

*AIX-MARSEILLE UNIVERSITE*  
*FACULTE DE MEDECINE DE MARSEILLE*  
*ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE*

## **THESE**

**Présentée et publiquement soutenue devant**

**LA FACULTE DE MEDECINE DE MARSEILLE**

Le 18 décembre 2017

Par Monsieur **SUCHON Pierre**

Né à La Ciotat le 9 juillet 1986

## **IDENTIFICATION DE VARIANTS GENETIQUES ASSOCIES A LA THROMBOSE VEINEUSE**

**Pour obtenir le grade universitaire de DOCTEUR de l'UNIVERSITE d'AIX MARSEILLE**

Discipline : **Pathologies Humaines**

Spécialité : **Pathologie Vasculaire et Nutrition**

### *Membres du Jury*

**Monsieur le Professeur Pierre MORANGE**

**DIRECTEUR DE THESE**

**Monsieur le Docteur David-Alexandre TREGOUET**

**CO-DIRECTEUR DE THESE**

**Monsieur le Professeur Yves GRUEL**

**RAPPORTEUR**

**Monsieur le Professeur Joseph EMMERICH**

**RAPPORTEUR**

**Madame le Professeur Marie-Christine ALESSI**

**EXAMINATEUR AMU**

# Remerciements

*Aux membres du jury,*

**Monsieur le Professeur Pierre-Emmanuel Morange,**

Merci d'avoir accepté de diriger mes travaux pour la seconde fois. Il me semble qu'il s'agit de la dernière. Cette thèse marque la fin de mon cursus universitaire en tant qu'étudiant. Grâce à vous, c'est en tant qu'enseignant que mon parcours se poursuivra peut-être au sein de l'Université. Je profite de cette occasion pour vous témoigner tout mon respect. Je vous remercie de la confiance que vous m'avez accordée il y a quatre ans déjà. Parce que vous formez un binôme indissociable, vous partagez ce paragraphe avec **Monsieur le Docteur David-Alexandre Tregouët**. Cher co-directeur, merci d'avoir apposé ta griffe sur ce travail. De ton lointain pays parisien tu as suivi ce travail pas à pas. Merci pour tes précieux conseils et ton œil de lynx.

**Messieurs les Professeurs Yves Gruel et Joseph Emmerich,**

Merci d'avoir accepté la lourde tâche d'être les rapporteurs de cette thèse auprès de l'école doctorale. Je suis très honoré de pouvoir vous compter parmi les membres de ce jury. Monsieur Gruel, cher Président, je vous suis reconnaissant d'entreprendre ce périple depuis Tours. Monsieur Emmerich, je suis fier de pouvoir vous présenter mes travaux, alors que nourrissez ma bibliographie.

**Mme le Professeur Marie-Christine Alessi,**

Je vous remercie de l'honneur que vous m'avez accordé en acceptant de juger ce travail. Je voudrais profiter de cette opportunité pour vous remercier pour l'aide extrêmement précieuse que vous m'avez apportée cette année quand je me débattais avec l'administration. Vous m'avez accordé beaucoup de temps et m'avez grandement aidé à affronter la terrible épreuve de l'Habilitation pédagogique. J'espère que vous ne m'en tiendrez pas rigueur, mais il est très peu question de plaquettes dans ce manuscrit.

Je voudrais par ailleurs adresser mes remerciements à mes maitres. **Madame le Professeur Laurence Camoin**, merci pour votre soutien et vos encouragements pendant l'écriture de ce manuscrit. Je voudrais également saluer les déserteurs : **Corinne Frère** et **Marie-Françoise Aillaud**. Corinne, en bonne Bobo que tu es, je suis certain que tu t'épanouis pleinement à Paris. Madame Aillaud, je sais que je n'ai pas profité assez de votre enseignement. J'essaye cependant de le suivre à la lettre. Une pensée pour votre voisine de bureau...

*A mes collaborateurs,*

Tout d'abord les médecins du CEHT, à l'origine de ces belles publications : **Dominique, Marie-Christine, Katia et la dernière arrivée Céline**. Aux cadres du service ensuite. Il faut ici considérer le terme de façon générique. Merci en particulier à **Marie B et Noémie S** pour leur rigueur et leur participation active à l'activité de recherche. Les résultats de ce manuscrit n'auraient pas pu être publiés sans vous. Et merci à mes techniciennes préférées. Elles se reconnaîtront. Le laboratoire ne serait pas le même sans vous.

A l'unité de recherche et en particulier à **Pablo, Matthias, Anna** (qui est maintenant Bretonne), **Marjorie, Dorsaf** (qui est maintenant ricaine), **Grace**. On passe quand même de chouettes congrès avec vous.

Merci à **Noémie R** avec qui nous avons, de façon étonnante, partagé un congrès également. Tu as activement participé au travail présenté dans ce manuscrit. Merci encore pour tes dernières corrections (j'ai été assez insistant, il est vrai).

**A Denis**. Tu es dans la case des collaborateurs mais tu glisses peu à peu dans la case suivante comme l'attestent nos joutes fifesques. Je ne pourrais pas valider de GPP sans toi. Mais il faut reconnaître que notre tâche la plus ardue concerne un certain TOP 30. J'ai une de ces envies de TRAP-6, moi !

**A Manal**. Tu es sans doute celle qui fait le mieux le lien entre les deux mondes. Voilà déjà quelques années que nous collaborons. Et que nous dégustons les bons petits plats de Georgette. Je suis content d'écrire ces quelques lignes après ta soutenance. Cela me permet de te féliciter de façon très officielle. Même si j'avoue que l'écriture de ton manuscrit ne m'a pas facilité la tâche, je reconnais que ça en valait la peine. Ta thèse était une superproduction. Et je ne parle pas que de la soirée dans les bureaux de Joe. Tu as été brillante. Merci pour ta générosité. Et merci à Joe. Je sais bien qu'il n'est pas étranger à cette dernière.

*A mes amis,*

**A Geoffroy**. Quelle histoire, oui ! Je réponds aux remerciements de ta propre thèse. Tu as eu la gentillesse de me confier ton manuscrit pour me faciliter la vie (merci !!). J'en ai profité pour relire ce que tu m'avais écrit. Je n'ai même pas tenté de vérifier les chiffres mais

d'après ce que tu dis cela doit maintenant faire 12 ans que nous avons copiné. C'est assez vertigineux. J'ai également ressorti le volume 1 de mes remerciements et je disais d'ailleurs à ton sujet que tu es mon plus vieil « ami médical ». Ce titre n'est donc pas usurpé. Comme l'exige notre pacte, tu es cité en premier et tu récoltes toi aussi 32 points SIGAPS. Quoique, il ne s'agit peut-être pas d'une revue de rang A. Après le 18 décembre, je serai aussi gradé que toi et je pourrai légitimement me dire ton aîné hospitalo-universitaire. Tu commences cette année ton assistanat. Ta vie va donc devenir un enfer. Toi qui te la coulais douce ces derniers temps. Finies les pauses manga au bord de la piscine en milieu de journée. Finies les saisons entières de SOA en 2 jours. Place au travail ! En tant qu'aîné, tu pourras toutefois compter sur moi pour te filer la main. Et tant qu'ami tu pourras bien sûr compter sur moi, tout court. Nous ne sommes pas du genre à nous épancher mais j'en profite pour te dire à quel point ton amitié m'est précieuse. Allez, un petit hug à la Joey et Chandler quand même ! Ce n'est pas ce que j'imaginai quand j'apercevais tes bouclettes blondes en amphi. Il faut dire que tu es un sacré personnage. Un personnage haut en couleurs. Adeptes des TV de 4m12, des shonen (je te rappelle que tu es MD-PhD) et des douches froides. Je ne vais pas parler de tes problèmes de « costipation », ni de tes contrefaçons (Liam Nelson, Jude Lao et compagnie). Je vais en revanche remercier **Roulie** pour son amitié, sa gentillesse, sa bonne humeur et son réconfort quand je suis sujet à une petite déprime (tu comprends parfaitement ces moments là). Merci aussi de supporter notre amitié réputée insupportable.

**A Jean Greg.** Bro. Dioudé. Bodhi. Mon evil twin qui a en ras la mèche. Je sais bien que nous nous croisons un peu moins souvent depuis que tu as emménagé dans l'IHU mais tu n'as pas quitté mon cœur. Ça manque sévèrement de beaux butts en ce moment. Mais je suis certain que nous allons nous rattraper. Je ne sais pas si je te l'ai déjà dit (peut-être à l'occasion d'une Grim) mais avant de devenir mon frangin et mon co-scénariste ninja, tu me faisais peur quand je croisais ton chemin à l'internat. Impressionnant le grand rouquin blanc. Depuis, tu m'impressionnes pour d'autres raisons. Mais je n'ai plus peur. Evidemment, je ne peux pas parler de toi sans écrire un mot sur Jen. Merci de m'avoir accepté aussi facilement dans ta vie. Merci de supporter les blagues kaamelottesques. Et surtout, merci pour ta bienveillance. C'est ce qui te caractérise le mieux à mes yeux. Et puis merci pour ta progéniture, le bébé zombie.

**A Yaya et Poa.** Dans la version 1 des remerciements je vous appelais « mes logeurs parisiens ». Depuis, les choses ont quelque peu changé et vous avez réquisitionné ma chambre pour la louer à l'année à un jeune métis. Bon, c'est vrai qu'il est mignon. C'est vrai



aussi qu'il mange dans une assiette à l'effigie du Petit Prince. Je ferais bien une petite partie de Uno après la côte de bœuf, moi ! Pas trop cuite, hein ?

**A Olivier.** Mon bon Olivier, tu es fidèle à mes représentations universitaires. Tu es d'ailleurs fidèle à toutes nos représentations universitaires. Tu es un ami fidèle, voilà tout. Je sais que je n'ai pas été à la hauteur dernièrement mais par chance, tu es magnanime. Merci pour ton amitié et ton espièglerie.

**A Jean-Camille et Sarah.** Je vous dois ma conversion au vidéoprojecteur.

**A Vava.** La première co-interne. La première co-assistante aussi. Tu es rentrée, avec **Amin**, dans le cercle fermé des faiseurs de coulis. J'espère que vous serez des nôtres pour les prochaines sessions. A vos boccas !

A la petite troupe des soirées mondaines (d'ailleurs, je rate celle du jour pour écrire ces lignes, et finaliser le manuscrit). **Anne K** la petite teigne (amatrice de courts métrages un peu particuliers...) et son **Omar**. **Pauline** la cuisinière 3 étoiles, actrice (C'est l'anniversaire dans tous les recoins !), pilote de chaises à roulettes sur le circuit le plus dangereux de l'AP-HM, et son **Roro** (et son **PL**). **Lakhi** et sa **Djé Djé**.

**A Nader et Pierre-Antoine.** Les partenaires de FIFA. Je crois que cette appellation tend à disparaître pour raison de famille.

**A Alexia.** Tu occupes une place particulière dans ces remerciements. Tu es responsable des prémices d'un article présenté dans ce manuscrit. A cette occasion, tu m'as permis de devenir directeur de thèse. Merci pour la confiance que tu m'as accordée alors. Et merci pour tout le travail que tu as fourni. Je vous embrasse Julien et toi. Et bon courage pour ce qui vous attend. Vous allez être parfaits.

**A Maguelonne et Marine.** Les fidèles collaboratrices de mon co-directeur. J'ai vécu un très bel ISTH 2017 grâce à vous. Vous avez insisté pour me soutenir aujourd'hui. Merci !

Allez, je crois que tu mérites un petit paragraphe perso, Lakhdar. Ça veut dire vert. Mais l'orthographe est soumise à débat. D'ailleurs, je sais que tu vas guetter la moindre faute. Je t'ai croisé lors de ton premier stage. Tu étais le binôme d'Antonin. En biochimie. Tu portais alors la même chemise que moi. Mais nous nous sommes surtout trouvés pendant notre assistantat. Une sorte de rencontre en deux temps. Mec, je sais que tu préfères les insultes mais comme le veut la coutume... On croise souvent la phrase suivante dans les thèses : « tu

as changé la face de mon internat ». Mais dans ma thèse de science tu pourras lire : « tu as changé la face de mon assistanat ». Pour tous les burgers trop cuits (ou pas assez), pour tous les lobs au millimètre (et leur contrepartie, les passings dans les couloirs), les dribbles chaloupés, les smooth pool, les pizze 16/9 à 17 euros et les bruschette mozza-pesto, merci. Tu vas me manquer gros. Je sais que je vais profiter encore un peu de ta hargne pendant quelques mois et après ciao mais ça aura été quand même bien sympa. Je sais qu'aujourd'hui tu es content parce que tu vas te faire rincer. Je vois venir le snake Tissot spécial mariage. C'est de bonne guerre. J'en profite pour saluer ta femme. Djé Djé, tu es une perle.

Et puis, je suis heureux d'avoir croisé la route de Julio à fleurs, Florence (qui a failli faire de l'hémostase), Plumelle, Guidur, Sophie, Perrine, Antonin, NicoF...

*A ma famille,*

J'ai profité de ma précédente thèse pour vous dire mon amour. Le noyau familial n'est plus très gros. Il n'est donc pas difficile de savoir à qui s'adressent ces lignes. Je vous embrasse.

Un petit mot, toujours, pour les absents. Je ne vous oublie pas. Pietro, j'ai trouvé une nouvelle partenaire de Scopa ! Je pense à toi à chaque fois que nous ressortons ton jeu de cartes. Auguste, merci pour tous les Lion malgré les menaces d'Anne Marie.

**A mes parents, évidemment.** Je vous dédie cette thèse et mon parcours. Je n'aurais pas pu faire un tel chemin sans votre soutien.

**A ma douce.** Je pourrais te consacrer un manuscrit mais je sais que des petits curieux vont s'empresse de lire les remerciements. Aussi, je vais m'octroyer un peu de pudeur. D'autant que je n'ai pas véritablement besoin de cette tribune pour te dire tout ce que tu représentes pour moi. Néanmoins... Je sais que tu n'aimes pas que je te le dise. C'est pourquoi je te l'écris. Tu auras ainsi le choix de ne pas lire. Tu es un être exceptionnel. Tu es la personne la plus forte que je connaisse. Je sais déjà que tu vas te moquer de la formule mais tu es l'architecte de mon bonheur. Merci d'être à mes côtés.

Et un petit clin d'œil à Catwoman pour finir.

Les amis, je vais rendre la carte étudiant que j'arborais fièrement (pour réclamer un deuxième burger) depuis 13 ans. Comme dirait ma douce : Pfiou !

*Toutes les grandes personnes ont d'abord été des enfants. (Mais peu d'entre elles s'en souviennent.)*

**Antoine de Saint-Exupéry**





# Sommaire

|   |            |
|---|------------|
| Liste des figures et tableaux .....   | 10         |
| Liste des abréviations .....  | 11         |
| Résumé .....  | 12         |
| Summary.....  | 13         |
| Avant-propos .....  | 14         |
| <b>Partie 1 : génétique de la protéine S et maladie thromboembolique veineuse.....</b>  | <b>16</b>  |
| <b>Introduction .....</b>   | <b>17</b>  |
| I. Considérations générales sur la maladie thromboembolique veineuse.....   | 18         |
| 1. Epidémiologie de la maladie thromboembolique veineuse .....  | 18         |
| 2. Physiopathologie de la maladie thromboembolique veineuse.....  | 20         |
| 3. Etiologie de la maladie thromboembolique veineuse .....  | 21         |
| II. Considérations générales sur la protéine S .....  | 48         |
| 1. Fonction de la protéine S .....  | 49         |
| 2. Protéine S et risque thrombotique, diagnostic biologique et problématique .....  | 51         |
| 3. Mutations situées sur PROS1 associées au déficit en protéine S.....  | 53         |
| III. Problématique .....  | 58         |
| <b>Résultats .....</b>  | <b>59</b>  |
| <b>Article 1 “PROS1 genotype phenotype relationships in a large cohort of adults with suspicion of inherited quantitative protein S deficiency” .....</b> | <b>60</b>  |
| <b>Article 2 “Protein S Heerlen mutation heterozygosity is associated with venous thrombosis risk” .....</b>  | <b>71</b>  |
| <b>Partie 2 : facteurs de risque fréquents et évaluation du risque individuel de thrombose ..</b>   | <b>77</b>  |
| <b>Introduction .....</b>   | <b>78</b>  |
| I. Identification de nouveaux variants génétiques associés à la maladie thromboembolique veineuse.....  | 79         |
| 1. Découverte de nouveaux variants par l’approche gène-candidat .....   | 80         |
| 2. Découverte de nouveaux variants par l’approche génome entier ou GWAS .....   | 82         |
| 3. Découverte de nouveaux variants par l’approche génome entier ou GWAS appliquée aux phénotypes intermédiaires.....                                      | 85         |
| II. Problématique .....   | 89         |
| <b>Résultats .....</b>  | <b>91</b>  |
| <b>Article 3 (en cours de soumission) “Common risk factors add to thrombophilia defects to predict venous thromboembolism risk in families” .....</b>     | <b>92</b>  |
| <b>Article 4 “Genetic risk factors for venous thrombosis in women using combined oral contraceptives: update of the PILGRIM study”.....</b>               | <b>114</b> |
| <b>Discussion et conclusion .....</b>   | <b>121</b> |
| <b>Références bibliographiques .....</b>  | <b>134</b> |
| <b>Liste des publications relatives au sujet de la thèse .....</b>  | <b>150</b> |

# Liste des figures et tableaux

**Figure 1 :** Incidence (%) des premiers évènements thrombotiques (TVP et EP) parmi les habitants de Brest, France, par tranches d'âges, en fonction du sexe, sur la période comprise entre le 1er avril 1998 et le 31 mars 1999.

**Figure 2 :** Voie physiopathologique de la thrombogénèse.

**Figure 3 :** Représentation schématique de l'organisation moléculaire des protéines S et C activée.

**Tableau 1 :** Facteurs de risque environnementaux de thrombose veineuse.

**Tableau 2 :** Facteurs de risque génétiques associés à la maladie thromboembolique veineuse.

# Liste des abréviations

**AT** : Antithrombine

**AUC** : Aire sous la courbe

**C4bBP** : C4b-Binding Protein

**EP** : Embolie pulmonaire

**FVL** : Facteur V Leiden

**GLA** : gamma-carboxyglutamique

**HR** : Hazard ratio

**IC95%** : Intervalle de confiance à 95%

**IMC** : Indice de masse corporelle

**KHPM** : Kininogène de haut poids moléculaire

**MLPA** : Multiplex ligation-dependent probe amplification

**MTEV** : Maladie thromboembolique veineuse

**OR** : Odds ratio

**PC** : Protéine C

**PCa** : Protéine C activée

**PS** : Protéine S

**PSL** : Protéine S libre

**PST** : Protéine S totale

**PTG20210A** : Mutation G20210A du gène de la prothrombine

**RR** : Risque relatif

**sEPCR** : Récepteur soluble de la PC

**SHBG** : Sex hormone-binding globulin

**TCA** : Temps de céphaline activée

**TFPI** : Tissue-factor pathway inhibitor

**TRALI** : Transfusion-related acute lung injury

**TVP** : Thrombose veineuse profonde



## Résumé

La maladie thromboembolique veineuse (MTEV) est une pathologie complexe résultant de l'interaction entre des facteurs de risque environnementaux et génétiques. Cinq anomalies constitutionnelles majeures de la coagulation constituent le classique bilan de thrombophilie (BT) : les déficits en antithrombine (AT), protéines C (PC) et S (PS), le facteur V Leiden (FVL) et la mutation G20210A du gène de la prothrombine. De nombreuses mutations situées sur les gènes *SERPINC1*, *PROC* et *PROS1* ont été décrites afin d'expliquer les phénotypes biologiques. Seule la moitié des déficits « présumés » en PS trouvent une explication moléculaire. Cette dernière anomalie du BT apparaît donc extrêmement difficile à appréhender et elle a fait l'objet de deux articles originaux présentés dans ce manuscrit. Dans le premier article, seuls les patients présentant une mutation délétère étaient à risque de MTEV : OR = 1,41 (IC 95% = 1,05-1,89). Seuls des taux inférieurs à 30%, soit bien inférieurs aux valeurs normales, permettaient de dépister les mutations délétères. La mutation Heerlen située sur *PROS1* est réputée neutre. Au demeurant, aucune étude majeure n'a évalué son impact sur le risque de MTEV. Dans le second article, la mutation Heerlen était associée à un risque de MTEV de 6,57 (IC 95% = 4,06-10,64). Le BT n'explique que 5% de l'héritabilité de la pathologie. Les récentes études d'association pangénomiques ont permis d'identifier une vingtaine de polymorphismes associés à la MTEV. Cependant, leur impact dans les familles présentant un facteur biologique de risque est méconnu. De même, l'impact de facteurs environnementaux tels que l'obésité et le tabagisme est mal évalué dans ces familles. Ainsi, nous avons testé dans le troisième article l'association entre 11 polymorphismes sélectionnés, l'obésité, le tabagisme et le risque de MTEV dans 651 familles avec thrombophilie diagnostiquée. La prise en compte de 5 facteurs de risque fréquents à effet faible (obésité, tabac, groupe sanguin et deux polymorphismes situés sur *F11* et *FGG*) en complément du dépistage de l'anomalie familiale permettait de mieux caractériser le risque individuel. Le bénéfice du BT étant principalement établi chez la femme en âge de procréer, nous avons testé la même stratégie dans une population spécifique de femmes sous contraceptifs oraux combinés. Trois facteurs de risque fréquents (groupe sanguin, obésité et un polymorphisme situé sur *F11*) étaient associés à un OR de 13 lorsqu'ils étaient combinés. Au total, la prise en compte de facteurs de risque fréquents à effet modéré, permettait une meilleure évaluation du risque individuel.

## Summary

Venous thromboembolism (VT) is a complex disease resulting from the interaction between environmental and genetic risk factors. Five major inherited hemostatic defects are part of the thrombophilia screening: antithrombin (AT), protein C and S deficiencies, factor V Leiden (FVL) and the prothrombin mutation. Many mutations located within *SERPINC1*, *PROC* and *PROS1* have been identified so far explaining observed phenotypic traits. A molecular defect is identified in only half of assumed PS deficiencies. This latter defect appears to be difficult to assess and has been the subject of 2 original articles presented here. In the first one, *PROS1* genotyping of 579 individuals with suspected inherited PS type I/III deficiency has been performed. Only detrimental mutations (DM) increased VT risk: OR= 1,41 (95% CI = 1.05-1.89). Only free PS levels below 30%, *id est* far below the current normal ranges, enabled the identification of DM. PS Heerlen mutation located within *PROS1* has been considered neutral for a long time. Although, no major study rigorously assessed its impact on VT risk. In the second article, the association between PS Heerlen and VT has been tested in a sample of 4173 patients with VT history and 5970 healthy individuals. PS Heerlen was associated with a 6.57 (95% CI = 4.06-10.64) increased risk of VT. The thrombophilia screening only explains 5% of VT heritability, itself estimated to be 50%. Recent genome wide association studies allowed the identification of nearly 20 polymorphisms associated with a lower risk of VT as compared with the 5 conventional defects (OR below 1.5). However, the impact of such polymorphisms in families with known defects is uncertain. Similarly, the impact of environmental risk factors such as obesity and smoking is unknown in such families. We therefore tested in a third article the association between 11 selected polymorphisms, obesity, smoking and VT in 651 families with known thrombophilia. Considering 5 common risk factors (obesity, smoking, ABO blood group, two polymorphisms located on *FGG* and *F11*) together with the thrombophilia screening resulted in a better assessment of VT risk in individuals from families with thrombophilia. The positive impact of a thrombophilia screening being mainly established in women of childbearing age, we then applied the same strategy in a sample of women using combined oral contraceptives. Three common risk factors (non-O blood groups, obesity and a polymorphism located on *F11*), when combined, were associated with a 13 OR. In conclusion, considering common risk factors improved the individual assessment of VT risk.

# Avant-propos

La maladie thromboembolique veineuse (MTEV) est une pathologie fréquente et potentiellement grave puisqu'il s'agit de la troisième cause de mortalité cardiovasculaire. Nos connaissances physiopathologiques reposent sur les travaux de Virchow conduits au cours du 19<sup>ème</sup> siècle. Il a en effet décrit pour la première fois la MTEV comme une pathologie complexe, c'est-à-dire aux facteurs explicatifs et causaux multiples. Ainsi nous savons que des facteurs environnementaux et génétiques agissant de concert sont responsables de la formation du caillot dans les veines. Ces deux types de facteurs contribuent de façon variable à l'apparition du phénotype maladie.

Les études épidémiologiques nous apprennent que la MTEV est une pathologie du sujet âgé (âge moyen au diagnostic = 60 +/- 20 ans). Ainsi, un diagnostic de MTEV avant l'âge de 40 ans, en l'absence de facteur de risque transitoire majeur, doit faire suspecter un facteur de risque génétique fort. A ce jour, cinq anomalies constitutionnelles de la cascade de la coagulation sont recherchées dans ce contexte : trois déficits en inhibiteurs de la coagulation, antithrombine (AT), protéine C (PC), protéine S (PS), et deux mutations ponctuelles situées sur les gènes des facteurs V et II, facteur V Leiden (FVL) et G20210A (PTG20210A). Ces cinq anomalies constituent ainsi le classique « bilan de thrombophilie » (constitutionnelle) recommandé par les sociétés savantes et remboursé par la Sécurité Sociale. Pour la suite du manuscrit, par souci de simplification, nous utiliserons le terme de « thrombophilie », communément admis par les sociétés savantes, pour désigner ces anomalies biologiques. Ces facteurs génétiques peuvent être classés en deux catégories en fonction de leur épidémiologie et du niveau de risque qu'ils confèrent. Ainsi, les déficits en inhibiteurs (AT, PC, PS) sont rares mais fortement associés à la MTEV (risque relatif (RR) voisin de 10) alors que les deux mutations ponctuelles sont plus fréquentes mais entraînent une augmentation plus modérée du risque de MTEV (RR = 2 à 5).

Contrairement au FVL et à PTG20210A, de nombreuses mutations situées sur les gènes *SERPINC1*, *PROC* et *PROS1* ont été décrites afin d'expliquer les phénotypes mesurés (déficits plasmatiques). Au demeurant, l'analyse systématique des gènes codant pour les protéines concernées ne permet pas d'identifier une mutation dans 100% des déficits constatés. Si une mutation est retrouvée dans la grande majorité des déficits en AT, et dans la majorité des déficits en PC, moins de la moitié des déficits « présumés » en PS trouvent une explication

moléculaire. Cette dernière anomalie du bilan de thrombophilie apparait donc extrêmement difficile à appréhender et elle a fait l'objet de deux articles originaux présentés dans ce manuscrit. Le premier article s'intéresse à la relation entre génotype et phénotype, et en particulier à la notion de seuil en dessous duquel un déficit constitutionnel doit être évoqué. Le second article traite de l'association entre un polymorphisme rare située sur *PROS1* et le risque de MTEV, malgré des taux de PS « normaux ».

Bien que considéré comme indéboulonnable depuis 1996 et la découverte de la dernière anomalie (PTG20210A), le bilan de thrombophilie n'est retrouvé positif que dans 30 à 40% des MTEV. De même, il n'explique que 5% de l'héritabilité de la pathologie, elle-même estimée à 50%. Afin de progresser dans la connaissance de la génétique de la MTEV, de nombreuses études d'associations pangénomiques ont été conduites au cours des dix dernières années. Elles ont ainsi permis d'identifier une vingtaine de nouveaux polymorphismes associés à la MTEV. Ces polymorphismes s'accompagnent d'un plus faible risque de MTEV que les cinq anomalies du bilan de thrombophilie puisque les Odds ratio (OR) qui leur sont associés sont inférieurs à 1,5. C'est pour cette raison qu'ils ne sont à ce jour pas pris en compte dans l'évaluation individuelle du risque de MTEV. Cependant, leur utilisation, en complément du bilan de thrombophilie, dans les familles présentant un facteur biologique de risque, pourrait permettre de mieux appréhender le risque individuel. En effet, à ce jour, seul le dépistage de l'anomalie familiale est utilisé pour évaluer le risque des apparentés, avec des résultats mitigés. En effet, la pénétrance de la MTEV est largement incomplète dans ces familles. A contrario, les apparentés ne présentant pas l'anomalie familiale sont exposés à un sur-risque de MTEV comparativement à la population générale. Ainsi, nous avons testé l'impact de polymorphismes récemment identifiés comme associés à la MTEV en population générale, dans une étude familiale, dont les résultats sont présentés dans un troisième article (en cours de révision). Enfin, parce que la question du bilan de thrombophilie chez les sujets asymptomatiques concerne principalement les femmes en âge de procréer, nous avons testé la même stratégie que celle précédemment énoncée dans une population spécifique de femmes sous contraceptifs oraux combinés (quatrième article).

Nous présenterons dans la première partie de ce manuscrit deux revues de la littérature, en français et en anglais, traitant de la génétique de la thrombose veineuse. Ces revues seront suivies des quatre articles brièvement annoncés ci-dessus. Les principaux résultats de ces travaux seront discutés dans un dernier chapitre.

# **Partie 1 : génétique de la protéine S et maladie thromboembolique veineuse**

# Introduction

# I. Considérations générales sur la maladie thromboembolique veineuse

## 1. *Epidémiologie de la maladie thromboembolique veineuse*

La MTEV regroupe deux entités cliniques liées que sont la thrombose veineuse profonde (TVP) et l'embolie pulmonaire (EP).

Les études rapportent une incidence d'environ 1 cas pour 1000 par an dans la population générale<sup>1,2</sup>. Alors que le sexe ratio global est voisin de 1<sup>3,4</sup>, l'analyse par tranches d'âge révèle que l'incidence est discrètement supérieure chez la femme jeune (avant 40 ans), et légèrement plus élevée chez les hommes au-delà de cet âge<sup>1</sup>.

De même, il existe une hétérogénéité ethnique. Dans une vaste étude épidémiologique portant sur près de 25000 américains, White et al. ont décrit des disparités en termes d'incidence selon l'origine ethnique. Les individus d'origine hispanique ou asiatique présentaient ainsi une incidence de MTEV plus faible que les individus d'origine caucasienne (respectivement 139 et 60 pour 1000000 versus 230 pour 1000000) alors que l'incidence chez les individus d'origine africaine était plus élevée (293 pour 1000000)<sup>5</sup>. Ces différences ne semblaient pas expliquées par les facteurs environnementaux qui étaient recueillis et pour lesquels il n'existait pas de différence significative. Une origine génétique a donc été suggérée. Il existe par exemple des différences importantes de prévalence pour le FVL : la mutation est quasiment absente des populations asiatique et africaine alors que sa prévalence est estimée à 5% chez les caucasiens<sup>6</sup>.

L'âge moyen de la maladie est de  $60 \pm 20$  ans. La figure 1 montre que son incidence augmente de façon importante après 75 ans, atteignant alors 1% par an<sup>4</sup>. Elle s'associe à une mortalité de 10%<sup>1</sup>. Environ 10000 morts sont probablement attribuables à l'EP chaque année en France. La morbidité de la MTEV représente par ailleurs un problème de santé publique. La MTEV se complique d'une part d'une maladie post-thrombotique dans environ 25% des cas après 5 ans de suivi<sup>7</sup>, et d'autre part les récurrences sont fréquentes avec une incidence annuelle de 5 à 7 % en cas de MTEV idiopathique<sup>8-11</sup>.

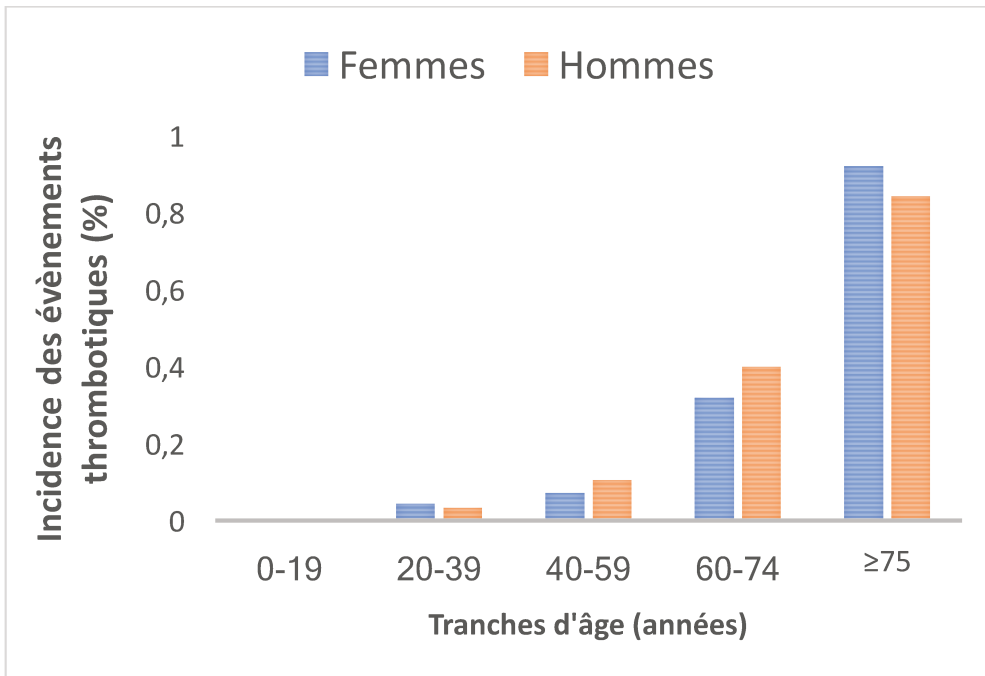


Figure 1 : Incidence (%) des premiers évènements thrombotiques (TVP et EP) parmi les habitants de Brest, France, par tranches d'âges, en fonction du sexe, sur la période comprise entre le 1er avril 1998 et le 31 mars 1999. D'après les données de Oger et al.<sup>4</sup>

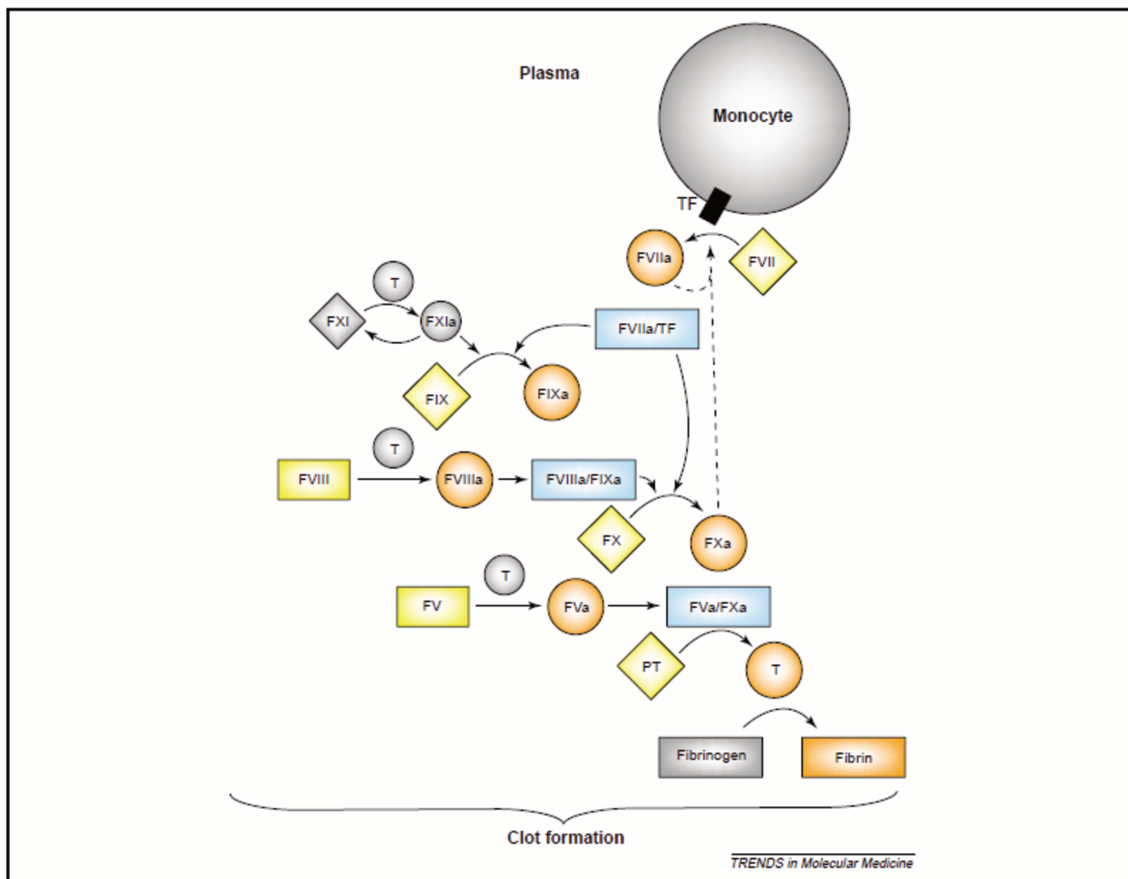


Figure 2 : Voie physiopathologique de la thrombogénèse. D'après Furie et Furie<sup>12</sup>.  
TF : Facteur tissulaire ; PT : Prothombine ; T : Thrombine.



## *2. Physiopathologie de la maladie thromboembolique veineuse*

La MTEV est une pathologie complexe résultant de l'interaction entre des facteurs génétiques et des facteurs environnementaux <sup>3</sup>.

C'est Rudolf Virchow qui a décrit le premier, en 1856, la complexité de cette pathologie à travers la triade qui porte son nom. La formation du thrombus veineux résulte de l'interaction entre les trois facteurs suivants :

- un ralentissement du flux sanguin veineux,
- une lésion pariétale,
- un état d'hypercoagulabilité.

Cette hypothèse physiopathologique s'est imposée pendant plus d'un siècle. En 1974, Simon Sevitt introduit le concept de « Nidus », lié à un flux turbulent qui se crée en aval des valvules veineuses et qui favoriserait l'agrégation plaquettaire, sans que l'altération de la paroi vasculaire soit nécessaire <sup>13</sup>. Le thrombus initial se formerait ainsi dans le repli valvulaire puis d'autres thrombi se juxtaposeraient par strates jusqu'à oblitérer complètement la veine. Contrairement à la physiopathologie de la thrombose artérielle, l'endothélium sous-jacent ne serait pas altéré et l'activation de la coagulation précéderait l'agrégation des plaquettes expliquant ainsi l'efficacité limitée des traitements antiagrégants plaquettaires dans la prévention de la MTEV.

Néanmoins, si l'endothélium n'est pas altéré dans sa structure, il l'est dans sa fonction. En effet, trois mécanismes sont aussi impliqués : l'activation des monocytes et des macrophages, les sélectines et leurs ligands, les microparticules et la sécrétion du facteur tissulaire.

La glycoprotéine P-selectine, transloquée à la surface des cellules endothéliales et des plaquettes activées, permet l'interaction de ces cellules avec les polynucléaires neutrophiles et les monocytes par l'intermédiaire de la liaison à son ligand, le PSGL-1 <sup>14</sup>. Les microparticules dérivées des monocytes ont à leur surface le facteur tissulaire et le PSGL-1. La fusion des microparticules et des plaquettes entraîne une activation de la coagulation à la surface des plaquettes et au niveau des cellules endothéliales en transférant le facteur tissulaire <sup>12</sup>, concourant ainsi à la formation du thrombus par activation de la cascade de la coagulation schématisée en figure 2. Outre les réactions d'activation successives se

déroulant depuis l'exposition du facteur tissulaire jusqu'à la formation de thrombine et de fibrine, deux voies d'activation complémentaires de cette cascade ont été récemment décrites.

Alors que le rôle du facteur tissulaire dans l'initiation de la coagulation via la voie extrinsèque est bien connu, son implication dans la voie intrinsèque a été démontrée récemment. Le complexe formé par le facteur tissulaire et les facteurs VII et X activés permettrait en effet d'activer le facteur VIII, indépendamment de la thrombine<sup>15</sup>.

Le système contact faisant intervenir 3 acteurs, le facteur XII, la prékallïcérine et le kininogène de haut poids moléculaire, n'est pas indispensable à la formation du caillot, comme l'atteste l'absence de syndrome hémorragique chez les patients déficitaires. Au demeurant, le système contact joue un rôle essentiel dans la stabilisation et la croissance du thrombus<sup>16</sup>. Le système contact est initié lorsque le facteur XII lie des composés polyanioniques, générant son activation autocatalytique. Le facteur XII activé active alors le facteur XI. Pendant longtemps, les seuls activateurs connus du système contact étaient les surfaces artificielles mais des activateurs naturels ont récemment été identifiés : les polyphosphates représentés par l'ADN et l'ARN libérés par les cellules lésées ou les « neutrophil extracellular traps » (NETs). Ces activateurs générés sur des sites lésions endothéliales renforcent le stimulus lié au facteur tissulaire<sup>16</sup>.

Le thrombus constitué évolue vers la recanalisation plus ou moins complète et le développement d'une circulation veineuse collatérale. La lyse du thrombus peut s'accompagner d'une lyse des valvules à l'origine de la maladie post-thrombotique par incontinence valvulaire chez 20 à 50% des patients<sup>17</sup>.

La complication immédiatement redoutée de la TVP est l'EP, qui correspond à la thrombose d'une artère pulmonaire après migration du thrombus veineux. Une TVP est ainsi retrouvée chez environ 50% des patients présentant une EP<sup>18</sup>.

### *3. Etiologie de la maladie thromboembolique veineuse*

De façon schématique, on considère que la MTEV résulte de l'interaction entre deux types de facteurs de risque, environnementaux et génétiques. Cette complexité a notamment été

illustrée par Rosendaal <sup>19</sup>. Nous détaillerons ci-dessous ces deux catégories de facteurs de risque.

#### a. Facteurs de risques environnementaux

On distingue classiquement les épisodes thromboemboliques dits « provoqués » par des facteurs transitoires ou persistants, des épisodes « non provoqués » qui représentent environ un tiers des épisodes et pour lesquels aucun facteur de risque transitoire n'a été identifié <sup>20</sup>.

Parmi les facteurs transitoires, le cancer et l'immobilisation/chirurgie sont des facteurs majeurs. Ils sont en effet retrouvés dans 20% et 15% des MTEV respectivement (25,26).

L'obésité et le tabagisme constituent des enjeux majeurs de santé publique. Le premier facteur (en partie environnemental) est depuis longtemps considéré comme un facteur de risque de MTEV. De nombreuses études ont permis d'estimer le risque associé à l'obésité, définie par un indice de masse corporelle (IMC) supérieur à 30 kg.m<sup>-2</sup>. Ainsi, le risque apparaît multiplié par 2 ou 3 chez les sujets obèses <sup>22-25</sup>. Dans une étude, le risque apparaissait indépendant des taux de fibrinogène, facteurs VIII et IX et de D-dimères <sup>24</sup>.

Le tabagisme a en revanche fait l'objet de controverses. En effet, la plupart des études ont décrit le tabagisme comme un facteur de risque modéré ou neutre de MTEV <sup>26-31</sup>. Quelques rares études ont même observé un effet protecteur chez les fumeurs <sup>32,33</sup>. Ce n'est que récemment que deux méta-analyses ont permis de conclure à une augmentation du risque de MTEV chez les fumeurs <sup>34,35</sup> : les fumeurs actifs présentaient des RR estimés à 1,23 (intervalle de confiance à 95% (IC95%) = 1,14-1,33) et 1,19 (IC95% = 1,08-1,32), respectivement dans les études de Cheng et de Mahmoodi <sup>34,35</sup>.

Les facteurs de risque environnementaux peuvent être classés selon la triade de Virchow. Ces facteurs de risque sont ainsi recensés selon cette classification dans le tableau 1 <sup>36</sup>.

| Hypercoagulabilité                               | Lésions vasculaires                 | Stase veineuse, immobilisation                 |
|--|-------------------------------------|--|
| Age avancé                                       | Chirurgie                           | Hospitalisation pour pathologie médicale aiguë |
| Cancer actif                                     | Traumatisme, fracture               | Alitement                                      |
| Syndrome des antiphospholipides                  | Cathéter veineux central, pacemaker | Voyage long courrier de plus de 4 heures       |
| Oestrogènes                                      |                                     | Parésie, paralysie                             |
| Grossesse et post-partum                         |                                     |  |
| Histoire personnelle ou familiale de MTEV        |                                     |  |
| Obésité  |                                     |  |
| Pathologie autoimmune ou inflammatoire chronique |                                     |  |
| Thrombopénie induite par l'héparine              |                                     |  |

Tableau 1 : Facteurs de risque environnementaux de thrombose veineuse. D'après les données de Di Nisio et al. <sup>36</sup>

#### b. Facteurs de risque génétiques : hérédité

La composante génétique de la MTEV est très forte puisque les analyses de ségrégation familiale ont montré que son hérédité est supérieure à 50 % <sup>37-39</sup>.

De nombreuses études épidémiologiques ont rapporté une augmentation du risque de MTEV chez les patients présentant une histoire familiale. De façon intéressante, ce risque persiste après prise en compte des facteurs de risque biologiques génétiques classiques <sup>40-45</sup>. La notion d'antécédent familial constitue ainsi un facteur de risque indépendant. Par ailleurs, la prise en compte de l'âge de survenue de l'épisode thrombotique ainsi que le nombre d'apparentés atteints permet de préciser le niveau de risque dans les familles. D'après Bezemer, comparativement aux sujets ne présentant pas d'antécédent familial de MTEV, l'existence d'un antécédent familial est associée à un OR de 2,2 (IC95% = 1,9-2,6) et ce risque apparaît augmenté si l'apparenté a présenté un épisode thrombotique avant l'âge de 50 ans (OR = 2,7 ; IC95% = 1,7-3,4) <sup>42</sup>. Le risque apparaît encore plus important si une histoire thrombotique est retrouvée chez plus de deux apparentés : OR = 3,9 (IC95% = 2,7-5,7) <sup>42</sup>.

A ce jour, seule une fraction de ces facteurs génétiques a pu être caractérisée. Une partie d'entre eux s'intègre dans le cadre du bilan de thrombophilie, comme l'indiquent les

experts du Groupe d'Etude en Hémostase et Thrombose (GEHT) et de la Société Française de Médecine Vasculaire (SFMV) dans leurs recommandations de 2009 pour la recherche de facteurs biologiques de risque de MTEV en pratique clinique <sup>46</sup> : déficits en AT, PC et PS, FVL et PTG20210A.

