr -/inter	1.56	8.82×10^{-9}
	rs2028385 rs2038227			0.16 0.38	12q23.1 16p13.3	*AC007513.1*/intr *RAB11FIP3*/intr	1.69	8.82×10^{-9}
	rs10476160 rs1707420			0.20 0.48	5q35.2 8p23.2	*SFXN1*/inter -/inter	0.62	9.09×10^{-9}
	rs971572 rs10828151			0.32 0.07	1q25.3 10p12.31	*TSEN15*/intr *NEBL*/intr	0.42	9.30×10^{-9}
	rs6858430 rs4800250			0.21 0.40	4q34.1 18q11.2	*ADAM29*/intr *TAF4B*/intr	1.62	9.67×10^{-9}
	rs467650 rs7153749			0.37 0.44	5q15 14q23.1	*RGMB*/inter *LINCO1500*/ intr	0.67	9.91×10^{-9}

++: The interactions did not reach the Bonferroni correction for the number of investigated interactions; **MAF** – minor allele frequency**; OR** – odds ratio

Appendix 3

Supplementary Table 3 - SNPs reported by VTE GWAS in European populations and their analysis in previously reported candidate gene studies or validation studies also in European populations

Gene	SNP	Type of study	No. cases/controls (combined)	MAF (cases)	OR (95% CI)	p-Value	References
F5	rs6025	Candidate gene approach	471/474	0.01 *	6.50 (1.80-23.00) (GG>AG)	<0.05	[377]
	rs4524	Candidate gene approach	1488/1439	0.25 **	0.77 (0.68-0.87)	2.51×10^{-5}	[138]
	rs1018827	Validation study	1040/16936	0.07 *	1.53 (1.29-1.79) (AA>AG)	6.53×10^{-6}	[136]
	rs6427196	Validation study	1040/16936	0.09 *	1.51 (1.28-1.78) (CC>CG)	9.21×10^{-6}	[136]
	rs2420371 ϫ	-	-	-	-	-	
F2	rs1799963	Candidate gene approach	471/474	<0.01 *	2.80 (1.40-5.60)	<0.05	[101]
	rs3136516	Candidate gene approach	428/795	0.28 *	1.50 (1.00-2.20)	<0.05	[378]
FGB/FGA/ FGG	rs2066865	Candidate gene approach	471/471	0.30 *	2.40 (1.50-3.90)	0.002	[104]
	rs6825454	Candidate gene approach	419/1228	0.31	-	2.80×10^{-4}	[106]
	rs7659024	Validation study	1040/16936	0.30 *	1.40 (1.09-1.78) (AA>GG)	3.03×10^{-2}	[136]
	rs6536024	Validation study	1040/16936	0.46 *	**-**	**0.23**	[136]
	rs7654093 Φ	-	-	-	-	-	
F11	rs3756008	Candidate gene approach	1837/2204	-	1.27 (1.16-1.38)	<0.05	[108]
	rs4253399	Candidate gene approach	1488/1439	0.41 **	1.28 (1.15-1.43)	6.33×10^{-6}	[138]
	rs4253417	-	-	-	-	-	
	rs4444878	-	-	-	-	-	
	rs4253416	-	-	-	-	-	
ABO	rs2519093	Candidate gene approach	1488/1439	0.24 **	1.68 (1.48-1.91)	8.08×10^{-16}	[138]
	rs505922	Validation study	1040/16936	0.35 *	1.78 (1.46-2.15) (CC>TT)	5.17×10^{-11}	[136]

Supplementary Table 3 –VTE related-SNPs reported by GWAS in European populations and their analysis in validation studies or previously reported candidate gene studies (cont.)