Ces différentes anomalies biologiques associées à la MTEV sont détaillées dans deux revues de la littérature présentées ci-dessous. La première est parue en 2016 dans une revue française à visée éducative. La seconde est parue en 2015 dans le journal international à comité de lecture *Thrombosis and Haemostasis*.

**Article de revue 1**  
**« Anomalies constitutionnelles de la  
coagulation prédisposant à la  
thrombose »**

*EMC – Hématologie 2016*

# Anomalies constitutionnelles de la coagulation prédisposant à la thrombose

P. Suchon, P.-E. Morange

La thrombose veineuse est une pathologie complexe qui résulte de l'interaction entre des facteurs génétiques et des facteurs environnementaux. Parmi les facteurs génétiques de risque, cinq anomalies constitutionnelles de l'hémostase sont classiquement reconnues et recherchées au cours du bilan de thrombophilie : les déficits en antithrombine, protéine C et protéine S ainsi que deux mutations (facteur V Leiden et mutation G20210A du gène de la prothrombine). La part d'héritabilité dans la thrombose veineuse est forte puisqu'elle atteint presque 60 %, mais ces cinq anomalies ne permettent d'expliquer qu'une faible partie de cette héritabilité. Les dix dernières années ont vu se multiplier les approches et les études génétiques pour l'identification de nouveaux marqueurs génétiques de risque. Elles ont ainsi permis d'identifier de nombreux polymorphismes associés à la thrombose veineuse. Comparativement aux anomalies du bilan de thrombophilie, ces polymorphismes confèrent une augmentation modérée du risque de thrombose veineuse, mais leur fréquence est très élevée dans la population générale.

© 2016 Elsevier Masson SAS. Tous droits réservés.

**Mots-clés** : Thrombose veineuse; Thrombophilie; Coagulation; Polymorphisme; Étude pangénomique

## Plan

|  |   |
|--|---|
| ■ Introduction   | 1 |
| ■ Thrombophilie classique                                    | 2 |
| Déficits en inhibiteurs                                      | 2 |
| Facteur V Leiden (FVL) et résistance à la protéine C activée | 3 |
| Mutation G20210A du gène de la prothrombine                  | 4 |
| Groupe sanguin   | 5 |
| Hyperhomocystéinémie   | 5 |
| ■ Nouveaux variants  | 6 |
| Approches gène-candidat                                      | 6 |
| Approches génome-entier ou GWAS                              | 6 |
| Variants rares   | 7 |
| ■ Perspectives   | 8 |

25 % des patients atteints de TV après 5 ans de suivi [2]. L'incidence des récurrences se situe autour de 5 à 7 % par an, en cas de TV idiopathique [3-6]. L'identification des patients à risque de TV et sa prévention constituent donc un réel enjeu en matière de santé publique, en particulier pour les populations vieillissantes.

## “ Point fort

La TV est une pathologie multifactorielle résultant de l'interaction entre des facteurs génétiques et environnementaux. De nombreux facteurs génétiques ont à ce jour été identifiés.

## ■ Introduction

La thrombose veineuse (TV), c'est-à-dire la thrombose veineuse profonde (TVP) ou l'embolie pulmonaire (EP), est une pathologie fréquente et potentiellement mortelle. En France, l'incidence des TVP et EP est respectivement de 1,2 et 0,6 pour 1000 par an [1]. L'âge moyen de la maladie est de 60 ± 20 ans et son incidence annuelle augmente de façon importante après 75 ans, atteignant alors 1 % [1]. La TV reste une cause majeure de décès en France puisque environ 10 000 morts sont probablement attribuables à l'EP chaque année. La maladie post-thrombotique est également un problème majeur de santé publique, survenant chez environ

La TV apparaît habituellement lors de situations à risque comme la chirurgie, l'hospitalisation, la grossesse, le post-partum, les traitements hormonaux (contraception et traitement hormonal substitutif de la ménopause), les cancers et les pathologies auto-immunes [7]. Cependant, les études familiales et de jumeaux indiquent que cette maladie présente une forte composante génétique (estimée à environ 60 %) [8, 9].

Depuis la description par Egeberg du déficit familial en antithrombine (AT), les progrès des techniques de biologie moléculaire ont permis d'identifier des dizaines de polymorphismes potentiellement associés à la TV. Les facteurs génétiques les plus récemment identifiés ne confèrent qu'une augmentation très

modérée du risque de thrombose. Ils s'intègrent cependant dans la compréhension globale de la physiopathologie de la TV, pathologie complexe qui résulte de l'interaction entre facteurs génétiques et environnementaux. Alors qu'il est clairement établi que les anomalies constitutionnelles de l'hémostase sont associées à la TV, les preuves de cette association sont beaucoup moins fortes pour la thrombose artérielle. C'est pour cette raison que nous ne traitons dans cet article que de la TV.

## ■ Thrombophilie classique

### Déficits en inhibiteurs

#### “ Point fort

Les déficits en inhibiteurs (AT, protéine C [PC], protéine S [PS]) sont rares mais associés à un risque élevé de TV.

### Déficit en antithrombine

Le déficit en AT est le premier déficit dont le lien avec la TV ait été prouvé. C'est Egeberg qui l'a décrit en 1965 : des taux diminués d'AT ont été retrouvés chez dix sujets apparentés, dont six avaient présenté une TV [10]. Le déficit était ainsi corrélé au phénotype maladie et de transmission verticale. L'AT est une glycoprotéine de 58 kDa synthétisée par les hépatocytes, appartenant à la famille des inhibiteurs de sérine protéase (ou serpins). Elle joue un rôle majeur dans la cascade de la coagulation en inhibant les facteurs activés de la coagulation. Ses deux principales cibles sont les facteurs Xa et IIa (ou thrombine). Les membres de la famille des serpins partagent des caractéristiques structurales leur permettant de former un complexe équimoléculaire irréversible avec leur cible sérine protéase. L'activité anticoagulante de l'héparine est liée à l'AT : l'héparine potentialise la réaction entre l'AT et la thrombine (action  $\times 1000$  environ). D'autres glycosaminoglycans tels que les héparanes sulfate présents à la surface endothéliale possèdent cette propriété. Deux sites fonctionnels sont décrits sur l'AT : le site réactif (arginine 393 et sérine 394) et le domaine de liaison à l'héparine (acides aminés [AA] 41 à 49 et 107 à 156). La concentration plasmatique en AT est d'environ 0,2 mg/ml et sa demi-vie est de 3 jours. Elle est diminuée par les estrogènes et l'héparine.

Le déficit en AT est de transmission autosomique dominante et s'accompagne d'une augmentation majeure du risque de TV. Sa prévalence dans la population générale est estimée à 0,02 % alors qu'elle est de 1 à 2 % chez les patients ayant présenté un épisode de TV. La présence d'un tel déficit confère un risque de TV estimé à 1 % par an, soit un risque relatif (RR) de 10 à 20 [11-14]. En dehors du déficit de type IIHBS, caractérisé par une mutation du site de liaison à l'héparine, il n'existe pas de déficits homozygotes (HMZ), ces derniers étant associés à des morts fœtales in utero [15].

Le diagnostic de déficit en AT repose sur son dosage plasmatique. La combinaison de deux techniques de dosage permet de distinguer deux types de déficit : le déficit quantitatif ou déficit de type I et le déficit qualitatif ou déficit de type II. Dans le déficit de type I, la protéine produite est fonctionnelle mais son taux est diminué. Cela se traduit par une diminution de son activité dosée par l'activité cofacteur de l'héparine et une diminution associée de la protéine en dosage immunologique. Dans les déficits de type II, les dosages immunologiques permettent de retrouver des concentrations protéiques normales (ou subnormales) alors que l'activité cofacteur de l'héparine est diminuée. La réalisation d'un test chromogénique non dépendant de l'héparine permet de distinguer le type de déficit qualitatif. L'anomalie fonctionnelle peut affecter le site réactif (IIRS), le site de liaison à l'héparine (IIHBS) ou les deux (effet pléiotropique IIPE) [16]. Dans le premier cas, l'activité de l'AT dosée par la technique chromogénique apparaît diminuée. La nécessité de typer le déficit est liée à la différence de risque

associée au type IIHBS. En effet, ce dernier déficit s'accompagne d'un risque thrombotique beaucoup plus faible [17], en dehors des déficits HMZ pour lesquels le tableau clinique est sévère.

### Bases moléculaires

Le gène de l'AT (*SERPINC1*) est porté par le bras court du chromosome 1 ; il comprend sept exons et s'étend sur 13,4 kpb. La protéine mature comporte 432 AA. Le déficit de type I résulte de la perte d'expression d'un des deux allèles. La séquence peptidique Arg393-Ser394 constitue le site réactif, elle est située à l'extrémité C-terminale de la protéine et codée par l'exon VI. Le site de liaison à l'héparine (HBS) est situé à l'extrémité N-terminale, il est codé par les exons II et IIIa. Les déficits de type IIRS sont expliqués par deux types de mutations. Lorsque les mutations affectent les AA 382 ou 384, l'AT se lie à la thrombine et est clivée par cette dernière, entraînant un changement conformationnel. L'AT n'agit ainsi plus comme un inhibiteur mais comme un substrat de la thrombine [18]. Lorsque la mutation concerne les AA 392, 393, 394, l'AT ne peut être reconnue par sa cible sérine protéase [19]. Le déficit de type IIHBS est le plus souvent expliqué par des mutations faux-sens des AA 47 et 129. Des mutations situées sur les AA 402, 404, 407 ou 429 se traduisent par un déficit de type IIPE. Au total, près de 400 mutations sont recensées dans la base de données de l'Imperial College Faculty of Medicine ([http : //www1.imperial.ac.uk/departementofmedicine/divisions/experimentalmedicine/haematology/coag/antithrombin](http://www1.imperial.ac.uk/departementofmedicine/divisions/experimentalmedicine/haematology/coag/antithrombin)). Les mutations associées au déficit de type IIHBS sont exclusivement des mutations faux-sens. Pour le reste des déficits, plus de 60 % des mutations décrites jusqu'alors sont des mutations faux-sens. Le risque de TV associé aux différents déficits en AT et aux types de mutations, à l'exclusion du type IIHBS qui est reconnu moins thrombogène, est mal évalué. Une étude a entrepris ce travail [20]. Elle comportait 377 patients déficitaires répartis comme suit : type I = 50,7 % ; type IIHBS = 30,0 % ; type IIHRS = 11,7 % ; type IIPE = 7,7 %. Le risque de TV était le plus important pour les déficits de type I et de type IIPE. L'étude confirme un risque plus faible de TV associé au déficit de type IIHBS (RR ajusté pour l'âge, le sexe et le statut familial = 0,23 ; intervalle de confiance [IC] à 95 % : 0,13-0,41) comparativement aux autres types de déficits en AT. Par ailleurs, le risque de TV était plus faible chez les patients présentant une mutation faux-sens comparativement aux mutations non-sens : RR ajusté = 0,39 (IC à 95 % : 0,27-0,58). La gravité de l'épisode semble également dépendre du type de déficit, puisque les déficits de type II se compliquent plus rarement d'une EP que les déficits de type I (RR = 0,60 ; IC à 95 % : 0,56-0,84). Ces résultats suggèrent que le déficit en AT ne doit plus être considéré comme une entité clinico-biologique unique, mais comme un ensemble de pathologies présentant des analogies mais des risques de TV propres. Dès lors, le recours à la biologie moléculaire pourrait permettre de mieux appréhender le risque de TV associé au déficit en AT diagnostiqué par les tests d'hémostase. Parallèlement aux types de mutation et de déficit, il existe une relation dose-réponse entre les taux plasmatiques d'AT et le risque de TV [21]. Ceci suggère l'implication de polymorphismes dans la modulation des taux d'AT. Conformément à cette idée, le polymorphisme rs2227589 du gène *SERPINC1* est associé aux taux d'AT et au risque de TV [22, 23]. Une étude de cohorte a toutefois évalué le risque de récurrence associé au taux d'AT. Le RR de récurrence était évalué dans deux groupes de patients comparativement aux patients avec des taux en AT supérieurs à 80 %. Le RR associé à des taux compris entre 70 et 80 % était estimé à 2,40 (IC 95 % : 1,51-3,80) alors que les patients qui présentaient des taux inférieurs à 70 % avaient un RR de récurrence de 3,48 (IC 95 % : 2,16-5,61) [24].

### Déficits en protéine C et protéine S

La PC est une glycoprotéine vitamine-K dépendante de 62 kDa synthétisée par l'hépatocyte. Sa demi-vie est très courte (6-8 h) et sa concentration plasmatique est de 3 à 5 mg/l. La PC possède une chaîne légère dont l'extrémité N-terminale contient neuf résidus d'acide gamma-carboxy-glutamique (GLA) nécessaires à sa fixation aux phospholipides membranaires en présence de calcium. Elle possède par ailleurs une chaîne lourde qui présente un site de clivage pour la thrombine et un site catalytique. Elle est activée



en PC activée (PCa) à la surface des cellules endothéliales par le complexe thrombine-thrombomoduline. La liaison de la PC à son récepteur endothélial (EPCR) majeure son activation. En présence de son cofacteur, PS, la PCa inactive le facteur V activé (FVa), en deux étapes. Elle clive tout d'abord le FVa au niveau d'un résidu arginine en position 506, puis le clivage d'un second résidu arginine en position 306 permet d'inactiver complètement le FVa. Cette seconde étape est fortement dépendante de la PS [25]. La PC inactive par ailleurs le FVIIIa au cours d'une double étape (clivage des arginines 562 puis 336) dépendantes de la PS et du FV qui agissent comme des cofacteurs [26].

La PS est également une glycoprotéine vitamine-K dépendante de 69 kDa, de synthèse majoritairement hépatique. Une partie de sa production est toutefois endothéliale et plaquettaire. Sa concentration plasmatique est de 20 à 25 mg/l et sa demi-vie est de 42 heures. La PS circule dans le plasma sous deux formes : libre (40 %) ou liée à la C4bBP (60 %). Seule la PS libre a une activité cofacteur de la PCa qui consiste à augmenter l'affinité de la PCa pour les phospholipides chargés négativement. La PS posséderait d'autres effets anticoagulants, indépendants de la PCa. La PS inhiberait en effet le complexe prothrombinase en se liant aux FVa et FXa [27, 28]. Par ailleurs, elle interviendrait dans la voie de l'inhibition du FXa par le *tissue factor pathway inhibitor* (TFPI) [29].

### Clinique

Les déficits en PC et PS sont associés à un risque majeur de TV. La transmission est autosomique dominante et les épisodes thrombotiques sont souvent précoces (< 45 ans). Chez les sujets hétérozygotes (HTZ), le risque de présenter un épisode de TV avant l'âge de 45 ans est estimé à 50 % [30, 31]. Plusieurs études ont évalué le risque de TV associé aux déficits en PC et PS dans des familles de déficitaires. L'incidence annuelle estimée était de 0,4 à 1 % et correspondait à des odds ratio (OR) compris entre 7 et 10 [11, 12, 14]. L'incidence annuelle de récurrence observée était de 5,1 et 6,5 %, respectivement pour les déficits en PC et PS, ce qui ne semble pas différent de l'incidence attendue en l'absence de thrombophilie [32]. Chez les sujets HMZ, le phénotype clinique est extrêmement sévère et précoce. Il se traduit classiquement par un purpura fulminans néonatal avec plaques de nécrose cutanée [33-35]. Lorsqu'il est impossible de doser les inhibiteurs chez le nouveau-né, le diagnostic pourra être évoqué en les dosant chez les parents, qui sont généralement HTZ et dont les taux sont voisins de 50 %.

Il existe une corrélation entre les taux plasmatiques en inhibiteurs et le risque de TV. Dans une étude de Spek et al. chez 240 patients ne présentant pas de déficit, trois polymorphismes identifiés dans la région du promoteur du gène de la PC semblaient associés aux taux de PC [36]. Le taux moyen de PC était de 103 %. Les patients HMZ pour le génotype CGT présentaient un taux moyen de 94 % alors que les HMZ pour le génotype TAA avaient des taux en moyenne plus élevés : 116 %. De façon intéressante, le risque de thrombose associé au génotype CGT et donc aux taux les plus élevés était plus important (OR = 1,6 ; IC 95 % : 1,0-2,5) comparativement au génotype TAA.

Concernant le déficit en PS, chez des patients non sélectionnés, il semble que seuls les déficits profonds soient associés à la TV. En effet, des résultats de l'étude cas-témoins MEGA montrent que des taux de PS inférieurs au 2,5<sup>e</sup> percentile ne permettaient pas d'identifier les patients à risque de TV. Seul un seuil fixé au 0,1<sup>e</sup> percentile, soit 33 UI/dl, permettait d'observer une augmentation non significative de risque (OR = 5,4 ; IC à 95 % : 0,61-48,8) [37].

### Diagnostic biologique

Trois types de tests sont disponibles et permettent de distinguer deux types de déficits en PC : les déficits quantitatifs de type I et les déficits qualitatifs de type II, eux-mêmes classés en déficits de type IIAC (activité anticoagulante) et IIAM (activité amidolytique). Seul le dosage chromométrique mesurant l'activité anticoagulante de la PC permet de dépister tous les déficits. C'est donc la technique à réaliser en première intention. En cas de taux diminué avec cette première technique, un dosage antigénique permet de déterminer s'il s'agit d'un déficit de type I (plus de 80 % des déficits en PC), où les deux techniques sont concordantes, ou de type II, où l'activité est diminuée alors que l'antigène est

normal. Dans ce dernier cas, le dosage chromométrique de l'activité amidolytique permet de typer le déficit qualitatif : l'activité amidolytique n'est diminuée que dans le type IIAM.

La classification des déficits en PS comprend trois catégories : les types I et III, quantitatifs et les plus fréquents (80 % et 15-20 % respectivement), et le déficit de type II, qualitatif et très minoritaire (0,1-5 %) [38, 39]. Comme pour le PC, trois tests permettent de typer le déficit : un test chromométrique, mesurant l'activité de la PS dépendante de la PC, et deux tests immunologiques, dosant respectivement les taux de PS libre et totale. Théoriquement, le premier test permet de dépister tous les types de déficit en PS. Mais s'il est vrai que sa sensibilité est très bonne, sa spécificité est médiocre (de l'ordre de 50 %), ce qui se traduit dans certaines séries par 10 à 15 % de faux positifs [40, 41]. En pratique, le dosage de la PS libre est utilisé en première intention, en tant que test fonctionnel. S'il est anormal, le dosage de l'activité est réalisé dans un second temps et permet de confirmer le diagnostic [41]. Le déficit de type I est défini par une diminution de la PS quelle que soit la technique de dosage. Le type II associe une PS libre antigène et une PS totale antigène normales et une activité diminuée. Enfin, le type III est caractérisé par une PS libre diminuée et une PS totale normale. Les déficits de types I et III sont en fait deux expressions phénotypiques d'une même entité génétique [31].

### Bases moléculaires

Le gène de la PC est situé sur le chromosome 2q13-q14 ; il s'étend sur plus de 11 kb et comprend neuf exons, dont les exons II à IX correspondent à la région codante. Plus de 160 mutations ont été décrites et sont recensées sur la base de donnée de l'International Society on Thrombosis and Haemostasis (ISTH [www.ISTH.org]) [42]. Les mutations à l'origine des déficits de type I sont principalement des mutations ponctuelles faux-sens. Une grande variété de mutations responsables de déficits de type II a été décrite. Parmi elles, celles qui affectent le domaine GLA se traduisent par une diminution de l'activité anticoagulante associée à une activité amidolytique normale. Des mutations du site de clivage de la thrombine et du site actif sont également responsables de déficits de type II.

La PS est codée par le gène *PROS1*, situé sur le chromosome 3p11, qui s'étend sur plus de 80 kb et comprend 15 exons [43]. Il possède un gène homologue (*PROS2*) dont les caractéristiques moléculaires semblent en faire un pseudogène. Plus de 200 mutations sont recensées dans la base de données du gène *PROS1* de l'ISTH. Il n'existe pas de point chaud de mutation (*hot spot*) et les mutations recensées sont réparties sur tout le gène. Il s'agit le plus souvent de mutations faux-sens mais des délétions/insertions, des mutations non-sens ainsi que des mutations du site d'épissage ont par ailleurs été décrites.

### “ Point fort

La pilule et la grossesse diminuent les taux de PS. Seuls les déficits profonds sont associés à la TV. Dans l'étude MEGA, seul un seuil fixé au 0,1<sup>e</sup> percentile, soit 33 UI/dl, a permis d'observer une augmentation du risque de TV.

### Facteur V Leiden (FVL) et résistance à la protéine C activée

#### Physiopathologie

Le FV est une glycoprotéine de 330 kDa dont la concentration plasmatique est de 20 nmol/l (0,007 g/l) [44, 45]. Il est au cœur de la cascade de la coagulation. Il constitue, avec le FX, le complexe prothrombinase, quand ils sont tous deux activés (en association avec les phospholipides situés sur la membrane plaquettaire). À ce titre, le FV possède une activité procoagulante. Mais il peut être anticoagulant en tant que cofacteur de l'activité de la PCa. L'identification de la mutation Leiden du FV (FVL) n'a été possible

qu'en deux étapes. En 1993, Dahlbäck a décrit la résistance à la PCa (RPCA), dont la base moléculaire ne fut établie que 1 an plus tard par Bertina [46, 47].

Le gène du FV est situé sur le chromosome 1 (1q23), il comporte 25 exons. Le FV possède six domaines répartis comme suit : la chaîne lourde, constituée des domaines A1 et A2, est séparée par la zone de connexion (domaine B) de la chaîne légère, elle-même constituée des domaines A3, C1 et C2.

Le FV ne devient actif (FVa cofacteur du FXa) qu'après avoir subi une série de coupures du domaine B, réalisées par la thrombine et/ou le FXa. Le domaine B est ainsi dissocié et des liaisons non covalentes permettent aux chaînes lourde et légère de s'assembler.

Le FVa peut dans un second temps être inactivé sous l'action de la PCa. Ce sont les résidus arginine en position 306, 506, 679 et 994 qui sont les cibles de la PCa et c'est l'arginine 506 qui est clivée la première. Pour être parfaitement inactivé, le FVa doit subir un deuxième clivage, celui de l'arginine 306, au cours d'une étape très dépendante de l'activité cofacteur de la PS.

La substitution d'une guanine par une adénine en position 1691 du gène du FV se traduit par la substitution de l'arginine 506 par une glutamine (FVL). La cible primordiale de la PCa disparaît alors et la dégradation du FVa, dont l'activité est procoagulante, est retardée. De façon réciproque, la forme clivée anticoagulante ne peut apparaître.

## Épidémiologie

L'épidémiologie du FVL se caractérise par un gradient nord-sud. La mutation est très fréquente en Europe du Nord alors qu'elle est rare, voire absente en Asie et en Afrique. Cette disparité en termes de prévalence semble trouver son origine dans la théorie de l'*Out of Africa*. Selon cette théorie, l'homme moderne serait sorti d'Afrique il y a 60 000 à 125 000 ans, notamment à la conquête de l'Europe. L'étude haplotypique est en faveur d'un effet fondateur. La mutation est apparue il y a 30 000 ans environ chez les Caucasiens [48]. La grande fréquence de la mutation s'explique probablement par le fait qu'elle confère un avantage sélectif. Elle aurait notamment un effet protecteur vis-à-vis de l'hémorragie de la délivrance [49]. Ainsi, la prévalence est estimée à 1 à 15 % en Europe, elle est inférieure à 1 % en Asie et en Afrique, et de 0,5 à 5,2 % aux États-Unis, avec des disparités importantes en fonction de l'origine ethnique [50, 51]. Les études épidémiologiques conduites dans des populations spécifiques de patients ayant présenté un épisode de TV retrouvent la mutation à une fréquence très élevée, puisqu'elle atteint 20 à 25 %. Le RR de TV associé à la présence de la mutation à l'état HTZ est voisin de 5 [52, 53]. Ce risque devient plus important chez les HMZ. Peu d'études ont réussi à évaluer ce risque ; celles qui l'ont fait l'ont estimé entre 30 et 80. Le FVL est par ailleurs associé à une augmentation d'environ 50 % du risque de récédive [54-56].

De façon intéressante, le risque de TV associé au FVL a été évalué chez les femmes sous contraceptifs oraux combinés. Il existe une interaction entre FVL (RR=5) et COC (RR=3-6) puisque l'association de ces deux facteurs de risque se traduit par un RR estimé à 30.

### “ Point fort

La prévalence du FVL est de 1 à 15 % en Europe. Elle suit un gradient nord-sud. Le FVL est associé à un RR de TV d'environ 5. Le RR peut atteindre 30 quand le FVL est associé à l'utilisation d'un contraceptif oral combiné.

Le FVL est associé à la TV, avec toutefois une plus faible incidence d'EP que de TVP du membre inférieur. On parle ainsi du paradoxe du FVL. Dans une méta-analyse, la prévalence de la mutation était de 19 % en cas de TVP isolée, alors qu'elle n'était que de 9 % chez les patients ayant présenté une EP isolée [57].

## Diagnostic

Historiquement, le diagnostic de FVL était réalisé en deux étapes. Dans un premier temps, un test fonctionnel permettait de dépister une résistance à la PCa. Pour cela, le temps de céphaline activé (TCA) était mesuré avant et après ajout de PCa. Le ratio TCA + PCa/TCA était ensuite calculé et un ratio anormalement bas faisait suspecter la présence de la mutation, qui était alors recherchée dans un second temps. La technique la plus fréquemment utilisée repose sur l'amplification de l'exon X du gène du FV et la détermination du statut mutationnel fait intervenir des *polymerase chain reaction* (PCR) spécifiques d'allèle ou des enzymes de restriction (*restriction fragment length polymorphism* [RFLP]). Depuis juin 2013 et l'inscription à la nomenclature de la recherche de la mutation, cette stratégie diagnostique a été revue et la recherche de la mutation doit être réalisée en première intention.

## Autres mutations du facteur V

D'autres polymorphismes du FV entraînant une résistance à la PCa ont été identifiés. Dans une méta-analyse publiée en 2009 (14 études et près de 13 000 patients), l'allèle R2 (ou polymorphisme HR2 A4070G) était associé à la TV avec un OR de 1,24 (IC à 95 % : 1,02-1,52) [52]. Il est présent chez environ 10 % des sujets d'origine caucasienne et n'est pas situé sur le même haplotype que le FVL. Par ailleurs, deux mutations ont été décrites au même locus : le FV Cambridge R306T et le FV Hong Kong R306C [58, 59]. Toutes les deux entraînent une perte d'un site de clivage du FVa par la PCa. Ce clivage est nécessaire pour l'inactivation complète du FVa. Des FV Cambridge et Hong Kong obtenus par recombinaison génétique ont été étudiés in vitro. Ils présentaient les mêmes caractéristiques fonctionnelles, c'est-à-dire une résistance à la PCa dont le niveau était intermédiaire entre le FV normal et le FVL [60]. Le FV Cambridge est une mutation extrêmement rare alors que le FV Hong Kong est fréquent dans certaines populations chinoises. Leur association avec la TV n'est pas démontrée [44]. En 2007, le variant rs4524, qui introduit un résidu arginine en position 858 au lieu d'un résidu lysine, apparaissait associé à la TV [61]. Depuis, trois études cas-témoins ont confirmé ces résultats et le risque de TV associé au polymorphisme a été évalué à 21 % après ajustement pour le FVL (OR=1,21 ; IC à 95 % : 1,10-1,34) [62]. Plus récemment, une équipe japonaise a identifié une nouvelle mutation du FV chez un jeune garçon ayant présenté une TVP massive. Cette mutation génère une arginine en position 1920 et a été nommée FV Nara (W1920R) [63]. De façon intéressante, elle confère une résistance à la PCa environ 4,5 fois plus importante que le FVL alors qu'elle n'intéresse pas l'un des sites de clivage du FV par la PCa.

## Mutation G20210A du gène de la prothrombine

La prothrombine est une glycoprotéine vitamine K-dépendante de 72 kDa. Le FXa active la prothrombine en la clivant au niveau des Arg271-Thr272 et Arg320-Ile321. La thrombine ainsi générée est une enzyme majeure de l'hémostase puisqu'elle est impliquée dans de nombreuses voies. Elle joue un rôle procoagulant en activant les FVIII, FV, FXI, FXIII et en clivant le fibrinogène en fibrine ; et anticoagulant en activant la PC au sein du complexe thrombine-thrombomoduline. Elle est par ailleurs un puissant activateur plaquettaire. Enfin, elle intervient dans la voie de la fibrinolyse en activant le *thrombin-activatable fibrinolysis inhibitor* (TAFI).

Le gène de la prothrombine est localisé sur le chromosome 11p11-q12. Il s'étend sur 21 kb et comprend 14 exons et des régions 5' et 3' non transcrites qui pourraient jouer un rôle dans l'expression du gène [64, 65].

La mutation G20210A du gène de la prothrombine (PTG20210A) constitue l'une des plus belles réussites des approches « gène candidat ». En 1996, dans la cohorte *Leiden Thrombophilia Study* (LETS), Poort a identifié ce polymorphisme chez 6,3 % des cas et 2,3 % des témoins. L'OR associé au polymorphisme était alors évalué à 2,8 (IC à 95 % : 1,4-5,6). Dans la même étude, la mutation était retrouvée chez 18 % des patients

qui présentaient à la fois un antécédent personnel et familial de TV. L'allèle 20210A était par ailleurs associé à une augmentation du taux de prothrombine, lui-même associé à la TV. C'est probablement le mécanisme de l'origine de l'augmentation du risque de TV. La mutation se situe dans la région 3' non codante, dans une région qui conditionne la maturation des acides ribonucléiques (ARN) messagers. L'augmentation de la synthèse protéique résulterait d'une plus grande stabilité de l'ARN messager et d'une transcription plus efficace.

## Épidémiologie

L'épidémiologie de PTG20210A est analogue à celle du FVL, avec une prévalence importante en Europe (1 à 6 %) et un gradient sud-nord (inverse de celui du FVL), alors qu'elle est très rare dans les populations africaines et asiatiques [66-68]. Cette grande fréquence de la mutation chez les Caucasiens s'explique, comme pour le FVL, par un effet fondateur et un avantage sélectif, notamment en termes de nidation de l'embryon [69].

La mutation PTG20210A est associée à une augmentation du risque de premier événement thrombotique veineux. De nombreuses études ont évalué ce risque et les méta-analyses de Gohil en 2009 et Simone en 2013 retrouvent des résultats comparables, avec un OR voisin de 3 [52, 53]. Parmi les méta-analyses qui ont évalué le risque de récurrence associé à la mutation PTG20210A, celle de Segal en 2009 dispose de l'effectif le plus important [56]. Elle n'a cependant pas permis de montrer un lien significatif entre PTG20210A et récurrence de TV (OR = 1,45 ; IC à 95 % : 0,96-2,21). Les résultats des différentes études sont contradictoires, mais s'il existe un lien avec la récurrence, celui-ci est très modéré.

## Diagnostic

Le diagnostic de PTG20210A repose exclusivement sur la biologie moléculaire. Les techniques utilisées sont superposables à celles du FVL. Les deux mutations peuvent d'ailleurs être recherchées simultanément (PCR duplex).

## Autres variants de la prothrombine

Le séquençage du gène de la prothrombine a permis d'identifier deux nouvelles mutations affectant le même AA qui s'associent à une résistance à l'AT. La première mutation, Arg596Leu, a été décrite dans une famille japonaise et porte le nom de mutation Yukuashi de la prothrombine [70] ; la seconde mutation, R596Q, a été identifiée dans deux familles serbes et a été baptisée mutation Belgrade de la prothrombine [71].

## Groupe sanguin

Le groupe sanguin non-O est un facteur de risque de TV identifié depuis 1969 [72]. Les patients de groupe O présentent des taux diminués de FVIII et de facteur de von Willebrand, facteurs reconnus de TV [73, 74]. L'augmentation de la clairance du facteur de von Willebrand est l'un des mécanismes physiopathologiques permettant d'expliquer les taux diminués de FVIII et en facteur de von Willebrand chez les sujets de groupe O [75]. Dans la cohorte LETS, l'analyse génotypique du groupe sanguin a permis d'évaluer le risque de TV associé aux différents génotypes. Il existait une augmentation du risque de TV associé aux groupes non-OO : l'OR était de 1,8 (IC à 95 % : 1,4-2,4). Il faut noter que l'allèle A2 n'était pas associé à une augmentation de ce risque. En effet, l'allèle A2 code une N-acétyl-galactosamine-transférase constituée de 21 AA de plus que celle codée par l'allèle A1. La N-acétyl-galactosamine-transférase ainsi produite possède une action incomplète, ce qui se traduit par la persistance de l'antigène H à la surface globulaire. Le niveau d'expression de l'antigène A est intermédiaire, entre les sujets A1 et O. Il existait par ailleurs une interaction entre FVL et groupe non-O. En effet, l'OR était évalué à 23,2 (IC à 95 % : 9,1-59,3) dans le groupe non-OO + FVL. En 2008, une méta-analyse de Wu a évalué le risque de TV associé aux groupes non-O [73]. Au total 21 études ont été incluses. Dix-huit d'entre elles rapportaient une augmentation du risque de TV, avec des OR allant de 1,26 à 3,92 ; et l'OR global était de 1,79 (IC à 95 % : 1,56-2,05). Le génotype sanguin n'était connu que pour trois études. Le risque

de TV associé aux groupes A1B/A1A1/BB était évalué à 2,44 (IC à 95 % : 1,79-3,33) comparativement aux groupes 00/A2A2/A2O. Une étude de cohorte a par ailleurs rapporté un risque de récurrence de TV augmenté chez les sujets de groupe non-OO après un premier épisode de TV non provoqué : RR = 1,98 (1,2-3,8) [76].

## “ Point fort

Le groupe sanguin est un facteur de risque de TV. Les groupes sanguins non-OO sont associés à un RR de premier épisode de TV de 1,8 et à un RR de récurrence de 2.

## Hyperhomocystéinémie

L'homocystéine est un AA formé à partir de la méthionine après déméthylation. Dans la cellule, l'homocystéine peut emprunter deux voies métaboliques : la méthylation vers la méthionine ou la trans-sulfuration vers la cystéine. La méthylation de l'homocystéine en méthionine fait intervenir deux voies. Dans la première, la méthylène tétrahydrofolate réductase (MTHFR) permet la régénération du méthyltétrahydrofolate, dont le groupement méthyle est transféré sur l'homocystéine par la méthionine synthase en présence de cobalamine (vitamine B<sub>12</sub>). Dans la seconde, le groupement méthyle est transféré par la bêtaïne-homocystéine méthyltransférase à partir de la bêtaïne. La trans-sulfuration transforme l'homocystéine en sulfate en trois étapes. La cystathionine bêtasynthase convertit l'homocystéine en cystathionine, elle-même convertie en cystéine par la cystathionine bêta-lyase. Ces deux étapes nécessitent la présence de pyridoxine (vitamine B<sub>6</sub>). Le métabolisme de l'homocystéine fait intervenir de nombreux acteurs enzymatiques et vitaminiques. Des déficits constitutionnels ou acquis d'un ou plusieurs de ces acteurs peuvent être à l'origine d'une hyperhomocystéinémie, qui confère une augmentation modérée du risque de thrombose artérielle et veineuse [77]. Une méta-analyse regroupant neuf études a retrouvé une association significative entre hyperhomocystéinémie et TV [78]. L'OR global était estimé à 2,95 (IC à 95 % : 2,08-4,17). De façon intéressante, le risque était plus fort quand l'analyse se concentrait chez les sujets de moins de 60 ans. L'OR était alors de 4,37 (IC à 95 % : 1,94-9,84).

Il convient de distinguer les hyperhomocystéinémies sévères des hyperhomocystéinémies faibles à modérées. Les premières sont toujours d'origine génétique et se traduisent par des taux plasmatiques supérieurs à 100 µmol/l, alors que les secondes peuvent être d'origine génétique ou acquise et les taux plasmatiques sont compris entre 16 et 100 µmol/l. Les déficits enzymatiques des hyperhomocystéinémies sévères se traduisent le plus souvent par des atteintes syndromiques. Le déficit HMZ en cystathionine bêtasynthase est le plus fréquent, bien que rare dans la population générale (1/200000) [79]. Il associe un retard mental sévère, une atteinte du squelette et des complications cardiovasculaires (maladie vasculaire artérielle et TV). Le déficit HMZ en MTHFR est plus rare. Le tableau neurologique est prédominant mais il s'associe à une atteinte vasculaire, avec notamment des TV. Le déficit HTZ en cystathionine bêtasynthase est fréquent dans la population générale (0,3 à 1,4 %) et peut se traduire par une hyperhomocystéinémie faible ou modérée. De même, la mutation C677T du gène MTHFR à l'état HMZ a une prévalence de 10 % dans la population générale. Elle peut s'associer à une hyperhomocystéinémie faible ou modérée et n'est pas en soi un facteur de risque de TV. Les hyperhomocystéinémies acquises sont fréquentes et sont le plus souvent expliquées par des déficits nutritionnels en vitamines B<sub>6</sub>, B<sub>9</sub>, B<sub>12</sub>.

Le rôle de l'environnement semble le plus important comme en atteste la faible héritabilité de l'hyperhomocystéinémie [80]. Par ailleurs, la correction de l'hyperhomocystéinémie ne permet pas de réduire le risque de thrombose [81, 82]. Il est alors difficile de déterminer si l'homocystéine explique ces événements thrombotiques ou s'il ne s'agit que d'un marqueur biologique indirect du risque de thrombose [83].



## ■ Nouveaux variants

L'existence d'un antécédent familial de TV est associée au risque de TV [84-86]. Ce risque persiste après prise en compte de la thrombophilie constitutionnelle [87, 88], ce qui suggère fortement l'implication d'autres facteurs génétiques de risque.

L'identification de facteurs génétiques associés à la TV a connu un tournant dans les années 1990 avec la généralisation de l'utilisation des outils de biologie moléculaire et la découverte de variants fréquents mais à effet modéré : les FVL et PTG20210A. Dans les années 2000, les approches gène-candidat, puis surtout les études d'association génome entier (*genome wide association study* [GWAS]) ont permis d'identifier de nouveaux variants fréquents potentiellement associés à la TV.

## Approches gène-candidat

Cette approche a été la stratégie privilégiée jusqu'au début des années 2000 pour l'identification de nouveaux variants génétiques. Elle repose sur la sélection de gènes potentiellement impliqués dans la physiopathologie de la TV (ou gènes-candidats), puis sur l'étude de l'association entre les polymorphismes de ces gènes et la TV.

### Gène *PROCR*

Le gène *PROCR* est situé sur le chromosome 20, il est constitué de quatre exons et de trois introns. *PROCR* code le récepteur endothélial de la protéine C (EPCR), protéine de 43 kDa et 238 AA. La liaison de la PC à son récepteur endothélial augmente son activation d'un facteur 20. Le polymorphisme rs867186, qui se traduit par la substitution d'une sérine en glycine dans le codon 19, explique 56 à 87 % des variations des taux circulants d'EPCR soluble [89-93] et 10 % des taux de PC [94]. Le rôle d'EPCR ne s'arrête pas à la PC, il limite la formation du thrombus par ses actions sur le FVII et sur le rétrocontrôle du FX [95, 96]. Si l'association de rs867186 avec les taux circulants d'EPCR ne fait aucun doute, ce n'est pas le cas du phénotype clinique. En effet, les résultats sont discordants quand il s'agit de relier le génotype à la pathologie. Une méta-analyse a cependant permis de mettre en évidence une association robuste entre le polymorphisme rs867186 et la TV, et a estimé l'OR à 1,22 pour l'haplotype H3 [97].

### Fibrinogène

Le fibrinogène est un composant essentiel de la cascade de la coagulation. Dans une étape finale, la thrombine clive les chaînes  $\alpha$  et  $\beta$ , libérant ainsi les fibrinopeptides A et B. Cette étape permet la conversion du fibrinogène soluble en fibrine insoluble, qui se polymérise spontanément. Le FXIII stabilise les monomères de fibrine en établissant des liaisons covalentes. La molécule de fibrinogène est constituée d'un dimère de trois chaînes polypeptidiques,  $\alpha$ ,  $\beta$  et  $\gamma$ , dont la liaison est assurée par un pont disulfure. Trois gènes (*FGA*, *FGB*, *FGG*), regroupés en cluster dans une région d'environ 50 kb du chromosome 4q31.3, codent les trois polypeptides. Des mutations situées sur ces différents gènes sont susceptibles de modifier les taux circulants de fibrinogène, normalement compris entre 2 et 4 g/l. Notamment, la production de la protéine (synthèse séparée des trois chaînes dans les hépatocytes puis assemblage et sécrétion) est liée à la synthèse de la chaîne  $\beta$ , qui constitue le facteur limitant. Des mutations situées sur *FGB* peuvent modifier la production de la molécule. La régulation du gène est par ailleurs sous l'influence de nombreux facteurs, dont l'interleukine 6 (IL6), médiateur de l'inflammation, qui vont se lier à la région proximale du promoteur [98]. De nombreux polymorphismes susceptibles de modifier quantitativement le fibrinogène ont été décrits, dans les régions codantes et non codantes. La structure de l'ARN messager de la chaîne  $\gamma$  du fibrinogène rend possibles l'épissage alternatif et la polyadénylation [99]. Cela se traduit par l'existence d'une forme minoritaire de la chaîne  $\gamma$ . Ainsi, la forme majoritaire  $\gamma A$  est constituée de 411 AA, et le gène comprend dix exons et neuf introns. Cette forme dépend d'un signal de polyadénylation situé sur l'exon 10 alors que la forme minoritaire  $\gamma'$  dépend d'un signal de polyadé-

nylation situé sur l'intron 9. Le risque vasculaire va dépendre des quantités relatives de ces deux formes. Le lien entre les haplotypes les plus fréquents et la TV ont été étudiés dans la cohorte LETS. Seul l'haplotype FGG-H2 était significativement associé à la TV, avec un OR de 2,4 (IC 95 % : 1,5-3,9) en cas d'homozygotie. Cet haplotype était associé au taux circulant de fibrinogène  $\gamma'$ . Le risque de TV était multiplié par 2 chez les sujets présentant un ratio  $\gamma'/$ fibrinogène total inférieur à 0,69 (10<sup>e</sup> percentile). Trois polymorphismes en déséquilibre de liaison complet définissent l'haplotype FGG-H2. Parmi eux, rs2066865 (C10034T) explique probablement les variations des taux de fibrinogène  $\gamma'$ . Son association avec la TV a pu être confirmée dans une cohorte américaine [99].

### Facteur XI

En 2000, une étude publiée dans le *New England Journal of Medicine* a révélé l'association des taux de FXI avec le risque de TV. L'OR associé à la TV était de 2,2 chez les patients de la cohorte LETS qui présentaient des taux de FXI supérieurs au 90<sup>e</sup> percentile [100]. Il existait par ailleurs une relation dose-réponse. Le FXI constitue le point de départ de la voie intrinsèque de la cascade de la coagulation puisque son activation (par les trois facteurs de la phase contact : FXII, prékallitrène, kininogène de haut poids moléculaire ; ou par la thrombine) permet de convertir le FIX en FIXa. Les GWAS ont permis dans un second temps d'identifier plusieurs polymorphismes du gène du FXI (*F11*) associés à la fois aux taux de FXI et au risque de TV : rs3756008, rs2036914, rs3822057, rs2289252 [23, 61]. Les variants fonctionnels doivent encore être identifiés.

## Approches génome-entier ou GWAS

### “ Point fort

Les GWAS ont permis d'identifier de nombreux polymorphismes fréquents à effet faible.

Le manque de résultats probants obtenus par ces approches gènes-candidats peut s'expliquer soit par l'existence de variants génétiques à effets modestes que ne pouvaient mettre en évidence les études réalisées en raison de leur taille relativement petite, soit par un choix inapproprié des gènes sélectionnés, basé sur des connaissances parcellaires d'un mécanisme biologique complexe. C'est pourquoi d'autres approches basées sur des stratégies de recherche agnostiques, c'est-à-dire indépendantes de la biologie traditionnelle, ont été proposées. Les approches les plus fréquemment utilisées sont celles dites d'association « génome-entier » ou « pangénomiques » (GWAS). Les avancées technologiques en matière de biologie moléculaire ont en effet permis de réaliser des études d'association non plus avec des polymorphismes de nucléotide unique (*single-nucleotide polymorphism* [SNP]) situés sur quelques gènes-candidats, mais avec des centaines de milliers, répartis sur l'ensemble du génome humain. L'intérêt de ces études est qu'elles ne nécessitent pas d'hypothèse mécanistique a priori. Ce type d'approche a donc l'énorme avantage d'identifier des associations insoupçonnées, permettant théoriquement d'effectuer des découvertes plus originales, impliquant des gènes non suspects au regard des connaissances actuelles sur la physiopathologie et impliqués dans de nouvelles voies physiologiques.

### Gène *GP6*

En 2008, une étude d'association pangénomique a établi un lien entre la TV et le gène *GP6* [23]. Ce dernier code la glycoprotéine VI, protéine transmembranaire de 58 kDa appartenant à la superfamille des Ig, qui agit comme récepteur du collagène à la surface membranaire plaquettaire. Cette étude a permis, pour la

première fois, d'identifier un facteur de risque potentiel de TV dont l'action ne passe pas par la cascade de la coagulation. En effet, jusqu'alors tous les facteurs identifiés étaient des acteurs directs de la coagulation. La plaquette, qui ne semblait impliquée que dans la thrombose artérielle, devenait alors un acteur possible de la TV. Cette association avec la TV a été répliquée 3 ans plus tard et, dans une méta-analyse regroupant cinq études cas-témoins, l'augmentation du risque de TV associée au polymorphisme rs1613662 était estimée à 14 %. Le mécanisme n'est pas clairement élucidé, mais le polymorphisme contribue à réduire l'activation plaquettaire médiée par le collagène<sup>[101]</sup>. L'idée que la plaquette est impliquée dans la TV a trouvé confirmation en 2012 dans deux études où l'utilisation de faibles doses d'aspirine, versus placebo, a permis de réduire le nombre de récurrences chez des patients ayant présenté un premier épisode de TV ambulatoire, avec un RR estimé à 0,55 (IC à 95 % : 0,33-0,92) et 0,74 (IC à 95 % : 0,52-1,05) respectivement.

### Approches GWAS et phénotypes intermédiaires

La recherche d'une association entre des phénotypes intermédiaires quantitatifs et des polymorphismes, en empruntant le schéma agnostique des études pangénomiques, a permis l'identification de nouveaux variants associés à la TV.

Les premiers résultats ont intéressé les taux plasmatiques de facteur de von Willebrand et de FVIII<sup>[102]</sup>. Parmi les polymorphismes modifiant les taux de facteur de von Willebrand ou de FVIII, ceux situés sur les gènes *VWF* et *STXBPS* étaient robustement associés au risque de TV. Le variant rs1063856, situé sur le gène *VWF* codant le facteur de von Willebrand, était associé aux taux de facteur de von Willebrand et au risque de TV, avec un OR estimé à 1,15<sup>[103]</sup>. S'il existe une plausibilité biologique pour ce polymorphisme, le variant rs1039084 situé sur le gène *STXBPS* était, de façon surprenante, associé à une diminution des taux plasmatiques de facteur de von Willebrand, et ainsi à un OR de 0,90 pour la TV. Le gène *STXPB5* code la *syntaxin binding protein 5*, dont le rôle dans la cascade de la coagulation ne semble pas évident. Au demeurant plusieurs études ont apporté des preuves indirectes de l'implication de *STXPB5* dans la cascade de la coagulation. En effet, deux études décrivaient une association entre le polymorphisme rs1039084 et le phénotype hémorragique chez des femmes d'une part<sup>[104]</sup>, et les taux de *tissue plasminogen activator* (t-PA) d'autre part<sup>[105]</sup>. *STXPB5* semble par ailleurs jouer un rôle dans le fonctionnement de deux acteurs essentiels de la coagulation : *STXPB5* favorise le relargage des granules plaquettaires alors qu'il exerce une action inhibitrice sur la sécrétion du contenu des corps de Weibel-Palade situés dans les cellules endothéliales<sup>[106, 107]</sup>.

### “ Point fort

La recherche d'une association entre des phénotypes intermédiaires quantitatifs et des polymorphismes, en empruntant le schéma agnostique des études pangénomiques, a permis l'identification de nouveaux variants associés à la TV, notamment situés sur les gènes *VWF*, *STXPB5* et *KNG1*.

Le raccourcissement de l'*activated partial thromboplastin time* (aPTT), test global de la coagulation, semble associé au risque de TV, comme l'ont montré plusieurs études<sup>[108, 109]</sup>. Plus récemment, une étude d'association génome-entier a identifié trois variants génétiques associés aux temps mesurés par *activated partial thromboplastin time* (aPTT) : rs27431672, situé sur le gène du FXII, rs9898, situé sur le gène *HRG*, et rs710446, situé sur le gène *KNG1*<sup>[110]</sup>. Une étude regroupant les cohortes MARTHA08, MARTHA10 et 3-City Study, totalisant 2652 patients, a repris ces résultats et recherché une association entre ces trois polymorphismes et la TV. Cette association n'a été observée que pour le polymorphisme rs710446 du gène *KNG1*. L'OR associé à rs710446

était de 1,196 (IC à 95 % : 1,071-1,336)<sup>[111]</sup>. Ce résultat a été répliqué au cours de la même étude dans la cohorte FARIVE chez 596 cas et 590 témoins et l'OR associé au polymorphisme de *KNG1* était de 1,171 (IC à 95 % : 0,889-1,541). L'implication de *KNG1* dans la physiopathologie de la TV est sous-tendue par plusieurs arguments. *KNG1* code le kininogène de haut poids moléculaire (KHPM), dont le rôle dans la cascade de la coagulation est bien défini. Il joue un rôle important dans l'initiation de la voie intrinsèque en rapprochant la prékallitrène et le FXI du FXII<sup>[112]</sup>. Un modèle murin KO pour *KNG1* a par ailleurs été décrit. Les souris présentaient un allongement de l'aPTT et un retard à la thrombose artérielle<sup>[113]</sup>. De façon similaire, toujours chez la souris, des anticorps dirigés contre le FXI, empêchant l'interaction entre ce dernier et le KHPM, prévenaient la thrombose artérielle<sup>[114]</sup>. Enfin, plus récemment, rs710446 était associé aux taux circulants de FXI dans une étude d'association génome-entier<sup>[115]</sup>.

### TSPAN15 et SLC44A2

Une méta-GWAS prenant en compte 12 GWAS et plus de 60 000 individus (7507 cas et 52 632 témoins) a permis d'identifier deux nouveaux loci associés à la TV : *TSPAN15* et *SLC44A2*<sup>[116]</sup>. Ces deux nouveaux variants génétiques ont été répliqués dans trois cohortes indépendantes totalisant 3009 cas et 2586 témoins. Ces deux variants ne sont pas liés aux marqueurs biologiques de la coagulation. Ils représentent probablement de nouvelles voies dans la compréhension de la physiopathologie de la TV.

Le polymorphisme rs78707713 de *TSPAN15* était associé à une augmentation modeste du risque de TV (OR=1,31). La mutation est intronique et l'étude in silico n'a pas retrouvé de rôle fonctionnel pour ce polymorphisme, qui semble donc en déséquilibre de liaison avec un variant fonctionnel encore inconnu. Le gène *TSPAN15* code la tétraspanine 15, protéine appartenant à la superfamille des tétraspanines. Il s'agit de protéines ubiquitaires à quatre domaines transmembranaires, impliquées dans de nombreux processus biologiques tels que l'adhérence, la migration, la transduction du signal. Le rôle fonctionnel de *TSPAN15* et son lien avec la TV ne sont pas connus mais d'autres tétraspanines ont un lien clairement établi avec l'hémostase. Ainsi, *TSPAN24* (CD151) et *TSPAN32* (TSSC6) sont impliquées dans la fonction plaquettaire, en particulier dans la signalisation faisant intervenir l'intégrine  $\alpha_{IIb}\beta_3$ . Les souris déficientes pour ces tétraspanines présentent un phénotype hémorragique<sup>[117, 118]</sup>. Un lien avec l'hémostase a par ailleurs pu être établi pour *TSPAN30* (CD63). Cette dernière est impliquée dans le relargage du facteur de von Willebrand contenu dans les cellules endothéliales en facilitant la fusion des corps de Weibel-Palade avec la membrane plasmique<sup>[119]</sup>. Le même mécanisme est mis en jeu pour le recrutement leucocytaire à la surface de la cellule endothéliale activée<sup>[120]</sup>.

Le second polymorphisme, rs2288904 situé sur le gène *SLC44A2*, était associé à un RR de TV de 1,21<sup>[116]</sup>. Il code une isoforme de *choline transporter-like protein 2* (CTL-2) dont l'association avec le *transfusion related acute lung injury* (TRALI) a été démontrée<sup>[121, 122]</sup>. La physiopathologie du TRALI fait intervenir d'une part les neutrophiles, qui adhèrent aux capillaires pulmonaires, et des anticorps anti-HNA-3a. L'antigène cible est un fragment peptidique de CTL-2 dont il existe deux allèles : HNA-3a et HNA-3b. Le polymorphisme rs2288904 code l'allèle HNA-3a. Ainsi, les patients victimes de TRALI sévère avec anticorps anti-HNA-3a sont HMZ HNA-3a et reçoivent des anticorps anti-HNA-3a du donneur, toujours HMZ HNA-3b.

### Variants rares

De plus en plus, les nouvelles techniques de biologie moléculaire nous autorisent à nous intéresser aux variants rares en augmentant la précision du génotypage ou en rendant accessible le séquençage de gènes cibles en routine (séquençage haut débit). Les mutations Yukuashi et Belgrade du gène de la prothrombine ainsi que le FV Nara, décrits plus haut, illustrent bien cette évolution.

Le FIX Padoue est un autre exemple de variant rare associé à la TV. Une équipe italienne a récemment mis en évidence une anomalie génétique chez un jeune patient de 23 ans ayant

**Tableau 1.**

Polymorphismes associés à la thrombose veineuse.

| Locus    | SNP                          | Allèle <sup>a</sup> | Prévalence <sup>b</sup> | OR <sup>c</sup> | Phénotype associé <sup>d</sup>          | Références bibliographiques |
|----------|------------------------------|---------------------|-------------------------|-----------------|---|-----------------------------|
| SERPINC1 | Nombreuses mutations privées | -                   | -                       | ~10             | ↓ AT                                    | [11-14, 20, 21, 24]         |
| PROC     | Nombreuses mutations privées | -                   | -                       | ~10             | ↓ PC                                    | [11, 12, 14, 42]            |
| PROS1    | Nombreuses mutations privées | -                   | -                       | ~10             | ↓ PS                                    | [11, 12, 14, 37]            |
| ABO      | [O,A2] versus [A1,B]         |                     | 0,30                    | 1,80            | ↑ VWF, ↑ FVIII                          | [73, 74]                    |
| F2       | rs1799963                    | G/ <u>A</u>         | 0,02                    | 3,00            | ↑ FII                                   | [52, 53, 56, 66-68]         |
| F5       | rs6025                       | G/ <u>A</u>         | 0,05                    | 5,00            | Résistance à la protéine C activée      | [50-53]                     |
| F11      | rs2036914                    | C/ <u>T</u>         | 0,52                    | 1,27            | ↑ FXI                                   | [23]                        |
|          | rs2289252                    | C/ <u>T</u>         | 0,41                    | 1,33            | ↑ FXI                                   | [61]                        |
| FGG      | rs2066865                    | C/ <u>T</u>         | 0,25                    | 1,47            | ↓ Fibrinogène γ'                        | [99]                        |
| GP6      | rs1613662                    | <u>A</u> /G         | 0,82                    | 1,15            | ↑ activation et agrégation plaquettaire | [23, 101]                   |
| KNG1     | rs710446                     | T/ <u>C</u>         | 0,45                    | 1,20            | ↓ aPTT                                  | [110, 111]                  |
| PROCR    | rs867186                     | A/ <u>G</u>         | 0,07                    | 1,22            | ↑ eEPCR, ↑ PC                           | [89-94, 97]                 |
| SLC44A2  | rs2288904                    | <u>T</u> /C         | 0,79                    | 1,21            | Inconnu                                 | [116]                       |
| STXBP5   | rs1039084                    | A/ <u>G</u>         | 0,46                    | 0,78            | ↓ VWF                                   | [103]                       |
| TSPAN15  | rs78707713                   | <u>G</u> /A         | 0,88                    | 1,31            | Inconnu                                 | [116]                       |
| VWF      | rs1063856                    | A/ <u>G</u>         | 0,37                    | 1,15            | ↑ VWF                                   | [103]                       |

SNP : *single-nucleotide polymorphism* ; OR : odd ratio ; AT : antithrombine ; PC : protéine C ; PS : protéine S ; VWF : facteur de von Willebrand ; FVIII : facteur VIII ; FII : facteur II ; FXI : facteur XI ; aPTT : *activated partial thromboplastin time* ; EPCR : récepteur endothélial à la protéine C.

<sup>a</sup> Les allèles à risque sont soulignés.

<sup>b</sup> Prévalence estimée de l'allèle à risque dans les populations de référence.

<sup>c</sup> Odds ratio (OR) estimés associés à l'allèle à risque.

<sup>d</sup> Phénotype associé à l'allèle à risque.

présenté une TVP proximale. Le bilan de thrombophilie réalisé chez ce patient était négatif mais l'existence d'un raccourcissement du TCA a conduit l'équipe à doser les facteurs de la voie endogène et une augmentation majeure de l'activité du FIX a été objectivée (× 8), alors que le FIX antigène était normal. La mutation gain de fonction responsable a été identifiée, il s'agit d'une substitution d'une arginine par une leucine en position 338. Cette mutation n'a pas été retrouvée dans une cohorte de 200 patients de la même région. Il s'agit d'une mutation rare et privée.

mutation, ce qui s'explique probablement par l'existence d'autres facteurs génétiques de risque de thrombose encore non identifiés. Enfin, les antécédents familiaux de TV restent associés au risque de TV après ajustement pour la thrombophilie classique [84]. Il a été récemment démontré que le groupe sanguin ABO et les taux plasmatiques de facteur de von Willebrand expliquent partiellement la pénétrance incomplète de la TV dans ces familles.

## ■ Perspectives

De façon générale, les tests de dépistage permettent de prévenir la survenue de pathologies ou de leurs complications chez des individus à risque. Ils doivent être réalisés en principe uniquement si leur positivité entraîne une modification de l'attitude thérapeutique, préventive ou curative.

Dans le contexte de la TV, le dépistage génétique, appelé bilan de thrombophilie, comprend à ce jour la recherche d'un déficit en AT, PC ou PS, le FVL et la mutation G20210A. Cette liste n'a pas changé depuis la découverte de la mutation G20210A en 1996. Malgré la publication de plusieurs recommandations, il n'existe pas de consensus sur les patients qui doivent bénéficier de ce dépistage. Chez les patients qui présentent un antécédent de TV, la présence d'une thrombophilie constitutionnelle n'influence pas la posologie ou la durée du traitement anticoagulant puisqu'ils n'ont pas, ou très peu, d'impact sur la réponse aux anticoagulants ou sur le taux de récurrence. Cependant, pour certains patients sélectionnés, la connaissance d'une thrombophilie constitutionnelle peut moduler la prise en charge, comme ce peut être le cas au cours de la grossesse. Le conseil génétique chez des familles avec TV est par ailleurs toujours débattu. En effet, l'identification de l'anomalie chez les apparentés d'un propositus avec antécédent de TV permettrait d'identifier ceux à haut risque pour leur proposer une prise en charge spécifique au cours de situations à risque. Cependant, même s'ils présentent la même mutation, il est fréquent que des apparentés ne présentent pas les mêmes manifestations cliniques que le propositus. Par exemple, la pénétrance du FVL est très faible puisque seulement 10 % des HTZ et 80 % des HMZ présenteront un jour un épisode de TV, avec une variabilité de la sévérité de l'épisode. De plus, dans ces familles, le risque de TV est augmenté chez les apparentés sans

## “ Point fort

De nouvelles approches basées sur les polymorphismes fréquents à effet faible identifiés ont été développées. L'utilisation des scores génétiques de risque prenant en compte les allèles à risque permet ainsi de classer les patients en fonction de leur risque de TV.

Les récentes études génétiques (approches gènes-candidats et surtout GWAS) ont permis d'identifier de nombreux polymorphismes associés à la TV (Tableau 1). De nouvelles approches basées sur ces polymorphismes fréquents à effet faible ont été développées. Ainsi, van Hylckama et al. ont récemment évalué la valeur prédictive de 31 polymorphismes sur le risque de récurrence de TV [123]. L'utilisation d'un score génétique de risque prenant en compte cinq SNP situés sur les gènes *ABO*, *F11*, *F2*, *F5* et *FGG* permettait de classer les patients en fonction de leur risque : après un suivi de 6 ans, l'incidence cumulée de la récurrence était élevée quand le score était supérieur ou égal à cinq allèles à risque (20,3 %) alors qu'elle était faible lorsque les patients présentaient un score inférieur ou égal à un allèle à risque (9,4 %).

Dans le futur, la combinaison de ces scores génétiques à des variables cliniques et plasmatiques dans le cadre d'algorithmes décisionnels validés par des études d'intervention devrait permettre une meilleure évaluation du risque de récurrence de TV.

**Déclaration d'intérêts** : les auteurs déclarent ne pas avoir de liens d'intérêts en relation avec cet article.



## ■ Références

- 900 ?
- 901 [1] Oger E. Incidence of venous thromboembolism: a community-  
902 based study in Western France. EPI-GETBP Study Group. Groupe  
903 d'étude de la thrombose de Bretagne occidentale. *Thromb Haemost*  
904 2000;**83**:657-60.
- 905 [2] Prandoni P, Lensing AW, Cogo A, Cuppini S, Villalta S, Carta M, et al.  
906 The long-term clinical course of acute deep venous thrombosis. *Ann*  
907 *Intern Med* 1996;**125**:1-7.
- 908 [3] Schulman S, Rhedin AS, Lindmarker P, Carlsson A, Lärffars G, Nicol P,  
909 et al. A comparison of six weeks with six months of oral anticoagulant  
910 therapy after a first episode of venous thromboembolism. Duration  
911 of Anticoagulation Trial Study Group. *N Engl J Med* 1995;**332**:  
912 1661-5.
- 913 [4] Kearon C, Gent M, Hirsh J, Weitz J, Kovacs MJ, Anderson DR, et al.  
914 A comparison of three months of anticoagulation with extended anti-  
915 coagulation for a first episode of idiopathic venous thromboembolism.  
916 *N Engl J Med* 1999;**340**:901-7.
- 917 [5] Agnelli G, Prandoni P, Santamaria MG, Bagatella P, Iorio A, Bazzan  
918 M, et al. Three months versus one year of oral anticoagulant therapy for  
919 idiopathic deep venous thrombosis. Warfarin Optimal Duration Italian  
920 Trial Investigators. *N Engl J Med* 2001;**345**:165-9.
- 921 [6] Pinede L, Ninet J, Duhaut P, Chabaud S, Demolombe-Rague S, Durieu  
922 I, et al. Comparison of 3 and 6 months of oral anticoagulant therapy  
923 after a first episode of proximal deep vein thrombosis or pulmonary  
924 embolism and comparison of 6 and 12 weeks of therapy after isolated  
925 calf deep vein thrombosis. *Circulation* 2001;**103**:2453-60.
- 926 [7] White RH. The epidemiology of venous thromboembolism. *Circula-*  
927 *tion* 2003;**107**:14-8.
- 928 [8] Souto JC, Almasy L, Borrell M, Blanco-Vaca F, Mateo J, Soria JM,  
929 et al. Genetic susceptibility to thrombosis and its relationship to phy-  
930 siological risk factors: the GAIT study. Genetic Analysis of Idiopathic  
931 Thrombophilia. *Am J Hum Genet* 2000;**67**:1452-9.
- 932 [9] Larsen TB, Sørensen HT, Skytthe A, Johnsen SP, Vaupel JW, Christen-  
933 sen K. Major genetic susceptibility for venous thromboembolism  
934 in men: a study of Danish twins. *Epidemiol Camb Mass* 2003;**14**:  
935 328-32.
- 936 [10] Egeberg O. Inherited antithrombin deficiency causing thrombophilia.  
937 *Thromb Diath Haemorrh* 1965;**13**:516-30.
- 938 [11] Bucciarelli P, Rosendaal FR, Tripodi A, Mannucci PM, De Stefano V,  
939 Palareti G, et al. Risk of venous thromboembolism and clinical mani-  
940 festations in carriers of antithrombin, protein C, protein S deficiency,  
941 or activated protein C resistance: a multicenter collaborative family  
942 study. *Arterioscler Thromb Vasc Biol* 1999;**19**:1026-33.
- 943 [12] Martinelli I, Mannucci PM, De Stefano V, Taioli E, Rossi V, Crosti F,  
944 et al. Different risks of thrombosis in four coagulation defects asso-  
945 ciated with inherited thrombophilia: a study of 150 families. *Blood*  
946 1998;**92**:2353-8.
- 947 [13] Seligsohn U, Lubetsky A. Genetic susceptibility to venous thrombosis.  
948 *N Engl J Med* 2001;**344**:1222-31.
- 949 [14] Simioni P, Sanson BJ, Prandoni P, Tormene D, Friederich PW, Giro-  
950 lami B, et al. Incidence of venous thromboembolism in families with  
951 inherited thrombophilia. *Thromb Haemost* 1999;**81**:198-202.
- 952 [15] Boyer C, Wolf M, Vedrenne J, Meyer D, Larrieu MJ. Homozygous  
953 variant of antithrombin III: AT III Fontainebleau. *Thromb Haemost*  
954 1986;**56**:18-22.
- 955 [16] Lane DA, Bayston T, Olds RJ, Fitches AC, Cooper DN, Millar DS, et al.  
956 Antithrombin mutation database: 2nd (1997) update. For the Plasma  
957 Coagulation Inhibitors Subcommittee of the Scientific and Standar-  
958 dization Committee of the International Society on Thrombosis and  
959 Haemostasis. *Thromb Haemost* 1997;**77**:197-211.
- 960 [17] Finazzi G, Caccia R, Barbui T. Different prevalence of thromboembo-  
961 lism in the subtypes of congenital antithrombin III deficiency: review  
962 of 404 cases. *Thromb Haemost* 1987;**58**:1094.
- 963 [18] Molho-Sabatier P, Aiach M, Gaillard I, Fiessinger JN, Fischer AM,  
964 Chadeuf G, et al. Molecular characterization of antithrombin III (ATIII)  
965 variants using polymerase chain reaction. Identification of the ATIII  
966 Charleville as an Ala 384 Pro mutation. *J Clin Invest* 1989;**84**:1236-42.
- 967 [19] Van Boven HH, Lane DA. Antithrombin and its inherited deficiency  
968 states. *Semin Hematol* 1997;**34**:188-204.
- 969 [20] Luxembourg B, Pavlova A, Geisen C, Spannagl M, Bergmann F,  
970 Krause M, et al. Impact of the type of SERPINC1 mutation and subtype  
971 of antithrombin deficiency on the thrombotic phenotype in hereditary  
972 antithrombin deficiency. *Thromb Haemost* 2014;**111**:249-57.
- 973 [21] Bucciarelli P, Passamonti SM, Biguzzi E, Gianniello F, Franchi F,  
974 Mannucci PM, et al. Low borderline plasma levels of antithrombin,  
975 protein C and protein S are risk factors for venous thromboembolism.  
976 *J Thromb Haemost* 2012;**10**:1783-91.
- [22] Antón AI, Teruel R, Corral J, Miñano A, Martínez-Martínez I, Ordóñez  
977 A, et al. Functional consequences of the prothrombotic SERPINC1  
978 rs2227589 polymorphism on antithrombin levels. *Haematologica*  
979 2009;**94**:589-92.
- [23] Bezemer ID, Bare LA, Doggen CJ, Arellano AR, Tong C, Rowland  
981 CM, et al. Gene variants associated with deep vein thrombosis. *JAMA*  
982 2008;**299**:1306-14.
- [24] Di Minno MN, Dentali F, Lupoli R, Ageno W. Mild antithrombin defi-  
984 ciency and risk of recurrent venous thromboembolism: a prospective  
985 cohort study. *Circulation* 2014;**129**:497-503.
- [25] Rosing J, Hoekema L, Nicolaes GA, Thomassen MC, Hemker HC,  
987 Varadi K, et al. Effects of protein S and factor Xa on peptide bond clea-  
988 vages during inactivation of factor Va and factor VaR506Q by activated  
989 protein C. *J Biol Chem* 1995;**270**:27852-8.
- [26] Shen L, Dahlbäck B. Factor V and protein S as synergistic cofactors  
991 to activated protein C in degradation of factor VIIIa. *J Biol Chem*  
992 1994;**269**:18735-8.
- [27] Heeb MJ, Mesters RM, Tans G, Rosing J, Griffin JH. Binding of  
994 protein S to factor Va associated with inhibition of prothrombinase  
995 that is independent of activated protein C. *J Biol Chem* 1993;**268**:  
996 2872-7.
- [28] Heeb MJ, Rosing J, Bakker HM, Fernandez JA, Tans G, Griffin JH.  
998 Protein S binds to and inhibits factor Xa. *Proc Natl Acad Sci USA*  
999 1994;**91**:2728-32.
- [29] Hackeng TM, Maurissen LF, Castoldi E, Rosing J. Regulation of TFPI  
1001 function by protein S. *J Thromb Haemost* 2009;**7**:165-8.
- [30] Allaart CF, Poort SR, Rosendaal FR, Reitsma PH, Bertina RM, Briët  
1003 E. Increased risk of venous thrombosis in carriers of hereditary protein  
1004 C deficiency defect. *Lancet* 1993;**341**:134-8.
- [31] Zöller B, García de Frutos P, Dahlbäck B. Evaluation of the rela-  
1006 tionship between protein S and C4b-binding protein isoforms in  
1007 hereditary protein S deficiency demonstrating type I and type III defi-  
1008 ciencies to be phenotypic variants of the same genetic disease. *Blood*  
1009 1995;**85**:3524-31.
- [32] Vossen CY, Walker ID, Svensson P, Souto JC, Scharrer I, Pres-  
1011 ton FE, et al. Recurrence rate after a first venous thrombosis in  
1012 patients with familial thrombophilia. *Arterioscler Thromb Vasc Biol*  
1013 2005;**25**:1992-7.
- [33] Dreyfus M, Magny JF, Bridey F, Schwarz HP, Planché C, Dehan M,  
1015 et al. Treatment of homozygous protein C deficiency and neonatal  
1016 purpura fulminans with a purified protein C concentrate. *N Engl J Med*  
1017 1991;**325**:1565-8.
- [34] Marciniak E, Wilson HD, Marlar RA. Neonatal purpura fulminans: a  
1019 genetic disorder related to the absence of protein C in blood. *Blood*  
1020 1985;**65**:15-20.
- [35] Gómez E, Ledford MR, Pegelow CH, Reitsma PH, Bertina RM. Homo-  
1022 zygous protein S deficiency due to a one base pair deletion that leads  
1023 to a stop codon in exon III of the protein S gene. *Thromb Haemost*  
1024 1994;**71**:723-6.
- [36] Spek CA, Koster T, Rosendaal FR, Bertina RM, Reitsma PH. Genoty-  
1026 pic variation in the promoter region of the protein C gene is associated  
1027 with plasma protein C levels and thrombotic risk. *Arterioscler Thromb*  
1028 *Vasc Biol* 1995;**15**:214-8.
- [37] Pintao MC, Ribeiro DD, Bezemer ID, Garcia AA, de Visser MC, Dog-  
1030 gen CJ, et al. Protein S levels and the risk of venous thrombosis: results  
1031 from the MEGA case-control study. *Blood* 2013;**122**:3210-9.
- [38] Ten Kate MK, van der Meer J. Protein S deficiency: a clinical perspec-  
1033 tive. *Haemophilia* 2008;**14**:1222-8.
- [39] Castoldi E, Hackeng TM. Regulation of coagulation by protein S. *Curr*  
1035 *Opin Hematol* 2008;**15**:529-36.
- [40] Johnston AM, Aboud M, Morel-Kopp MC, Coyle L, Ward CM. Use  
1037 of a functional assay to diagnose protein S deficiency; inappropriate  
1038 testing yields equivocal results. *Intern Med J* 2007;**37**:409-11.
- [41] Marlar RA, Gausman JN. Protein S abnormalities: a diagnostic night-  
1040 mare. *Am J Hematol* 2011;**86**:418-21.
- [42] Reitsma PH, Bernardi F, Doig RG, Gandrille S, Greengard JS, Ireland  
1042 H, et al. Protein C deficiency: a database of mutations, 1995 update.  
1043 On behalf of the Subcommittee on Plasma Coagulation Inhibitors of  
1044 the Scientific and Standardization Committee of the ISTH. *Thromb*  
1045 *Haemost* 1995;**73**:876-89.
- [43] Aiach M, Borgel D, Gaussem P, Emmerich J, Alhenc-Gelas M,  
1047 Gandrille S. Protein C and protein S deficiencies. *Semin Hematol*  
1048 1997;**34**:205-16.
- [44] Nicolaes GA, Dahlbäck B. Factor V and thrombotic disease: des-  
1050 cription of a janus-faced protein. *Arterioscler Thromb Vasc Biol*  
1051 2002;**22**:530-8.
- [45] Castoldi E, Rosing J. APC resistance: biological basis and acquired  
1053 influences. *J Thromb Haemost* 2010;**8**:445-53.
- 1054

- [46] Dahlbäck B, Carlsson M, Svensson PJ. 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. *Proc Natl Acad Sci USA* 1993;**90**:1004–8.
- [47] Bertina RM, Koelman BP, Koster T, Rosendaal FR, Dirven RJ, de Ronde H, et al. Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* 1994;**369**:64–7.
- [48] Zivelin A, Griffin JH, Xu X, Pabinger I, Samama M, Conard J, et al. A single genetic origin for a common Caucasian risk factor for venous thrombosis. *Blood* 1997;**89**:397–402.
- [49] Lindqvist PG, Svensson PJ, Dahlbäck B, Marsál K. Factor V Q506 mutation (activated protein C resistance) associated with reduced intrapartum blood loss—a possible evolutionary selection mechanism. *Thromb Haemost* 1998;**79**:69–73.
- [50] Dahlbäck B. Resistance to activated protein C caused by the factor V R506Q mutation is a common risk factor for venous thrombosis. *Thromb Haemost* 1997;**78**:483–8.
- [51] Ridker PM, Miletich JP, Hennekens CH, Buring JE. Ethnic distribution of factor V Leiden in 4047 men and women. Implications for venous thromboembolism screening. *JAMA* 1997;**277**:1305–7.
- [52] Gohil R, Peck G, Sharma P. The genetics of venous thromboembolism. A meta-analysis involving approximately 120 000 cases and 180 000 controls. *Thromb Haemost* 2009;**102**:360–70.
- [53] Simone B, De Stefano V, Leoncini E, Zacho J, Martinelli I, Emmerich J, et al. Risk of venous thromboembolism associated with single and combined effects of Factor V Leiden, prothrombin 20210A and methylenetetrahydrofolate reductase C677T: a meta-analysis involving over 11 000 cases and 21 000 controls. *Eur J Epidemiol* 2013;**28**:621–47.
- [54] Ho W, Hankey GJ, Quinlan DJ, Eikelboom JW. Risk of recurrent venous thromboembolism in patients with common thrombophilia: a systematic review. *Arch Intern Med* 2006;**166**:729–36.
- [55] Marchiori A, Mosena L, Prins MH, Prandoni P. The risk of recurrent venous thromboembolism among heterozygous carriers of factor V Leiden or prothrombin G20210A mutation. A systematic review of prospective studies. *Haematologica* 2007;**92**:1107–14.
- [56] Segal JB, Brotman DJ, Necochea AJ. Predictive value of factor V Leiden and prothrombin G20210A in adults with venous thromboembolism and in family members of those with a mutation: a systematic review. *JAMA* 2009;**301**:2472–85.
- [57] Bounameaux H. Factor V Leiden paradox: risk of deep-vein thrombosis but not of pulmonary embolism. *Lancet* 2000;**356**:182–3.
- [58] Williamson D, Brown K, Luddington R, Baglin C, Baglin T. Factor V Cambridge: a new mutation (Arg306→Thr) associated with resistance to activated protein C. *Blood* 1998;**91**:1140–4.
- [59] Chan WP, Lee CK, Kwong YL, Lam CK, Liang R. A novel mutation of Arg306 of factor V gene in Hong Kong Chinese. *Blood* 1998;**91**:1135–9.
- [60] Norström E, Thorelli E, Dahlbäck B. Functional characterization of recombinant FV Hong Kong and FV Cambridge. *Blood* 2002;**100**:524–30.
- [61] Smith NL, Hindorff LA, Heckbert SR, Lemaitre RN, Marcianti KD, Rice K, et al. Association of genetic variations with nonfatal venous thrombosis in postmenopausal women. *JAMA* 2007;**297**:489–98.
- [62] Bezemer ID, Bare LA, Arellano AR, Reitsma PH, Rosendaal FR. Updated analysis of gene variants associated with deep vein thrombosis. *JAMA* 2010;**303**:421–2.
- [63] Nogami K, Shinozawa K, Ogiwara K, Matsumoto T, Amano K, Fukutake K, et al. Novel FV mutation (W1920R, FVNara) associated with serious deep vein thrombosis and more potent APC resistance relative to FVLeiden. *Blood* 2014;**123**:2420–8.
- [64] Poort SR, Rosendaal FR, Reitsma PH, Bertina RM. 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**:3698–703.
- [65] Degen SJ, Davie EW. Nucleotide sequence of the gene for human prothrombin. *Biochemistry* 1987;**26**:6165–77.
- [66] Leroyer C, Mercier B, Oger E, Chenu E, Abgrall JF, Férec C, et al. Prevalence of 20210 A allele of the prothrombin gene in venous thromboembolism patients. *Thromb Haemost* 1998;**80**:49–51.
- [67] Tosetto A, Missiaglia E, Frezzato M, Rodeghiero F. The VITA project: prothrombin G20210A mutation and venous thromboembolism in the general population. *Thromb Haemost* 1999;**82**:1395–8.
- [68] Folsom AR, Cushman M, Tsai MY, Heckbert SR, Aleksic N. Prospective study of the G20210A polymorphism in the prothrombin gene, plasma prothrombin concentration, and incidence of venous thromboembolism. *Am J Hematol* 2002;**71**:285–90.
- [69] Hundsdoerfer P, Vetter B, Stöver B, Bassir C, Scholz T, Grimmer I, et al. Homozygous and double heterozygous Factor V Leiden and Factor II G20210A genotypes predispose infants to thromboembolism but are not associated with an increase of foetal loss. *Thromb Haemost* 2003;**90**:628–35.
- [70] Miyawaki Y, Suzuki A, Fujita J, Maki A, Okuyama E, Murata M, et al. Thrombosis from a prothrombin mutation conveying antithrombin resistance. *N Engl J Med* 2012;**366**:2390–6.
- [71] Djordjevic V, Kovac M, Miljic P, Murata M, Takagi A, Pruner I, et al. A novel prothrombin mutation in two families with prominent thrombophilia—the first cases of antithrombin resistance in a Caucasian population. *J Thromb Haemost* 2013;**11**:1936–9.
- [72] Jick H, Slone D, Westerholm B, Inman WH, Vessey MP, Shapiro S, et al. Venous thromboembolic disease and ABO blood type. A cooperative study. *Lancet* 1969;**1**:539–42.
- [73] Wu O, Bayoumi N, Vickers MA, Clark P. ABO(H) blood groups and vascular disease: a systematic review and meta-analysis. *J Thromb Haemost* 2008;**6**:62–9.
- [74] Jenkins PV, Rawley O, Smith OP, O'Donnell JS. Elevated factor VIII levels and risk of venous thrombosis. *Br J Haematol* 2012;**157**:653–63.
- [75] O'Donnell J, Laffan MA. The relationship between ABO histo-blood group, factor VIII and von Willebrand factor. *Transfusion* 2001;**11**:343–51.
- [76] Gándara E, Kovacs MJ, Kahn SR, Wells PS, Anderson DA, Chagnon I, et al. Non-OO blood type influences the risk of recurrent venous thromboembolism. A cohort study. *Thromb Haemost* 2013;**110**:1172–9.
- [77] Eldibany MM, Caprini JA. Hyperhomocysteinemia and thrombosis: an overview. *Arch Pathol Lab Med* 2007;**131**:872–84.
- [78] Ray JG. Meta-analysis of hyperhomocysteinemia as a risk factor for venous thromboembolic disease. *Arch Intern Med* 1998;**158**:2101–6.
- [79] Malinow MR. Homocyst(e)line and arterial occlusive diseases. *J Intern Med* 1994;**236**:603–17.
- [80] Martinelli I, Bucciarelli P, Zighetti ML, Cafro A, Mannucci PM. Low risk of thrombosis in family members of patients with hyperhomocysteinemia. *Br J Haematol* 2002;**117**:709–11.
- [81] Lonn E, Yusuf S, Arnold MJ, Sheridan P, Pogue J, Micks M, et al. Homocysteine lowering with folic acid and B vitamins in vascular disease. *N Engl J Med* 2006;**354**:1567–77.
- [82] Cacciapuoti F. Lowering homocysteine levels with folic acid and B-vitamins do not reduce early atherosclerosis, but could interfere with cognitive decline and Alzheimer's disease. *J Thromb Thrombolysis* 2013;**36**:258–62.
- [83] Cacciapuoti F. Hyper-homocysteinemia: a novel risk factor or a powerful marker for cardiovascular diseases? Pathogenetic and therapeutic uncertainties. *J Thromb Thrombolysis* 2011;**32**:82–8.
- [84] Bezemer ID, van der Meer FJ, Eikenboom JC, Rosendaal FR, Doggen CJ. The value of family history as a risk indicator for venous thrombosis. *Arch Intern Med* 2009;**169**:610–5.
- [85] Zöller B, Li X, Sundquist J, Sundquist K. Parental history and venous thromboembolism: a nationwide study of age-specific and sex-specific familial risks in Sweden. *J Thromb Haemost* 2011;**9**:64–70.
- [86] Sørensen HT, Riis AH, Diaz LJ, Andersen EW, Baron JA, Andersen PK. Familial risk of venous thromboembolism: a nationwide cohort study. *J Thromb Haemost* 2011;**9**:320–4.
- [87] Bucciarelli P, De Stefano V, Passamonti SM, Tormene D, Legnani C, Rossi E, et al. Influence of proband's characteristics on the risk for venous thromboembolism in relatives with factor V Leiden or prothrombin G20210A polymorphisms. *Blood* 2013;**122**:2555–61.
- [88] Rossi E, Ciminello A, Za T, Betti S, Leone G, De Stefano V. In families with inherited thrombophilia the risk of venous thromboembolism is dependent on the clinical phenotype of the proband. *Thromb Haemost* 2011;**106**:646–54.
- [89] Reiner AP, Carty CL, Jenny NS, Nievergelt C, Cushman M, Stearns-Kurosawa DJ, et al. PROC, PROCr and PROS1 polymorphisms, plasma anticoagulant phenotypes, and risk of cardiovascular disease and mortality in older adults: the Cardiovascular Health Study. *J Thromb Haemost* 2008;**6**:1625–32.
- [90] Saposnik B, Reny JL, Gaussem P, Emmerich J, Aiach M, Gandrille S. 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**:1311–8.
- [91] Uitte de Willige S, Van Marion V, Rosendaal FR, Vos HL, de Visser MC, Bertina RM. Haplotypes of the EPCR gene, plasma sEPCR levels and the risk of deep venous thrombosis. *J Thromb Haemost* 2004;**2**:1305–10.



- 1210 [92] Medina P, Navarro S, Estellés A, Vayá A, Woodhams B, Mira Y, 1211 et al. Contribution of polymorphisms in the endothelial protein C 1212 receptor gene to soluble endothelial protein C receptor and circulating 1213 activated protein C levels, and thrombotic risk. *Thromb Haemost* 1214 2004;**91**:905–11.
- 1215 [93] Ireland H, Konstantoulas CJ, Cooper JA, Hawe E, Humphries SE, 1216 Mather H, et al. EPCR Ser219Gly: elevated sEPCR, prothrombin F1+2, 1217 risk for coronary heart disease, and increased sEPCR shedding in vitro. 1218 *Atherosclerosis* 2005;**183**:283–92.
- 1219 [94] Tang W, Basu S, Kong X, Pankow JS, Aleksic N, Tan A, et al. 1220 Genome-wide association study identifies novel loci for plasma levels 1221 of protein C: the ARIC study. *Blood* 2010;**116**:5032–6.
- 1222 [95] Ireland HA, Cooper JA, Drenos F, Acharya J, Mitchell JP, Bauer KA, 1223 et al. FVII, FVIIa, and downstream markers of extrinsic pathway activation 1224 differ by EPCR Ser219Gly variant in healthy men. *Arterioscler* 1225 *Thromb Vasc Biol* 2009;**29**:1968–74.
- 1226 [96] Ghosh S, Pendarthi UR, Steinoe A, Esmon CT, Rao LV. Endothelial 1227 cell protein C receptor acts as a cellular receptor for factor VIIa on 1228 endothelium. *J Biol Chem* 2007;**282**:11849–57.
- 1229 [97] Dennis J, Johnson CY, Adediran AS, de Andrade M, Heit JA, Morange 1230 PE, et al. The endothelial protein C receptor (PROCR) Ser219Gly 1231 variant and risk of common thrombotic disorders: a HuGE review 1232 and meta-analysis of evidence from observational studies. *Blood* 1233 2012;**119**:2392–400.
- 1234 [98] Green F, Humphries S. Genetic determinants of arterial thrombosis. 1235 *Baillières Clin Haematol* 1994;**7**:675–92.
- 1236 [99] Uitte de Willige S, Standeven KF, Philippou H, Ariëns RAS. The 1237 pleiotropic role of the fibrinogen gamma' chain in hemostasis. *Blood* 1238 2009;**114**:3994–4001.
- 1239 [100] Meijers JC, Tekelenburg WL, Bouma BN, Bertina RM, Rosendaal 1240 FR. High levels of coagulation factor XI as a risk factor for venous 1241 thrombosis. *N Engl J Med* 2000;**342**:696–701.
- 1242 [101] Snoep JD, Gaussem P, Eikenboom JC, Emmerich J, Zwaginga JJ, 1243 Holmes CE, et al. The minor allele of GP6 T13254C is associated with 1244 decreased platelet activation and a reduced risk of recurrent cardiovascular 1245 events and mortality: results from the SMILE-Platelets project. 1246 *J Thromb Haemost* 2010;**8**:2377–84.
- 1247 [102] Smith NL, Chen MH, Dehghan A, Strachan DP, Basu S, Soranzo N, 1248 et al. Novel associations of multiple genetic loci with plasma levels 1249 of factor VII, factor VIII, and von Willebrand factor: The CHARGE 1250 (Cohorts for Heart and Aging Research in Genome Epidemiology) 1251 Consortium. *Circulation* 2010;**121**:1382–92.
- 1252 [103] Smith NL, Rice KM, Bovill EG, Cushman M, Bis JC, McKnight 1253 B, et al. Genetic variation associated with plasma von Willebrand 1254 factor levels and the risk of incident venous thrombosis. *Blood* 1255 2011;**117**:6007–11.
- 1256 [104] Van Loon JE, Leebeek FW, Deckers JW, Dippel DW, Poldermans D, 1257 Strachan DP, et al. Effect of genetic variations in syntaxin-binding 1258 protein-5 and syntaxin-2 on von Willebrand factor concentration and 1259 cardiovascular risk. *Circ Cardiovasc Genet* 2010;**3**:507–12.
- 1260 [105] Huang J, Huffman JE, Yamakuchi M, Yamkauchi M, Trompet S, 1261 Asselbergs FW, et al. Genome-wide association study for circulating 1262 tissue plasminogen activator levels and functional follow-up implicates 1263 endothelial STXBP5 and STX2. *Arterioscler Thromb Vasc Biol* 1264 2014;**34**:1093–101.
- 1265 [106] Ye S, Huang Y, Joshi S, Zhang J, Yang F, Zhang G, et al. Platelet 1266 secretion and hemostasis require syntaxin-binding protein STXBP5. *J* 1267 *Clin Invest* 2014;**124**:4517–28.
- [107] Zhu Q, Yamakuchi M, Ture S, de la Luz Garcia-Hernandez M, Ko 1268 KA, Modjeski KL, et al. Syntaxin-binding protein STXBP5 inhibits 1269 endothelial exocytosis and promotes platelet secretion. *J Clin Invest* 1270 2014;**124**:4503–16.
- [108] Lowe GD, Haverkate F, Thompson SG, Turner RM, Bertina RM, 1271 Turpie AG, et al. Prediction of deep vein thrombosis after elective 1272 hip replacement surgery by preoperative clinical and haemostatic 1273 variables: the ECAT DVT Study. European Concerted Action on 1274 Thrombosis. *Thromb Haemost* 1999;**81**:879–86.
- [109] Tripodi A, Chantarangkul V, Martinelli I, Bucciarelli P, Mannucci PM. 1275 A shortened activated partial thromboplastin time is associated with 1276 the risk of venous thromboembolism. *Blood* 2004;**104**:3631–4.
- [110] Houlihan LM, Davies G, Tenesa A, Harris SE, Luciano M, Gow AJ, 1277 et al. Common variants of large effect in F12, KNG1, and HRG are 1278 associated with activated partial thromboplastin time. *Am J Hum Genet* 1279 2010;**86**:626–31.
- [111] Morange PE, Oudot-Mellakh T, Cohen W, Germain M, Saut N, Antoni 1280 G, et al. KNG1 Ile581Thr and susceptibility to venous thrombosis. 1281 *Blood* 2011;**117**:3692–4.
- [112] Gailani D, Renné T. Intrinsic pathway of coagulation and arterial 1282 thrombosis. *Arterioscler Thromb Vasc Biol* 2007;**27**:2507–13.
- [113] Merkulov S, Zhang WM, Komar AA, Schmaier AH, Barnes E, Zhou 1283 Y, et al. Deletion of murine kininogen gene 1 (mKng1) causes loss of 1284 plasma kininogen and delays thrombosis. *Blood* 2008;**111**:1274–81.
- [114] Cheng Q, Tucker EI, Pine MS, Sisler I, Matafonov A, Sun MF, et al. 1285 A role for factor XIIa-mediated factor XI activation in thrombus 1286 formation in vivo. *Blood* 2010;**116**:3981–9.
- [115] Sabater-Lleal M, Martínez-Perez A, Buil A, Folkersen L, Souto JC, 1287 Bruzelius M, et al. A genome-wide association study identifies KNG1 1288 as a genetic determinant of plasma factor XI Level and activated partial 1289 thromboplastin time. *Arterioscler Thromb Vasc Biol* 2012;**32**:2008–16.
- [116] Germain M, Chasman DI, de Haan H, Tang W, Lindström S, Weng 1290 LC, et al. Meta-analysis of 65 734 individuals identifies TSPAN15 and 1291 SLC44A2 as two susceptibility loci for venous thromboembolism. *Am* 1292 *J Hum Genet* 2015;**96**:532–42.
- [117] Lau L-M, Wee JL, Wright MD, Moseley GW, Hogarth PM, Ash- 1293 man LK, et al. The tetraspanin superfamily member CD151 regulates 1294 outside-in integrin alphaIIb beta3 signaling and platelet function. *Blood* 1295 2004;**104**:2368–75.
- [118] Goschnick MW, Lau LM, Wee JL, Liu YS, Hogarth PM, Robb LM, 1296 et al. Impaired “outside-in” integrin alphaIIb beta3 signaling and throm- 1297 bus stability in TSSC6-deficient mice. *Blood* 2006;**108**:1911–8.
- [119] Bailey RL, Herbert JM, Khan K, Heath VL, Bicknell R, Tomlinson 1298 MG. The emerging role of tetraspanin microdomains on endothelial 1299 cells. *Biochem Soc Trans* 2011;**39**:1667–73.
- [120] Poeter M, Brandherm I, Rossaint J, Rosso G, Shahin V, Skryabin BV, 1300 et al. Annexin A8 controls leukocyte recruitment to activated endothelial 1301 cells via cell surface delivery of CD63. *Nat Commun* 2014;**5**:3738.
- [121] Greinacher A, Wesche J, Hammer E, Füll B, Völker U, Reil A, et al. 1302 Characterization of the human neutrophil alloantigen-3a. *Nat Med* 1303 2010;**16**:45–8.
- [122] Curtis BR, Cox NJ, Sullivan MJ, Konkashbaev A, Bowens K, Hansen 1304 K, et al. The neutrophil alloantigen HNA-3a (5b) is located on choline 1305 transporter-like protein 2 and appears to be encoded by an R>Q154 1306 amino acid substitution. *Blood* 2010;**115**:2073–6.
- [123] Van Hylckama Vlieg A, Flinterman LE, Bare LA, Cannegieter SC, 1307 Reitsma PH, Arellano AR, et al. Genetic variations associated with 1308 recurrent venous thrombosis. *Circ Cardiovasc Genet* 2014;**7**:806–13. 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332

P. Suchon.

P.-E. Morange (pierre.morange@ap-hm.fr).