Gene	SNP	Type of study	No. cases/controls (combined)	MAF	OR (95% CI)	p-Value	Reference
ABO	rs630014	Validation study	1040/16936	0.42 **	0.75 (0.67-0.84)	2.67×10^{-7}	[138]
	rs8176719	Validation study	1040/16936	0.42 **	1.47 (1.32-1.64)	5.68×10^{-12}	[138]
		Validation study	96/148	0.48	1.62 (1.09-2.38)	0.015	[137]
	rs687621	Validation study	1040/16936	0.38 *	1.74 (1.43-2.10) (AA>GG)	5.45×10^{-10}	[136]
	rs495828	Validation study	1040/16936	0.16 *	2.09 (1.64-2.63) (GG>TT)	1.72×10^{-10}	[136]
	rs8176750 Ф	-	-	-	-	-	
	rs657152	-	-	-	-	-	
	rs529565	-	-	-	-	-	
	rs8176645 Ж	-	-	-	-	-	
C4BPB	rs3813948	Validation study	1433/1402	0.07	-	**0.25**	[379]
NME7	rs16861990	Validation study	1040/16936	0.06 *	4.11 (2.14-7.33) (CC>AA)	2.90×10^{-7}	[136]
PROCR	rs6087685	Validation study	1040/16936	0.39 *	-	**0.92**	[136]
	rs34234989 Ł	-	-	-	-	-	
TSPAN15	rs78707713	Validation study	1040/16936	0.05 *	0.77 (0.66-0.91) (TT>TC)	6.22×10^{-3}	[136]
	rs17490626 Ʊ	-	-	-	-	-	
ZFPM2	rs4602861	-	-	-	-	-	
SLC44A2	rs2288904	Validation study	1040/16936	0.18 *	0.63 (0.44-0.89) (AA>GG)	2.42×10^{-2}	[136]
	rs9797861 ¥	-	-	-	-	-	
SLC19A2	rs2038024	-	-	-	-	-	
CCDC181	rs1208134	-	-	-	-	-	
CNTN6	rs6764623	-	-	-	-	-	
SUSD1	rs4979078	-	-	-	-	-	
OTUD7A	rs7164569	-	-	-	-	-	
SV2C	rs3733860	-	-	-	-	-	
FUNDC2	rs114209171	-	-	-	-	-	
COX7A2L	rs72798544	-	-	-	-	-	

Supplementary Table 3 –VTE related-SNPs reported by GWAS in European populations and their analysis in validation studies or previously reported candidate gene studies (cont.)

Gene	SNP	Type of study	No. cases/controls (combined)	MAF	OR (95% CI)	p-Value	Reference
-	rs113092656	-	-	-	-	-	-
EPHA3	rs60942712	-	-	-	-	-	-

MAF: minor allele frequency**; OR**: odds ratio

*****: MAF values obtained from "Ensembl" database; ******: Total MAF in the study (cases and controls)

Ⱡ: SNP in high LD with rs6427196, particularly for European ancestry populations (r^2>0.81), according to "Ensembl" database; **ф:** SNP in high LD with rs2066865 for all populations according to "Ensembl" database (r^2>0.81); **Ƀ:** This SNP was only validated in oral contraceptive users [380]; **Ж:** SNP in high LD with rs8176719, particularly for European ancestry populations (r^2>0.90), according to "Ensembl" database; **ł:** SNP in high LD with rs6087685 for all populations according to "Ensembl" database (r^2>0.86, except in Kenya population); **Ʊ:** SNP in high LD with rs78707713 for most populations, particularly the European ancestry populations (r^2=1), according to "Ensembl" database; **¥:** SNP in high LD with rs2288904 for most populations, particularly the European ancestry populations (r^2>0.90), according to "Ensembl" database.

Appendix 4

Supplementary Table 4 - VTE related-genes reported by GWAS and their putative links with cancer hallmarks