Inserm, UMR1062, « Nutrition, obésité et risque de thrombose », Faculté de médecine, Aix-Marseille Université, 27, boulevard Jean-Moulin, 13385 Marseille cedex 5, France.

Service d'hématologie biologique, Hôpital de la Timone, 264, rue St-Pierre, 13005 Marseille, France.

Toute référence à cet article doit porter la mention : Suchon P, Morange PE. Anomalies constitutionnelles de la coagulation prédisposant à la thrombose. EMC - Hématologie 2016;0(0):1-11 [Article 13-022-B-60].

Disponibles sur [www.em-consulte.com](http://www.em-consulte.com)Arbres  
décisionnelsIconographies  
supplémentairesVidéos/  
AnimationsDocuments  
légauxInformation  
au patientInformations  
supplémentairesAuto-  
évaluationsCas  
clinique

**Article de revue 2**

**“Genetics of Venous Thrombosis: update  
in 2015”**

*Thrombosis and Haemostasis 2015*

## Genetics of Venous Thrombosis: update in 2015

Pierre-Emmanuel Morange<sup>1,2,3</sup>; Pierre Suchon<sup>1,2,3</sup>; David-Alexandre Trégouët<sup>4,5,6</sup>

<sup>1</sup>Laboratory of Haematology, La Timone Hospital, Marseille, France; <sup>2</sup>Institut National pour la Santé et la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR\_S) 1062, Nutrition Obesity and Risk of Thrombosis, Marseille, France; <sup>3</sup>Aix-Marseille University, UMR\_S 1062, Nutrition Obesity and Risk of Thrombosis, Marseille, France; <sup>4</sup>INSERM UMR\_S 1166, Paris, France; <sup>5</sup>Sorbonne Universités, Université Pierre et Marie Curie (UPMC Univ Paris 06), UMR\_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases, Paris, France; <sup>6</sup>Institute for Cardiometabolism and Nutrition (ICAN), Paris, France

### Summary

Venous thrombosis (VT) is a common multifactorial disease with a genetic component that was first suspected nearly 60 years ago. In this review, we document the genetic determinants of the disease, and update recent findings delivered by the application of high-throughput genotyping and sequencing technologies. To date, 17 genes have been robustly demonstrated to harbour genetic variations associated with VT risk: *ABO*, *F2*, *F5*, *F9*, *F11*, *FGG*, *GP6*, *KNG1*, *PROC*, *PROCR*, *PROS1*, *SERPINC1*, *SLC44A2*, *STXBP5*, *THBD*, *TSPAN15* and

*VWF*. The common polymorphisms are estimated to account only for a modest part (~5%) of the VT heritability. Much remains to be done to fully disentangle the exact genetic (and epigenetic) architecture of the disease. A large suite of powerful tools and research strategies can be deployed on the large collections of patients that have already been assembled (and additional are ongoing).

### Keywords

Venous thrombosis, single nucleotide polymorphisms, gene mutations

### Correspondence to:

Dr. Pierre E. Morange  
Inserm – Hematology  
264 Rue Saint-Pierre  
13385 Marseille, France  
Tel.: +33 491386049, Fax: +33 491942332  
E-mail: pierre.morange@ap-hm.fr

Received: May 15, 2015

Accepted after minor revision: July 14, 2015

Epub ahead of print: September 10, 2015

<http://dx.doi.org/10.1160/TH15-05-0410>

Thromb Haemost 2015; 114: 910–919

## Introduction

Venous thrombosis (VT) is the pathologic expression of the development of a blood clot (i.e. thrombus), usually in the deep veins of the leg which can further obstruct venous circulation (deep-vein thrombosis, DVT) or subsequently embolise, travel to the lung and lead to pulmonary embolism (PE). VT is a multifactorial disease with a major contribution of genetic factors. The recognition that VT has a strong genetic risk component dates back to the 1960s (1, 2). Modern methodologies in genetic epidemiology estimate that the heritability of the disease is around 50% (3–5) and characterised by a sibling relative risk of ~2.50 (6).

Jordan and Nandorff (1) were the first, in 1956, to describe a familial clustering for VT. But this was only a decade later, in 1965, that the first genetic cause of VT was discovered: a partial anti-thrombin (AT) deficiency in a Norwegian family presenting familial VT aggregation (2). A few years later, in 1969, the *ABO* locus coding for the ABO blood group was mentioned for the time as a susceptibility gene for VT (7). But the impact of *ABO* on disease risk will be really acknowledged 40 years later (see thereafter). In the early 1980s, deficiencies in the other two natural coagulation inhibitors, protein C (PC) (8) and protein S (PS) (9), were also identified as additional culprit origins of VT. These deficiencies are the consequence of private mutations in their structural genes (*SERPINC1*, *PROC* and *PROS1* for AT, PC and PS deficiencies, respectively) that increase the risk of VT 10-fold in heterozygous

carriers. These plasmatic defects hamper the two main regulatory pathways of coagulation: the inhibition of serine proteases by anti-thrombin and that of the non-enzymatic cofactors VIIIa and Va by activated protein C (APC) and its cofactor protein S.

## Candidate gene approach

The 1990s were the advent of the candidate gene approach (10). Of course, genes coding molecules of the yin/yang coagulation/fibrinolysis pathways were natural candidates and subject to intensive research for identifying genetic polymorphisms associated with VT. Compared to the number of candidate genes assessed, unquestionable successes were relatively few. In 1994, the Q506 allele of the *F5* R506Q mutation (rs6025), also known as the Factor V Leiden (FVL) mutation, was discovered as a rather common (~5% allele frequency in the general population of European descent origin) susceptibility allele (11, 12). This allele is associated with a ~3-fold increased risk and influences VT risk through resistance to APC. More recently, another non-synonymous *F5* variant, K858R (rs4524), was also found to influence VT risk independently of the rs6025 variant (13, 14). Its risk allele has frequency ~0.75 and associates with an increase of the odds ratio (OR) for VT of ~1.20 (13–15). The FVL mutation's discovery was quickly followed in 1996 by the identification of the *F2* G20210A polymorphism (rs1799963) as another genetic risk factor for VT (16). This

polymorphism is located in the 3'UTR region of the prothrombin gene and the risk allele, rs1799963-A, increases plasma prothrombin concentrations (16). The risk allele is rare, ~2% of the general population, and is associated with a ~2.5 fold increased risk of developing VT. About 10 years then passed before the discovery of a novel genetic risk factor for VT. The *PROCR* gene coding for the endothelial protein C receptor (EPCR), a major regulator of PC activation (17), was an obvious candidate to investigate through the genetic prism. In 2004, Saposnik et al. (18) were the first to report the association of *PROCR* Ser219Gly (rs867186) with VT risk and the variability of plasma soluble EPCR. In a recent systematic review, the rs867186-G allele whose frequency is ~8% in the European ancestry is associated with an OR for VT of 1.22 (19). The last candidate gene that provided robust positive findings was the fibrinogen gamma gene (*FGG*) belonging to the fibrinolysis cascade. The T allele of the rs2066865 polymorphism, with frequency ~0.25, was found in 2005 to reduce gamma' fibrinogen plasma levels and to increase VT risk by a factor of ~1.50 (20).

## Common and low-frequency susceptibility alleles for VT – contribution of high-throughput genotyping

The early 2000s witnessed the development of micro-array based high throughput genotyping technologies, allied to a deeper understanding of the pattern of human sequence variations. This offered the opportunity to perform association studies with hundreds of thousand of single nucleotide polymorphisms (SNPs) expected to cover the whole genome without any biological and functional a priori on the tested SNPs. Genome-wide association studies (GWAS) represented an important advance compared to 'candidate gene' studies in which sample size were generally smaller and the variants assayed were limited to a selected few, often on the basis of imperfect understanding of biological pathways and often yielding associations that are difficult to replicate (21). Genome-wide association studies (GWAS) and meta-analysis of GWAS have identified lots of loci that contribute to complex diseases in humans. The National Human Research Institute Catalog of published GWAS indicates that by 2013, 1751 curated publications reported 11, 912 SNPs associations with  $P < 1 \times 10^{-5}$  (22).

The first attempt to perform agnostic association scan on large number of SNPs in the field of VT genetics was the work of Bezemer et al. in 2008 (23). It concentrated on approximately 20,000 SNPs, mainly non-synonymous of known genes with minor allele frequency (MAF) >5%, that were assessed for their association with VT through a multi-stage study including up to ~3,000 cases and 5,000 controls. Two novel susceptibility loci were identified using this approach, *GP6* and *CYP4V2*. *GP6* encodes glycoprotein VI that has a major role in collagen signalling and the identified VT-associated SNP was the Pro219Ser (rs1613662) variant. The common form, Pro219 (rs1613662-A), with frequency ~0.80, was associated with a 1.15 fold increased risk. This association was further confirmed in four independent populations of

French and American origins (24, 25). The association observed at the *CYP4V2* locus was later refined and explained by the presence of linkage disequilibrium (LD) between the *CYP4V2* gene and the nearby *F11* gene coding for Factor XI (26). Even though the true functional variants at the *F11* locus remain to be identified, strong evidence suggest that they influence VT risk through a modulation of FXI levels (26).

## Insights from HapMap2 based imputation

Following this first proof-of-concept study, several GWAS projects were launched to genotype between 300K and 600K SNPs and test their association with VT risk (24, 27–29). The number of SNPs was bolstered through statistical inference of variants not molecularly typed on the genotyping array but inferred through phasing and imputation guided by publicly available reference panels (30). These first GWAS imputed data were derived from the HapMap2 dataset containing 2,557,252 autosomal SNPs. This genetic resource was particularly efficient for inferring and assessing the impact on a phenotype of all untyped SNPs of the genome with MAF greater than 5% (31). No new robust VT-associated locus emerged from these studies that mainly confirmed the effects of the common alleles at *F5*, *F11*, *FGG*, *GP6* and *PROCR* genes discussed above. However, these studies permitted to highlight and to revisit the major role of the *ABO* locus in VT susceptibility that was let aside since the pioneer work of Jick (7). These GWAS studies not only confirmed that the A1 and B *ABO* blood groups were at a 1.5-fold increased risk of VT, but also suggested that a novel common SNP located in inton 1, rs2519093 could associate with VT (OR ~1.60) independently of the major *ABO* blood groups (28, 29). However, others (32) have suggested that *ABO* rs2519093, in strong LD ( $r^2 = 0.96$ ) with *ABO* rs579459, is tagging for *ABO* A1 blood group. Further studies are still mandatory to dissect whether the effect on VT of *ABO* rs2519093 is really independent of the *ABO* blood group effect.

The application of GWAS strategy to quantitative intermediate traits for the disease has also brought several new findings to the field. The general idea of such strategy is to apply GWAS principles to identify loci that associate with plasma levels of procoagulant proteins and then to look at the influence of the identified SNPs on the risk of VT. Genome wide association scans for SNPs associated with quantitative risk factors are usually performed in large samples of unrelated healthy individuals. But patients and family members can also be handled. Quantitative traits-associated SNPs that reached genome-wide significance (i.e. association p-value  $< 5 \times 10^{-8}$ ) can then be considered as good candidates for association with the disease. Since few SNPs are generally tested at this step, there is no need to apply stringent genome-wide significance level but just correct for the number of tested SNPs, enhancing the power to detect disease-SNP associations compared to a complete scan over the genome. The first major successes were brought by the application of HapMap2 imputation based GWAS on von Willebrand Factor (VWF) and Factor VIII (FVIII) plasma levels (33). From the eight loci that were found influencing VWF and/or



FVIII plasma levels, two apart from *ABO* were further observed to robustly associate with the risk of VT, *VWF* and *STXBP5*. *VWF* is the structural gene for VWF and the Ala789 allele of its coding Thr789Ala variant (rs1063856-G) associates with both increased VWF levels and VT risk (OR ~ 1.15) (34). While the association of *VWF* with VT-susceptibility was not surprising due to the key role of VWF in the coagulation cascade, that observed with *STXBP5* was rather unexpected. The *STXBP5* encodes syntaxin binding protein 5 (STXBP5; also known as Tomosyn-1). The less common *STXBP5* rs1039084-G allele of the coding Asn436Ser variant that was found to decrease VWF plasma levels was observed to be associated with an OR for VT of 0.90 (34). Of note, the same allele that is protective for VT was observed associated with an increased bleeding phenotype in women (35). Another interesting result is the recent observation that the *STXBP5* rs1039084 also associates with t-PA levels (36) suggesting a role of STXBP5 in the regulation of the secretory status of endothelium. Two studies have recently addressed the role of STXBP5 on the secretion of granular contents from platelets and endothelial cells (37, 38). These studies demonstrate different functions for STXBP5 in these two cell types. While STXBP5 facilitates granule release from platelets, it inhibits secretion from the Weibel-Palade bodies of endothelial cells. Two other genes identified in the GWAS analysis for VWF, *TC2N* and *STAB2*, also exhibited suggestive evidence for association with VT risk (27, 39) but additional support is needed to definitively declare them as susceptibility genes for the disease.

Another successful example of this GWAS approach on quantitative traits that further translates to the disease is the identification of the *KNG1* gene as a new VT disease locus. *KNG1* gene encodes high-molecular-weight kininogen (HK), an obvious candidate for VT pathophysiology. Among three SNPs that were reported to influence activated partial thromboplastin time (aPTT) in a GWAS conducted by Houlihan et al. (40), only one, the *KNG1* Ile581Thr (rs710446), was further confirmed to associate with VT (41). The less frequent rs710446-C allele, with frequency ~0.45, is associated with decreased aPTT levels and increased risk (OR~1.15) of the disease, as expected from the known traditional biology relating aPTT to coagulation and thrombosis.

In addition to aPTT, FVIII, VWF phenotypes aforementioned, other quantitative intermediate phenotypes to VT have been investigated through the GWAS strategy relying on HapMap2 imputed data. These include fibrinogen (42), D-dimers (43), PC (44), PS (45), AT (45) and PAI-1 (46) levels. None of the novel phenotype-SNP associations discovered in these investigations led so far to novel genetic associations with respect to VT risk.

## Insights from 1000 Genomes-based imputation

The HapMap2 reference database has been a valuable resource for assessing the impact of SNPs with MAF > 5% in numerous GWAS projects. However, the tagging for common SNPs was far from being complete even in known functional regions (47–50). The resequencing of 500 genes had revealed that a relatively high pro-

portion of common SNPs, ranging from 50% to 20% according to populations, could not be tagged by SNPs from HapMap (47). The same study estimated that only 30% of non synonymous SNPs were in high LD with any HapMap SNPs. This, with the increasing recognition of the importance of SNPs with lower MAF but larger genetic effects (50) allied to the development of new high-throughput sequencing technologies led to the constitution of international projects, such as the 1000 genomes projects (<http://www.1000genomes.org>) aiming at characterising all human polymorphisms with very low MAF. To overcome the HapMap2 limitation, the 1000G database (51) has become a standard reference for imputation based association analyses and make possible the inference of low-frequency SNPs with MAF between 1 and 5%. VT was one of the first trait that benefited from the application of the 1000G based imputation strategy (52). The analysis of 1,961 VT cases and 2,338 controls well imputed for ~6.7 millions with MAF higher than 1% enabled to re-discover the strong association at the *F2* rs1799963. This variant, with MAF ~2% in the control population, was missed by the previous GWAS because it could not be tagged by more frequent SNPs. This work nicely illustrated the relevance and the accuracy of the method to extend the spectrum of genetic variants that can be assessed in a GWAS context.

Following this work, all VT GWAS datasets that were previously imputed for HapMap2 (28, 29) were re-imputed, together with some additional newly set up GWAS studies, using the 1000G reference as part of the International Network on VENous Thrombosis (INVENT) consortium. Altogether, 12 GWAS studies for VT totalling 7,507 cases and 52,632 controls with imputed genotyped data at ~7 millions of SNPs were assembled and entered a meta-analysis of study-specific GWAS results with the aim to discover novel variants in novel or known genes that associate with VT risk (15). The INVENT analysis has produced a wealth of results whose in-depth investigation is still at its beginning. While no new low-frequency disease-associated allele was discovered, this meta-GWAS yielded new findings at known and unknown genes with respect to common alleles. First, it suggested that the association of *PROCR* and *F11* loci with VT is more complex than initially thought. A second *PROCR* SNP, rs6088735, independently added to the rs867186 discussed above to influence VT risk. The rs867186 has been previously reported to influence PC levels (44). The association at the *F11* locus was even more complex with three SNPs, rs1218625, rs2036914 and rs2289252, observed at this locus to additively influence VT risk. A preliminary haplotype analysis revealed that a unique at-risk haplotype can be derived from these three SNPs. Its frequency was ~0.40 and it associated with an increased OR of 1.40 (15, 27). As no low-frequency variant demonstrated suggestive statistical evidence at the *F11* locus, it is unlikely that the observed common haplotype tags for an uncommon variant and suggests that the functionality of this haplotype should be envisaged. More novel, and then opening novel perspectives with respect to VT pathophysiology, are the robust associations detected at the *TSPAN15* and *SLC44A2* loci for which no link with VT has ever been reported. Two common polymorphisms, *TSPAN15* rs78707713 with MAF ~0.10 and *SLC44A2*

rs2288904 with MAF ~0.20, were found to associate, and replicate in independent samples of ~3,000 cases and 2,600 controls, with VT. In both cases, the common alleles were those that associated with an increase of VT risk, OR~ 1.30 for *TSPAN15* and OR~1.20 for *SLC44A2*, respectively. No association with known haemostatic plasma markers was observed for these variants. The challenge is now to understand the functional consequences of these SNPs and to accurately elucidate the biological mechanism by which these genes and SNPs act on the risk of VT. The lead rs78707713 SNP at the *TSPAN15* gene, that codes for tetraspanin 15, a member of the tetraspanin superfamily acting scaffolding proteins, anchoring multiple proteins to the cell membrane (53), is intronic. No evidence has been collected so far to suggest that it could functionally influence pathophysiological process of VT. Conversely, the *SLC44A2* rs2288904 SNP is a non-synonymous (Arg154Gln) variant previously shown to associate with transfusion related acute lung injury (TRALI), a life-threatening complication of blood transfusion and the leading cause of transfusion-associated mortality in developed countries (54). It has just been proposed that the choline transporter-like protein-2 (CTL-2) encoded by *SLC44A2* could represent a new binding partner of VWF (55).

Similar to what has been done with the HapMap2 imputation strategy, one could expect that the application of the 1000G based imputation to quantitative risk factors for VT would lead to novel discoveries. So far, only one study reporting the results of a 1000G based GWAS on such a trait has been published (56) and it related to thrombin generation. This study identified the *ORM1* gene which encodes orosomucoid as a new gene contributing to thrombin generation. However the lead *ORM1* rs150611042 SNP did not demonstrate any evidence for association with the disease per se. Of note, in this study, the two main genetic factors influencing

thrombin generation were the *F2* rs1799963 and rs3136516 SNPs. While the association of the rs1799963 variant with VT is well established and was discussed above, it is worthy of note that rs3136516 demonstrates strong, but not genome-wide significance, statistical association ( $p = 5.65 \cdot 10^{-6}$ ) with VT in the INVENT meta-analysis. Larger sample size would be needed to confirm the association observed with the second SNP, that is relatively common and associated with a modest increase VT risk (OR~1.10). This association would be independent of that observed at rs1799963 and would suggest that two SNPs could exert additive effects at the *F2* locus to impact on VT risk. Several projects are ongoing for revisiting previous HapMap2 GWAS on quantitative traits using 1000G reference. The impact on VT risk of the SNPs that will emerge from these analyses would deserve further attention. For this purpose, the data and results assembled by the INVENT consortium could be of high value.

To sum up, 12 genes have been robustly demonstrated to harbour common and low-frequency susceptibility alleles for VT: *ABO*, *F11*, *F2*, *F5*, *FGG*, *GP6*, *KNG1*, *PROCR*, *SLC44A2*, *STXBP5*, *TSPAN15* and *VWF* (► Table 1). Even though the following estimates need to be revised in view of the INVENT results, common polymorphisms are expected to account only for a modest part (~5%) of the VT heritability (27) and of the sibling relative risk (~1.11) (29) mentioned in the introductory paragraph. They cannot therefore explain the whole genetic susceptibility to VT nor its familial aggregation. Family history still remains a strong risk factor for VT after adjusting for all common VT SNPs and coagulation defects (57, unpublished data). The increased risk for VT in relatives is incompletely explained by the presence of known VT defects, as the risk for VT in first degree relatives is increased even if patients do not have a detectable defect (58, 59). The question of

**Table 1: VT-associated variants identified by GWAS conducted in individuals of Caucasian origin.**

| Locus          | SNP                | Alleles <sup>a</sup> | Frequency <sup>b</sup> | OR <sup>c</sup> | Associated phenotype <sup>d</sup>     | References <sup>e</sup> |
|----------------|--------------------|----------------------|------------------------|-----------------|---------------------------------------|-------------------------|
| <i>ABO</i>     | [O, A2] vs [A1, B] |                      | 0.30                   | 1.50            | ↑ VWF, ↑ VIII                         | 28, 29                  |
|                | rs2519093          | G/A                  | 0.24                   | 1.68            | Unknown                               | 28                      |
| <i>F2</i>      | rs1799963          | G/A                  | 0.02                   | 2.50            | ↑ FII                                 | 16                      |
| <i>F5</i>      | rs6025             | G/A                  | 0.05                   | 3.00            | Resistance to activated protein C     | 12                      |
| <i>F11</i>     | rs2036914          | C/T                  | 0.52                   | 1.35            | ↑ FXI                                 | 26                      |
|                | rs2289252          | C/T                  | 0.41                   | 1.35            | ↑ FXI                                 | 26                      |
| <i>FGG</i>     | rs2066865          | C/T                  | 0.25                   | 1.47            | ↓ Fibrinogen $\gamma\gamma'$          | 20                      |
| <i>GP6</i>     | rs1613662          | A/G                  | 0.82                   | 1.15            | ↑ platelet activation and aggregation | 23                      |
| <i>KNG1</i>    | rs710446           | T/C                  | 0.45                   | 1.20            | ↓ aPTT                                | 41                      |
| <i>PROCR</i>   | rs867186           | A/G                  | 0.07                   | 1.22            | ↑ sEPCR, ↑ PC                         | 18                      |
| <i>SLC44A2</i> | rs2288904          | T/C                  | 0.79                   | 1.19            | Still unknown                         | 15                      |
| <i>STXBP5</i>  | rs1039084          | A/G                  | 0.46                   | 1.11            | ↑ VWF                                 | 34                      |
| <i>TSPAN15</i> | rs78707713         | G/A                  | 0.88                   | 1.28            | Still unknown                         | 15                      |
| <i>VWF</i>     | rs1063856          | A/G                  | 0.37                   | 1.15            | ↑ VWF                                 | 34                      |

<sup>a</sup>Underlined are the at risk alleles. <sup>b</sup>Estimates of the risk allele frequency observed in References population. <sup>c</sup>Estimates of the odds ratio (OR) associated with the risk allele. <sup>d</sup>Phenotype associated with the at risk allele. <sup>e</sup>According to References numbering.

missing or hidden heritability (31, 60) that underlies complex diseases also holds for VT. An army of sources could contribute to this missing heritability including extremely rare variants (61).

## Rare variants for VT

The 1000G based imputation strategy is often proposed as an efficient tool for discovering novel disease-associated variants with MAF as low as 1%. However, despite being based on the largest GWAS resources on VT, the INVENT meta-analysis failed to detect such associations. The 1000G INVENT analysis did discover two newly VT-associated SNPs, but the discovery of these two SNPs that were common was likely due to the increase in sample size compared to previous GWAS on VT rather than to the use of a more comprehensive panel of genetic variations used to impute non-typed variants. The INVENT meta-analysis had a 80% power to detect, at genome-wide significance level, the effect of a variant with MAF 0.01 as soon as it was associated with an increased risk of ~1.80. The power fell down to 25% for an OR of ~1.65. Without a dramatic increase in the number of VT patients included in GWAS, it is unlikely that low-frequency variants associated with VT risk could be discovered using a standard GWAS framework. Other strategies should then be envisaged.

## Genotyping arrays for low-frequency coding variants

Novel original discoveries in the genetics of human complex traits are expected from the application to large epidemiological studies of DNA genotyping arrays dedicated to known coding variants including low-frequency ones with  $0.5\% < \text{MAF} < 5\%$ , that may not be well imputed by statistical inference. The Illumina Exome chip is a popular array that is currently used by many projects (62). It has been developed to include on a dedicated array all rare variants identified through the sequencing of the coding part of the genome (i.e. exome) of ~12,000 individuals representing diverse populations and a range of common conditions such as asthma, stroke, myocardial infarction, extreme LDL, extreme blood pressure. The general idea under the use of such array is that common variations might be less important than rare variations because selective pressure has had more time to eliminate deleterious alleles and that coding variants can more easily pinpoint the causal gene in regions enriched for several loci. The analysis of such genotype data requires dedicated statistical tools that jointly test all variants of the same gene for association with the trait of interest rather than testing them individually as it is done in standard GWAS. This grouping is made necessary by the rarity of individuals carrying the low-frequency alleles. Because variants of the same gene are tested together, these methodologies are generally referred to as gene-based test approaches (63). The deployment of the exome array has already produced a few successes with the identification of novel variants at known and unknown loci associated with complex traits, such as psoriasis (64), insulin-related traits (65),

glucose (66), haematological traits (67), lipid levels (68, 69), myocardial infarction (MI) (68, 69) and non-alcohol fatty liver disease (NAFLD) (70). However, not all discovered variants were with low frequency. For example, the same coding *TM6FS2* rs58542926 (Gly167Lys) that associated with lipids, MI and NAFLD (68–70) was in fact a rather frequent SNP (MAF ~9%) whose associations with studied phenotypes were missed by previous GWAS analyses. As part of the French GENMED project (<http://www.genmed.fr/>), 2,700 VT patients and 3,000 healthy individuals have recently been typed with this array and similar international VT projects are ongoing. One can reasonably anticipate that the application of this array to VT will lead in the next coming years to new discoveries in the field.

## Next-generation sequencing for extremely rare variants

By design, genotyping arrays are not able to assess unknown genetic variations. Sequencing is then the method of choice for identifying unknown, or extremely rare (<5%), mutations causing diseases. The next-generation sequencing (NGS) technologies have radically revolutionised the research landscape in the field of human disease genetics and offer a powerful tool to characterise the spectrum of very rare disease-causing variants. NGS of DNA samples can be applied to a set of candidate genes (target gene resequencing), to the whole coding part of the genome (the so-called exome) or in a complete agnostic manner by covering the whole genome. Only the latter two strategies afford the discovery of unsuspected genes and novel biological pathways related to the disease under study (71, 72).

## Targeted gene sequencing

Four candidate genes for VT have been investigated with success through targeted gene re-sequencing. In a thrombophilic patient with factor IX (FIX) activity eight times higher than the normal physiological levels, a gain-of-function mutation (R338L) was discovered in the *F9* gene (73). This mutation, named as the FIX Padua mutation in reference to the Italian city where the patient was hospitalised, is anticipated to be a private mutation as it has not been reported in the Exome Aggregation Consortium (ExAC) database (<http://exac.broadinstitute.org>) cataloguing all variants detected in the coding sequences of more than 60,000 individuals. Worthy of note, Finn et al used the sequence of the abnormally potent FIX for construction of improved gene therapy vectors to treat haemophilia B (74).

Two different extremely rare mutations at the same position of the *F2* gene, R596L and R596Q, have recently been discovered causing inherited thrombophilia due to resistance to AT. The former, named the prothrombin Yukuashi mutation, was detected in a Japanese family (75) and the latter, known as the prothrombin Belgrade mutation, in two Serbian families (76).

By re-sequencing the entire genomic sequence of the *THBD* gene encoding thrombomodulin in 60 VT patients and 60 controls



from China, the c-151G>T variant in the 5'UTR region of the gene was discovered as a low-frequency SNP affecting VT risk (77). Its risk-allele frequency was estimated to be ~1% in the Chinese population and associated with a 2.8-fold increased risk of VT. It remains to be elucidated whether this genetic risk factor is specific to the Chinese population and/or what could be its prevalence in other populations.

The last example of the efficiency of the candidate gene re-sequencing strategy is the recent discovery of the FV Nara mutation (F5 W1920R) in a 13-year-old Japanese boy who developed recurrent VT during oral anticoagulant treatment (78). The patient has reduced FV levels and pronounced APC resistance due to homozygosity at the detected missense mutation. In line with the observations made in the patient, recombinant FV Nara showed reduced expression in conditioned media (~50% of wild-type FV) and conferred APC resistance to reconstituted FV-deficient plasma. Because the FV Nara mutation is located in the light chain of FV(a), far from the known APC-cleavage sites, the molecular mechanism by which it causes APC resistance remains unclear (79).

Altogether, these examples strongly advocate for extensive re-sequencing of the validated VT genes aforementioned and clearly demonstrate that both rare and common variants at a given gene can act to influence VT risk.

### Whole exome sequencing (WES)

In presence of familial aggregation of VT that cannot be explained by a culprit mutation in known VT associated genes, the WES approach seems rather intuitive given the assumption that the causal mutation(s) lie(s) in (a) coding region(s). Numerous WES projects have demonstrated the power of the approach to identify very rare variants responsible for familial forms of human diseases. For instance, a major step forward was made during the last four years using this approach that resulted in the discovery of novel genes responsible for inherited platelet bleeding disorders (80–84). So far, its successful application to VT has never been reported. As part of the French INNOVTE network (<http://www.fcrin.org/en/support-tools/innovte-thrombosis>), a national initiative has recently been launched to make an inventory of all families with at least three family members who had unprovoked VT across three generations with the aim of performing WES in families excluded for mutations in known VT genes. The exome sequencing of several families has already started.

The WES strategy can of course be applied to unrelated idiopathic patients without any familial information. However, the outstanding wealth of rare single-nucleotide variants in individual human genomes calls for stringent criteria and large sample size to establish association, and hence a causal relationship, between rare single-nucleotide variants and a disease. So far, the WES strategy applied to large sample of independent disease patients have produced little success in identifying novel disease genes. It has rather led to the discovery of new rare mutations in genes already known to be involved in the pathophysiology of the disease under study (e.g. 85, 86).

### Whole genome sequencing (WGS)

A large number of instances have been accumulated to highlight the role of non coding variations in the molecular aetiology of complex diseases. WGS appears to be the method of choice for identifying extremely rare non coding variants that cannot be properly addressed by WES. The WGS strategy has also the advantages over WES to be more sensitive to detect variations in coding regions where target capture is less efficient (87) and to detect a larger spectrum of genetic variants including structural variants (88). Both familial cases and unrelated disease patients can be assessed by WGS. But again, due to the complexity induced by the high number of detected variations, a family-based strategy relying on the information about the segregation of variants within affected and unaffected relatives substantially enhanced the power to detect the culprit variants (89). So far, WGS has been particularly efficient to detect causal variations in Mendelian disorders (e.g. 90) and few complex diseases such autism (91, 92) and some cancers (93, 94). At its early stage, WGS was preferentially advocated for families or patients for which WES had remained unfruitful. However, as the cost of the WGS dramatically decreased up to 1,300\$ (for a mean coverage of 30X), and as WGS presents several merits over WES, the utility of WES (whose cost is ~450\$ for the same coverage) starts to be questioned.

### Perspectives

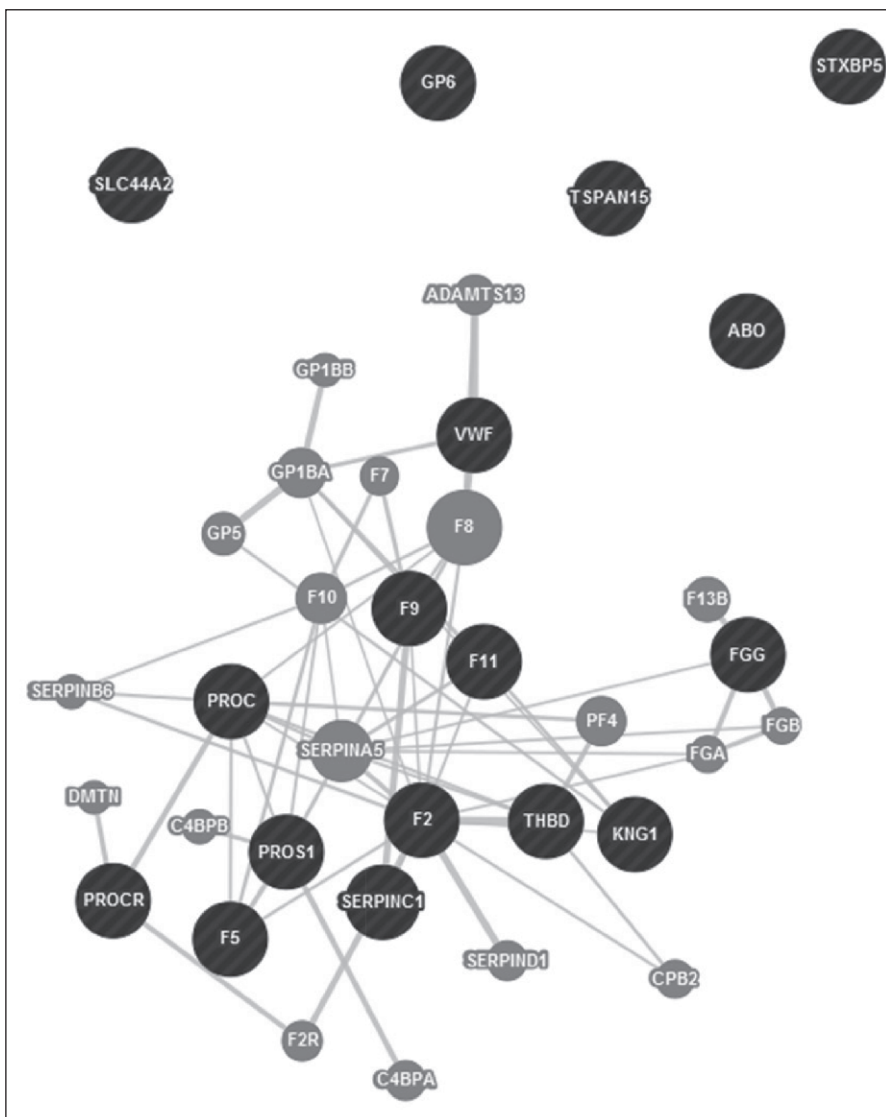
If we add to the list of genes with common and low-frequency susceptibility alleles for VT (see above) those known to carry extremely rare mutations responsible for VT (*F9*, *PROC*, *PROS1*, *SERPINC1*, and *THBD*), the total number of established VT associated genes raises to 17. ► Figure 1 displays a schematic network of the physical interactions between these 17 genes as produced by the GeneMania (95) tool. GeneMANIA interrogates several protein-protein interactions databases to identify whether input proteins have been reported to be functionally related through direct physical interaction. Similar networks can also be derived using other functional information such as co-expression of genes/proteins. ► Figure 1 well illustrates how most of these genes are strongly interconnected. These connections can be direct or involve others partners among which one can easily notice some genes that have previously been suspected to associate with VT but without achieving sufficient level of evidence. Examples include *ADAMTS13* (96), *CPB2* (97) and *C4BPA* (98). It has been proposed that SNP x SNP interactions could contribute to the missing heritability of complex diseases (99). Rather than performing a SNP x SNP interaction scan over the whole genome which would require tremendous sample size to achieve statistical power while controlling for the multiplicity of the tests (100), the information of known physical interactions between genes as in ► Figure 1 could serve as an a priori to test for the presence of interaction only between predefined subsets of polymorphisms mapping to genes known to physically interact with each other.



Not all the missing heritability is hiding in DNA genetic variations and epigenetic mechanisms are also proposed as a possible source that could contribute to it (101). Although the role of epigenetic mechanisms in some rare developmental syndromes and in cancer is well established, systematic examination of their contribution to common non malignant disease is only beginning. For this purpose, specific high throughput technologies are available to quantify non coding RNAs expression and methylation profiles from cells, tissues and blood. A primary concern in this new era of epigenetic epidemiology is the tissue-specific nature of the epigenome as in large well-phenotype sample cohorts, DNA from whole blood and plasma or citrate are often the only source of biological material available. While assessment of DNA methylation in whole blood has been shown able to identify robust and biologically relevant epigenetic variation related to cardiometabolic quantitative risk factors (102, 103), it remains to be determined whether it can be relevant for specifically VT-related traits. A first

attempt was undertaken in this spirit by searching for blood DNA methylation marks associated with thrombin generation markers using the dedicated Illumina HumanMethylation450 array in a sample of 350 individuals (104). No conclusive results were obtained from this pioneer work. This would suggest that DNA methylation measured in whole blood DNA may not be a good mirror of what happens in specific cell type more relevant to VT hemostasis regulation and/or that methylation marks in peripheral blood cells have tiny effects on plasma levels of thrombin generation that would require much larger sample size in order to be statistically detected.

High-throughput technologies also offered the opportunity to simultaneously quantify the expression levels of several thousand genes using microarrays. Comparing expression profiles between samples of disease patients and controls enables the identification of genes and/or set of genes simultaneously up or downregulated in patients, then pointing out functional modules that could be



**Figure 1: Physical interaction map derived from established VT associated genes.** Map produced by the GeneMANIA online tool (95) with the 17 established VT-associated genes (in black circles) as input. Linked circles represent pairs of genes that have been reported to physically interact with each other in the various protein-protein interaction sources implemented in GeneMANIA. The very recently suggested *SLC44A2* – *VWF* physical interaction (discussed in the text) is not yet indicated in this map. Grey genes are additional genes identified by GeneMANIA to physically interact with VT-associated genes. The stronger evidence for interaction, the thicker the link. This map is not exhaustive and must be viewed as a simplified illustration of the main physical connections linking together, as much as possible, the 17 VT genes. The lack of connections at few genes does not preclude their co-regulation with other shown genes. For examples, *ABO* and *STXBP5* have been shown to influence *vWF* plasma levels (see text).

implicated in the disease aetiology. A recent application to VT has been proposed to identify in whole blood expression signatures differentiating different clinical phenotypes of VT (105). The profiles obtained distinguish patients with recurrent, unprovoked VT from healthy controls and patients with provoked VT only.

The annual incidence of VT is ~1‰ (106) but it substantially varies across populations (107), and the prevalence of known genetic risk factors of VT can differ according to ethnicity such as FVL and *F2 G20210A* mutation that are very rare or absent among nonwhites. Until now all GWAS have been performed in individuals of European ancestry (whites) where it is known that LD can extend over large genomic regions compared to other populations, including Africans. Thoroughly designed epidemiological research (including GWAS) should be expanded to non-white populations with higher genetic diversity to facilitate the identification of additional genetic risk factors for VT.

Fifty percent of VT events occur without any precipitating conditions such as cancer, surgery, trauma, immobilisation, pregnancy, long haul flights, and are designed as unprovoked VT. The annual rate of recurrence is 10% in unprovoked patients stopping the anti coagulant treatment and ~10% of VT patients will die from fatal PE. Of survivors, 25–50% will have lasting debilitating health problems such as post-thrombotic syndrome, severely hampering mobility and quality of life. Up to now, much efforts have been made to characterise the genetics of prevalent VT without focused attention on the distinction between DVT and PE. There are unmet needs for identifying (epi-)genetic factors predisposing more to PE than to DVT and those predicting the recurrence of the event (108). A few initiatives have recently been conducted to address the later issue. van Hylckama Vlieg et al. assessed the predictive value of multiple SNPs testing for the risk of recurrence using VT-associated polymorphisms at the *ABO*, *F11*, *F2*, *F5* and *FGG* loci (109). Using a 5-SNPs genetic risk score (GRS), they were able to stratify patients on their recurrence risk: the six-year cumulative incidence of recurrence was high for individuals with  $\geq 5$  (20.3%) and low for individuals with  $\leq 1$  risk allele (9.4%) (110). The predictive ability was similar in individuals with first provoked or unprovoked event. The clinical utility of this GRS must be now assessed in a prospective randomised study where the duration of anticoagulant treatment would be individually adapted according to each patient's score.

To conclude, VT is the third leading cause of cardiovascular mortality after coronary artery disease and stroke and is responsible for a large public health burden (105). Novel major genetic discoveries have recently been obtained using high-throughput genotyping and sequencing technologies but, as discussed in this review, much remains to be done to fully disentangle the exact genetic (and epigenetic) architecture of the disease. A large suite of powerful tools and research strategies can be deployed on the large collections of patients that have already been assembled (and additional are ongoing).

### Acknowledgements

Pierre-Emmanuel Morange and David-Alexandre Trégouët's work forms part of the research programs on venous thromboembolism

genetics which are supported and funded by the the GenMed LABEX (ANR-10-LABX-0013) and the French INvestigation Network On Venous Thrombo-Embolism (INNOVTE).

### Conflicts of interest

None declared.

### References\*

- Jordan FL, Nandorff A. The familial tendency in thrombo-embolic disease. *Acta Med Scand* 1956; 156: 267–275.
- Egeberg O. Inherited antithrombin deficiency causing thrombophilia. *Thromb Diath Haemorrh* 195; 13: 516–530.
- Souto JC, et al. Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation* 2000; 101: 1546–1551.
- Larsen TB, et al. Major genetic susceptibility for venous thromboembolism in men: a study of Danish twins. *Epidemiology* 2003; 14: 328–332.
- Heit JA, et al. Familial segregation of venous thromboembolism. *J Thromb Haemost* 2004; 2: 731–736.
- Zöller B, et al. Age- and gender-specific familial risks for venous thromboembolism: a nationwide epidemiological study based on hospitalizations in Sweden. *Circulation* 2011; 124: 1012–1020.
- Jick H, et al. Venous thromboembolism disease and ABO blood type. A cooperative study. *Lancet* 1969; 1: 539–542.
- Griffin JH, et al. Deficiency of protein C in congenital thrombotic disease. *J Clin Invest* 1981; 68: 1370–1373.
- Schwarz HP, et al. Plasma protein S deficiency in familial thrombotic disease. *Blood* 1984; 64: 1297–1300.
- Clerget-Darpoux F. Overview of strategies for complex genetic diseases. *Kidney Int* 1998; 53: 1441–1445.
- Dahlbäck B, et al. 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. *Proc Natl Acad Sci USA* 1993; 90: 1004–1008.
- Bertina RM, et al. Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* 1994; 369: 64–67.
- Smith NL, et al. Association of genetic variations with non fatal venous thrombosis in postmenopausal women. *J Am Med Assoc* 2007; 297: 489–498.
- Bezemer ID, et al. Updated analysis of gene variants associated with deep vein thrombosis. *J Am Med Assoc* 2010; 303: 421–422.
- Germain M, et al. Meta-analysis of 65, 734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism. *Am J Hum Genet* 2015; 96: 532–542.
- Poort SR, 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: 3698–3703.
- Stearns-Kurosawa DJ, et al. The endothelial cell protein C receptor augments protein C activation by the thrombin-thrombomodulin complex. *Proc Natl Acad Sci USA* 1996; 93: 10212–10216.
- 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: 1311–1318.
- Dennis J, et al. The endothelial protein C receptor (PROCR) Ser219Gly variant and risk of common thrombotic disorders: a HuGE review and meta-analysis of evidence from observational studies. *Blood* 2012; 119: 2392–2400.
- Uitte de Willige S, et al. Genetic variation in the fibrinogen gamma gene increases the risk for deep venous thrombosis by reducing plasma fibrinogen gamma' levels. *Blood* 2005; 106: 4176–4183.
- Zondervan Kt, Cardon LR. Designing candidate gene and genome-wide case-control association studies. *Nat Protoc* 2007; 2: 2492–2501.
- Welter D, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 2014; 42: D1001–1006.

\* Some earlier investigations on genetic variants of candidate genes are not cited due to space constraints.

23. Bezemer ID, et al. Gene variants associated with deep vein thrombosis. *J Am Med Assoc* 2008; 299: 1306–1314.
24. Tregouet DA, 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: 5298–5303.
25. Austin H, et al. New gene variants associated with venous thrombosis: a replication study in White and Black Americans. *J Thromb Haemost* 2011; 9: 489–495.
26. Li Y, et al. Genetic variants associated with deep vein thrombosis: the F11 locus. *J Thromb Haemost* 2009; 7: 1802–1808.
27. Germain M, et al. Genetics of venous thrombosis: insights from a new genome wide association study. *PLoS One* 2011; 6: e25581.
28. Heit JA, et al. A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. *J Thromb Haemost* 2012; 10: 1521–1531.
29. 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. *Genet Epidemiol* 2013; 37: 512–521.
30. Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet* 2010; 11: 499–511.
31. Manolio TA, et al. Finding the missing heritability of complex diseases. *Nature* 2009; 461: 747–753.
32. Bruzelius M, et al. Influence of coronary artery disease-associated genetic variants on risk of venous thromboembolism. *Thromb Res* 2014; 134: 426–432.
33. Smith NL, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. *Circulation* 2010; 121: 1382–1392.
34. Smith NL, et al. Genetic variation associated with plasma von Willebrand factor levels and the risk of incident venous thrombosis. *Blood* 2011; 117: 6007–6011.
35. van Loon JE, et al. Effect of genetic variations in syntaxin-binding protein-5 and syntaxin-2 on von Willebrand factor concentration and cardiovascular risk. *Circ Cardiovasc Genet* 2010; 3: 507–512.
36. Huang J, et al. Genome-wide association study for circulating tissue plasminogen activator levels and functional follow-up implicates endothelial STXBP5 and STX2. *Arterioscler Thromb Vasc Biol* 2014; 34: 1093–1101.
37. Ye S, et al. Platelet secretion and hemostasis require syntaxin-binding protein STXBP5. *J Clin Invest* 2014; 124: 4517–4528.
38. Zhu Q, et al. Syntaxin-binding protein STXBP5 inhibits endothelial exocytosis and promotes platelet secretion. *J Clin Invest* 2014; 124: 4503–4516.
39. Morange PE, et al. Impact on venous thrombosis risk of newly discovered gene variants associated with FVIII and VWF plasma levels. *J Thromb Haemost* 2011; 9: 229–231.
40. Houlihan LM, et al. Common variants of large effect in F12, KNG1, and HRG are associated with activated partial thromboplastin time. *Am J Hum Genet* 2010; 86: 626–631.
41. Morange PE, et al. KNG1 Ile581Thr and susceptibility to venous thrombosis. *Blood* 2011; 117: 3692–3694.
42. Sabater-Lleal M, et al. Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease. *Circulation* 2013; 128: 1310–1324.
43. Smith NL, et al. Genetic predictors of fibrin D-dimer levels in healthy adults. *Circulation* 2011; 123: 1864–1872.
44. Tang W, et al. Genome-wide association study identifies novel loci for plasma levels of protein C: the ARIC study. *Blood* 2010; 116: 5032–5036.
45. Oudot-Mellakh T, et al. Genome wide association study for plasma levels of natural anticoagulant inhibitors and protein C anticoagulant pathway: the MARTHA project. *Br J Haematol* 2012; 157: 230–239.
46. Huang J, et al. Genome-wide association study for circulating levels of PAI-1 provides novel insights into its regulation. *Blood* 2012; 120: 4873–4881.
47. Xu Z, et al. Tag SNP selection for candidate gene association studies using HapMap and gene resequencing data. *Eur J Hum Genet* 2007; 15: 1063–1070.
48. McPherson R. A gene-centric approach to elucidating cardiovascular risk. *Circ Cardiovasc Genet* 2009; 2: 3–6.
49. Evans DM, et al. To what extent do scans of non-synonymous SNPs complement denser genome-wide association studies? *Eur J Hum Genet* 2008; 16: 718–723.
50. Cambien F, Tiret L. Genetics of cardiovascular diseases: from single mutations to the whole genome. *Circulation* 2007; 116: 1714–1724.
51. The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* 2010; 467: 1061–1073.
52. 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: e38538.
53. Berditchevski F. Complexes of tetraspanins with integrins: more than meets the eye. *J Cell Sci* 2001; 114: 4143–4151.
54. Mair DC, Eastlund T. The pathophysiology and prevention of transfusion-related acute lung injury (TRALI): a review. *Immunohematology* 2010; 26: 161–173.
55. Bayat B, et al. Choline Transporter-Like Protein-2: New von Willebrand Factor-Binding Partner Involved in Antibody-Mediated Neutrophil Activation and Transfusion-Related Acute Lung Injury. *Arterioscler Thromb Vasc Biol* 2015; Epub ahead of print.
56. Rocanin-Arjo A, et al. A meta-analysis of genome-wide association studies identifies ORM1 as a novel gene controlling thrombin generation potential. *Blood* 2014; 123: 777–785.
57. Cohen W, et al. Risk assessment of venous thrombosis in families with known hereditary thrombophilia: the MARseilles-Nimes prediction model. *J Thromb Haemost* 2014; 12: 138–146.
58. Sørensen HT, et al. Familial risk of venous thromboembolism: a nationwide cohort study. *J Thromb Haemost* 2011; 9: 320–324.
59. Couturaud F, et al. Factors that predict risk of thrombosis in relatives of patients with unprovoked venous thromboembolism. *Chest* 2009; 136: 1537–1545.
60. Sadee W, et al. Missing heritability of common diseases and treatments outside the protein-coding exome. *Hum Genet* 2014; 133: 1199–1215.
61. Zuk O, et al. Searching for missing heritability: designing rare variant association studies. *Proc Natl Acad Sci USA* 2014; 111: E455–E464.
62. Guo Y, et al. Illumina human exome genotyping array clustering and quality control. *Nat Protoc* 2014; 9: 2643–2662.
63. Lee S, et al. Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet* 2014; 95: 5–23.
64. Zuo X, et al. Whole-exome SNP array identifies 15 new susceptibility loci for psoriasis. *Nat Commun* 2015; 6: 6793.
65. Huyghe JR, et al. Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. *Nat Genet* 2013; 45: 197–201.
66. Wessel J, et al. Low-frequency and rare exome chip variants associate with fasting glucose and type 2 diabetes susceptibility. *Nat Commun* 2015; 6: 5897.
67. Auer PL, et al. Rare and low-frequency coding variants in CXCR2 and other genes are associated with hematological traits. *Nat Genet* 2014; 46: 629–634.
68. Peloso GM, et al. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56, 000 whites and blacks. *Am J Hum Genet* 2014; 94: 223–232.
69. Holmen OL, et al. Systematic evaluation of coding variation identifies a candidate causal variant in TM6SF2 influencing total cholesterol and myocardial infarction risk. *Nat Genet* 2014; 46: 345–351.
70. Kozlitina J, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2014; 46: 352–356.
71. Metzker ML. Sequencing technologies – the next generation. *Nat Rev Genet* 2010; 11: 31–46.
72. Koboldt DC, et al. The next-generation sequencing revolution and its impact on genomics. *Cell* 2013; 155: 27–38.
73. Simioni P, et al. X-linked thrombophilia with a mutant factor IX (factor IX Padua). *N Engl J Med* 2009; 361: 1671–1675.
74. Finn JD, et al. The efficacy and the risk of immunogenicity of FIX Padua (R338L) in hemophilia B dogs treated by AAV muscle gene therapy. *Blood* 2012; 120: 4521–4523.
75. Miyawaki Y, et al. Thrombosis from a prothrombin mutation conveying anti-thrombin resistance. *N Engl J Med* 2012; 366: 2390–2396.
76. Djordjevic V, et al. A novel prothrombin mutation in two families with prominent thrombophilia—the first cases of antithrombin resistance in a Caucasian population. *J Thromb Haemost* 2013; 11: 1936–1939.
77. Tang L, et al. Common genetic risk factors for venous thrombosis in the Chinese population. *Am J Hum Genet* 2013; 92: 177–187.

78. Nogami K, et al. Novel FV mutation (W1920R, FVNara) associated with serious deep vein thrombosis and more potent APC resistance relative to FVLeiden. *Blood* 2014; 123: 2420–2428.
79. Castoldi E. FV and APC resistance: the plot thickens. *Blood* 2014; 123: 2288–2289.
80. Albers CA, et al. Exome sequencing identifies NBEAL2 as the causative gene for gray platelet syndrome. *Nat Genet* 2011; 43: 735–737.
81. Albers CA, et al. Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome. *Nat Genet* 2012; 44: 435–439.
82. Kunishima S, et al. ACTN1 mutations cause congenital macrothrombocytopenia. *Am J Hum Genet* 2013; 92: 431–438.
83. Canault M, et al. Human CalDAG-GEFI gene (RASGRP2) mutation affects platelet function and causes severe bleeding. *J Exp Med* 2014; 211: 1349–1362.
84. Noetzi L, et al. Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. *Nat Genet* 2015; 47: 535–538.
85. Do R, et al. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature* 2015; 518: 102–106.
86. SIGMA Type 2 Diabetes Consortium Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. *J Am Med Assoc* 2014; 311: 2305–2314.
87. Belkadi A, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad Sci USA* 2015; 112: 5473–5478.
88. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med* 2010; 61: 437–455.
89. Ott J, et al. Genetic linkage analysis in the age of whole-genome sequencing. *Nat Rev Genet* 2015; 16: 275–284.
90. Lupski JR, et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med* 2010; 362: 1181–1191.
91. Jiang YH, et al. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. *Am J Hum Genet* 2013; 93: 249–263.
92. Michaelson JJ, et al. Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell* 2012; 151: 1431–1442.
93. Guo G, et al. Whole-genome and whole-exome sequencing of bladder cancer identifies frequent alterations in genes involved in sister chromatid cohesion and segregation. *Nat Genet* 2013; 45: 1459–1463.
94. Gartner JJ, et al. Whole-genome sequencing identifies a recurrent functional synonymous mutation in melanoma. *Proc Natl Acad Sci USA* 2013; 110: 13481–13486.
95. Warde-Farley D, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 2010; 38 (Suppl): W214–20.
96. Lotta LA, et al. Next-generation sequencing study finds an excess of rare, coding single-nucleotide variants of ADAMTS13 in patients with deep vein thrombosis. *J Thromb Haemost* 2013; 11: 1228–1239.
97. Martini CH, et al. The effect of genetic variants in the thrombin activatable fibrinolysis inhibitor (TAFI) gene on TAFI-antigen levels, clot lysis time and the risk of venous thrombosis. *Br J Haematol* 2006; 134: 92–94.
98. 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; 115: 4644–4650.
99. Zuk O, et al. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci USA* 2012; 109: 1193–1198.
100. Greliche N, et al. A genome-wide search for common SNP x SNP interactions on the risk of venous thrombosis. *BMC Med Genet* 2013; 14: 36.
101. Monte E, Vondriska TM. Epigenomes: the missing heritability in human cardiovascular disease? *Proteomics Clin Appl* 2014; 8: 480–487.
102. Dick KJ, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet* 2014; 383: 1990–1998.
103. Gagnon F, et al. Robust validation of methylation levels association at CPT1A locus with lipid plasma levels. *J Lipid Res* 2014; 55: 1189–1191.
104. Rocañín-Arjó A, et al. Thrombin generation potential and whole-blood DNA methylation. *Thromb Res* 2015; 135: 561–564.
105. Lewis DA, et al. Whole blood gene expression profiles distinguish clinical phenotypes of venous thromboembolism. *Thromb Res* 2015; 135: 659–665.
106. ISTH Steering Committee for World Thrombosis Day. Thrombosis: a major contributor to global disease burden. *Thromb Haemost* 2014; 112: 843–852.
107. Bounameaux H, Rosendaal FR. Venous thromboembolism: why does ethnicity matter? *Circulation* 2011; 123: 2189–2191.
108. Rosendaal FR. Once and only once. *Circulation* 2010; 121: 1688–1690.
109. de Haan HG, et al. Multiple SNP testing improves risk prediction of first venous thrombosis. *Blood* 2012; 120: 656–663.
110. van Hylckama Vlieg A, et al. Genetic variations associated with recurrent venous thrombosis. *Circ Cardiovasc Genet* 2014; 7: 806–813.



## II. Considérations générales sur la protéine S

La PS a été décrite pour la première fois en 1979 à Seattle (qui lui donne son nom) par DiScipio *et al* comme une nouvelle glycoprotéine comprenant des résidus gamma-carboxyglutamiques (GLA)<sup>47</sup>. Contrairement aux autres protéines de la coagulation vitamine k-dépendantes, la PS ne présente pas d'activité sérine protéase.

La PS est majoritairement synthétisée par les hépatocytes, mais aussi par les mégacaryocytes, les cellules endothéliales, les cellules de Leydig et les cellules musculaires lisses vasculaires. Elle circule dans le plasma à une concentration de 20-25 mg/l avec une demi-vie de 42 heures<sup>48</sup>. Dans le plasma, la PS forme un complexe équimoléculaire très stable avec la C4b-Binding Protein (C4bBP)<sup>49</sup>.

Le précurseur de la PS contient 676 acides aminés répartis en 8 feuillettes. Quinze exons et 14 introns constituent le gène de la PS.

L'exon 1 code pour le codon initiateur de la traduction et le peptide signal (résidus 1 à 24).

L'exon 2 code pour le propeptide (résidus 25 à 41) contenant la séquence reconnue par la gamma-carboxylase. Le peptide signal et le propeptide subissent une protéolyse avant sécrétion pour donner la protéine mature constituée de 635 acides aminés.

L'extrémité 3' de l'exon 2 et l'exon 3 codent pour le domaine GLA (résidus 42 à 86).

L'exon 4 code pour une boucle sensible à l'action de la thrombine (« thumb loop ») (résidus 87 à 113) dont le clivage par la thrombine abolit l'activité cofacteur de la PC activée (PCa). Ce domaine stabilise la structure du domaine GLA et augmente son affinité pour l'ion calcium, mais il ne semble pas impliqué dans le contact direct avec les phospholipides.

Les exons 5 à 8 codent respectivement pour les domaines EGF-like 1 à 4. Les 3 domaines EGF-like C-terminaux jouent un rôle important en liant les ions calcium avec une haute affinité<sup>50</sup>.

Les exons 9 à 14 et l'extrémité 5' de l'exon 15 codent pour des structures (résidus 284 à 676) uniques dans les protéines de la coagulation, des domaines LG (« globular domains of laminin A) retrouvés dans diverses protéines de la matrice extracellulaire et des protéines plasmatiques telles que Gas6 et la sex hormone-binding globulin (SHBG)<sup>51</sup>. C'est en raison de leur grande similarité avec cette dernière protéine que ces domaines sont également appelés région SHBG-like.

## 1. Fonction de la protéine S

La PS est un anticoagulant naturel. Cette activité est principalement liée à son activité cofacteur de la PCa. Cette dernière a pour fonction essentielle d'inactiver les facteurs V et VIII activés par protéolyse. La PC circule dans le plasma sous la forme d'une sérine-protéase inactive. Elle est activée par le complexe thrombine-thrombomoduline présent à la surface des cellules endothéliales<sup>52</sup>. La PCa libre a un effet anticoagulant très faible, mais son activité est renforcée par la présence de la PS. En effet, il a été démontré que la PS déplace le site actif de la PCa. Si ce déplacement est effectué par mutagenèse sur la PCa, cette dernière ne nécessite plus l'activité cofacteur de la PS<sup>53</sup>. Ainsi, la PS accélère principalement le clivage du facteur V activé (résidu arginine en position 306) par la PCa<sup>54,55</sup>. La PS favorise par ailleurs le clivage du facteur V activé par la PCa en déplaçant le facteur X activé du complexe qu'il forme avec le facteur V activé. Cette opération rend l'arginine 506 plus accessible à la PCa.

De nombreuses études ont permis de démontrer que cette activité cofacteur est médiée par la « thumb loop » et le premier module « EGF-like »<sup>56</sup>. Le rôle capital de la « thumb loop » est renforcé par le fait que la PS perd son activité cofacteur après clivage par la thrombine ou le facteur X activé<sup>57</sup>. Par ailleurs, l'utilisation d'anticorps monoclonaux dirigés contre la « thumb loop » et EGF1 permet d'obtenir l'inhibition la plus puissante de cette activité cofacteur<sup>58</sup>. De façon complémentaire, les domaines LG sont nécessaires pour obtenir une activité cofacteur de la PCa complète. En effet, une PS tronquée, amputée de ses domaines LG, permet d'obtenir une activité cofacteur de la PCa normale dans des systèmes purifiés, mais pas dans le plasma (figure 3)<sup>59</sup>.

La PS présente par ailleurs une activité anticoagulante indépendante de la PCa, que l'on caractérisera de directe. La PS fixe les phospholipides membranaires avec une haute affinité, ainsi que les facteurs V, X et VIII activés. Ainsi, la PS présenterait une activité anticoagulante directe par deux mécanismes : d'une part, elle occuperait la surface phospholipidique entrant en compétition avec les complexes procoagulants ; d'autre part, elle inhiberait directement les facteurs X et V activés présents dans le complexe prothrombinase en se fixant à eux<sup>60-62</sup>. De plus, La PS exerce un rôle de cofacteur du « tissue-factor pathway inhibitor » (TFPI) dans l'inhibition du facteur X activé<sup>63</sup>. Bien que faisant l'objet de débats,

les preuves d'une activité anticoagulante de la PS non médiée par la PCa sont multiples. Tout d'abord, dans un modèle thrombotique de babouin la PS présentait une activité anticoagulante après utilisation d'anticorps permettant le blocage de l'activation de la PC<sup>64</sup>. De plus, de nombreux patients hétérozygotes présentent des taux effondrés de PSL. Pourtant leur phénotype clinique n'est pas aussi sévère que les patients homozygotes non traités, suggérant que leurs complexes PS-C4bBP sont protecteurs, en partie en raison de leur activité anticoagulante directe. Enfin, une mutation responsable d'un déficit de type II située sur le domaine C-terminal de la PS a été rapportée. Il s'agit d'un domaine qui ne joue pas un rôle majeur dans l'activité cofacteur de la PCa mais il présente une activité anticoagulante directe via sa liaison au facteur V activé<sup>65</sup>.

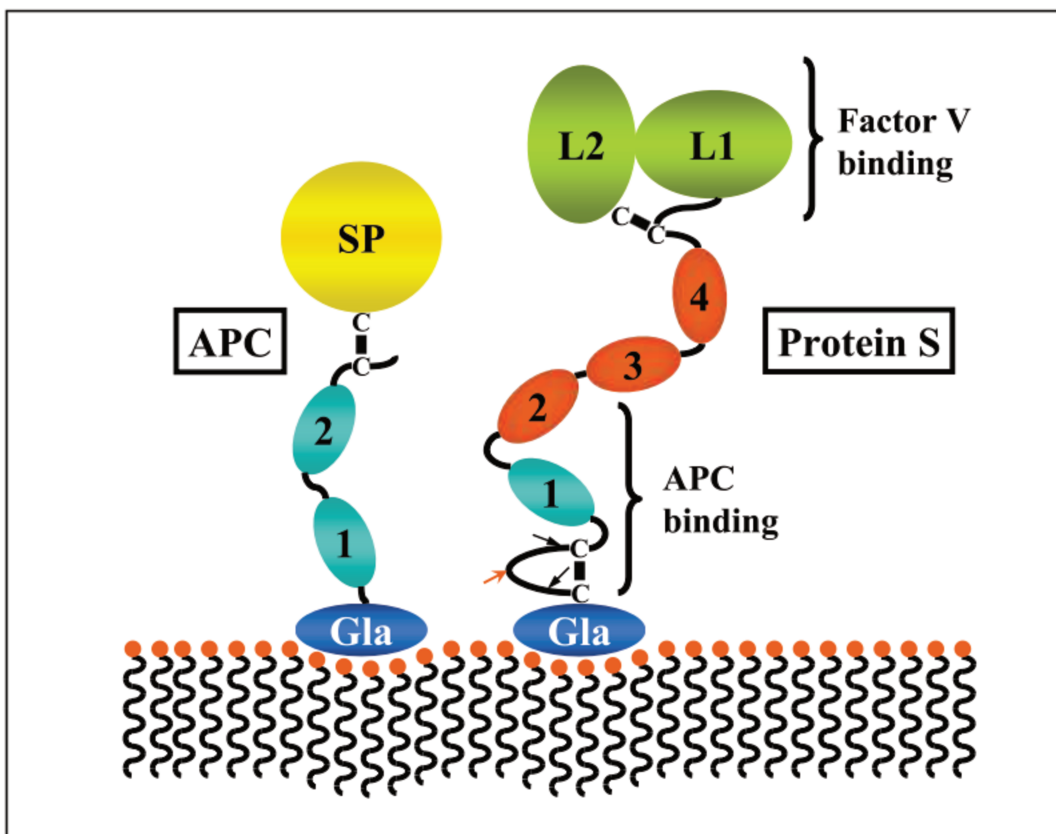


Figure 3. Représentation schématique de l'organisation moléculaire des protéines S et C activées. D'après de Frutos et al<sup>66</sup>. De 1 à 4 : domaines EGF. L1 et L2 : modules LG.

## *2. Protéine S et risque thrombotique, diagnostic biologique et problématique*

Parmi les 5 anomalies recherchées dans le bilan de thrombophilie (constitutionnelle), le déficit en PS est un facteur de risque majeur. L'importance de la PS dans l'anticoagulation est attestée par la sévérité des tableaux cliniques associés aux déficits homozygotes<sup>67</sup>. Un déficit congénital en PS est retrouvé chez 2 à 12% des patients ayant présenté un épisode de MTEV. Cette prévalence dépend notamment de l'origine géographique<sup>68-70</sup>.

Les mutations hétérozygotes situées sur *PROS1* responsables d'un déficit en PS entraînent une augmentation importante du risque de MTEV. En effet, la prévalence de la MTEV est près de 10 fois plus élevée chez les patients porteurs d'une mutation comparativement aux apparentés ne présentant pas la mutation<sup>70-73</sup>. Dans les études conduites en population générale, en revanche, la présence d'un déficit constitutionnel en PS ne semble conduire qu'à une augmentation modérée du risque thrombotique. En effet, le RR pour la présence d'un déficit en PS est estimé à 2 environ<sup>70,73</sup>. Ces résultats paradoxaux soulèvent deux problèmes spécifiquement liés au diagnostic du déficit en PS (contrairement aux déficits en AT et PC) : les techniques de dosage ne permettent pas actuellement d'identifier spécifiquement tous les patients à risque de MTEV ; et il existe une mauvaise corrélation entre les taux bas de PS, définis à partir du calcul des valeurs normales, et l'existence d'une mutation de *PROS1*<sup>74-76</sup>. L'évaluation du risque conféré par un marqueur biologique dépend d'une part de la capacité à mesurer de façon fiable sa prévalence dans la population générale, et d'autre part de la détermination des valeurs normales. Ces concepts expliquent les difficultés rencontrées quant à l'évaluation du risque de MTEV associé au déficit en PS. En effet, les taux plasmatiques de PS ne sont pas uniquement génétiquement déterminés mais également fortement influencés par l'environnement. Ainsi, le sexe, l'imprégnation hormonale (grossesse, cycle menstruel, contraception), le tabagisme, l'âge et certains états pathologiques constituent autant de facteurs environnementaux influençant les taux de PS<sup>73,77-80</sup>. Il faut toutefois noter que d'après les données issues de la détermination des taux de PS chez des volontaires sains, la prévalence du déficit en PS est estimée à moins de 0,5% et à environ 1,5% dans les populations européenne et japonaise respectivement<sup>80,81</sup>. Outre la problématique de la variabilité intra-individuelle des taux de PS (en rapport avec les



différentes situations environnementales énumérées ci-dessus), la détermination du taux de PS se heurte à une problématique technique. L'activité anticoagulante de la PS étant principalement une activité cofacteur, c'est cette dernière qui est mesurée par la réalisation de tests fonctionnels. On parle alors d'activité de la PS. Cependant, en raison de la faible spécificité des premières trousse commerciales, pendant de nombreuses années, cette activité n'a pas été réalisée en première intention dans un grand nombre de laboratoires qui préféraient alors utiliser des techniques antigéniques<sup>82</sup>. En effet, la PS circule dans le plasma sous deux formes : libre et liée à la C4bBP. L'activité cofacteur de la PS est liée à sa fraction plasmatique libre. De nombreux kits commerciaux permettent son dosage antigénique. Cette technique était alors majoritairement utilisée en première intention pour dépister des déficits en PS. Le dosage complémentaire de la PS totale (PST) permet ensuite de typer le déficit. Cette dernière technique mesure la PS « dans tous ses états » : libre et liée à la C4bBP. Deux types de déficits quantitatifs sont rapportés : les types I et III dont la signification clinique et biologique a longtemps été débattue<sup>83,84</sup>. En effet, alors que les profils plasmatiques diffèrent (PSL et PST diminuées dans le type I ; PSL diminuée et PST normale dans le type III), ce n'est pas le cas des profils moléculaires. De nombreuses études ont rapporté les deux profils biologiques au sein d'une même famille, et pour une même mutation<sup>85-88</sup>. Ces déficits quantitatifs représentent la grande majorité des déficits diagnostiqués. Un troisième déficit, beaucoup plus rare, complète le tableau. Il s'agit du déficit de type II, qualitatif. Ce dernier déficit ne s'accompagne pas d'une diminution des taux antigéniques de PS. Il n'est donc pas dépisté par les dosages des PSL et PST. Seuls les dosages fonctionnels permettent d'observer une diminution de l'activité de la PS. Au demeurant, cette étape ne permet de confirmer un déficit en PS, « preuve moléculaire » à l'appui, que dans environ 50% des familles avec déficit plasmatique en PS<sup>66</sup>. Une récente étude a entrepris de tester les seuils permettant de dépister les déficits constitutionnels en PS dans une étude cas-témoins totalisant 5317 sujets<sup>89</sup>. Les valeurs normales ont tout d'abord été calculées sur la base des 2,5<sup>ème</sup> et 97,5<sup>ème</sup> percentiles. Ces dernières étaient ainsi comprises entre 53 et 139 UI/dL pour la PSL. Les patients présentant des taux de PSL situés entre ces bornes constituaient le groupe de référence. Après exclusion des femmes en situation hormonale (œstrogènes ou grossesse/post-partum), il n'existait pas de différence significative : OR = 1,51 (IC95% = 0,82-2,78) pour les taux de PSL inférieurs à 53 UI/dL comparativement aux patients présentant des taux normaux. Il faut toutefois noter que

l'analyse restreinte aux épisodes ambulatoires permettait d'identifier une augmentation du risque pour les taux bas de PSL : OR = 2,31 (IC95% = 1,06-5,05). L'analyse de sensibilité pour les taux bas de PSL, inférieurs au 2,5<sup>ème</sup> percentile, ne permettait pas d'identifier de groupe à risque. En effet, seuls des taux inférieurs au 0,1<sup>ème</sup> percentile, soit inférieurs à 33 UI/dL semblaient associés à une augmentation du risque. Cependant, les faibles effectifs se traduisaient par des intervalles de confiance très importants ne permettant pas de conclure définitivement à une association. Les auteurs ont sélectionné dans un second temps 48 sujets (24 cas et 24 témoins) présentant des taux de PS inférieurs à 46 UI/dL (1<sup>er</sup> percentile) pour la réalisation du séquençage de *PROS1*. Cette sélection n'a permis d'identifier une mutation que chez 20% des patients (5 cas et 5 témoins). Ainsi, les auteurs concluent que le déficit en PS est rare en population générale et que le dépistage (et l'analyse génétique de *PROS1*) d'un déficit en PS ne devrait pas être réalisé chez des sujets ayant présenté un épisode de MTEV en l'absence de sélection préalable. Deux études confirment par ailleurs les résultats préalablement discutés. L'utilisation de taux bien inférieurs aux valeurs normales permet d'identifier une association entre le déficit en PS et le risque thrombotique. Ainsi, pour la première étude, chez des apparentés de familles déficitaires en PS, il existait une augmentation du risque de MTEV pour les sujets présentant des taux inférieurs au 5<sup>ème</sup> percentile, soit inférieurs à 41 UI/mL : hazard ratio (HR) = 5,6 (IC95% = 2,7-11,5) comparativement au quartile supérieur (> 91 UI/mL). Le risque était augmenté quand le seuil utilisé était le 2,5<sup>ème</sup> percentile (<33 UI/mL) : HR = 11,3 (IC95% = 5,4-23,6)<sup>90</sup>. Dans la seconde étude, l'utilisation d'un seuil à 30 UI/mL (au lieu de 65 UI/mL définissant la valeur inférieure de la normale) permettait d'améliorer la spécificité pour l'identification des sujets à risque de MTEV. La spécificité passait ainsi de 50% pour un seuil à 65 UI/mL à 62% pour un seuil à 30 UI/mL, cette augmentation des performances s'accompagnant d'une perte minimale de sensibilité (85% au lieu de 88%)<sup>83</sup>.

### 3. Mutations situées sur *PROS1* associées au déficit en protéine S

La PS est codée par un gène situé sur le chromosome 3, près du centromère en position 3q11.2<sup>91</sup>. Il s'agit du gène *PROS1* (GeneID : 5627 ; MIM # 176880). Ce dernier présente un pseudogène hautement homologue, *PROSP*, ou *PROS2* (GeneID : 5628), situé de l'autre côté

du centromère, en position 3p11.1<sup>92</sup>. *PROS1* s'étend sur 101 Kb et comporte 15 exons et 14 introns. *PROSP* s'étend sur 34 Kb et partage 96,5% et 95,4% d'homologie de séquence avec les exons et les introns de *PROS1* respectivement. L'exon 1 est notamment manquant. Ce pseudogène n'est pas actif en raison de l'absence de cadre de lecture ouvert.

Plus de 200 mutations ont été décrites à ce jour et sont recensées dans les bases de données de l'International Society on Thrombosis and Haemostasis (<http://www.hgmd.cf.ac.uk>)<sup>93</sup> et HGMD (Human Gene Mutation Database, <http://www.hgmd.cf.ac.uk/ac/>). La majorité d'entre elles sont des mutations ponctuelles ou de courtes insertions ou délétions identifiées chez des patients déficitaires pour la PS. La plupart de ces mutations sont associées à des phénotypes quantitatifs (déficits de type I ou III). Elles sont distribuées tout le long du gène. Il n'existe pas de « hotspot ». Les mutations associées aux déficits de type II sont des mutations faux-sens qui affectent principalement les domaines GLA à EGF4, dont le rôle est primordial dans l'activité cofacteur de la PCa. Parmi les déficits de type II, une mutation est fréquente dans la population japonaise : la mutation Lys196Glu. Cette mutation, qui entraîne une diminution isolée du taux de PS activité (PSL et totale normales), n'existe pas en dehors de l'archipel nippon. Elle est associée à une augmentation du risque de MTEV (OR voisin de 5)<sup>94,95</sup>.

L'analyse moléculaire du gène *PROS1* chez des patients déficitaires issus de familles thrombophiles permet de retrouver une mutation dans 50 à 90% des cas<sup>70,78,96</sup>. Cette grande variabilité dépend probablement des critères de sélection. Il faut toutefois noter que les techniques de séquençage utilisées dans ces études ne permettent pas la détection de larges délétions/insertions. Une étude a permis d'estimer à 38% le nombre de larges délétions dans des familles préalablement séquencées pour *PROS1* sans succès. Sur 17 familles déficitaires en PS initialement génétiquement explorées, seules 9 présentaient une mutation avec les techniques classiques de séquençage. L'utilisation d'une carte dense de 9 marqueurs microsatellites et 14 polymorphismes d'une seule base a permis d'identifier une grande délétion dans 3 familles analysées sur 8, soit 38%. Il existait en effet une coségrégation entre les haplotypes observés et le déficit en PS<sup>97</sup>. Une autre étude a entrepris d'évaluer le bénéfice que pourrait apporter l'utilisation d'une technique permettant la détection de grandes délétions/insertions : la « multiplex ligation-dependent probe amplification » (MLPA). Dix-huit sujets issus de familles thrombophiles avec déficit en

PS ont ainsi été analysés. Au total, 6 anomalies ont été identifiées : 3 délétions complètes du gène, 2 délétions partielles et 1 duplication d'une partie de *PROS1*<sup>98</sup>.

Les mutations les plus fréquemment identifiées par l'étude du gène *PROS1* sont de type faux-sens (45%), non-sens (19%) et d'épissage (11%)<sup>66</sup>. Peu d'études ont permis d'estimer le taux de larges délétions/insertions, mais les résultats discutés plus haut semblent indiquer qu'elles occupent une part non négligeable : l'existence de grandes délétions/insertions pourrait expliquer environ un tiers des déficits constitutionnels en PS non expliqués par le séquençage de *PROS1*.

L'identification d'une mutation doit être accompagnée d'une caractérisation fonctionnelle et/ou d'une étude *in silico* pour déterminer si elle est potentiellement causale. Parmi les mutations faux-sens décrites, la mutation S501P est relativement fréquente. Elle est plus communément appelée protéine S Heerlen. Sa première description est attribuée à Bertina *et al* en 1990. Le variant Heerlen est associé à une diminution de la masse moléculaire de la PS : 71 versus 73 Kd. L'hypothèse alors admise était la disparition d'un site de glycosylation. La séquence d'acides aminés de la PS humaine présente 3 sites putatifs de N-glycosylation sur les résidus 457, 468 et 489 de la protéine mature<sup>87</sup>. L'analyse moléculaire a ainsi permis aux auteurs d'identifier la mutation ponctuelle responsable de la perte d'un site de N-glycosylation. Il s'agit de la transition d'un T en C sur le codon Ser<sup>460</sup> (Ser<sup>510</sup> selon la nomenclature HGVS). Ainsi le codon TCC codant pour un résidu Sérine devient CCC et code alors pour un résidu Proline. La Sérine faisant partie de la séquence consensus (Asn<sup>458</sup>-Val<sup>459</sup>-Ser<sup>460</sup>) de la potentielle N-glycosylation de l'Asn<sup>458</sup>, sa disparition empêche la N-glycosylation de l'Asn<sup>458</sup><sup>87</sup>. L'utilisation d'un test fonctionnel (allongement du temps de céphaline activée (TCA) en présence de PCa) a par ailleurs permis aux auteurs de conclure à l'absence de retentissement fonctionnel de la mutation. D'après les auteurs, le polymorphisme devait donc être considéré comme neutre<sup>87</sup>. L'étude épidémiologique a permis de préciser que le variant apparaissait relativement fréquent dans la population étudiée, et il n'existait pas de différence entre les sujets présentant un antécédent de MTEV non provoquée et les donneurs de sang sains qui présentaient respectivement une prévalence pour l'allèle à risque de 0,67% et 0,52%. Ces résultats semblaient donc confirmer l'absence de retentissement fonctionnel de la mutation. L'étude de l'association potentielle entre ce variant fréquent et la MTEV a pendant longtemps reposé sur cette seule étude et la comparaison des prévalences de la PS Heerlen entre 96 donneurs sains et 1182 patients avec

antécédents thrombotiques. En revanche, de nombreuses publications ont évalué le retentissement fonctionnel de cette mutation, avec des résultats contradictoires. Certaines études ont ainsi apporté des preuves indirectes de la contribution de la mutation au phénotype thrombotique, tandis que pour d'autres la mutation semblait neutre.

Dans le premier groupe de publications, une étude a tout d'abord mis en évidence une association entre la mutation et des déficits en PS de type III <sup>99</sup>. Pour cela, le variant a été génotypé chez 85 propositi présentant un déficit en PS et 118 sujets sains. Sa prévalence était de 18,8% et 0,8% respectivement pour les propositi et les sujets sains. Parmi les propositi, 15 patients sur 16 présentaient une diminution du taux de PSL, et 15 présentaient des taux normaux ou limites de PST. Au total, la majorité des propositi mutés pour la PS Heerlen présentaient un déficit de type III <sup>99</sup>. Alors que physiologiquement la C4bBP lie une molécule de PS, l'analyse de la courbe de liaison entre la PS et la C4bBP était biphasique, suggérant ainsi que la C4bBP peut lier 2 molécules de PS Heerlen. Ceci permettrait d'expliquer les taux bas de PSL chez les patients mutés (la PS Heerlen est liée à 80% à la C4bBP). Une seconde publication a révélé que la mutation Heerlen était associée à des déficits en PS de type III à l'état hétérozygote et à un déficit de type I à l'état homozygote <sup>100</sup>. Les auteurs ont génotypé le variant Heerlen dans une famille de déficit en PS. Le propositus était une femme de 53 ans ayant présenté de multiples épisodes thrombotiques veineux depuis l'âge de 38 ans. Le bilan de thrombophilie était positif chez cette patiente et retrouvait la présence d'un FVL et une PTG20210A à l'état hétérozygote, ainsi qu'un déficit en PS de type III (PSL = 40% et PST = 79%). L'étude familiale a permis d'identifier la mutation Heerlen à l'état hétérozygote chez le propositus ainsi que chez tous les apparentés présentant des taux bas ou limites de PSL. Un apparenté présentait un phénotype différent : la nièce du propositus avait un déficit de type I (PSL = 34% ; PST = 50%). De façon intéressante, l'analyse génétique a révélé une mutation Heerlen à l'état homozygote. Les sujets hétérozygotes pour la mutation Heerlen présentaient des taux significativement diminués en PSL comparativement aux sujets non mutés (56% vs 87%,  $p = 0,005$ ). Bien que les taux de PST restaient normaux chez tous les sujets, les taux étaient significativement diminués chez les patients hétérozygotes comparativement aux apparentés non mutés (86% vs 102%,  $p = 0,048$ ) <sup>100</sup>. Une autre étude a rapporté un cas de déficit en PS chez un patient homozygote pour la mutation Heerlen ayant présenté un épisode de MTEV à l'âge de 42 ans.

Le bilan de thrombophilie avait alors révélé un déficit profond en PSL, dosée à deux reprises à des taux inférieurs à 35%<sup>101</sup>.

Un second groupe de publications conforte les résultats de Bertina et sa première description de l'anomalie concernant l'apparente neutralité du variant Heerlen. Dans une première étude, notamment, le variant Heerlen était plus fréquent chez les sujets déficitaires en PS comparativement aux volontaires sains (fréquence allélique = 0,3 vs 12,2% respectivement pour les volontaires sains et les patients déficitaires respectivement,  $p < 0,001$ ). L'analyse en sous-groupes a permis de démontrer que cette différence était expliquée par les déficits de type III (fréquence allélique = 19,6%). Il n'existait pas de différence significative entre les volontaires sains et les patients présentant un déficit de type I ou un déficit de type I/III (déficits mal classés dans les familles qui exprimaient les 2 phénotypes). Cependant, l'analyse de ségrégation n'a pas permis de démontrer la coségrégation de la mutation Heerlen avec le déficit en PS de type III dans 3 familles<sup>85</sup>. Une étude a par ailleurs entrepris la caractérisation fonctionnelle de cette mutation en étudiant l'activité cofacteur de la PCa d'une protéine recombinante. Il n'existait pas de différence en termes d'activité cofacteur de la PCa et de liaison aux phospholipides<sup>102</sup>. La constante de dissociation  $K_d$  du complexe C4bBP $\beta$ /PS n'apparaissait pas non plus modifiée par la mutation soulignant la normalité de l'interaction entre la PS Heerlen recombinante et la chaîne  $\beta$  de la C4bBP<sup>102</sup>. Une autre étude a évalué l'activité anticoagulante directe de la PS Heerlen<sup>103</sup>. Les auteurs ont tout d'abord confirmé les résultats précédemment publiés en termes d'équivalence d'interaction entre la PS Heerlen ou la PS sauvage et la C4bBP (constantes de dissociation = 0,33 et 0,27nM respectivement). L'étude de l'activité inhibitrice du complexe prothrombinase de la PS Heerlen, libre ou complexée au C4bBP, en l'absence de PCa, était par ailleurs identique à celle de la PS sauvage<sup>103</sup>. Les propriétés anticoagulantes (inactivation du facteur V activé, inactivation du facteur V Leiden activé et inactivation du facteur VIII activé catalysées par la PCa) de trois préparations de PS ont enfin été testées, sans différence : PS sauvage, PS Heerlen recombinante et PS dérivée du plasma

<sup>103</sup>.

### III. Problématique

Le déficit en PS est associé de manière robuste au risque de MTEV. Au demeurant, le risque semble varier en fonction des groupes étudiés. En effet, alors que le risque de MTEV associé au déficit en PS identifié dans des familles est estimé à 10 environ, comparativement à des apparentés sains, le déficit en PS ne semble conférer qu'une augmentation modérée du risque en population générale (risque multiplié par 2 à 3). Ce risque différentiel trouve probablement son explication dans la définition même du déficit en PS. En effet, la diversité des techniques utilisées pour le dosage de la PS, la variabilité intra-individuelle et l'absence de seuil clairement établi compliquent le diagnostic. De même, la corrélation entre phénotype et génotype est incertaine. Ainsi, pour une définition du déficit en PS classiquement établie au 2,5<sup>ème</sup> percentile, environ la moitié des patients ne présentent pas de mutation de *PROS1*. La diminution de ce seuil permettrait d'augmenter la spécificité du dosage pour le diagnostic d'une anomalie moléculaire. Il apparaît donc nécessaire de redéfinir la notion de seuil dans le déficit en PS afin d'améliorer l'identification des déficits constitutionnels. Cette problématique fait l'objet de la première publication présentée dans ce manuscrit. Cette dernière est parue dans le journal international à comité de lecture *Thrombosis and Haemostasis*. Elle s'intitule « *PROS1* genotype phenotype relationships in a large cohort of adults with suspicion of inherited quantitative protein S deficiency ». Le génotypage de *PROS1* chez 579 individus présentant un déficit en PS de type I ou III a permis d'identifier une anomalie moléculaire dans 334 cas. Parmi les patients mutés, 79 présentaient un variant relativement fréquent permettant d'expliquer un nombre important de déficits de type III : le variant Heerlen. Les conséquences fonctionnelles d'un tel polymorphisme sont mal évaluées, de même que son association éventuelle avec la MTEV. Seule une étude mobilisant de faibles effectifs a tenté d'apporter une réponse à cette dernière question. Nous avons ainsi, au cours d'une seconde publication, testé l'association entre la mutation Heerlen et la MTEV dans 4 études cas-témoins indépendantes françaises totalisant 4173 cas et 5970 témoins. Les résultats de cette étude ont été publiés dans la revue internationale à comité de lecture *Scientific reports* sous le titre « Protein S Heerlen mutation heterozygosity is associated with venous thrombosis risk ».

# Résultats



## **Article 1**

**“*PROS1* genotype phenotype relationships in a large cohort of adults with suspicion of inherited quantitative protein S deficiency”**

*Thrombosis and Haemostasis 2016*

# *PROS1* genotype phenotype relationships in a large cohort of adults with suspicion of inherited quantitative protein S deficiency

Martine Alhenc-Gelas<sup>1</sup>; Genevieve Plu-Bureau<sup>2</sup>; Marie H el ene Horellou<sup>3</sup>; Antoine Rauch<sup>4</sup>; Pierre Suchon<sup>5</sup>; for the GEHT genetic thrombophilia group

<sup>1</sup>H ematologie biologique, AP-HP, H opital Europ een G Pompidou, H opitaux Universitaires Paris Ouest, Paris, France; <sup>2</sup>Unit  d'h emostase et Unit  de Gyn ecologie endocrinienne, AP-HP, H opital Cochin, H opitaux universitaires Paris Centre, Facult  de M decine Paris Descartes, Paris, France; <sup>3</sup>Laboratoire d'H ematologie, AP-HP H opital Cochin, H opitaux universitaires Paris Centre Facult  de M decine Paris Descartes, Paris, France; <sup>4</sup>P le Biologie Pathologie G n tique, Institut d'H ematologie-Transfusion, CHRU, Lille, France; <sup>5</sup>Laboratoire d'h ematologie, CHU La Timone, INSERM UMR S1062, Universit  Aix Marseille, Marseille, France

## Summary

Inherited protein S deficiency (PSD) is an established risk factor for venous thromboembolism (VTE). However, data are conflicting concerning risk of VTE associated with decreased free PS level (FPS) and information on *PROS1* genotype-phenotype relationship is sparse. In a retrospective cohort of 579 patients with inherited type I/III deficiency suspicion, *PROS1* genotyping was performed and the effect of genotype on FPS and on VTE risk was investigated. We found 116 (including 65 novel) detrimental mutations (DM) in 222 (type I/III in 194, type II in 28), PS Heerlen in 74, possibly non DM in 38 and no mutation in 245 subjects. Among DMs, type I/IIIDMs only were found in subjects with FPS < 30%. Prevalence of type I/III DM decreased with increasing FPS level. Risk of VT associated with FPS level and genotype was studied in the 467 subjects with personal or family history of thrombo-

sis. Only type I/IIIDM carriers presented with an increased risk of VTE [1.41 (95%CI 1.05–1.89)] compared to subjects with no mutation. Among the group of type I/IIIDM heterozygotes and subjects with no mutation, the optimal FPS cut-off point for identifying subjects at increased VTE risk was searched for. We found that only subjects with FPS < 30% and type I/IIIDM presented with an increased risk [1.48 (95%CI 1.08–2.04)]. Our findings confirm the value of a cut-off FPS level for identifying subjects at increased VTE risk far below the lower limit of the normal range and suggest a place for *PROS1* genotyping in PSD diagnosis strategy.

## Keywords

Coagulation inhibitor, protein S, *PROS1*, thrombophilia, venous thrombosis

## Correspondence to:

M. Alhenc-Gelas  
Laboratoire d'H ematologie, HEGP  
20 Rue Leblanc, 75015 Paris, France  
Tel.: +33 156093901, Fax: +33 156093393  
E-mail: martine.alhenc-gelas@aphp.fr

## Financial support:

This study was in part supported by a DHOS program „Soutien financier en faveur des laboratoires pratiquant le diagnostic par g n tique mol culaire des maladies rares”.

Received: May 11, 2015

Accepted after major revision: October 8, 2015

Epub ahead of print: October 15, 2015

<http://dx.doi.org/10.1160/TH15-05-0391>

Thromb Haemost 2016; 115: 570–579

## Introduction

Protein S (PS) acts as a cofactor for activated protein C (APC) in proteolytic inactivation of blood coagulation factors Va and VIIIa, thus providing a negative feedback on thrombin generation and making the APC/PS pathway essential for normal haemostasis. The anticoagulant activity of PS might also be direct, i.e. independent of APC, as PS is also a non-enzymatic cofactor for tissue factor pathway inhibitor during factor Xa inhibition. PS also directly limits thrombin generation by binding to and inhibiting FXa and FVa in the prothrombinase complex. Plasma PS can circulate as a free glycoprotein or bound to the complement regulatory factor C4b-binding protein. Free PS (FPS) accounts for 40% of total PS and is the most active form (1).

Hereditary PS deficiency (PSD) (MIM:#176880) is a recognized genetic risk factor for venous thrombosis (VT) (2). It usually occurs as a partial (heterozygous) deficiency and segregates as an

autosomal dominant trait. Based on PS antigen and APC cofactor activity levels, PSD is classified as type I (low total and free antigen, reduced activity), type III (normal total antigen, reduced free antigen and activity) and type II (normal total and free antigen, reduced activity). Quantitative (type I/III) deficiencies account for 95% of cases of PSD and often occur together in families as phenotypic variants of the same genetic defect (3).