Genes	HUGO nomenclature	Molecular processes that promote carcinogenesis	Cancer hallmarks
F5	*Coagulation Factor V*	Generation of thrombin [118, 381]	Metastasis [352] Angiogenesis [350, 351] Immune evasion [348, 349] Apoptosis [346, 347]
CCDC181 (*C1orf114*)	*Coiled-Coil Domain Containing 181*	Despite the unknown role in carcinogenesis, this gene is frequently methylated in patients with prostate cancer [382]	Genome instability and mutation [383-386]
ABO	*ABO Blood Group*	Activation of adhesion molecules [387]	Inflammation, immune evasion and metastasis [387, 388]
		Regulation of plasmatic levels of von Willebrand factor (vWF) [389]	Angiogenesis and apoptosis [390]
C4BPB	*Complement Component 4 Binding Protein Beta*	Inactivation of protein S, which is an important cofactor to activated protein C and constitutes a ligand for the Axl family of receptor tyrosine kinases [391, 392]	Inflammation and apoptosis [391] Proliferation signalling, invasion and apoptosis through Axl receptor tyrosine kinase signalling [392-395]
NME7	*NME/NM23 Family Member 7*	Embryonic Stem Cell Renewal [396]	Metastasis [397]
FGB/FGG/ FGA	*Fibrinogen Beta Chain/ Fibrinogen Gamma Chain/ Fibrinogen Alpha Chain*	Formation of fibrin clot	Angiogenesis [121, 351]
		Immune response [398]	Immune evasion [398] Inflammation [399]
		Augmentation of the proliferative effect of fibroblast growth factor-2 (FGF-2) [400]	Proliferative signalling and angiogenesis [400]
F11	*Coagulation Factor XI*	Generation of Factor Xa [345]	Apoptosis [343]
		Generation of thrombin [344, 345]	Metastasis [352] Angiogenesis [350, 351] Immune evasion [348, 349] Apoptosis [346, 347]
SLC19A2	*Solute Carrier Family 19 Member 2*	Metabolism	Cancer metabolism [301]
F2	*Coagulation Factor II, thrombin*	Generation of thrombin [1]	Metastasis [352] Angiogenesis [350, 351] Immune evasion [348, 349] Apoptosis [346, 347]
CNTN6	*Contactin 6*	Activating of Notch signalling pathway [323] Mediation of cell surface interactions	Proliferative signalling and metastasis [401-403]
OTUD7A	*OTU Deubiquitinase 7A*	Modulation of nuclear factor kappa B (NF-κB) expression through interaction with TNF receptor associated factor 6 (TRAF6)	Cell survival, invasion, migration and metastasis [336] Inflammation [335]
SV2C	*Synaptic Vesicle Glycoprotein 2C*	Modulation of dopamine release [404]	Apoptosis and inflammation [405]

Supplementary Table 4 – VTE related-genes reported by GWAS and their putative links with cancer hallmarks (cont.)

Genes	HUGO nomenclature	Molecular processes that promote cancer growth and progression	Cancer hallmarks
SUSD1	*Sushi Domain Containing 1*	Unknown role in carcinogenesis	unknown
PROCR	*Protein C Receptor*	Protein C pathway	Proliferative signalling, angiogenesis, and metastasis [256, 406] Apoptosis [358] Immune evasion [260]
ZFPM2 (*FOG2*)	*Zinc Finger Protein, FOG Family Member 2*	GATA transcriptional network	Apoptosis [407] Inflammation [407, 408] Invasion [407] Angiogenesis [279]
		Genetics variants in *ZFPM2* have been associated with plasmatic levels of VEGF [409]	Angiogenesis [290, 409]
TSPAN15	*Tetraspanin 15*	Mediates signal transduction events that play a role in the regulation of cell activation, growth, development and motility.	Metastasis [410]
SLC44A2	*Solute Carrier Family 44 Member 2*	Cellular intake of thiamine	Cancer metabolism [411]
FUNDC2	*FUN14 Domain Containing 2*	Modulation of platelet survival [412]	Metastasis, angiogenesis and immune evasion [349, 413]
COX7A2L	*Cytochrome C Oxidase Subunit 7A2 Like*	Regulation of oxidative phosphorylation	Cancer metabolism [414]
EPHA3	*EPH Receptor A3*	Regulation of developmental events Regulation of cytoskeletal organization, cell-cell adhesion and cell migration	Invasion [415] Angiogenesis [416, 417] Metastasis [418, 419]
B3GAT2	*Beta-1,3-Glucuronyltransferase 2*	Mismatch repair deficiency [420]	Genome instability and mutation [420-424]
THBD	*Thrombomodulin*	Protein C pathway Regulation of adhesion molecules [425, 426]	Invasion and metastasis [425-428] Angiogenesis [429]
LEMD3 (*MAN1*)	*LEM Domain Containing 3*	Regulation of transforming growth factor-beta (TGF-beta) signalling at the inner nuclear membrane	Proliferative signalling [430-433] Invasion [432, 433] Apoptosis [432, 433] Immune evasion [433-435]
LY86 (*MD-1)*	*Lymphocyte Antigen 86*	Innate Immune System	Inflammation [436, 437]
LOC100130298	*HCG1816373-Like*	Unknown role in carcinogenesis	Unknown

8. Appendix

The data shown in the **Supplementary Table 4** concerning the HUGO nomenclature and the molecular process involved in carcinogenesis were obtained from "Genecards" database (exceptions are referenced)

Appendix 5

A paper entitled *Venous thromboembolism GWAS reported genetic makeup and the hallmarks of cancer: linkage to ovarian tumour behaviour* has been submitted to the scientific journal *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer.*