More than 200 mutations impairing protein synthesis and/or function have been identified in the PS gene (*PROS1*). They are distributed along the entire gene with no region showing higher prevalence of mutations. More frequent gene defects are missense mutations, but other types of gene defects, such as nonsense mutations, splice-site mutations, deletions and insertions have also been reported (2).

Data are conflicting concerning the risk of venous thromboembolism (VTE) associated with PSD. Although family-based studies show a consistent four- to ten-fold higher life-time risk of VTE in

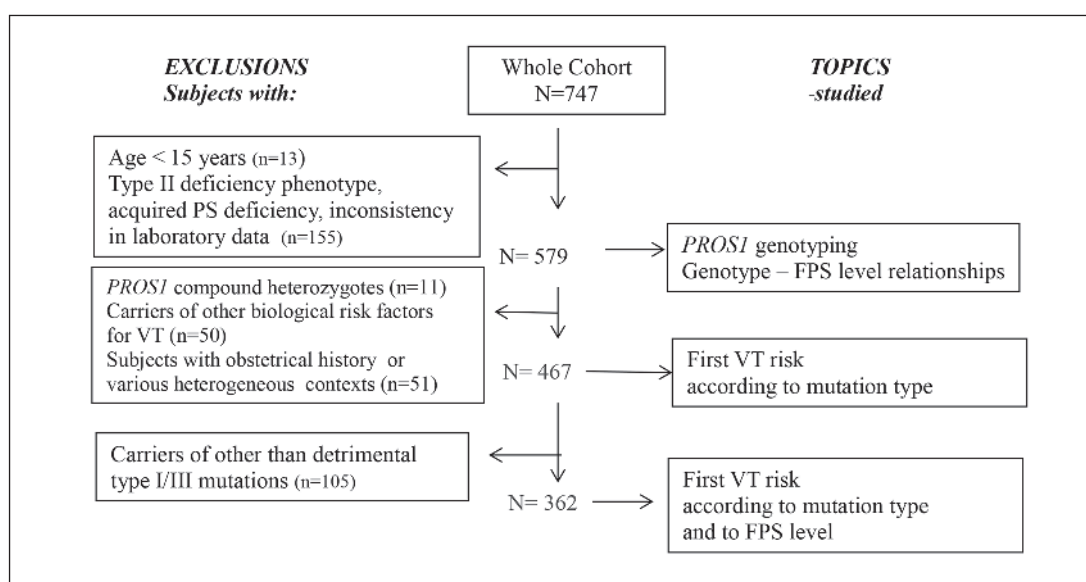


Figure 1: Flow-chart of the study.

affected family members when compared with their non-affected relatives (3–9), most population-based studies point toward a low-risk (no more than 2.5) (10–13). Several studies tried to define a cut-off point in FPS level for identifying subjects who are at risk of VT. A cut-off level of 30–40% far below the lower limit of normal range (50–60%) has been suggested in most of these studies (5, 7, 13). In only one study, low borderline levels (61–75%) were associated with a two-fold increased risk (14)

Finally, only limited information on risk of VTE associated with PSD has been provided by studies in which *PROS1* genotyping was performed (4, 6, 9).

We therefore decided to use data from a large retrospective cohort of patients with suspicion of inherited quantitative PSD having had *PROS1* gene analysis performed in our centre in order to gain more insight into *PROS1* mutations and their influence on FPS plasma level and VTE risk.

## Material and methods

### Study population

Between 2001 and 2011, we received 747 DNA samples for *PROS1* gene analysis. The patients were initially referred to 36 specialised French haematology laboratories for thrombophilia screening. Recording of clinical characteristics of the subjects and family members, including thromboembolic events, and screening for biological risk factors for VT (antiphospholipid antibodies, anti-thrombin, protein C or protein S deficiency, the FV Leiden and the F2 G20210A mutations) were completed locally. Thrombophilia screening was performed at distance from thromboembolic events, anticoagulant or hormonal oestrogenic treatments. Free PS (FPS) was measured in most centres by using Stago Asse-rachrom or Liatest reagent. When an inherited PSD had been suspected locally, DNA was sent to our laboratory for *PROS1* geno-

typing. After excluding subjects younger than 15 years of age, subjects with clear type II deficiency phenotype (ratio PS activity/free PS < 0.50), subjects with known causes of acquired deficiency and subjects with missing laboratory data, 579 subjects remained in the study cohort, belonging to 571 different families. Influence of *PROS1* genotype on FPS level was studied in this population. Influence of genotype on risk of VT was studied in the 467 subjects remaining after excluding compound heterozygotes, carriers of other risk factors for VT, and subjects studied in clinical contexts other than thrombosis. Finally, relationships among type I/III mutations carrying, FPS level and VTE were studied in the sub-group of 362 subjects comprising type I/III mutation carriers and subjects carrying no mutation (► Figure 1).

### *PROS1* analysis

Informed consent for genetic analysis was obtained from all patients. The exons, intron-exon boundaries and 700 bp of the proximal promoter of *PROS1* (Gene ID: 5627, MIM#176880) were studied by direct sequencing as previously described (15, 16). In samples from subjects with FPS level below 40% in which no detrimental mutation was detected by sequencing, we searched for large deletion/insertion by multiplex ligation-dependent probe amplification (MLPA) (SALSA MLPA kit P112, MRC-Holland, Amsterdam, The Netherlands). The test uses probes for 12 among the 15 exons of *PROS1*. Mutations were numbered according to the HUGO recommendations for mutation nomenclature (<http://www.hgvs.org>). Amino-acid numbering according to the ISTH database can be obtained by subtracting 41 for residues of the mature protein or 42 for residues belonging to the signal peptide or pro-peptide. Any *PROS1* sequence variation identified was noted, and the published literature was checked to see if the corresponding variant had been previously recorded and its molecular mechanism elucidated. For each novel sequence variation, the

likelihood that the change could be pathogenic was checked by using different *in silico* bioinformatic tools, an evolution-based method, Sorting Intolerant From Tolerant (SIFT, available at <http://sift.jcvi.org>) (17) available in Alamut v2.3.6 ([alamut.interactive-biosoftware.com](http://alamut.interactive-biosoftware.com)), a structure-based method, Polymorphism Phenotyping (Polyphen2, available at <http://genetics.bwh.harvard.edu/pph>) (18) and the splicing prediction module of Alamut which uses the Human Splicing Finder (HSF) splice site prediction algorithm and a number of other prediction methods (NNSplice, Splice Site finder-like, MaxEntScan, Gene splicer) (19). For missense mutations, SIFT scores less than 0.05 predict a detrimental character, those greater than or equal to 0.05 predict that mutations are probably tolerated. The output of the Polyphen-2 prediction pipeline is a prediction of "probably damaging" (PbD), "possibly damaging" (PoD) or "benign" (B), along with a numerical score ranging from 1 to 0.

### Statistical analysis

Quantitative variables were expressed as mean values and standard deviations (SD), and qualitative variables as count numbers and percentages. Statistical analyses were performed using SAS software, version 9.2 (SAS-Institute Inc, Cary, NC, USA). Variance analysis and Chi<sup>2</sup> test were performed to compare quantitative or qualitative variables among groups. The t-test was used to compare quantitative variables between two groups. A two-tailed p-value below 0.05 was considered to indicate statistical significance. Risk of VTE was studied in the population remaining after excluding subjects with obstetric complications and various, heterogeneous other clinical contexts. In subjects with personal or family history of VT, observation time was defined as the period from the age of 15 years until the end of the observation period, or until the first episode of VT. In subjects of the arterial thrombosis (AT) group, the end of the observation period was the date of the arterial event (in order to avoid biases due to possible use of drugs with influence on thrombosis risk after the arterial event). The cut-off point determination method in survival analysis developed by Contal et al. (20) (SAS Findcut macro) was used to search for the optimal FPS cut-off point which could identify subjects at in-

creased risk in the sub-group consisting of patients with type I/III DM and of subjects with no mutation. Relative risks (RR) for VT and 95 % confidence intervals (CI) associated with FPS level or *PROS1* mutation were estimated by using a multivariate Cox proportional hazard model.

## Results

### Characteristics of the study population

Among the 579 patients included, 359 were screened for thrombophilia because they had a personal history of VTE, 122 because they had a family history of VTE in the absence of personal VTE antecedent. Some other contexts led to the screening in 98 patients (including unexplained AT suspicions, obstetrical complications, or various heterogeneous contexts in 43, 27 and 28 subjects, respectively). In 89 % of subjects with personal VTE, the first venous thrombosis event was a pulmonary embolism and/or a deep-VT while in the remaining 11 % the first episode was a superficial VT. Other biological risk factors for VTE, (mainly FV Leiden or *F2G20210A* mutation, present at the heterozygous state), were present in 10.6 % subjects, with similar frequencies in both sexes. There were 435 females, mean age 38.9 years (SD 13.9), and 144 males, mean age 45.2 years (SD 16.2), all with decreased FPS level (i.e. <50 % in premenopausal or <55 % in menopausal women and <60 % in men) or borderline (-up to 80 %). Mean (SD) FPS level was 43.8 % (SD 15.6) (females 44.8 % (SD 14.9); males 40.6 % (SD 17.2) (p=0.005)). Characteristics of the study population and mutation profile are reported in ► Table 1.

### Mutation profile

A detrimental mutation (DM) was found in 222 among the 579 subjects studied (38.3 %) with 116 different mutations (51 already reported, 65 novel). These mutations included null mutations (large deletions, non-sense point mutations, splice-site substitutions, small insertions or deletions <40bp) and missense mutations, which were found in 11, 45, 22, 29 or 115 subjects respectively. DMs were type I/III mutations in 194 subjects and type II

|   | Personal VTE | Family history of VT | Other contexts | P-value |
|---|--------------|----------------------|----------------|---------|
| n (%)   | 359 (62)     | 122 (21.1)           | 98 (16.9)      |         |
| Females (%)   | 253 (70.5)   | 100 (82)             | 82 (84)        | 0.004   |
| Age (years) (m, SD)                                     | 43.3 (15.1)  | 33.6 (12.6)          | 38.9 (12.9)    | <0.0001 |
| FPS % (m, SD)   | 43.2 (16.7)  | 42.1 (14.9)          | 47.9 (11)      | 0.02    |
| Carriers of other biological risk factors for VT (n, %) | 38 (10.6)    | 7 (5.7)              | 6 (6.1)        | NS      |
| Mutation (n carriers, %)                                |              |                      |                |         |
| -Detrimental  | 154 (42.9)   | 53 (43.4)            | 15 (15.3)      | 0.0001  |
| -Non detrimental  | 26 (7.3)     | 5 (4.1)              | 7 (7.1)        |         |
| -PS Heerlen   | 46 (12.8)    | 13 (10.7)            | 15 (15.3)      |         |

**Table 1: Characteristics of the study population and mutation profile.**

mutations in 28 subjects. PS Heerlen (p.Ser 501 Pro) was present in 74 (12.8%) subjects (21). Thirty-eight (6.6%) subjects carried possibly non-detrimental mutations (substitutions in non-coding regions or missense mutations) according to results of *in silico* prediction analysis. Twenty-two of these mutations were, to the best of our knowledge, novel mutations. No mutation was found in 245 (42.3%) subjects.

DMs were present mainly in patients with a personal or family history of VT. Frequency of PS Heerlen was independent from clinical context of thrombophilia screening.

Novel mutations are reported in ► Table 2.

Sixty-five novel mutations were considered as probably detrimental. These mutations comprised 18 small deletions or insertions (17 among them being frameshifts and the last one leading to loss of one amino-acid), 15 point mutations leading to stop-codons, nine splice-site mutations in canonical positions, mutations for which probability of a detrimental character is very high; in addition, there were 23 missense mutations for which SIFT score and Polyphen 2 score and interpretation are shown in ► Table 2. Two among these novel missense DMs (p.Arg41Leu, p.Arg41Ser), could possibly be type II mutations due to their position at the P1 position of the PS maturation cleavage site. Given the narrow specificity of subtilisin-like protein convertase for positively charged Arg/Lys residues at positions P4, P2 and P1, mutations at position 41 may indeed impair maturation of the full-length cofactor with retained propeptide, being therefore susceptible to lead to abnormal carboxylation of the GLA domain residue (22). Moreover, another mutation located at this position (p.Arg41His) was previously reported to be associated with qualitative PSD (15).

Novel mutations for which detrimental character was doubtful were one small deletion and five substitutions in non-coding regions, 15 missense mutations with SIFT score and/or Polyphen 2 score not in favour of a detrimental character, and c.1707C>T in the coding sequence which does not change the amino-acid. For this last mutation, the splicing predicting module from Alamut did not predict significant change in splice site.

### ***PROS1* mutation type – FPS level relationships**

Mutation profile according to FPS level among the 579 subjects studied is reported in ► Table 3 and ► Figure 2.

Mean FPS level was clearly low only in carriers of type I/IIIDM and borderline in the other groups (global comparison between the five groups:  $p < 0.0001$ ). As could be expected, prevalence of type I/IIIDM decreased when FPS level increased. Interestingly, presence of detrimental type I/IIIDM explained low FPS level in most patients (86%) with FPS < 30%, and in half of the patients with FPS of 30–40% approximately. PS Heerlen and detrimental type II mutations were found in subjects with FPS > 30% only. Possibly non-detrimental mutations were found in subjects with moderately decreased FPS level, as could be expected. There were 11 compound heterozygotes with moderately decreased FPS level (mean 33.8%, SD 10.9). Among them, six carried PS Heerlen on one allele, three others carried type II mutations already reported on one allele (p.Thr78Met, p.Arg40Leu) (23), and another one car-

ried a novel mutation, p.Pro441Leu, for which detrimental character is doubtful. Mean levels by mutation type did not significantly change after exclusion of these subjects.

Null mutations were present in 107 subjects (FPS mean level 24.4% (SD 9.9) and missense mutations in 87 subjects (FPS mean level 35.3% (SD 13.3). Frequency of null mutations decreased when FPS level increased and conversely missense mutations were more frequent in subjects with the highest FPS levels ( $p < 0.0001$ ).

### **Venous thrombosis risk**

Risk of 1<sup>st</sup> VT associated with *PROS1* mutations or FPS level was studied in 467 (314 symptomatic, 153 asymptomatic) among the 579 subjects of the cohort, remaining after excluding compound heterozygotes (n=11), carriers of other biological risk factors for VT (n=50) and subjects with obstetrical complications or heterogeneous clinical contexts (n= 51). Carrying type I/IIIDM was associated with a significantly increased risk of 1<sup>st</sup> VT, with a similar effect of null or missense mutations. For all other mutation types a non-significant increase in risk was observed (► Table 4).

A significant level related effect of FPS level on VTE risk was observed ( $p$  trend= 0.04). Study by genotype showed a significant effect only in type I/III DM carriers ( $p$  trend= 0.004). In addition, we studied the combined influence of genotype and FPS level in the population consisting of patients with type I/III DM and subjects with no mutation (i.e. the 362 remaining patients after excluding all other *PROS1* gene defects carriers) and found that only DM carriers with FPS < 30% presented with a significantly increased risk (► Table 5).

After excluding non-spontaneous VT, leaving 237 subjects, the risk associated with detrimental mutation carrying and FPS < 30% remained significant [2.0 (95%CI 1.20–3.32)].

## **Discussion**

In the present cohort, we identified a large number of gene alterations, including 65 novel, probably detrimental mutations.

The mutation profile was like in previous studies, with mutations of all types distributed throughout the gene (2, 24).

Prevalence of detrimental mutation was high (94% for subjects with FPS level < 25% and still 34% for subjects with PS level 25–55%). PSD of genetic origin cannot be excluded in the few patients with very low FPS level and no identified mutation, since our analysis was limited to copy number variations in 12 among the 15 exons of *PROS1*, and sequencing analysis of exons and their flanking regions. Prevalence of detrimental mutation was lower than previously reported by Dubgen et al. (25) in a cohort of similar design but higher than prevalence reported by Caspers et al in a cohort of patients with personal history of VT and reduced PS activity (75% for FPS level < 25% and 29% for the range 25–55%) (24).

Low detrimental mutation detection in subjects with the highest PS levels is in accordance with previous publications suggesting that 'true' PS deficiencies are much less frequent than the

Table 2A: Novel mis-sense mutations.

| cDNA: c.                 |     | Mutation position: p. | Mutation: p. | SIFT comparison score | Polyphen comparison score |
|--------------------------|-----|-----------------------|--------------|-----------------------|---------------------------|
| probably detrimental     |     |                       |              |                       |                           |
| 121                      | C>A | 41                    | Arg 41 Ser   | 0                     | 1 (PbD)                   |
| 122                      | G>T | 41                    | Arg 41 Leu   | 0                     | 1 (PbD)                   |
| 338                      | G>A | 113                   | Cys 113 Tyr  | 0                     | 0.998 (PbD)               |
| 428                      | G>T | 143                   | Cys 143 Phe  | 0                     | 0.999 (PbD)               |
| 565                      | G>T | 189                   | Gly 189 Cys  | 0.1                   | 1 (PbD)                   |
| 736                      | G>A | 246                   | Glu 246 Lys  | 0.06                  | 0.999 (PbD)               |
| 794                      | G>A | 265                   | Cys 265 Tyr  | 0.03                  | 0.999 (PbD)               |
| 832                      | G>T | 278                   | Asp 278 Tyr  | 0.03                  | 0.999 (PbD)               |
| 988                      | C>T | 330                   | Arg 330 Trp  | 0                     | 1 (PbD)                   |
| 992                      | C>A | 331                   | Tyr 331 Lys  | 0                     | 1 (PbD)                   |
| 995                      | A>G | 332                   | Tyr 332 Cys  | 0                     | 1 (PbD)                   |
| 1079                     | A>G | 360                   | Glu 360 Gly  | 0                     | 1 (PbD)                   |
| 1078                     | G>A | 360                   | Glu 360 Lys  | 0                     | 1 (PbD)                   |
| 1331                     | C>T | 444                   | Pro 444 Leu  | 0                     | 1 (PbD)                   |
| 1343                     | G>A | 448                   | Gly 448 Glu  | 0                     | 1 (PbD)                   |
| 1345                     | T>G | 449                   | Cys 449 Gly  | 0                     | 0.999 (PbD)               |
| 1478                     | T>C | 493                   | Phe 493Ser   | 0                     | 0.998(PbD)                |
| 1490                     | A>G | 497                   | Tyr 497 Cys  | 0.1                   | 1 (PbD)                   |
| 1547                     | C>T | 516                   | Pro 516 Leu  | 0                     | 0.967 (PbD)               |
| 1870                     | G>A | 624                   | Asp 624 Asn  | 0.04                  | 0.968 (PbD)               |
| 1934                     | A>G | 645                   | Asn 645 Ser  | 0                     | 0.997 (PbD)               |
| 1955                     | A>T | 652                   | Asp 652 Val  | 0                     | 1 (PbD)                   |
| 1996                     | T>A | 666                   | Cys 666 Ser  | 0.08                  | 0.998 (PbD)               |
| possibly non detrimental |     |                       |              |                       |                           |
| 71                       | C>T | 24                    | Ala 24 Val   | 0.35                  | 0.427 (B)                 |
| 224                      | A>T | 75                    | Asp 75 Val   | 0.01                  | 0.013 (B)                 |
| 671                      | T>C | 224                   | Phe 224 Ser  | 0.22                  | 0.78 (PoD)                |
| 688                      | G>A | 230                   | Glu 230 Lys  | 0.23                  | 0.028 (B)                 |
| 925                      | G>T | 309                   | Val 309 Phe  | 0.02                  | 0.718 (PoD)               |
| 1126                     | G>T | 376                   | Asp 376 Tyr  | 0.03                  | 0.391 (B)                 |
| 1247                     | C>T | 416                   | Pro 416 Leu  | 0.02                  | 0.015 (B)                 |
| 1286                     | G>T | 429                   | Gly 429 Val  | 0                     | 0.853 (PoD)               |
| 1322                     | C>T | 441                   | Pro 441 Leu  | 0.08                  | 0.904 (PoD)               |
| 1532                     | C>T | 511                   | Thr 511 Ile  | 0.02                  | 0.001 (B)                 |
| 1544                     | G>A | 515                   | Arg 515 His  | 0.23                  | 0.068 (B)                 |
| 1673                     | T>C | 558                   | Val 558 Ala  | 0.28                  | 0.656 (PoD)               |
| 1707                     | C>T | 569                   | Ser 569 Ser  |                       |                           |
| 1847                     | C>T | 616                   | Ala 616 Val  | 0.4                   | 0.002 (B)                 |
| 1931                     | T>C | 644                   | Ile 644 Thr  | 0.01                  | 0.322 (B)                 |
| 1939                     | G>A | 647                   | Val 647 Ile  | 0.2                   | 0.258 (B)                 |



Table 2B: Novel mutations of other types.

| cDNA: c.                        |     | Mutation position:<br>p. | Mutation:<br>p.   |
|---------------------------------|-----|--------------------------|-------------------|
| Probably detrimental            |     |                          |                   |
| <i>small ins or del</i>         |     |                          |                   |
| 25_34                           | del | 9                        | Gly 9 Trp fs*6    |
| -6_-34                          | dup | 13                       | Ala 13 Arg fs*39  |
| 89_95?                          | del | 30                       | Gln 30 Leu fs*55  |
| 195                             | del | 66                       | Glu 66 Lys fs*21  |
| 200_202                         | del | 67                       | Glu 67 del        |
| 348                             | del | 117                      | Ile 117 Phe fs*16 |
| 400                             | del | 134                      | Cys 134 Ala fs*25 |
| 432_433                         | dup | 145                      | Cys 145 Phe fs*15 |
| 545                             | dup | 182                      | Tyr 182*          |
| 545                             | del | 182                      | Tyr 182 Ser fs*26 |
| 901_905                         | del | 301                      | Tyr 301 Gly fs*26 |
| 918                             | dup | 307                      | Ala307 Cys fs*22  |
| 1303_1309                       | del | 435                      | Glu 435 Asn fs*10 |
| 1638                            | dup | 547                      | Ser 547 Ile fs*9  |
| 1728                            | del | 576                      | Arg 576 Asp fs*17 |
| 1761_1762                       | dup | 588                      | Thr 588 Arg fs*6  |
| 1767_1788                       | dup | 597                      | His 597 Thr fs*2  |
| 2015                            | dup | 673                      | Thr 673 Asp fs*26 |
| <i>non sense</i>                |     |                          |                   |
| 169                             | G>T | 57                       | Glu 57*           |
| 194                             | G>T | 67                       | Glu 67*           |
| 494                             | C>G | 165                      | Ser 165*          |
| 940                             | A>T | 314                      | Lys 314*          |
| 955                             | G>T | 319                      | Glu 319*          |
| 964                             | A>T | 322                      | Arg 322*          |
| 1249                            | G>T | 417                      | Glu 417*          |
| 1402                            | C>T | 468                      | Gln 468*          |
| 1425                            | C>A | 475                      | Cys 475*          |
| 1455                            | T>A | 485                      | Tyr 485*          |
| 1528                            | del | 510                      | Val 510*          |
| 1655                            | T>A | 552                      | Leu 552*          |
| 1711                            | C>T | 571                      | Arg571*           |
| 1854                            | C>G | 618                      | Tyr 618*          |
| 1984                            | A>T | 662                      | Arg 662*          |
| <i>Splice site substitution</i> |     |                          |                   |
| 235-1                           | G>A |                          |                   |
| 260-1                           | G>A |                          |                   |
| 727+2                           | T>C |                          |                   |
| 728-2                           | A>G |                          |                   |
| 850-1                           | G>T |                          |                   |
| 965+1                           | G>A |                          |                   |
| 965+2                           | T>G |                          |                   |
| 966-1                           | G>T |                          |                   |
| 1156-1                          | G>A |                          |                   |
| Possibly non detrimental        |     |                          |                   |
| 347-32_347-25                   | del |                          |                   |
| -62                             | T>G |                          |                   |
| -48                             | C>A |                          |                   |
| 76+31                           | C>A |                          |                   |
| 260-15                          | T>A |                          |                   |
| 346+5                           | G>A |                          |                   |

Table 3: FPS level according to mutation type.

| FPS%      | FPS%<br>(m, SD) | n   | Mutation<br>(n, % carriers by FPS level) |                           |                        |                     |            |
|-----------|-----------------|-----|--|---------------------------|------------------------|---------------------|------------|
|           |                 |     | No<br>mutation                           | Detrimental<br>Type I/III | Detrimental<br>Type II | Non<br>detrimental? | PS Heerlen |
|           |                 |     | n= 245                                   | n= 194                    | n= 28                  | n= 38               | n= 74      |
| 0–20 %    | 16.1 (3.4)      | 60  | 3 (5)                                    | 57 (95)                   | 0                      | 0                   | 0          |
| 21–30 %   | 25.3 (3.1)      | 74  | 14 (18.9)                                | 58 (78)                   | 0                      | 2 (2.7)             | 0          |
| 31–40 %   | 35.9 (3)        | 83  | 19 (22.9)                                | 39 (50)                   | 4 (4.8)                | 8 (9.6)             | 13 (15.7)  |
| 41–50 %   | 46.2 (2.9)      | 146 | 70 (47.9)                                | 24 (16.4)                 | 10 (6.8)               | 15 (10.3)           | 27 (18.5)  |
| 51–60 %   | 55.1 (2.6)      | 145 | 88 (60.7)                                | 14 (9.7)                  | 10 (6.9)               | 8 (5.5)             | 25 (17.2)  |
| 61–80 %   | 67.4 (4.6)      | 71  | 51 (71.8)                                | 2 (2.8)                   | 4 (5.6)                | 5 (7)               | 9 (12.7)   |
| Mean (SD) |                 |     | 52 (11.7)                                | 29.3 (12.8)               | 51.1 (9.6)             | 46.0 (11.7)         | 50.5 (9.3) |

proportion of the population with PS level below the lower limit of normal range (13). Differences in recruitment criteria among studies can explain differences in genotyping performance. Sex ratio in the previous cohort studies was not reported (24, 25). Our study population comprised a high percentage of women (75%) and 55% of these women presented with FPS level at the lower limit of normal range (40–60%) versus 36% of men. It is well-known that women have lower PS levels than men, due to hormonal influence (13, 26). This could explain why there were more women than men in this cohort and also why detrimental mutations were clearly less frequent in women with FPS level 40–60% (17%) than in men (34%).

Previous publications demonstrated that mutations have quantitatively distinct effects (2, 27). Null mutations result in severe defects, most of them causing PSD through synthesis reduction leading to clearly low FPS level. Missense mutations can have severe, intermediate or mild consequences. Missense mutations with severe consequences result in very low to undetectable levels of an unstable or misfolded mutated protein which is not secreted in the extracellular medium. Other missense mutations are associated with higher PS plasma level due to milder effect on PS expression or metabolism. This explains the distribution of null and missense mutations according to FPS level observed in the present study. PS Heerlen is an example of a mild mutation previously reported, associated at the heterozygous state with borderline FPS levels due to reduced half-life (28). This explains why we did not find PS Heerlen in subjects with very low FPS levels. In some cases, PS coded by the mutant allele, which contributes to the plasma pool, can have altered functional properties therefore leading to combined quantitative/qualitative deficiencies. This could explain why type II mutations were found in subjects of this cohort selected on the basis of suspicion of inherited PS quantitative deficiency. Moreover, type II mutations are not systematically associated with a clear type II phenotype in heterozygous carriers (23). Concerning mutations at position 41, the first type II mutation Ala41His heterozygous carrier reported in the literature presented

with a clearly decreased PS activity /FPS ratio at 0.6 (15), but we recently found heterozygosity for this mutation in 3 other patients presenting with ratios at 0.71, 0.72, 0.76, respectively. Ratios found in the three heterozygous carriers of Arg 41 Leu or Arg 41 Ser reported here (0.76, 0.79, 0.84) do not therefore allow to exclude that these novel mutations could be type II mutations.

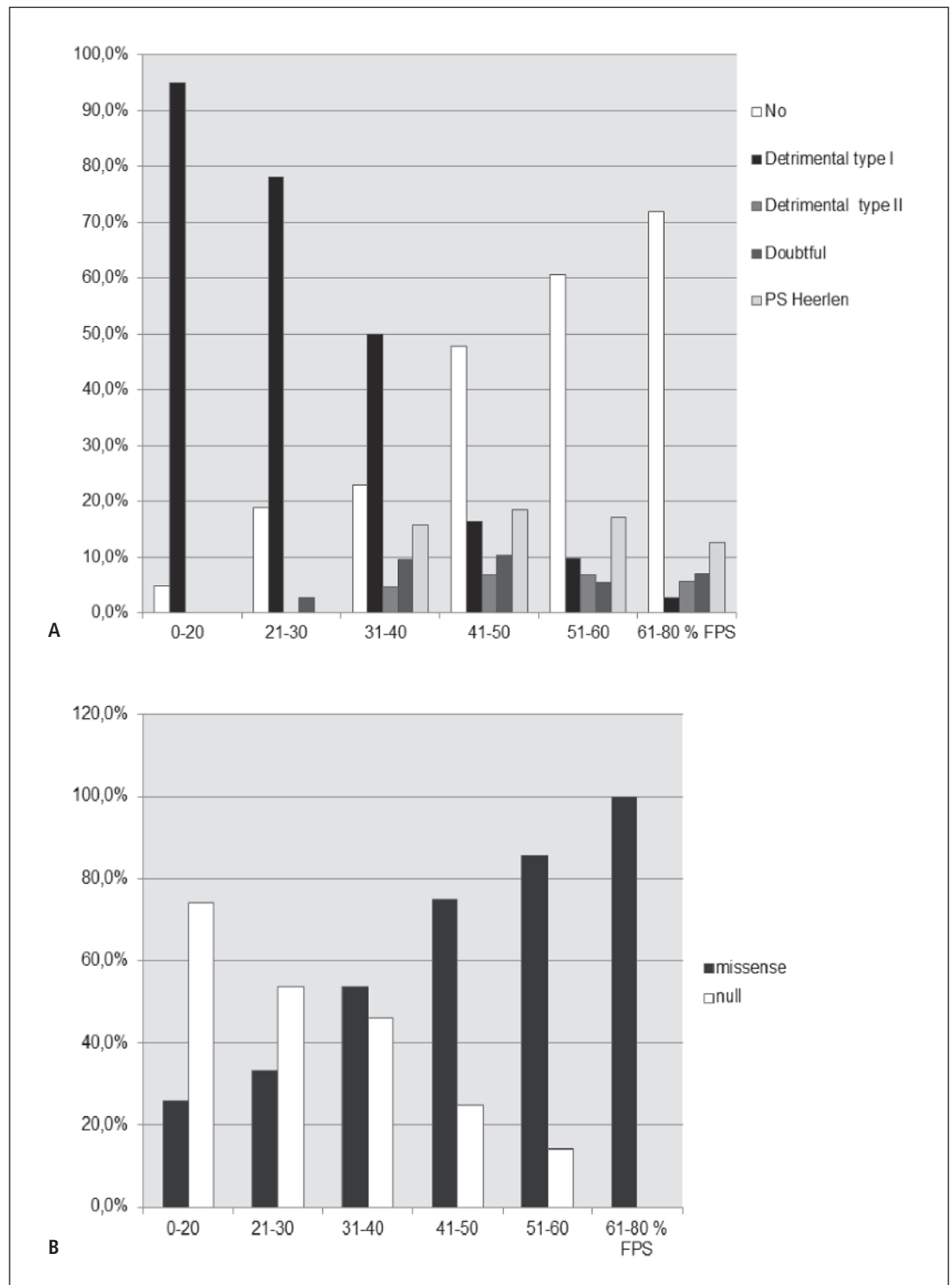
At last, presence of mild mutation on one allele explained moderate decrease in FPS level observed in compound heterozygotes reported here.

We studied VTE risk in the population comprising all subjects studied in the context of thrombosis history. We did not exclude subjects with unexplained AT suspicions since it has been reported that hereditary PS deficiency could confer increased risk of AT events (29).

Concerning risk of VTE, our findings are in accordance with previous studies reporting a significant association between FPS level and VTE, and a cut-off point for FPS level at 30–40% for identifying subjects at increased risk (3–8, 13). However, whether this cut-off value applies to all genetic status was not clear in these studies, because in most of them no genotyping was performed. In the two cohorts of thrombophilic family members in which gene analysis was performed, detrimental mutation carrying was associated with a five-fold increased risk of VT (4, 6). It was suggested that the risk may be variable and dependent on the functional properties of causative variants but no information on the influence of PS level on risk of thrombosis among detrimental variant carriers was provided. In the present study, we observed a significant risk of VT associated with type I/IIIDM heterozygous carrying, in accordance with the previous publications. However, the level of risk was lower than previously reported. This may be due to the clinical profile of the population. PS Heerlen, or type II mutation heterozygous carrying non associated with clear PS activity/FPS discrepancy did not increase the risk of VTE, in accordance with previous reports (21, 23).

Moreover, we found no significant association between FPS level and VTE in subjects with no mutation and we observed that,





**Figure 2: Mutation profile according to FPS level. A) All mutation types. B) Missense and null mutations in carriers of detrimental type I/III mutations.**

among type I/IIIDM heterozygous carriers, only those with FPS level under a cut-off level of 30% were at increased risk.

Main potential limitations of the study are its retrospective nature, and the fact that FPS measurements were not centralised, albeit most laboratories used reagent from the same manufacturer.

Our findings confirm the value of a cut-off FPS level far below the lower limit of normal range for identifying subjects at increased risk of VT. Moreover, pending confirmation in other studies, our findings suggest a place for genotyping in diagnosis strategy of quantitative PSD, for example when results of plasma

**Table 4: Risk of thrombosis according to mutation profile.** \*Reference category \*\* adjusted on sex and PS level. RR: relative risk.

| Mutation type     | Whole population (n) | Symptomatic (n) | Crude RR (95 %CI) | Adjusted RR ** (95 %CI) |
|-------------------|----------------------|-----------------|-------------------|-------------------------|
| No                | 194                  | 120             | 1.0*              |                         |
| Non-detrimental?  | 26                   | 20              | 1.06 (0.66–1.72)  | 1.09 (0.68–1.77)        |
| PS Herleen        | 58                   | 42              | 1.14 (0.78–1.65)  | 1.23 (0.84–1.79)        |
| Detrimental (all) | 189                  | 132             | 1.42 (1.10–1.85)  | 1.41 ((1.05–1.89)       |
| -Null             | 95                   | 66              | 1.49 (1.08–2.05)  | 1.42 (0.96–2.09)        |
| -Missense type I  | 73                   | 51              | 1.40 (1.00–1.97)  | 1.45 (1.02–2.07)        |
| -Missense type II | 21                   | 15              | 1.25 (0.72–2.19)  | 1.29 (0.73–2.25)        |

\*Reference category \*\* adjusted on sex and PS level.

phenotyping are doubtful, or in patients with clearly low FPS level when no family study is feasible.

#### Acknowledgements

Hospitals in which initial thrombophilia screening was performed are located in Paris (AP-HP), Argenteuil, Beziers, Bordeaux, Lille, Limoges, Lyon, Marseille, Montpellier, Nantes, Poissy, Quimper, Reims, Rouen, St Etienne, Tours, Valenciennes. We thank the following colleagues of the GEHT genetic thrombophilia group for giving us the opportunity to study their patients: Dr N Ajzenberg, A Ankri, S Bellucci, JY Borg, C Biron-Andreani, C Boiteux, L Boudaoud, D Brunet, L Darnige, MD Dautzenberg, E de Raucourt, B Delahousse, M Donnard, M Dreyfus, A Dupont, J Emmerich, R Favier, Y Gruel, MH Horellou, MF Hurtaud, M Gouault-Heilman,

#### What is known about this topic?

- Inherited PS deficiency is an established risk factor for venous thromboembolism. Several papers suggested a cut-off level of free PS around 30–40 % for identifying subjects at increased risk of venous thrombosis.
- However, whether this cut-off can apply across all *PROS1* genotypes is not clear

#### What does this paper add?

- Subjects with suspicion of inherited quantitative PS deficiency having increased risk of thrombosis are those with clearly low free PS level (<30 %) who carry a detrimental *PROS1* type I/III mutation. Our findings suggest interest for genotyping in diagnosis strategy of quantitative PS deficiency.

**Table 5: Risk of VTE according to presence or absence of a detrimental type I/III mutation and FPS level.**

|                           | Number of subjects (number of events) | RR (95 %CI) for FPS cut-off 30 % |
|---------------------------|---------------------------------------|----------------------------------|
| No mutation and FPS >30 % | 177 (109)                             | 1 (reference category)           |
| No mutation and FPS <30 % | 17 (11)                               | 0.58 (0.31–1.09)                 |
| Mutation and FPS >30 %    | 70 (45)                               | 0.90 (0.62–1.29)                 |
| Mutation and FPS <30 %    | 98 (72)                               | 1.48 (1.08–2.04)*                |

\*: p=0.008 as compared with mutation and FPS >30 %.

B Jude, F Lellouche, E Mazoyer, P Morange, P Nguyen, JL Reny, A Robert, J Roussi, L Rugeri, B Tardy, C Ternisien, C Trichet, N Trillot, V Siguret. We thank I Presot F Anezmi and ML Aubry for the excellent technical assistance. We thank Pr E Oger (Pharmacologie Clinique, CRPV and Université Rennes 1, France) for help in improving the paper.

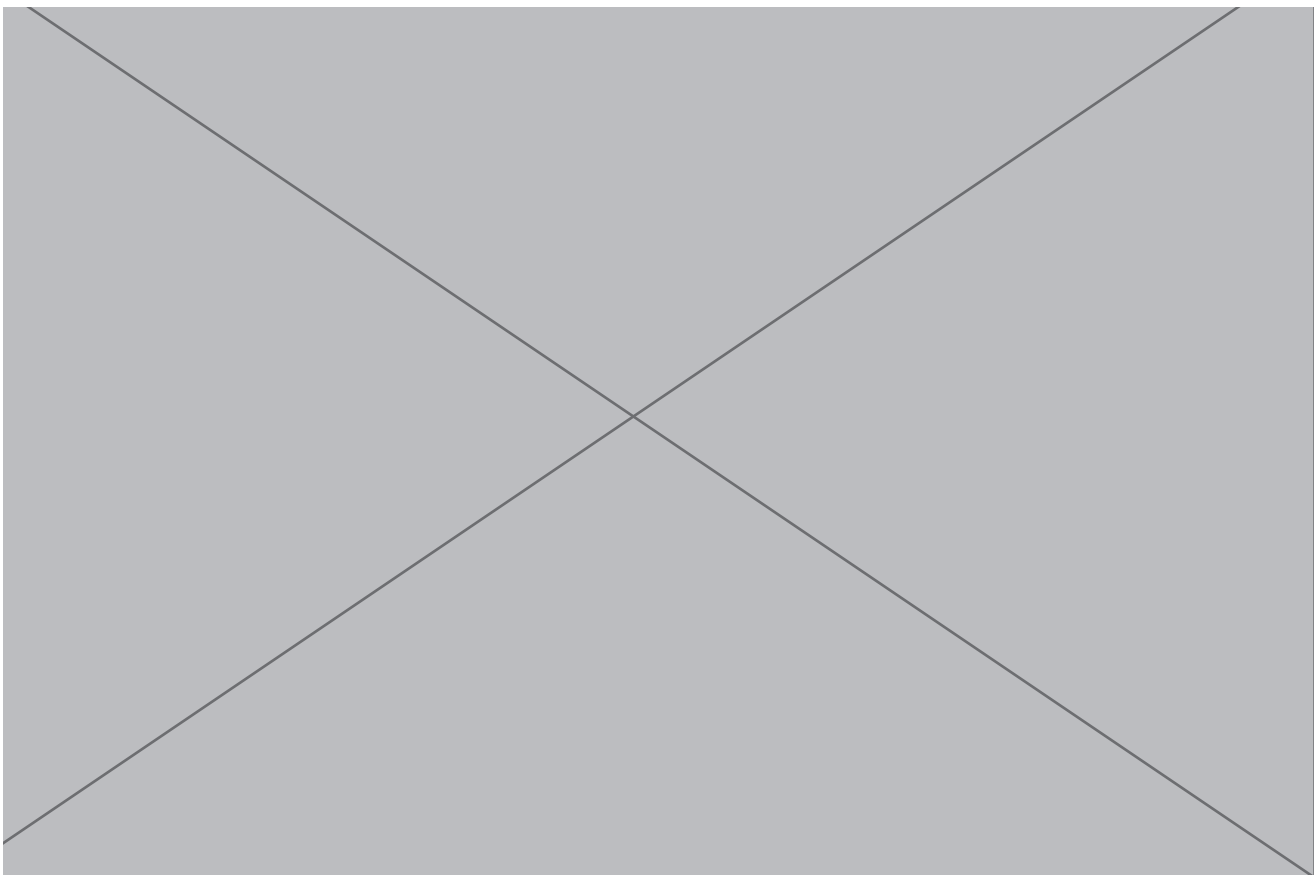
#### Conflicts of interest

None declared.

#### References

1. Castoldi E, Hackeng TM. Regulation of coagulation by protein S. *Curr Opin Hematol* 2008; 15: 529–536.
2. Garcia de Frutos P, Fuentes-Prior P, Hurtado B, et al. Molecular basis of protein S deficiency. *Thromb Haemost* 2007; 98: 543–545.
3. Castoldi E, Maurissen LFA, Tormene D, et al. Similar hypercoagulable state and thrombosis risk in type I and type III protein S-deficient individuals from families with mixed type I/III protein S deficiency. *Haematologica* 2010; 95: 1563–1571.
4. Makris M, Leach M, Beauchamp NJ, et al. Genetic analysis, phenotypic diagnosis, and risk of venous thrombosis in families with inherited deficiencies of PS. *Blood* 2000; 95: 1935–1941.
5. Brouwer JLP, Veeger NJGM, Van der Schaaf W, et al. Difference in absolute risk of venous and arterial thrombosis between familial protein S deficiency type I and type III. Results from a family cohort study to assess the clinical impact of a laboratory test-based classification. *Brit J Haematol* 2005; 128: 703–710.
6. Biguzzi E, Razzari C, Lane DA, et al. and protein S Italian team (PROSIT). Molecular diversity and thrombotic risk in protein S deficiency: the PROSIT study. *Hum Mut* 2005; 25: 259–269.
7. Lijfering WM, Mulder R, Ten Kate MK, et al. Clinical relevance of decreased free protein S levels: results from a retrospective family cohort study involving 1143 relatives. *Blood* 2009; 113: 1225–1230.
8. Mahmoodi BK, Brouwer JLP, Ten Kate MK, et al. A prospective cohort study on absolute risks of venous thromboembolism and predictive value of screening asymptomatic relatives of patients with hereditary deficiencies of protein S, protein C or antithrombin. *J Thromb Haemost* 2010; 8: 1193–1200.
9. Simmonds RE, Ireland H, Lane DA, et al. Clarification on the risk of venous thrombosis associated with hereditary protein S deficiency by investigation of a large kindred with characterized gene defect. *Ann Intern Med* 1998; 128: 8–14.
10. Koster T, Rosendaal FR, Briet E, et al. Protein C deficiency in a controlled series of unselected patients: an infrequent but clear risk factor for venous thrombosis (Leiden thrombophilia study). *Blood* 1995; 10: 2756–2761.

11. Faioni EM, Valsecchi C, Taioli E, et al. Free protein S deficiency is a risk factor for venous thrombosis. *Thromb Haemost* 1997; 78: 1343–1346.
12. Pintao MC, Ribeiro DD, Bezemer ID, et al. Protein S levels and the risk of venous thrombosis. Results from the MEGA case-control study. *Blood* 2013;122: 3210–3219.
13. Liberti G, Bertina RM, Rosendaal FR. Hormonal state rather than age influences cut-off values of protein S: reevaluation of the thrombotic risk associated with protein S deficiency. *Thromb Haemost* 1999; 82: 1093–1096.
14. Bucciarelli P, Passamonti SM, Biguzzi E, et al. Low borderline plasma levels of antithrombin, protein C and protein S are risk factors for venous thromboembolism. *J Thromb Haemost* 2012; 10: 1783–1791.
15. Gandrille S, Borgel D, Eschwege-Gufflet V, et al. Identification of 15 different candidate causal point mutations and three polymorphisms in 19 patients with protein S deficiency using a scanning method for the analysis of the protein S active gene. *Blood* 1995; 85: 130–138.
16. Hall AJ, Peake IR, Winship PR. Regulation of the human protein S gene promoter by liver enriched transcription factors. *Brit J Haematol* 2006; 135: 538–546.
17. Ng PC, Henikoff S. SIFT: Predicting amino-acid changes that affect protein function. *Nucleic Acids Res* 2003; 30: 3812–3814.
18. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nature Methods* 2010, 7: 248–249.
19. Houdayer C. In silico prediction of splice-affecting nucleotide variants. In silico tools for gene discovery. *Methods Mol Biol* 2011, 760: 269–281.
20. Contal C, O'Quigley J. An application of change point methods in studying the effect of age on survival in breast cancer. *Comput Statistics Data Anal* 1999; 30: 253–270.
21. Bertina RM, Ploos Van Amstel HK, Van Wunjaarden A, et al. Heerlen polymorphism of protein S, an immunologic polymorphism due to dimorphism of residue 406. *Blood* 1990;75: 538–548.
22. Rockwell NC, Thorner JW. The kindest cuts of all crystal structures of Kex2 and furin reveal secrets of precursor processing. *Trends Biochem Sci* 2004; 29: 80–87.
23. Alhenc Gelas M, Canonico M, Morange PE, et al.; GEHT Genetic Thrombophilia Group. Protein S inherited qualitative deficiency: novel mutations and phenotypic influence. *J Thromb Haemost* 2010; 8: 2718–2726.
24. Caspers M, Pavlova A, Driesen J, et al. Deficiencies of antithrombin, protein C and protein S – Practical experience in genetic analysis of a large patient cohort. *Thromb Haemost* 2012; 108: 247–257.
25. Duebgen S, Kauke T, Marschall C, et al. Genotype and laboratory and clinical phenotypes of protein S deficiency. *Am J Clin Pathol* 2012; 137: 178–184.
26. Suzuki A, Sanda N, Miyawaki Y, et al. Down-regulation of PROS1 gene expression by 17 $\beta$ -estradiol via estrogen receptor  $\alpha$  (ER $\alpha$ )-Sp1 interaction recruiting receptor-interacting Protein 140 and the corepressor-HDAC3 complex. *J Biol Chem* 2010; 285: 13444–13453.
27. Rezende SM, Lane D, Zöller B, et al. Genetic and phenotypic variability between families with hereditary protein S. *thromb Haemost* 2002; 87: 258–265.
28. Denis CV, Roberts SJ, Hackeng TM, et al. In vivo clearance of human protein S in a mouse model. Influence of C4b-binding protein and the Heerlen polymorphism. *Arterioscl Thromb Vasc Biol* 2005; 25: 2209–2215.
29. Mahmoodi BK, Brouwer JLP, Veeger NJGM, et al. Hereditary deficiency of protein C or protein S confers increased risk of arterial thromboembolic events at a young age. Results from a large family study cohort. *Circulation* 2008; 118: 1659–1167.




## **Article 2**

**“Protein S Heerlen mutation heterozygosity is associated with venous thrombosis risk”**

*Scientific reports 2017*

# SCIENTIFIC REPORTS



OPEN

## Protein S Heerlen mutation heterozygosity is associated with venous thrombosis risk

P. Suchon<sup>1,2</sup>, M. Germain<sup>3,4</sup>, A. Delluc<sup>5</sup>, D. Smadja<sup>6,7</sup>, X. Jouven<sup>8,9,10</sup>, B. Gyorgy<sup>3,4</sup>, N. Saut<sup>1</sup>, M. Ibrahim<sup>1,2</sup>, J. F. Deleuze<sup>11,12</sup>, M. C. Alessi<sup>1,2</sup>, P. E. Morange<sup>1,2,\*</sup> & D. A. Trégouët<sup>3,4,\*</sup>

Received: 31 October 2016

Accepted: 27 February 2017

Published: 04 April 2017

Hereditary Protein S (PS) deficiency is a rare coagulation disorder associated with an increased risk of venous thrombosis (VT). The PS Heerlen (PSH) mutation is a rare S501P mutation that was initially considered to be a neutral polymorphism. However, it has been later shown that PSH has a reduced half-life *in vivo* which may explain the association of PSH heterozygosity with mildly reduced levels of plasma free PS (FPS). Whether the risk of VT is increased in PSH carriers remains unknown. We analyzed the association of PSH (rs121918472 A/G) with VT in 4,173 VT patients and 5,970 healthy individuals from four independent case-control studies. Quantitative determination of FPS levels was performed in a subsample of 1257 VT patients. In the investigated populations, the AG genotype was associated with an increased VT risk of 6.57 [4.06–10.64] ( $p = 1.73 \times 10^{-14}$ ). In VT patients in whom PS deficiency was excluded, plasma FPS levels were significantly lower in individuals with PSH when compared to those without [72 + 13 vs 91 + 21 U/dL;  $p = 1.86 \times 10^{-6}$ , mean + SD for PSH carriers ( $n = 21$ ) or controls ( $n = 1236$ ) respectively]. We provide strong evidence that the rare PSH variant is associated with VT in unselected individuals.

Protein S (PS) is a potent anticoagulant protein that down regulates thrombin formation via two mechanisms<sup>1</sup>. It can stimulate the proteolytic inactivation of coagulation factors Va and VIIIa by activated protein C (APC), and can also enhance FXa inhibition by interacting with tissue factor pathway inhibitor (TFPI). Although free PS (FPS) is the most active PS form, recent reports indicate that the PS-C4BP complex also exhibits APC- and TFPI-cofactor activities<sup>2–4</sup>.

Hereditary PS deficiency is a rare coagulation disorder associated with an increased risk of venous thrombosis (VT)<sup>5</sup>. As for other deficiencies in natural anticoagulant inhibitors (antithrombin and protein C), PS deficiency is due to rare or private mutations within the *PROS1* gene (database available online: <http://www.hgmd.cf.ac.uk>). Among these, the rs121918472 variation—referred to as the PS Heerlen mutation—is detected at a high frequency in the general population (0.5%). PS Heerlen is a S501P amino acid change polymorphism that causes a loss of N-linked glycosylation at position 499<sup>6</sup>, and was initially considered a neutral polymorphism. However, it was later shown that 1) PS Heerlen has a reduced half-life *in vivo*<sup>7</sup>, 2) and that heterozygous and homozygous PS Heerlen genotypes were associated with mildly reduced or low levels of plasma FPS, respectively<sup>8</sup>. Moreover, PS Heerlen was found to increase thrombin generation in the presence of APC compared to wild type PS<sup>9</sup>. Thus whether VT risk is increased in PS Heerlen heterozygous subjects remains uncertain. Indeed, Bertina *et al.* found no differences in mutation prevalence in a panel of 96 uncharacterized normal individuals as compared to 1,182

<sup>1</sup>Laboratory of Haematology, La Timone Hospital, Marseille, France. <sup>2</sup>Institut National pour la Santé et la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR\_S) 1062, Nutrition Obesity and Risk of Thrombosis, Aix-Marseille University, Marseille, France. <sup>3</sup>Sorbonne Universités, UPMC Univ. Paris 06, INSERM, UMR\_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases, Paris, France. <sup>4</sup>ICAN Institute for Cardiometabolism and Nutrition, Paris, France. <sup>5</sup>Université de Brest, EA3878 and CIC1412, 29238 Brest, France. <sup>6</sup>Service d'hématologie biologique, AP-HP, Hôpital Européen Georges Pompidou, Paris, France. <sup>7</sup>Université Paris Descartes, Sorbonne Paris Cité, France, Inserm UMR-S1140, Paris, France. <sup>8</sup>INSERM, UMR-S970, Department of Epidemiology, Paris, France. <sup>9</sup>Université Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine, Paris, France. <sup>10</sup>APHP, Georges Pompidou European Hospital, Department of Cardiology, Paris, France. <sup>11</sup>Centre National de Génotypage, Institut de Génétique, CEA, 91057 Evry, France. <sup>12</sup>CEPH, Fondation Jean Dausset, Paris, France. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to P.E.M. (email: pierre.morange@ap-hm.fr)

|             | rs121918472  | AG         | GG | Cochran Armitage Trend P value | Age- and sex- adjusted Odds Ratio            |
|-------------|--------------|------------|----|--------------------------------|--|
|             | AA           |            |    |                                |  |
| All         |              |            |    |                                |  |
| Controls    | 5970 (99.6%) | 22 (0.4%)  | 0  | 4.40 10 <sup>-18</sup>         | 6.57 [4.06–10.64] p = 1.73 10 <sup>-14</sup> |
| Cases       | 4173 (97.8%) | 94 (2.2%)  | 0  |                                |  |
| EDITH       |              |            |    |                                |  |
| Controls    | 1156 (99.6%) | 5 (0.4%)   | 0  | 7.54 10 <sup>-5</sup>          | 5.59 [2.145–14.57] p = 4.29 10 <sup>-4</sup> |
| Cases       | 1117 (97.6%) | 27 (2.4%)  | 0  |                                |  |
| FARIVE      |              |            |    |                                |  |
| Controls    | 657 (98.2%)  | 12 (1.8%)  | 0  | 0.555                          | 0.72 [0.30–1.72] p = 0.4565                  |
| Cases       | 706 (98.6%)  | 10 (1.4%)  | 0  |                                |  |
| MARTHA12    |              |            |    |                                |  |
| Controls    | 768 (99.4%)  | 5 (0.6%)   | 0  | 3.22 10 <sup>-3</sup>          | 3.98 [1.48–10.72] p = 6.28 10 <sup>-3</sup>  |
| Cases       | 734 (97.5%)  | 19 (2.5%)  | 0  |                                |  |
| EOVT/MARTHA |              |            |    |                                |  |
| Controls    | 3379 (100%)  | 0 (–)      | 0  | 8.52 10 <sup>-20</sup>         | Not applicable                               |
| Cases       | 1603 (97.6%) | 40 (2.4%)* | 0  |                                |  |

**Table 1. Association of the *PROSI* rs121918472 (PS Heerlen mutation) with VT in four French case-control sample studies.** \*The frequency of AG genotype was 2.7% and 2.4% in EOVT and MARTHA VT patients, respectively.

Dutch patients with unexplained VT (0.52% vs 0.68% respectively)<sup>6</sup>. Apart from this study with poorly described patients, the impact of PS Heerlen on VT has only been studied in families either with PS deficiency or with other combined genetic or acquired risk factors. In these situations, PS Heerlen was thought to increase VT susceptibility<sup>10</sup>.

Although PS Heerlen is more frequent than any other *PROSI* mutation, it is still not considered a common variant. Consequently, previous genome-wide association studies (GWASs) which mainly interrogate common genetic variations (i.e. frequencies higher than 1%) were not able to identify the mutation<sup>11</sup>.

In the present study, we aimed to systematically determine if there was an association between PS Heerlen and VT by combining four French case-control DNA samples totaling 4,173 patients and 5,970 healthy individuals.

## Results

Genotype distributions of the rs121918472 variant within the four case-control studies are shown in Table 1. No homozygous carrier of the rare rs121918472-G allele was observed. When samples were combined, the AG genotype was more frequent in VT patients than in healthy controls, 2.2% vs 0.4% respectively, and was associated with an increased VT risk (Odds Ratio OR) of 6.57 [4.06–10.64] ( $p = 1.73 \times 10^{-14}$ ). The association was significantly heterogeneous across studies ( $p < 10^{-3}$ ) due to strong associations in the EDITH cohort (2.4% vs 0.4%), the MARTHA12 cohort (2.5% vs 0.6%), the EOVT/MARTHA cohort (2.4% vs 0%) samples, but not in FARIVE study samples (1.1% vs 1.8%). Note that the observed AG genotype frequency was higher in FARIVE controls with hypertension (2.9%,  $n = 246$ ) than in those without (0.9%,  $n = 343$ ) in whom the frequency was similar to the other cohort controls. Excluding the FARIVE study from the combined analysis increased the significance of the association between rs121918472 and VT (OR = 14.58 [7.42–28.65],  $p = 7.66 \times 10^{-15}$ ).

We also examined whether the rs121918472-G allele frequency varied in specific VT subgroups, such as in patients with provoked or unprovoked VT, patients with deep vein thrombosis or pulmonary embolism, and patients with F5 or F2 G20210A mutations. We did not observe preferential associations in any specific strata.

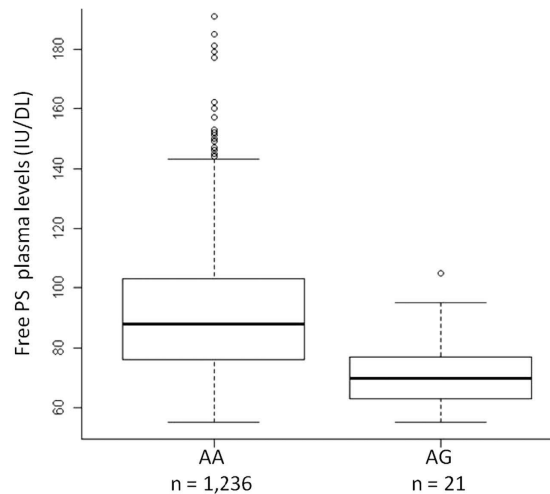
Interestingly, the rs121918472 variant was not identified as a VT-associated polymorphism in the two most recent VT GWAS projects by INVENT<sup>11</sup> and 23 and Me consortia<sup>12</sup>. In the INVENT GWAS, the Heerlen mutation imputation quality was extremely low,  $r^2 = 0.15^{11}$ , which is far lower than the gold standard of  $r^2 > 0.8$  for imputation quality, and the increased risk associated with the imputed rs121918472-G allele was OR = 1.46 [0.28–7.58],  $p = 0.655$ . Because 865 MARTHA/EOVT participants examined in the current study also took part in the INVENT study, we were able to calculate the Spearman correlation between the true rs121918472 genotype as provided by Illumina Exome array results, and the imputed dose-derived from GWAS arrays. The resulting low correlation of  $\rho = 0.25$  illustrated how imputation techniques are not reliable for inferring rare variants such as rs121918472.

To provide additional evidence implicating rs121918472 in increased VT risk, we assessed its association with plasma FPS levels in the MARTHA study samples. Even if by design FPS deficiency was excluded in this study, plasma FPS levels were significantly lower in individuals with PS Heerlen ( $p = 1.86 \times 10^{-6}$ ) compared to those without (mean  $\pm$  standard deviation:  $72 \pm 13$  vs  $91 \pm 21$  UI/dL, respectively) (Fig. 1).

## Discussion

In the present study we systematically analyzed four French case-control collections totalling 4,173 patients and 5,970 healthy individuals and provide strong evidence that the rare PS Heerlen variant is associated with VT.

In the investigated populations, the AG genotype was significantly more frequent in VT patients than in healthy controls, 2.2% vs 0.4%, respectively, and was also associated with an increased VT risk of 6.57. This



**Figure 1. Free PS plasma levels (IU/dL) in MARTHA cases according to the rs121918472 polymorphism (age and sex adjusted P-value:  $1.86 \times 10^{-6}$ ).**

association pattern was evident in EDITH, MARTHA12, and EOVT/MARTHA samples, but not in FARIVE study samples. There is no apparent explanation for the surprisingly high PS Heerlen mutation prevalence in apparently healthy individuals from the FARIVE study (1.8%). This observation could be due to the study design, as controls were actually matched cardiology patients without VT, and not healthy individuals from the general population. The significantly increased thrombosis risk associated with this mutation supports the concept that PS Heerlen has deleterious effects on the disease. However PS Heerlen is usually classified as a neutral polymorphism, both via bioinformatics prediction tools or by functional studies where it only slightly affects FPS levels in plasma. Of note, no homozygous carrier was identified in any of the four case-control studies. This observation was entirely compatible with data from the ExAC (Exome Aggregation Consortium) database, in which only one individual was homozygous for PS Heerlen among the ~33,000 studied subjects of European origin. This low frequency is likely due to the fact that homozygous carriers are considered to be protein S deficient, a condition with high morbidity. There is one PS Heerlen homozygous patient documented in the literature, with very low levels of free protein S (34%)<sup>13</sup>. In the MARTHA study, even if FPS deficiency was excluded by design, plasma FPS levels were significantly lower in individuals with PS Heerlen compared to those without. Assessing whether the disease effect is completely mediated through a direct effect on decreased FPS levels would require further investigations, including epidemiological studies where both the PSH mutation and FPS levels are measured in case-control samples. However, due to the extremely low frequency of the PSH variation in well matched control population, it will be nearly impossible to conduct mediation analysis to properly assess whether the effect of the mutation on disease risk is completely due to its effect on lowering FPS levels. However, in an additional population based study, only very low levels of FPS (<1/10<sup>th</sup> percentile or <33 UI/dl) were associated with increased VT risk<sup>14</sup>, suggesting that the effect of PS Heerlen on VT risk might thus be related to mechanisms other than solely half-life reduction. However, recombinant PS Heerlen does not differ in either binding to, or dissociation to or from, immobilized C4BP *in vitro*, nor does it show any differences in the inactivation of FVa or FVIIIa<sup>15</sup>. PS is also a cofactor for TFPI, and accelerates the inhibition of activated factor X (FXa). The PS Heerlen mutation is located within the SHBG-like domain (Val243-Ser635) critical for TFPI binding and enhancement<sup>16</sup>. Thus, one might hypothesize that PS Heerlen could affect thrombosis risk by impairing TFPI function.

The recent GWAS on VT has enabled the identification of several new VT susceptibility variants. Very recently, the INVENT consortium gathered 7,507 VT subjects and 52,632 controls and identified two new loci (*SLC44A2* and *TSPAN15*) and robustly replicated six previously identified loci<sup>11</sup>. PS Heerlen was not among the identified disease polymorphisms, likely because this GWAS relied on imputation techniques that were not optimal for inferring rare variants<sup>17</sup>. The exome chip used in the present study was enriched by rare coding variants predicted to be causal. The results presented herein as well as those from the recently published paper on key hemostatic traits<sup>18</sup> emphasize the promise of this new chip in identifying low-frequency variants associated with VT.

In conclusion, we provide evidence that PS Heerlen is associated with increased VT risk in individuals of the general population. Further investigations are required to better understand the surprisingly high prevalence of PS Heerlen mutation in apparently healthy individuals from the FARIVE study. Functional studies investigating the potential influence of PS Heerlen on VT risk would also be of benefit. These results also raise the question of whether PS Heerlen should always be genotyped irrespective of plasma PS levels in thrombophilia screening.

## Methods

The present work utilized four case-control sample cohorts (EDITH, FARIVE, EOVT/MARTHA)<sup>11</sup> where VT events (pulmonary embolism or deep vein thrombosis) were objectively diagnosed<sup>11</sup>. Detailed descriptions of the four case-control sample groups are included in the supplementary text.

In the FARIVE and EDITH studies, the PS Heerlen variation was genotyped by High Resolution Melting (HRM). All EOVT and MARTHA patients were typed using the Illumina Human Exome Beadchip 12v1-2-A



array and compared to control participants from the Paris Prospective Study<sup>19</sup> typed with the Human Exome Beadchip 12v1-A chip (Illumina, Inc., Sand Diego, CA). Both of these genotyping arrays were designed to assess coding variants across the whole genome. Data analysis revealed 239,569 single nucleotide polymorphisms common to both arrays, including PS Heerlen rs121918472. Individuals with less than 98% genotyping success or high heterozygosity were excluded from the analyses, as were individuals demonstrating previously unknown close relatedness. The latter were separately assessed via pairwise clustering of identity by state distance (IBS) and multi-dimensional scaling (MDS) using PLINK software<sup>20</sup>. Genetic outlier analysis was performed using IBS statistics and principal component analysis to detect individuals of non-European origin, who were then removed from the analysis. This led to a selected cohort of 2,630 VT patients and 3,414 controls typed with Illumina exome chips for final association analysis.

The overall rs121918472 genotype call rate was 0.98, with 0.97 in the FARIVE and EDITH studies typed with HRM, and 0.99 in individuals typed with the Illumina Human Exome chips.

In MARTHA, fasting blood was drawn for FPS measurement. Quantitative FPS determination was performed by enzyme-linked immunosorbent assay (ELISA) using the Asserachrom FPS assay (Diagnostica Stago). PS deficiency was defined as FPS plasma antigen levels  $<55$  IU/dL<sup>-1</sup> using the specific ELISA assay. FPS data were available in 1,257 MARTHA individuals.

Genotype distribution of the rs121918472 was tested for deviation from Hardy-Weinberg (HW) equilibrium via the HW exact test. Association of rs121918472 with VT was analyzed via the Cochran-Armitage trend test and by logistic regression analysis adjusting for age and sex. Association of rs121918472 with FPS was assessed using linear regression analysis. Analyses were first performed separately in each study, and were then performed on all samples combined and further adjusted for the study group.

Informed consent was obtained from all participants, and the study met all institutional ethics requirement. The procedures employed were reviewed and approved by the *Assistance Publique des Hôpitaux de Marseille* institutional review committee.

## References

- Castoldi, E. & Hackeng, T. M. Regulation of coagulation by protein S. *Curr Opin Hematol* **15**, 529–36 (2008).
- Maurissen, L. F., Thomassen, M. C., Nicolaes, G. A., Dahlbäck, B., Tans, G., Rosing, J. & Hackeng, T. M. Re-evaluation of the role of the protein S-C4b binding protein complex in activated protein C-catalyzed factor Va-inactivation. *Blood* **111**, 3034–41 (2008).
- Dahlbäck, B. C4b-binding protein: a forgotten factor in thrombosis and hemostasis. *Semin Thromb Hemost* **37**, 355–61 (2011).
- Hackeng, T. M., Seré, K. M., Tans, G. & Rosing, J. Protein S stimulates inhibition of the tissue factor pathway by tissue factor pathway inhibitor. *Proc Natl Acad Sci USA* **103**, 3106–11 (2006).
- Schwarz, H. P., Fischer, M., Hopmeier, P., Batard, M. A. & Griffin, J. H. Plasma protein S deficiency in familial thrombotic disease. *Blood* **64**, 1297–300 (1984).
- Bertina, R. M. *et al.* Heerlen polymorphism of protein S, an immunologic polymorphism due to dimorphism of residue 460. *Blood* **76**, 538–48 (1990).
- Denis, C. V., Roberts, S. J., Hackeng, T. M. & Lenting, P. J. *In vivo* clearance of human protein S in a mouse model: influence of C4b-binding protein and the Heerlen polymorphism. *Arterioscler, Thromb Vasc Biol.* **25**, 2209–15 (2005).
- Duchemin, J. *et al.* The Ser 460 to Pro substitution of the protein S alpha (PROS1) gene is a frequent mutation associated with free protein S (type IIa) deficiency. *Blood* **86**, 3436–43 (1995).
- Alhenc-Gelas, M., Canonico, M., Morange, P. E. & Emmerich, J. Geht Genetic Thrombophilia Group. Protein S inherited qualitative deficiency: novel mutations and phenotypic influence. *J Thromb Haemost.* **8**, 2718–26 (2010).
- Giri, T. K., Yamazaki, T., Sala, N., Dahlbäck, B. & de Frutos, P. G. Deficient APC-cofactor activity of protein S Heerlen in degradation of factor Va Leiden: a possible mechanism of synergism between thrombophilic risk factors. *Blood* **96**, 523–31 (2000).
- Germain, M. *et al.* Meta-analysis of 65,734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism. *Am J Hum Genet.* **96**, 532–42 (2015).
- 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. *Hum Mol Genet.* **25**, 1867–74 (2016).
- Wypasek, E., Potaczek, D. P., Alhenc-Gelas, M. & Undas, A. Heerlen polymorphism associated with type III protein S deficiency and factor V Leiden mutation in a Polish patient with deep vein thrombosis. *Blood Coagul Fibrinolysis* **25**, 84–5 (2014).
- Pintao, M. C. *et al.* Protein S levels and the risk of venous thrombosis: results from the MEGA case-control study. *Blood* **122**, 3210–9 (2013).
- Koenen, R. R., Gomes, L., Tans, G., Rosing, J. & Hackeng, T. M. The Ser460Pro mutation in recombinant protein S Heerlen does not affect its APC-cofactor and APC-independent anticoagulant activities. *Thromb Haemost.* **91**, 1105–14 (2004).
- Reglińska-Matveyev, N. *et al.* TFPI cofactor function of protein S: essential role of the protein S SHBG-like domain. *Blood* **123**, 3979–87 (2014).
- Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–53 (2009).
- Huffman, J. E. *et al.* Rare and low-frequency variants and their association with plasma levels of fibrinogen, FVII, FVIII, and vWF. *Blood* **126**, 19–29 (2015).
- Proust, C. *et al.* Contribution of Rare and Common Genetic Variants to Plasma Lipid Levels and Carotid Stiffness and Geometry: A Substudy of the Paris Prospective Study 3. *Circ Cardiovasc Genet.* **8**, 628–36 (2015).
- Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* **81**, 559–75 (2007).

## Acknowledgements

We are grateful to Rachel Peat for English editing and grammatical corrections. This research was partially supported by the GENMED Laboratory of Excellence on Medical Genomics (ANR-10-LABX-0013), the French Clinical Research Infrastructure Network on Venous Thrombo-Embolic (F-CRIN INNOVTE) and the ICAN Institute for Cardiometabolism and Nutrition (ANR-10-IAHU-05), three research programs managed by the National Research Agency (ANR) as part of the French Investment for the Future initiative.



### Author Contributions

P.S., M.G. and B.G. were in charge of statistical data analysis. N.S. and M.I. were in charge of the genotyping work. A.D., D.S., X.J., J.F.D., M.C.A., P.E.M. and D.A.T. designed the participating studies and the overall project. P.S. drafted the manuscript that was further revised by A.D., D.S., X.J., P.E.M. and D.A.T.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing Interests:** The authors declare no competing financial interests.

**How to cite this article:** Suchon, P. *et al.* Protein S Heerlen mutation heterozygosity is associated with venous thrombosis risk. *Sci. Rep.* 7, 45507; doi: 10.1038/srep45507 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

# **Partie 2 : facteurs de risque fréquents et évaluation du risque individuel de thrombose**

# Introduction

# I. Identification de nouveaux variants génétiques associés à la maladie thromboembolique veineuse

Les 5 anomalies du bilan de thrombophilie n'expliquant qu'une faible part de l'héritabilité de la MTEV<sup>104</sup>, d'autres facteurs de risque génétiques restent donc à identifier. De nombreuses études ont entrepris ce travail. Deux méthodologies ont prédominé : les approches gène-candidat et pangénomique ou *genome wide association study* en anglais (GWAS). Toutes deux constituent des études d'association. La première a été largement utilisée au cours des années 90 et 2000. Elle nécessite la compréhension préalable de la physiopathologie du phénotype étudié. Le séquençage de gènes d'intérêt associé au génotypage des polymorphismes alors identifiés dans un groupe de sujets atteints et un groupe de sujets sains pour le phénotype étudié permet alors de mettre en évidence une association entre un marqueur génétique et la pathologie. Le second type d'étude repose sur le génotypage de centaines de milliers de polymorphismes dans ces mêmes groupes de patients. Il s'agit donc d'une étude cas-témoins à très grande échelle. La très grande quantité de polymorphismes étudiés et donc de tests générés nécessite d'utiliser des seuils de significativité bien inférieurs à ceux habituellement utilisés (en règle générale 5%) pour ainsi maîtriser le risque d'erreur de type I (risque alpha). Seuls des effectifs importants, le plus souvent plusieurs centaines voire milliers de patients, permettent alors d'obtenir une puissance statistique suffisante. Ces études nécessitent au préalable l'identification des polymorphismes qui seront génotypés. Cette étape a été rendue possible par le séquençage complet du génome humain (projet génome humain), et le projet HapMap achevés respectivement en 2003 et 2005. Depuis, de nombreux projets ont vu le jour et ont permis d'augmenter le recensement des polymorphismes. Notamment, le « projet 1000 génomes » a entrepris en 2008 le séquençage de 2504 génomes (participants anonymes issus de 26 populations d'Afrique, Amérique, Asie, Europe). Il a permis d'élaborer un catalogue de plus de 95% des variations génétiques existants dans le génome humain avec une fréquence de l'allèle mineur  $\geq 1\%$ <sup>105</sup>. C'est aujourd'hui plus de 15000 génomes et leurs données correspondantes qui sont disponibles grâce à la base de données gnomAD. Contrairement à l'approche gène candidat, cette approche est agnostique. Elle permet alors de découvrir de nouvelles voies physiopathologiques. Elle ne permet toutefois pas de conclure à la causalité. En effet,

l'identification d'une association entre un polymorphisme et un phénotype doit d'une part être vérifiée (étape de réplique dans une cohorte indépendante), et d'autre part, le polymorphisme identifié ne permet que de localiser une région d'intérêt. En effet, l'analyse pangénomique repose sur le déséquilibre de liaison. L'association entre polymorphisme et phénotype signifie donc que le polymorphisme causal est en déséquilibre de liaison avec celui étudié.

Nous allons passer en revue les variants récemment identifiés à l'aide de ces deux méthodologies.

## 1. Découverte de nouveaux variants par l'approche gène-candidat

Le gène *PROCR*, qui code pour le récepteur à la PC (238 AA), est situé sur le chromosome 20. La PC contribue par plusieurs voies, notamment par son interaction avec son récepteur, à la formation du thrombus. Une étude cas-témoins conduite chez des patients coronariens avait préalablement montré une association entre l'augmentation de la génération de thrombine et le variant du gène codant pour le récepteur à la PC, *PROCR* Ser219Gly<sup>106</sup>. L'équipe de Reiner a mis en évidence en 2008 une modulation des taux de PC et des taux de récepteur soluble de la PC (sEPCR) par ce même polymorphisme *PROCR* Ser219Gly (rs867186)<sup>107</sup>. D'autres études ont ensuite étayé ces données. Ce polymorphisme du codon 19 explique 56 à 87 % des variations des taux circulants de sEPCR<sup>108-110</sup> et 10% des taux de PC<sup>111</sup>. De façon intéressante, le facteur VIIa se lie avec la même affinité que la PC à EPCR. Le facteur VIIa ainsi lié ne peut activer le facteur X. Cette liaison facilite par ailleurs l'endocytose du facteur VIIa, suggérant ainsi que la liaison EPCR/facteur VIIa est un mécanisme d'élimination du facteur VIIa. Ainsi, une augmentation des taux d'EPCR pourrait avoir un effet anticoagulant<sup>112</sup>. Chez des sujets porteurs de l'allèle à risque Ser219Gly, des taux augmentés de sEPCR sont observés, de même qu'une augmentation des taux de facteur VII ainsi que des marqueurs de la coagulation<sup>113</sup>. Les auteurs suggèrent que l'augmentation des taux de sEPCR s'accompagne d'une diminution des taux d'EPCR présents à la surface des cellules endothéliales et ainsi d'une diminution de la liaison EPCR/facteur VII (activé). Malgré l'association indéniable entre ce polymorphisme et ces phénotypes intermédiaires, l'association entre le polymorphisme et le risque de MTEV est encore incertaine. Toutefois,

une méta-analyse d'études observationnelles a récemment démontré une association entre ce polymorphisme et la MTEV avec un OR estimé à 1,22 pour l'allèle à risque rs867186-G<sup>114</sup>. Un autre polymorphisme du gène *PROCR*, rs2069951, semble d'intérêt compte tenu de sa présence dans un haplotype associé à une augmentation des taux de sEPCR et du risque de MTEV<sup>115</sup>. La fréquence de l'allèle de l'haplotype à risque était de 13,8% et l'OR de 1,8 en analyse multivariée prenant en compte le sexe, l'âge, la présence d'un FVL ou de PTG20210A.

Les gènes *FGA*, *FGB*, *FGG* codent pour les chaînes polypeptidiques  $\alpha\alpha$ ,  $\beta\beta$  et  $\gamma$  ( $\gamma A$  est la forme majoritaire et  $\gamma'$  la forme minoritaire) de la molécule de fibrinogène. Ils sont organisés en cluster sur le chromosome 4q31.3. Lors de la dernière étape permettant d'obtenir la formation du caillot, la thrombine clive les chaînes  $\alpha$  et  $\beta$  générant les fibrinopeptides A et B qui se polymérisent spontanément pour former ainsi la fibrine insoluble. Le facteur XIII établit des liaisons covalentes entre les monomères de fibrine ainsi stabilisés. Une étude cas-témoins d'envergure a montré dans la cohorte LETS (*Leiden Thrombophilia Study*) qu'une augmentation des taux de fibrinogène augmentait le risque de MTEV<sup>116,117</sup>. C'est en 2005, dans cette même cohorte, qu'Uitte de Willige et son équipe ont expliqué les variations des taux de fibrinogène  $\gamma'$  par le polymorphisme rs2066865 (C10034T) du gène *FGG*<sup>118</sup>. Quatre ans plus tard, la même équipe a mis en évidence son association avec la MTEV. Le risque de MTEV idiopathique apparaissait ainsi multiplié par 2 chez les porteurs de l'allèle 10034TT dans une population caucasienne<sup>119</sup>.

Le polymorphisme rs3136516 (A19911G) situé sur le gène de la prothrombine a été rapporté comme associé à la MTEV. Une publication a notamment évalué l'impact du génotype rs3136516-AG chez les sujets porteurs de la mutation PTG20210A à l'état hétérozygote. Le polymorphisme rs3136516 modulait le risque chez ces patients. Ainsi, comparativement aux sujets non mutés pour les 2 variants, le risque de MTEV était estimé à 5,86 (3,22-10,86) pour la combinaison PTG20210A à l'état hétérozygote/rs3136516-AG alors qu'il n'était que de 3,34 (1,99-5,64) chez les sujets hétérozygotes pour PTG20210A non mutés pour rs3136516 (génotype AA)<sup>120</sup>. Le risque de MTEV des porteurs du polymorphisme rs3136516-G était par ailleurs multiplié par 1,5 chez les patients sans thrombophilie et par 2,1 lorsqu'ils étaient porteurs d'un FVL<sup>121</sup>. De façon intéressante, le polymorphisme était associé à l'augmentation des taux de prothrombine<sup>122</sup>.

Le polymorphisme rs4524 est situé sur le gène du facteur V. Il existait une association avec la MTEV dans une étude cas-témoins de femmes ménopausées ou en péri-ménopause <sup>123</sup>. Le polymorphisme entraîne la substitution d'une lysine (K) par une arginine (R) en position 858 (K858R) du facteur V, dans son domaine B. Trois études ont par ailleurs confirmé cette association après ajustement pour le FVL <sup>124</sup>. La fréquence de l'allèle à risque était de 74% et le risque était multiplié par 1,21. Dans une récente méta-analyse, la fréquence de l'allèle à risque (rs4524-T) était estimée à 0,74 et le polymorphisme conférait une augmentation du risque de MTEV de 1,20 (IC95% = 1,14-1,26) <sup>125</sup>.

Comme pour la PS, les mutations responsables de déficits en AT sont généralement privées. Toutefois, une récente étude a mis en évidence une association entre le polymorphisme rs2227624 situé sur *SERPINC1* et la MTEV sans que ce polymorphisme ne soit associé à un déficit plasmatique en AT. Le polymorphisme porte également le nom d'antithrombine Dublin. Il s'agit d'un variant non synonyme, Val30Glu, décrit pour la première fois en 1987 <sup>126</sup>. Dans l'étude de Navarro-Fernandez et al, le polymorphisme, génotypé chez plus de 4000 patients était associé au risque de MTEV avec un OR estimé à 2,9 (IC95% = 1,07-8,09). Il faut noter qu'un autre polymorphisme de *SERPINC1*, a été identifié comme associé au risque de MTEV dans une étude dont nous reparlerons dans le chapitre suivant. Il s'agit du polymorphisme intronique rs2227589 <sup>127</sup>. Cependant, cette association initialement décrite n'a pas été reproduite dans les études de réplication qui ont suivi. Au demeurant, le polymorphisme a été décrit comme étant fonctionnel et associé à des taux modérément diminués de l'activité anticoagulante de l'AT <sup>128</sup>.

## *2. Découverte de nouveaux variants par l'approche génome entier ou GWAS*

En 2008, une étude a évalué pour la première fois l'association entre 20000 polymorphismes environ et la MTEV <sup>127</sup>. Il ne s'agissait pas à proprement d'une GWAS parce que les polymorphismes étudiés ont été sélectionnés en raison de leur capacité à affecter la fonction ou l'expression du gène, mais la méthodologie employée en termes de génotypage,

d'effectifs et de tests statistiques se rapprochait d'une étude d'association pangénomique. Trois polymorphismes étaient robustement associés à la MTEV : rs13146272 situé sur le gène *CYP4V2* (OR=1,24 ; IC95% = 1,11-1,37), rs2227589 situé sur le gène *SERPINC1* (OR=1,29 ; IC95% = 1,10-1,49) et rs1613662 situé sur le gène *GP6* (OR=1,15 ; IC95% = 1.01-1.30).

*CYP4V2* code pour une protéine de la famille des CYP450. La fréquence de l'allèle à risque (rs13146272-A) est estimée à 64%. Il se situe sur le chromosome 4, dans une région contenant des gènes codant pour des protéines de la coagulation : la prékallikréine (*KLKB1*) et le facteur XI (*F11*). L'association observée entre rs13146272 et la MTEV s'expliquait par le déséquilibre de liaison qui existe entre ce polymorphisme et ceux du *F11* <sup>129</sup>.

Le gène *SERPINC1* code pour l'AT. La fréquence de l'allèle à risque (rs2227589-A) dans l'étude de Bezemer et al. était de 10%. Le polymorphisme se traduit par la substitution d'une guanine par une adénine en position 786 dans l'intron 1. D'après une étude menée dans une population de donneurs de sang espagnols, ce polymorphisme semblerait diminuer discrètement le taux d'AT (respectivement 94,8% +/-5,6% versus 99,5% +/-5,8%) comparativement aux sujets rs2227589-GG <sup>130</sup>. Dans une étude conduite chez 188 patients hétérozygotes pour le FVL, l'haplotype contenant rs2227589-A diminuait le taux d'AT et augmentait la génération de thrombine <sup>131</sup>. Toutefois, ces résultats n'ont pu être confirmés dans une étude de réplcation incluant 1076 cas et 1239 témoins <sup>132</sup>.

*GP6* code pour une glycoprotéine présente à la surface des plaquettes : la GpVI. Cette protéine joue un rôle majeur dans l'activation et l'agrégation plaquettaire en tant que récepteur du collagène <sup>133</sup>. La fréquence de l'allèle à risque (rs1613662-A) était estimée à 84%. Le polymorphisme entraîne la substitution d'une sérine par une proline en position 219. Il se situe sur le chromosome 19 dans un locus de gènes codant pour des immunorécepteurs de la superfamille des immunoglobulines. Le polymorphisme entraînerait la diminution de l'activation plaquettaire médiée par le collagène <sup>134</sup>. Il jouerait ainsi un rôle dans la constitution du thrombus veineux. Toutefois, les preuves de cette contribution au phénotype thrombotique sont mitigées. En effet, une association entre le polymorphisme rs1613662-A et la MTEV a été mise en évidence dans la GWAS de Trégouët et al. <sup>135</sup>, mais n'a pas été reproduite dans l'étude de réplcation d'Austin et al. <sup>132</sup>.

*F11* code pour le facteur XI (607 AA). Il est situé sur le chromosome 4q34-35. En 2000 a été mis en évidence pour la première fois un lien entre les taux de facteur XI et la MTEV. En



effet, il existait une augmentation du risque de MTEV estimée à 2,2 (IC95% = 1,5-3,2) chez les patients de la cohorte LETS qui présentaient des taux de facteurs XI supérieurs au 90<sup>ème</sup> percentile <sup>136</sup>. Le facteur XI, dont la demi-vie est de 24-48 heures, est synthétisé par le foie. Il circule sous la forme d'un homodimère non activé, complexé par des liaisons non covalentes au kininogène de haut poids moléculaire. Chaque monomère est activé après clivage des résidus Arg369-Ile370 par ses activateurs, la thrombine, le facteur XIIIa et par lui-même une fois activé (facteur XIa). Ce clivage génère une chaîne légère qui contient le site catalytique et une chaîne lourde qui contient 4 séquences répétées en tandem appelées domaines «apple». Ces dernières sont nécessaires au processus de dimérisation initial. In vivo, le facteur XI est essentiel pour le maintien et la propagation de la coagulation. Il permet d'amplifier la génération de thrombine et de diminuer la fibrinolyse. Plusieurs études ont mis en évidence une association entre la MTEV et les polymorphismes rs3756008, rs2289252 et rs2036914 <sup>123,127,129</sup>, situés sur *F11*. Ces 2 derniers polymorphismes sont en déséquilibre de liaison avec rs10029715, lui-même associé à la MTEV dans une récente GWAS <sup>137</sup>. Cette même étude a permis de décrire l'organisation en haplotypes des polymorphismes rs925451, rs10029715, rs1008728 et rs13133050 situés sur *F11*. L'haplotype ATTC, était associé à une augmentation du risque de MTEV estimé à 1,218 alors que l'haplotype GCCA était protecteur (OR = 0,493).

En 2015, une méta-analyse de 12 GWAS a identifié deux nouveaux facteurs de risque génétiques de MTEV sans lien évident avec la coagulation. Il s'agit des polymorphismes rs2288904 et rs78707713 situés respectivement sur *SLC44A2* et *TSPAN15*. Le premier était associé à un OR de 1,28, et le second à un OR de 1,42 après répliation dans trois études indépendantes <sup>125</sup>.

*SLC44A2*, situé sur le chromosome 19, code pour CTL-2, un transporteur de choline transmembranaire, essentiellement connu pour son rôle dans le TRALI (*Transfusion-related acute lung injury*) ou syndrome de détresse respiratoire post-transfusionnel <sup>138</sup>. La fréquence de l'allèle à risque (rs2288904-G) est estimée à 79%. Ce polymorphisme correspond à la substitution d'une adénine en position 461 par une guanine et ainsi apparition d'un résidu arginine en position 154 à l'origine de l'expression d'un épitope particulier appelé HNA-3a. Cet épitope jouerait un rôle essentiel dans la physiopathologie du TRALI. Son implication dans la physiopathologie de la MTEV reste à découvrir.

*TSPAN15*, situé sur le chromosome 10, code pour la tétraspanine 15, une protéine de signalisation appartenant à la superfamille des tétraspanines. La fréquence de l'allèle à risque (rs78707713-A) est estimée à 88%. Il s'agit d'une variation intronique. La tétraspanine 15 interagit spécifiquement avec l'ADAM10 et permet sa progression depuis le réticulum endoplasmique jusqu'à sa forme activée vers la surface cellulaire<sup>139</sup>. Or, l'ADAM10 clive la glycoprotéine VI (GpVI) qui est impliquée dans l'activation, l'adhésion et l'agrégation plaquettaire<sup>140</sup>. Le rôle de *TSPAN15* dans la MTEV n'est cependant pas clairement établi.

### *3. Découverte de nouveaux variants par l'approche génome entier ou GWAS appliquée aux phénotypes intermédiaires*

Les méthodologies utilisées au cours des 10 dernières années ont permis d'identifier des variants de plus en plus fréquents et associés à des effets de plus en plus faibles. Pour cela, les études d'association génome entier n'ont eu de cesse d'augmenter leurs effectifs. Malgré cette stratégie purement mathématique, l'héritabilité de la MTEV demeure largement inexpliquée. De nouvelles stratégies ont ainsi été utilisées pour permettre l'identification de nouveaux variants associés à la MTEV à moindre « coût-patients ». L'utilisation de phénotypes intermédiaires préalablement définis comme impliqués dans la physiopathologie de la MTEV est un moyen d'augmenter la puissance de détection de nouveaux gènes de susceptibilité.

Les premiers résultats émanant de cette stratégie concernaient l'étude des taux plasmatiques de facteur Willebrand<sup>141</sup>. Les résultats ainsi obtenus ont ensuite été confirmés dans une étude cas-témoins<sup>142</sup>. Dans cette dernière étude, il existait une association entre l'allèle rs1063856-G, situé sur *VWF*, et l'augmentation des taux de facteur Willebrand ainsi que le risque de MTEV (OR = 1,15) et entre l'allèle rs1039084-G, situé sur *STXBP5*, et la diminution des taux plasmatiques de facteur Willebrand ainsi que le risque de MTEV (OR de 0,90).

*VWF*, situé sur le chromosome 12, code pour un pré-pro-facteur Willebrand (2813 AA) synthétisé dans les cellules endothéliales et les mégacaryocytes. Il subit plusieurs processus de modifications structurelles. Il est stocké sous forme de grands multimères dans les corps

de Weibel-Palade des cellules endothéliales et dans les granules plaquettaires. Il circule dans le plasma sous forme de multimères de poids moléculaires différents obtenus par le clivage d'une métalloprotéase, l'ADAMTS13. Le facteur Willebrand est sécrété vers la lumière vasculaire et vers le sous endothélium, en partie au moment de sa synthèse (voie constitutive) et en partie en réponse à des substances sécrétagogues (voie régulée). Le facteur Willebrand stocké dans les plaquettes n'intervient dans le processus de coagulation qu'après activation plaquettaire et sécrétion. Il est constitué de plusieurs domaines nécessaires à son implication dans l'hémostase. Il protège et transporte le facteur VIII par ses 272 premiers AA, se lie au collagène et aux molécules adhésives par ses domaines A1 et A3 et interagit avec les glycolipides sulfatés et la Gplb $\alpha$  plaquettaire par son domaine A1 et avec l' $\alpha$ IIb $\beta$ 3 et l' $\alpha$ V $\beta$ 3 par sa séquence RGD. Lorsque l'endothélium est lésé, le facteur Willebrand ralentit les plaquettes par interaction avec la Gplb $\alpha$ , permettant une adhésion irréversible des plaquettes.

*STXBP5*, situé sur le chromosome 6, code pour la *syntaxin binding protein 5* (1151 AA). Son rôle dans le processus thrombotique n'est pas aussi évident que celui du facteur Willebrand. Deux études récentes ont mis en évidence une association entre l'allèle rs1039084-G (protecteur), et la majoration des saignements chez les femmes présentant une maladie de Willebrand de type 1<sup>143</sup> ainsi que la diminution des taux de t-PA<sup>144</sup>. Le lien entre *STXBP5* et hémostase est conforté par d'autres études. *STXBP5* favoriserait en effet la libération des granules plaquettaires et inhiberait la sécrétion du contenu des corps de Weibel-Palade des cellules endothéliales<sup>145,146</sup>.

*KNG1* code pour le kininogène de haut poids moléculaire (KHPM). L'approche génome entier appliquée aux phénotypes intermédiaires a mis en évidence son association à un raccourcissement du TCA. Le TCA est un test global de la coagulation dont le raccourcissement a été décrit comme associé au risque de MTEV dans plusieurs études<sup>147,148</sup>. En 2010 une étude d'association génome-entier a démontré une association du polymorphisme rs710446, situé sur *KNG1*, avec une diminution du TCA<sup>149</sup>. Il existait par ailleurs une association entre le polymorphisme et la MTEV dans plusieurs cohortes (OR voisin de 1,18)<sup>150</sup>. Physiologiquement, le KHPM permet l'initiation de la voie intrinsèque en

rapprochant la prékallicréine et le facteur XI du facteur XII <sup>151</sup>. Son implication dans la formation du thrombus veineux est donc aisément envisageable.

Le tableau 2 présente les principaux polymorphismes décrits comme associés à la MTEV <sup>152</sup>.

Tableau 2. Facteurs de risque génétiques associés à la maladie thromboembolique veineuse.

| Locus           | Polymorphisme | Localisation génomique | Allèle <sup>1</sup> | Fréquence de l'allèle à risque | Risque estimé | Degré de certitude <sup>2</sup> |
|-----------------|---------------|------------------------|---------------------|--------------------------------|---------------|---------------------------------|
| <i>ABO</i>      | rs579459      | Intron                 | T/C                 | 0,22                           | 1,50          | +++                             |
|                 | rs8176749     | Leu310Leu              | G/A                 | 0,10                           | 1,50          | +++                             |
| <i>F2</i>       | rs1799963     | 3'UTR                  | G/A                 | 0,008                          | 2,29          | +++                             |
|                 | rs3136516     | Intron                 | G/A                 | 0,40                           | 1,12          | ++                              |
| <i>F5</i>       | rs6025        | Arg506Gln              | G/A                 | 0,03                           | 3,25          | +++                             |
|                 | rs4524        | Lys858Arg              | A/G                 | 0,75                           | 1,20          | +++                             |
| <i>F8</i>       | rs114209171   | Séquence régulatrice   | T/C                 | 0,80                           | 1,11          | ++                              |
| <i>F11</i>      | rs2289252     | Intron                 | C/T                 | 0,41                           | 1,35          | ++                              |
|                 | rs2036914     | Intron                 | C/T                 | 0,60                           | 1,35          | ++                              |
| <i>FGG</i>      | rs2066865     | 3'UTR                  | C/T                 | 0,25                           | 1,25          | +++                             |
|                 | rs2227421     | 3'UTR                  | A/C                 | 0,34                           | 1,13          | +                               |
| <i>KNG1</i>     | rs7100446     | Ile581Thr              | Ile/Thr             | 0,40                           | 1,20          | +++                             |
| <i>PROCR</i>    | rs6088735     | Séquence régulatrice   | C/T                 | 0,23                           | 1,11          | ++                              |
|                 | rs867186      | Ser219Gly              | Ser/Gly             | 0,10                           | 1,22          | +++                             |
| <i>PROS1</i>    | rs121918472   | Ser501Pro              | Ser/Pro             | 0,002                          | 6,57          | +++                             |
|                 | rs121918474   | Lys196Glu              | Lys/Glu             | 0,009                          | 5,00          | +++                             |
| <i>SERPINC1</i> | rs2227624     | Val30Glu               | Val/Glu             | 0,004                          | 2,30          | +++                             |
|                 | rs2227589     | Intron                 | C/T                 | 0,10                           | 1,05          | ++                              |
| <i>SLC44A2</i>  | rs2288904     | Arg154Gln              | Arg/Gln             | 0,76                           | 1,28          | ++                              |
| <i>STXBP5</i>   | rs1039084     | Asn436Ser              | A/G                 | 0,55                           | 1,11          | ++                              |
| <i>TSPAN15</i>  | rs78707713    | Intron                 | T/C                 | 0,90                           | 1,42          | +                               |
| <i>VWF</i>      | rs1063856     | Thr789Ala              | The/Ala             | 0,40                           | 1,20          | +++                             |

Les mutations privées, notamment situées sur les gènes *SERPINC1*, *PROC* et *PROS1*, ne sont pas listées. <sup>1</sup>Allèles communs/mineurs. En gras l'allèle associé au risque de MTEV. <sup>2</sup>+++ Variants répondant aux trois critères suivants : preuve statistique définitive –fonctionnalité établie – rôle physiopathologique établi ; ++ variants répondants à deux critères sur trois ; + variants répondants à un critère sur trois.

## II. Problématique

Les anomalies du bilan de thrombophilie sont généralement recherchées chez deux types de patients. La première indication concerne les sujets jeunes ayant présenté un épisode de MTEV en l'absence de facteur déclenchant. Il s'agit alors d'évaluer le risque de récurrence afin d'identifier les patients qui bénéficieront d'un traitement anticoagulant au long cours. Au demeurant, les facteurs biologiques de risque recherchés par le bilan de thrombophilie sont faiblement associés au risque de récurrence<sup>153</sup> et la stratégie qui reposerait sur la réalisation d'un bilan de thrombophilie pour décider de la poursuite du traitement anticoagulant n'a pas fait ses preuves<sup>154</sup>. Ainsi, le résultat du bilan de thrombophilie ne doit pas influencer la durée du traitement anticoagulant d'après les différentes recommandations<sup>155</sup>. La seconde indication concerne les apparentés, lorsque le bilan pratiqué dans la première catégorie de patients est positif. Dans ce cas, l'idée est d'évaluer le risque de premier épisode de MTEV chez des sujets encore indemnes. L'identification d'un facteur biologique de risque permet alors de proposer une prise en charge adaptée en situation à risque (immobilisation, chirurgie, post-partum...). Cependant, cette stratégie, bien que largement partagée, apparaît beaucoup trop simpliste. En effet, le résultat du dépistage de l'anomalie familiale ne permet pas de classer de façon manichéenne les apparentés : le risque n'est pas nul (ou en tout cas comparable à la population générale) chez les apparentés à bilan négatif, et maximal lorsque l'anomalie est retrouvée. Dans le premier cas, les apparentés présentent un risque de MTEV multiplié par 2 ou 3 comparativement à la population générale. Dans le second, une proportion variable (qui dépend de l'anomalie biologique) de sujets présentera un jour un épisode thrombotique. Celle-ci est estimée à 50% avant l'âge de 50 ans pour les déficits en inhibiteurs et à 10% environ pour les mutations du FVL ou PTG20210A à l'état hétérozygote. On comprend donc aisément que le bilan de thrombophilie seul ne permet pas d'évaluer précisément le risque de MTEV à l'échelle individuelle dans ces familles. L'utilité du bilan de thrombophilie est ainsi débattue par les spécialistes<sup>155,156</sup>. Depuis la description de PTG20210A en 1996, de nombreux polymorphismes ont été décrits comme associés à la MTEV. Ces polymorphismes confèrent une augmentation modérée du risque, bien inférieure aux facteurs génétiques conventionnels. A ce titre, ils ne font pas partie du bilan de thrombophilie actuellement reconnu. Au demeurant, ils pourraient contribuer à mieux caractériser le risque individuel dans ces familles. Peu d'études ont entrepris ce travail. De

même, alors que la MTEV est une maladie complexe dont la part environnementale est majeure, le risque associé à des facteurs tels que le tabagisme ou l'obésité est mal évalué dans les familles avec bilan de thrombophilie positif. L'article présenté ci-après, en cours de soumission et intitulé « Venous thromboembolism risk assessment in 2214 relatives from 651 families with known thrombophilia defects » tente ainsi d'apporter une réponse à ces questionnements.

L'identification des sujets à risque de MTEV ne permet à ce jour que de prévenir la survenue d'un épisode thrombotique en situation à risque. En effet, la faible prévalence de la pathologie ne permet pas d'envisager une prévention primaire avec la prescription d'une anticoagulation au long cours chez des sujets certes à risque, mais encore asymptomatiques. Parmi les situations à risque, outre les situations transitoires principalement représentées par l'immobilisation et la chirurgie, la grossesse et le post-partum exposent les femmes à un risque relatif majeur pour une période durable. Ainsi, la question du dépistage des facteurs biologiques de risque se pose principalement chez les femmes en âge de procréer. Ces dernières sont par ailleurs exposées à un facteur de risque environnemental extrêmement fréquent, les contraceptifs oraux combinés. La France compte plus de 4 millions d'utilisatrices. Dans certaines tranches d'âge, près de 80% des femmes utilisent cette méthode contraceptive. Alors que la présence d'une anomalie biologique (parmi les 5 explorées par le bilan de thrombophilie) contre-indique de façon absolue l'utilisation des contraceptifs oraux combinés, d'après les recommandations françaises et internationales, la majorité des utilisatrices ne présenteront jamais de MTEV. Une récente revue de la littérature a même conduit ses auteurs à conclure à la possible utilisation des contraceptifs oraux combinés chez les femmes présentant une thrombophilie modérée en l'absence de facteur de risque supplémentaire <sup>157</sup>. Il apparaît donc nécessaire de disposer d'outils complémentaires au bilan de thrombophilie pour améliorer l'identification des patientes à risque. Nous avons ainsi transposé la démarche qui consistait à intégrer des facteurs de risque fréquents à effet faible (environnementaux et génétiques) dans l'évaluation du risque individuel de MTEV chez les apparentés issus de familles avec facteur biologique de risque, dans un groupe particulier de femmes sous contraceptifs oraux combinés.

# Résultats



**Article 3 (en cours de soumission)**  
**“Common risk factors add to  
thrombophilia defects to predict venous  
thromboembolism risk in families”**

## **Common risk factors add to thrombophilia defects to predict venous thromboembolism risk in families**

Short title: Common risk factors in families with thrombophilia

Pierre Suchon,<sup>1,2</sup> Noémie Resseguier,<sup>3,4</sup> Manal Ibrahim,<sup>1,2</sup> Alexia Robin,<sup>1</sup> Geoffroy Venton,<sup>5,6</sup> Marie-Christine Barthet,<sup>1</sup> Dominique Brunet,<sup>1</sup> Noémie Saut,<sup>1</sup> Marie-Christine Alessi,<sup>1,2</sup> David-Alexandre Trégouët,<sup>7,8\*</sup> Pierre Emmanuel Morange<sup>1,2\*</sup>

<sup>1</sup>Laboratory of Haematology, La Timone Hospital, Marseille, 13385, France

<sup>2</sup>Institut National pour la Santé et la Recherche Médicale (INSERM), Unité Mixte de Recherche en Santé (UMR\_S) 1062, Nutrition Obesity and Risk of Thrombosis, Aix-Marseille University, Marseille, 13385, France

<sup>3</sup>Support Unit for clinical research and economic evaluation, Assistance Publique - Hôpitaux de Marseille, Marseille, 13385, France

<sup>4</sup>EA3279 Self-perceived Health Assessment Research Unit, Aix-Marseille University, Marseille, 13385, France

<sup>5</sup>Aix-Marseille Université, INSERM, UMR1090, TAGC Campus de Luminy, Marseille, F\_13288, France

<sup>6</sup>Hematology and cellular therapy department, AP-HM, Conception Hospital, Marseille, 13385, France

<sup>7</sup>Sorbonne Universités, UPMC Univ. Paris 06, INSERM, UMR\_S 1166, Team Genomics & Pathophysiology of Cardiovascular Diseases, Paris, France

<sup>8</sup>ICAN Institute for Cardiometabolism and Nutrition, Paris, France

\* Equal contribution to the present work

### **Corresponding author:**

Pierre Emmanuel Morange

Service d'hématologie biologique, Hôpital de la Timone

264 rue Saint Pierre

13005 MARSEILLE

pierre.morange@ap-hm.fr

Tel: +33491386048; Fax: +33491389155

Text word count: 3821

Abstract word count: 250

Number of tables: 3

Number of figures: 1

Number of references: 30

#### KEY POINTS

- Venous thromboembolism incidence in relatives from families with thrombophilia is highly modulated by common risk factors
- Smoking is associated with a higher venous thromboembolism risk in families with thrombophilia than in the general population

## ABSTRACT

Clinical assessment of venous thromboembolism (VTE) risk mainly relies on five biomarkers: antithrombin, protein C and S deficiencies, factor V Leiden and prothombin mutation. However, the clinical VTE pattern often shows wide heterogeneity within relatives of a VTE affected family despite they carry the same thrombophilia defect. This study aims to assess whether common environmental and genetic risk factors for VTE contribute to explain this heterogeneity. A thrombophilia screening, ABO blood group determination and the genotyping of 11 common polymorphisms known to associate with VTE in the general population were performed in 2,214 relatives from 651 families with multiple affected VTE individuals. According to the severity of the thrombophilia defect, individuals were split into 3 groups: VTE incidence was 1.74, 3.64 and 6.40 per 1,000 person-years (p-y) in individuals with no defect, mild thrombophilia and severe thrombophilia respectively. In addition, the main common risk factors modulating VTE incidence were obesity, smoking, ABO blood group and *F11\_rs2036914* and *FGG\_rs2066865* polymorphisms. These common factors were then included into a risk score allowing a better stratification of at-risk individuals. Three categories were then defined to associate with low, medium and high risk, with incidence rate of 1.30, 2.94 and 4.95 per 1,000 p-y respectively. The score displayed high efficiency in assessing VTE risk in mild thrombophilia patients as it identified 2 groups of individuals harboring the same risk as no defect and severe thrombophilia individuals respectively. Integrating common environmental and genetic risk factors improved VTE risk assessment in relatives from families with thrombophilia.

## INTRODUCTION

Venous thromboembolism (VTE) including deep vein thrombosis (DVT) and pulmonary embolism (PE) is the third cause of cardiovascular death<sup>1</sup>. VTE incidence widely depends on age. It dramatically increases after 60. It is then considered as a disease of the elderly. When a VTE episode occurs in younger patients, in the absence of triggering situation, two underlying situations are usually investigated: a neoplasm and/or a genetic thrombophilia. The genetic part plays an important role as the heritability of VTE has been assessed to be 50-60%<sup>2-4</sup>. In clinical practice, when an inherited disorder is suspected a thrombophilia testing consisting of the search for the 5 following defects is performed: antithrombin (AT), protein C (PC) and S (PS) deficiencies, factor V Leiden (FVL) and the G20210A prothrombin mutation (PTM)<sup>5</sup>. This testing is positive in about 40% of VTE patients<sup>6</sup>. The 5 defects could be classified into 2 groups: severe thrombophilia including the AT, PC and PS deficiencies, homozygous for FVL or PTM and combined defects associated with a high risk of VTE (relative risk (RR) x 10) and mild thrombophilia including FVL or PTM heterozygosity associated with a moderate increased risk of VTE (RR x 3-5). When a defect is diagnosed in a patient with a personal history of VTE, testing of family members is often performed to aid decision making regarding future VTE prophylaxis. However, usefulness of thrombophilia testing in assessing individual risk in families with inherited thrombophilia is largely debated as it does not properly assess the individual risk within a family<sup>7,8</sup>. Indeed, the test results do not clearly dichotomize carriers and non-carriers in terms of their risk for a first episode of VTE.

The assessment of VTE risk in families with known inherited thrombophilia needs to be improved by acquiring more insight into genetic and environmental risk factors. However, the thrombophilia testing has remained the same since the identification of the PTM in 1996<sup>9</sup>. Other factors may then be taken into consideration when assessing individual risk. These could include both additional genetic risk factors much more common than those involved in the thrombophilia testing but also environmental factors, the two mains being obesity and smoking. Obesity, is a well-established risk factor for VTE<sup>10,11</sup> whereas smoking has been debated for many years. Recently, two large meta-analyses have been performed in the general population and demonstrated that smoking is associated with VTE<sup>12,13</sup>. Assessment of obesity and smoking on the risk of VTE in families with known thrombophilia has never been performed yet.

The aim of the present study is to study the influence of smoking, obesity and newly identified common genetic factors on VTE risk in families with known thrombophilia with ultimate goal of improving VTE risk assessment.

## **METHODS**

### **Patients**

Investigated families were part of the MARseilles FAmily Study on venous Thrombosis (MARFAST). The original MARFAST cohort has been extensively described previously <sup>14</sup>. Briefly, families were recruited between 1986 to 2008 at the referral center for thrombophilia in Marseilles, France. A family was included in the MARFAST cohort if it comprised at least two members with the following characteristics: one individual with a personal history of confirmed VTE and a positive thrombophilia testing (the proband) and a relative harboring the same defect. A VTE episode was considered to be confirmed if objectively diagnosed by medical imaging: compression ultrasound, venography, ventilation/perfusion lung scan, spiral computed tomography or pulmonary angiography; or if the patient received full-dose anticoagulation for at least 3 months. All family members underwent a thrombophilia testing.

Patients were classified into three classes according to the severity of the thrombophilia: no thrombophilia, mild thrombophilia (FVL heterozygosity or PTM heterozygosity) and severe thrombophilia (AT, PC, PS deficiencies, FVL homozygosity, PTM homozygosity and combined defects).

The present study is an extension of the MARFAST cohort. The inclusions, using the same criteria, have been extended until 2013 at the same center. This MARFAST update included a total of 2,214 relatives (including 1,492 first degree relatives) from 651 families representing 151 more families and 940 more relatives than the previous analysis.

### **Variables**

A standardized questionnaire allowed the collection of clinical data during the consultation by a trained physician. The age, localization and the triggering circumstances of VTE episodes were collected. A VTE episode was considered provoked if associated with an exposure within 3 months to exogenous risk factors: surgery, immobilization for at least 7 days, trauma, combined oral contraceptive use, pregnancy, puerperium, malignancy. Otherwise the episode was defined as unprovoked. The body mass index (BMI) was calculated during the

consultation. Relatives were classified into 3 classes of BMI:  $<30 \text{ kg.m}^{-2}$ ,  $30\text{-}35 \text{ kg.m}^{-2}$  and  $>35 \text{ kg.m}^{-2}$  based on the WHO classification of obesity. The smoking status was determined either at the time of the episode for VTE relatives or at the time of the consultation for relatives with no personal history of VTE.

A thrombophilia testing was systematically performed in all included relatives: AT, PC and PS levels were measured as previously described<sup>14</sup> and FVL and PTM were genotyped using light cycler technology (Roche Diagnostics, Indianapolis, INUSA). ABO blood group was genotyped and classified into three categories according to the level of risk previously estimated: A/B, AB and O.

The previous MARFAST investigation mainly relied on well-established risk factors for VTE and measurable coagulation-related biomarkers. Only 3 single nucleotide polymorphisms (SNPs) (rs2036914 and rs2289252 located on *F11* and rs2066865 located on *FGG*) were previously tested. We here investigated the impact of 8 additional SNPs whose association with VTE was supported by a recent genome wide association studies (GWAS): *F5\_rs4524*, *F2\_rs3136516*, *PROCR\_rs867186* and rs2069951, *SERPINC1\_rs2227589*, *SLC44A2\_rs2288904*, *KNG1\_710446* and *TSPAN15\_rs78707713*.

### **Statistical analysis**

Probands were excluded for the statistical analysis. First, a descriptive analysis was performed. Quantitative variables were described using means and standard deviations, and categorical variables were described using numbers and percentages. Comparisons of major clinical characteristics between VTE+ and VTE- groups were performed using a naïve approach: chi-square test or Fisher test when appropriate for categorical variables, and using Student's t test or Mann-Whitney test when appropriate for quantitative variables.

The incidence rate of VTE was estimated by dividing the number of VTE by the sum of the observation time for each participant. The observation time was defined as the period from birth until the first episode of VTE for those who presented a VTE, or until the age at the visit for those who did not present a VTE. The incidence rate was expressed as number per 1,000 person-years (p-y).

Univariate time-to-event analysis was first performed to assess the association between the risk of VTE over time and clinical and biological characteristics. A Cox regression model was used, including a frailty term to handle the family structure of the data (family identifier as a random effect).

Multivariate analyses were then performed to simultaneously integrate all characteristics that were associated with the risk of VTE with a p-value <0.15 in the univariate analysis. A multivariate frailty Cox model was built using a backward selection procedure to retain significant variables (p<.05), and estimate adjusted hazard ratios (HR) with their 95% confidence interval (CI). The proportional hazards assumption was assessed by testing covariate interactions with quadratic function of time and checked graphically using Schoenfeld-type residuals. According to the value of the adjusted parameters (log hazard ratios), a score of VTE risk was established by attributing 0, 1 or 2 points for each modality of the characteristics included in the multivariate model. The incidence rate of VTE and its 95% CI was estimated according to this score to assess its validity.

All tests were two-sided. P-values <.05 were considered significant. All analyses were performed with the R software.

All participants gave written informed consent, and the study met all institutional ethics requirement. The procedures employed were reviewed by the *Assistance Publique des Hopitaux de Marseille* institutional review committee.

All authors had full access to all of the data and take responsibility for the integrity of the data and the accuracy of the data analysis.

## **RESULTS**

More than 60% of relatives had a positive thrombophilia testing, from whom 43.8% and 17.2% harbored a mild and a severe thrombophilia respectively (table S1). As expected, FVL was the most prevalent thrombophilia defect with a prevalence of 30.8% in relatives.

Main characteristics of the relatives according to VTE status are shown in table 1. Two hundred and forty-six (11.1%) presented a VTE history. The mean age at the end of the follow-up was 41.7 and 31.9 years ( $p < 10^{-4}$ ) for VTE and VTE free relatives respectively.

### **Univariate analysis**

#### *Environmental risk factors*

Grade 1 obesity defined as a BMI between 30 and 35 and grade 2 obesity defined as a BMI above 35 were associated with an increased risk of VTE. The incidence rate of VTE was 3.03 per 1,000 p-y for non-obese relatives (BMI<30 Kg.m<sup>-2</sup>), 6.38 per 1,000 p-y for relatives with



grade 1 obesity and 6.05 per 1,000 p-y for relatives with grade 2 obesity. As grade 1 and grade 2 incidences were comparable, the two categories were merged into one category of obese relatives ( $BMI > 30 \text{ Kg.m}^{-2}$ ) for further analyses. The overall incidence rate for obesity was thus 6.30 per 1,000 p-y (95% CI 4.63-8.36).

The incidence rate was 3.19 per 1,000 p-y (95% CI 2.74-3.69) and 4.00 per 1,000 p-y (95% CI 3.11-5.07) for non-smokers and smokers.

#### *Conventional genetic risk factors*

A thrombophilia was diagnosed in 76.1% of VTE+ relatives whereas only 59% of VTE- relatives harbored a thrombophilia, the difference holding mainly in severe thrombophilia (30.5% vs 15.5%). The incidence rate of VTE increased with the severity of the thrombophilia defect: 1.74, 3.64 and 6.40 per 1,000 p-y in relatives with no defect, mild and severe thrombophilia respectively.

The incidence rate of VTE greatly varied according to the blood group, the highest incidence rate being observed for AB blood group relatives. Incidence rates were 2.29 per 1,000 p-y (95% CI 1.71-2.99), 3.58 per 1,000 p-y (95% CI 3.03-4.19) and 6.15 per 1,000 p-y (95% CI 3.98-9.06) for non-O, A/B and AB relatives respectively.

#### *Genetic polymorphisms recently identified*

Table 2 shows the results of the genotyping of the 11 selected polymorphisms. Three SNPs, *F11\_rs2036914* ( $p = 0.003$ ), *FGG\_rs2066865* ( $p = 0.015$ ) and *F11\_rs2289252* ( $p = 0.036$ ) were significantly associated with VTE.

#### **Multivariate analysis**

The multivariate analysis comprised the thrombophilia status, BMI analyzed in two classes, smoking status, ABO blood group, *FGG\_rs2066865*, *F11\_rs2036914*, *F11\_rs2289252* and a systematic adjustment for gender was decided. The results are shown in table 3.

In the multivariate model, the highest HR was observed for severe thrombophilia (HR=3.78, 95% CI 2.49-5.73). Among common risk factors, blood group AB and current smoking were associated with the highest levels of risk (HR= 2.51 (95% CI 1.50-4.22) and 2.08 (95% CI 1.51-2.86) respectively). Obesity was associated with a 1.78-fold increase of VTE (95% CI 1.24-2.55). Only two SNPs remained significant in the multivariate model, *FGG\_rs2066865* and *F11\_rs2036914*. The presence of at least one at risk allele was associated with an increased risk of VTE: HR= 1.55 (95% CI 1.06-2.26) and 1.61 (95% 1.21-2.13) for *F11\_rs2036914* ( $p = 0.001$ ) and *FGG\_rs2066865* ( $p = 0.023$ ) respectively.

A risk stratification was then performed taking into account the 5 common risk factors for VTE: *F11\_rs2036914*, *FGG\_rs2066865*, ABO blood group, BMI and smoking status. Observed HRs for the different components of the score (except for ABO blood group) being equally important, a simplified scoring method was applied: relatives were scored 0 or 1 point for each parameter but ABO. The scoring system was slightly different for ABO blood group because of the variety of ABO profiles and their associated HR: 0 for O blood group, 1 for A or B blood group and 2 for AB blood group. After calculating the overall score, theoretically spanning from 0 to 6, relatives were classified into 3 groups: 0-1 point; 2 points; 3 points or more. This stratification was chosen because of its ability to homogeneously divide relatives in terms of numbers. The VTE incidence was calculated in the 3 categories of risk in all relatives or according to the thrombophilia status (figure 1/ table S2).

A trend toward an increased incidence rate was observed according to the score: 1.30, 2.94 and 4.95 per 1,000 p-y in low, medium and high risk relatives respectively ( $p= 1.06 \cdot 10^{-8}$ ) (figure 1/ table S2). This observation held in all thrombophilia categories ( $p= 0.04$ ,  $1.78 \cdot 10^{-5}$  and  $0.002$  for no defect, mild and severe thrombophilia respectively). The combination of the information on this risk score together with the result of the thrombophilia testing led to an annual incidence rate ranging from 0.60 per 1,000 p-y for the combination of no defect and low risk score (score= 0 or 1) to 9.85 per 1,000 p-y for the combination of severe thrombophilia and high risk score (score $\geq$ 3) (figure 1/ table S2). It is noteworthy that the observed incidences in the mild thrombophilia group were quite heterogeneous in terms of risk. On the one hand, relatives with combination of mild thrombophilia/low risk score had a 0.91 per 1,000 p-y incidence which was comparable to the combination of no thrombophilia/low risk score (0.60 per 1,000 p-y). On the other hand, relatives with combinations of mild thrombophilia/high risk score and severe thrombophilia/medium risk score displayed very similar incidences (4.98 and 4.83 per 1,000 p-y respectively).

We then performed an overall disease-free survival analysis taking into account both score and thrombophilia status, indicating that the observed effects were independent. Medium and high risk scores were associated with HR of 2.31 (1.35-3.97) and 4.23 (2.54-7.04) respectively (global p-value  $<.0001$ ) as compared with low risk score. Similarly, mild and severe thrombophilia were associated with HR of 2.02 (1.38-2.95) and 3.88 (2.58-5.87) respectively (global p-value=  $<.0001$ ) as compared with no defect.

## DISCUSSION

It is well established that obesity and smoking modulate VTE risk in the general population. Although, few studies have assessed their impact in families with inherited thrombophilia. In the present study, we demonstrated that both common environmental and genetic factors modified the risk of VTE in those families.

Two major issues remain in the assessment of VTE risk in relatives of probands with known thrombophilia: relatives with no defect harbor a 2 to 3-fold increased risk<sup>15-17</sup> of VTE compared to the general population and the vast majority of patients with a mild thrombophilia won't ever undergo a VTE episode. This suggests that considering the familial defect testing only for assessing the individual risk in relatives is probably coarse and would necessitate more insight. The assessment of VTE risk in families with inherited thrombophilia needs to be improved. This could be done by assessing the impact of common environmental and genetic risk factors. Although VTE is a complex disease, its environmental part (at least the impact of obesity and smoking which are the 2 conventional arterial risk factors robustly associated with VTE) hasn't been assessed in families with heritable thrombophilia. To the best of our knowledge our study reports for the first time the annual incidence associated with environmental risk factors such as obesity and smoking in families with thrombophilia. Obesity and smoking appeared to be strong predictors for VTE with associated HR ~2. Obesity, is a well-established risk factor for VTE<sup>10,11</sup> whereas the relation between smoking and VTE has been debated for many years. Indeed, most of studies described a modest or a null effect of smoking on the risk of VTE.<sup>18-23</sup> Some rare studies even observed a protective effect in smokers.<sup>24,25</sup> During the past five years, two large meta-analyses have been performed in the general population and reached the same conclusion that smoking is associated with VTE.<sup>12,13</sup> Compared with never smokers, the RR for current smokers was found to be ~1.2 both in the Cheng and Mahmoodi studies respectively<sup>12,13</sup>. We here confirmed the association between current smoking and increased VTE risk, with a slightly higher risk (~2.1) observed in our thrombophilia families. Very interestingly, the risk associated with those two environmental and thus modifiable factors was comparable to the risk associated with the presence of a mild inherited thrombophilia (HR~ 1.9).

Among common genetic risk factors, we confirmed the previous results observed in the MARFAST study<sup>26</sup>. In the present extended MARFAST study, ABO blood group, *F11\_rs2036914* and *FGG\_rs2066865* were associated with HR ranging from 1.5 for *F11\_rs2036914* to 2.5 for

blood group AB. Interestingly, *F11\_rs2036914* and *FGG\_rs2066865* are located in genes related to the coagulation cascade. They were previously reported to associate with plasma levels of the corresponding factors: the factor XI<sup>27</sup> and the fibrinogen, respectively<sup>28</sup>. Of note, the latter actually associated with a subunit of fibrinogen, the fibrinogen gamma', which low levels were further demonstrated to associate with venous thrombosis. This opens the question of the clinical interest of measuring factor XI and fibrinogen gamma' in families with thrombophilia.

To improve the assessment of VTE risk in families with inherited thrombophilia we propose here an effective additional tool based on the 5 common risk factors associated with VTE within the MARFAST families: smoking, obesity, ABO blood group, *F11\_rs2036914* and *FGG\_rs2066865*. Those 5 factors were then combined into a predictive score which allowed the better stratification of the risk in relatives. Three categories of relatives were individualized after calculating the number of risk factors. Relatives with 0 or 1 risk factor were considered at low risk, relatives with 2 risk factors harbored medium risk and relatives with more than 2 risk factors displayed a high risk. When considering the thrombophilia status only, the incidence rate was 1.74 per 1,000 p-y, 3.64 per 1,000 p-y and 6.40 per 1,000 p-y for no defect, mild thrombophilia and severe thrombophilia groups respectively. When applying the stratification in relatives after the thrombophilia testing, the incidence rate appeared to widely depend on the score. In no defect relatives, the global incidence rate was 1.74 per 1,000 p-y. However, the incidence ranged from 0.60 to 2.62 per 1,000 p-y according to the stratification. This stratification seemed to identify a group (23% of relatives with no defect) with a risk similar to that of the general population (0.60 per 1,000 p-y versus 1 per 1,000 p-y)<sup>29,30</sup>. This observation held in the mild thrombophilia group in which the global annual incidence rate was 3.64 and ranged from 0.91 to 4.98 according to the stratification. In this group, 21% of the relatives harbored a low risk after stratification and an incidence rate comparable to the combination of no defect and low risk score (and to the general population). Of note, relatives presenting a mild thrombophilia and a high risk score according to the stratification displayed a high incidence rate (4.98 per 1,000 p-y). Very interestingly, this latter group of relatives seemed to present the same risk as relatives harboring a severe thrombophilia and a medium risk score: incidence rate being 4.98 per 1,000 p-y and 4.83 per 1,000 p-y respectively. Similarly, relatives with mild thrombophilia and medium risk score and relatives with severe thrombophilia and low risk score displayed a comparable incidence rate

(3.47 per 1,000 p-y and 3.44 per 1,000 p-y respectively). Of note, even if relatives with severe thrombophilia displayed a great amplitude of risk (incidence rate ranging from 3.44 to 9.85 per 1,000 p-y) according to the score, this stratification couldn't discriminate a group of relatives at low risk. As a consequence, those relatives with severe thrombophilia should be considered at high risk and benefit from prevention in at risk situations. We further showed that the proposed score based on common risk factors was independent of the result of the thrombophilia testing suggesting that it adds some valuable information for assessing the risk in relatives from families with known thrombophilia.

To the best of our knowledge, the present study is the largest family study on inherited thrombophilia published so far. However, this study shows some limits. In first instance, missing data led to a loss of power. Indeed, only 197 relatives with VTE could have been tested out of the 242 because of missing data, predominant on BMI (21 relatives). This lack of power could explain the absence of association between SNPs recently identified in GWASs and VTE in the families. Furthermore, we used cross-sectional data, and the use of prospective data would have been more relevant for risk estimates.

In conclusion, we have demonstrated the impact of smoking and obesity of the risk of VTE in families with inherited thrombophilia. Taking into account common risk factors (both environmental and genetics) for VTE in families with inherited thrombophilia could be very helpful for assessing individual risk, per se, or in addition to the thrombophilia testing.

**Acknowledgements:** This work was supported by a Programme Hospitalier de Recherche Clinique (PHRC 2003).

**Authorship Contributions:** PS, AR, MCA, DAT and PEM conceived the study. PS, NR, GV, MI, AR, MCB, DB and NS collected and analyzed the data. PS, DAT, MCA, NR and PEM wrote the first draft of the paper and all authors revised it critically for important intellectual content.

**Disclosure of Conflicts of Interest:** All authors declare no competing financial interests.

## References

1. Di Nisio M, van Es N, Büller HR. Deep vein thrombosis and pulmonary embolism. *Lancet Lond. Engl.* 2016;388(10063):3060–3073.
2. Souto JC, Almasy L, Borrell M, et al. Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study. Genetic Analysis of Idiopathic Thrombophilia. *Am. J. Hum. Genet.* 2000;67(6):1452–1459.
3. Heit JA, Phelps MA, Ward SA, et al. Familial segregation of venous thromboembolism. *J. Thromb. Haemost. JTH.* 2004;2(5):731–736.
4. Larsen TB, Sørensen HT, Skytthe A, et al. Major genetic susceptibility for venous thromboembolism in men: a study of Danish twins. *Epidemiol. Camb. Mass.* 2003;14(3):328–332.
5. Morange P-E, Suchon P, Trégouët D-A. Genetics of Venous Thrombosis: update in 2015. *Thromb. Haemost.* 2015;114(5):910–919.
6. De Stefano V, Rossi E. Testing for inherited thrombophilia and consequences for antithrombotic prophylaxis in patients with venous thromboembolism and their relatives. A review of the Guidelines from Scientific Societies and Working Groups. *Thromb. Haemost.* 2013;110(4):697–705.
7. Middeldorp S. Inherited thrombophilia: a double-edged sword. *Hematol. Am. Soc. Hematol. Educ. Program.* 2016;2016(1):1–9.
8. Connors JM. Thrombophilia Testing and Venous Thrombosis. *N. Engl. J. Med.* 2017;377(12):1177–1187.
9. Poort SR, Rosendaal FR, Reitsma PH, Bertina RM. 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):3698–3703.
10. Cushman M. Epidemiology and Risk Factors for Venous Thrombosis. *Semin. Hematol.* 2007;44(2):62–69.
11. Allman-Farinelli M. Obesity and Venous Thrombosis: A Review. *Semin. Thromb. Hemost.* 2011;37(08):903–907.
12. Cheng Y-J, Liu Z-H, Yao F-J, et al. Current and former smoking and risk for venous thromboembolism: a systematic review and meta-analysis. *PLoS Med.* 2013;10(9):e1001515.

13. Mahmoodi BK, Cushman M, Anne Næss I, et al. Association of Traditional Cardiovascular Risk Factors With Venous Thromboembolism: An Individual Participant Data Meta-Analysis of Prospective Studies. *Circulation*. 2017;135(1):7–16.
14. Cohen W, Castelli C, Alessi M-C, et al. ABO blood group and von Willebrand factor levels partially explained the incomplete penetrance of congenital thrombophilia. *Arterioscler. Thromb. Vasc. Biol*. 2012;32(8):2021–2028.
15. Bezemer ID, van der Meer FJM, Eikenboom JCJ, Rosendaal FR, Doggen CJM. The value of family history as a risk indicator for venous thrombosis. *Arch. Intern. Med*. 2009;169(6):610–615.
16. Zöller B, Ohlsson H, Sundquist J, Sundquist K. Familial risk of venous thromboembolism in first-, second- and third-degree relatives: a nationwide family study in Sweden. *Thromb. Haemost*. 2013;109(3):458–463.
17. Sørensen HT, Riis AH, Diaz LJ, et al. Familial risk of venous thromboembolism: a nationwide cohort study. *J. Thromb. Haemost. JTH*. 2011;9(2):320–324.
18. Goldhaber SZ, Grodstein F, Stampfer MJ, et al. A prospective study of risk factors for pulmonary embolism in women. *JAMA*. 1997;277(8):642–645.
19. Klatsky AL, Armstrong MA, Poggi J. Risk of pulmonary embolism and/or deep venous thrombosis in Asian-Americans. *Am. J. Cardiol*. 2000;85(11):1334–1337.
20. Lindqvist PG, Epstein E, Olsson H. The relationship between lifestyle factors and venous thromboembolism among women: a report from the MISS study. *Br. J. Haematol*. 2009;144(2):234–240.
21. Hippisley-Cox J, Coupland C. Development and validation of risk prediction algorithm (QThrombosis) to estimate future risk of venous thromboembolism: prospective cohort study. *BMJ*. 2011;343:d4656.
22. Wattanakit K, Lutsey PL, Bell EJ, et al. Association between cardiovascular disease risk factors and occurrence of venous thromboembolism. A time-dependent analysis. *Thromb. Haemost*. 2012;108(3):508–515.
23. Sweetland S, Parkin L, Balkwill A, et al. Smoking, surgery, and venous thromboembolism risk in women: United Kingdom cohort study. *Circulation*. 2013;127(12):1276–1282.
24. Ray JG, Langman LJ, Vermeulen MJ, et al. Genetics University of Toronto Thrombophilia Study in Women (GUTTSI): genetic and other risk factors for venous thromboembolism in women. *Curr. Control. Trials Cardiovasc. Med*. 2001;2(3):141–149.

25. Yamada N, Ota S, Liu Y, et al. Risk factors for nonfatal pulmonary embolism in a Japanese population: A hospital-based case-control study. *Angiology*. 2010;61(3):269–274.
26. Cohen W, Castelli C, Suchon P, et al. Risk assessment of venous thrombosis in families with known hereditary thrombophilia: the MARseilles-Nimes prediction model. *J. Thromb. Haemost. JTH*. 2014;12(2):138–146.
27. Li Y, Bezemer ID, Rowland CM, et al. Genetic variants associated with deep vein thrombosis: the F11 locus. *J. Thromb. Haemost. JTH*. 2009;7(11):1802–1808.
28. Uitte de Willige S, de Visser MCH, Houwing-Duistermaat JJ, et al. Genetic variation in the fibrinogen gamma gene increases the risk for deep venous thrombosis by reducing plasma fibrinogen gamma' levels. *Blood*. 2005;106(13):4176–4183.
29. Silverstein MD, Heit JA, Mohr DN, et al. Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. *Arch. Intern. Med.* 1998;158(6):585–593.
30. Naess IA, Christiansen SC, Romundstad P, et al. Incidence and mortality of venous thrombosis: a population-based study. *J. Thromb. Haemost. JTH*. 2007;5(4):692–699.



**Table 1. Main characteristics of the relatives according to the VTE status.**

| Variable                       | VTE +<br>(n=246) | VTE -<br>(n=1968) | p                 |
|--------------------------------|------------------|-------------------|-------------------|
| Female sex                     | 153 (62.2)       | 189 (60.4)        | 0.59              |
| Mean follow-up (SD)            | 41.7 (16.8)      | 31.9 (17.3)       | <10 <sup>-4</sup> |
| Thrombophilia screening        |                  |                   |                   |
| No defect                      | 58 (23.6)        | 807 (41.0)        | <10 <sup>-4</sup> |
| Moderate thrombophilia         | 113 (45.6)       | 856 (43.5)        |                   |
| Severe thrombophilia           | 75 (30.5)        | 305 (15.5)        |                   |
| ABO blood group                |                  |                   |                   |
| A or B                         | 154 (66.4)       | 1134 (60.0)       | <10 <sup>-4</sup> |
| AB                             | 25 (10.8)        | 99 (5.2)          |                   |
| O                              | 53 (22.8)        | 657 (34.8)        |                   |
| BMI <30 kg.m <sup>-2</sup>     | 178 (79.1)       | 1621 (93.0)       | <10 <sup>-4</sup> |
| BMI = 30-35 kg.m <sup>-2</sup> | 36 (16.0)        | 91 (5.2)          |                   |
| BMI >35 kg.m <sup>-2</sup>     | 11 (4.9)         | 31 (1.8)          |                   |
| Current smoking                | 68 (27.6)        | 453 (23.1)        | 0.11              |
| First VTE episode=PE (+/- DVT) | 53 (21.5)        | -                 | -                 |

**Table 2. Genotyping of 11 selected SNPs reported to associate with VTE.**

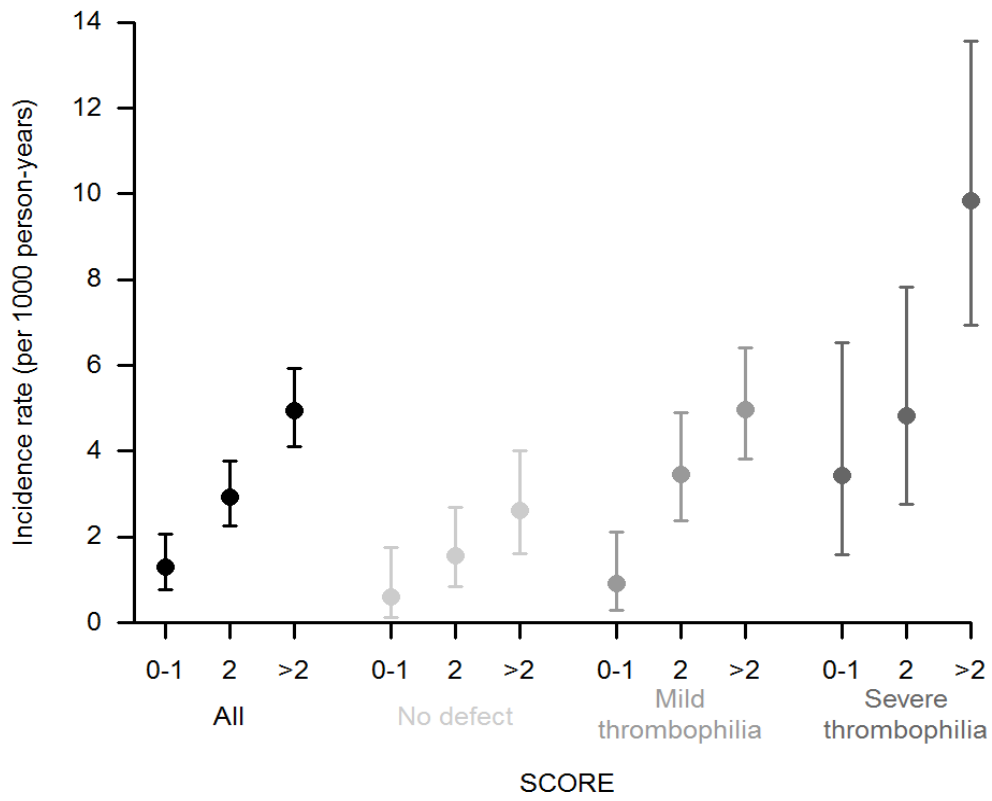
SNP= Single nucleotide polymorphism. VTE= venous thromboembolism. p= p value calculated using a Chi<sup>2</sup> test.

| SNP                                  | Risk allele | Genotype | VTE + | VTE - | P     |
|--------------------------------------|-------------|----------|-------|-------|-------|
| <b><i>FGG</i><br/>rs2066865</b>      | T           | CC       | 102   | 1005  | 0.015 |
|                                      |             | CT-TT    | 112   | 777   |       |
| <b><i>F11</i><br/>rs2036914</b>      | C           | TT       | 37    | 470   | 0.003 |
|                                      |             | CT-CC    | 178   | 1310  |       |
| <b><i>F11</i><br/>rs2289252</b>      | T           | CC       | 48    | 522   | 0.036 |
|                                      |             | CT-TT    | 162   | 1230  |       |
| <b><i>F5</i><br/>rs4524</b>          | A           | GG       | 5     | 65    | 0.307 |
|                                      |             | GA-AA    | 214   | 1729  |       |
| <b><i>SLC44A2</i><br/>rs2288904</b>  | G           | AA       | 12    | 103   | 0.882 |
|                                      |             | AG-GG    | 204   | 1671  |       |
| <b><i>TSPAN15</i><br/>rs78707713</b> | T           | CC       | 0     | 22    | 0.161 |
|                                      |             | CT-TT    | 224   | 1799  |       |
| <b><i>PROCR</i><br/>rs867186</b>     | G           | AA       | 180   | 1478  | 0.881 |
|                                      |             | AG-GG    | 38    | 321   |       |
| <b><i>F2</i><br/>rs3136516</b>       | G           | AA       | 86    | 628   | 0.211 |
|                                      |             | AG-GG    | 131   | 1150  |       |
| <b><i>PROCR</i><br/>rs2069951</b>    | A           | GG       | 191   | 1585  | 0.886 |
|                                      |             | GA-AA    | 26    | 209   |       |
| <b><i>SerpinC1</i><br/>rs2227589</b> | A           | GG       | 154   | 1364  | 0.491 |
|                                      |             | GA-AA    | 44    | 344   |       |
| <b><i>KNKI</i><br/>rs710446</b>      | T           | CC       | 41    | 357   | 0.702 |
|                                      |             | CT-TT    | 183   | 1486  |       |

**Table 3. Disease-free survival multivariate analysis with adjustment for age and sex and attributed score.**

| Variables                    | n (%)       | Hazard ratio<br>(95% CI) | p-value     | Attributed<br>score |
|------------------------------|-------------|--------------------------|-------------|---------------------|
| No thrombophilia             | 647 (37.4)  | 1                        | -           | -                   |
| Mild thrombophilia           | 794 (45.9)  | 1.91 (1.30-2.80)         | $9.10^{-4}$ | -                   |
| Severe thrombophilia         | 290 (16.8)  | 3.78 (2.50-5.73)         | $<10^{-4}$  | -                   |
| BMI < 30kg.m <sup>-2</sup>   | 1589 (91.8) | 1                        | -           | 0                   |
| BMI > 30kg.m <sup>-2</sup>   | 142 (8.2)   | 1.78 (1.24-2.55)         | 0.002       | 1                   |
| No smoking                   | 1306 (75.4) | 1                        | -           | 0                   |
| Current smoking              | 425 (24.6)  | 2.08 (1.51-2.86)         | $<10^{-4}$  | 1                   |
| Blood group O                | 565 (32.6)  | 1                        | -           | 0                   |
| Blood group A or B           | 1068 (61.7) | 1.37 (0.98-1.93)         | 0.067       | 1                   |
| Blood group AB               | 98 (5.7)    | 2.51 (1.50-4.22)         | $5.10^{-3}$ | 2                   |
| <i>FII</i> rs2036914 – TT    | 418 (24.1)  | 1                        |             | 0                   |
| <i>FII</i> rs2036914 – CT/CC | 1313 (75.9) | 1.55 (1.06-2.26)         | 0.023       | 1                   |
| <i>FGG</i> rs2066865 – CC    | 957 (55.3)  | 1                        | -           | 0                   |
| <i>FGG</i> rs2066865 – CT/TT | 774 (44.7)  | 1.61 (1.21-2.13)         | 0.001       | 1                   |

**Figure 1: VTE incidence according to the score and the thrombophilia status.**



**Table S1. Results of the thrombophilia screening in relatives.** <sup>1</sup>Factor V Leiden. <sup>2</sup>G20210A prothrombin mutation. <sup>3</sup>FVL heterozygous or PTG20210A heterozygous. <sup>4</sup>Antithrombin, protein C, protein S deficiencies, FVL homozygous, PTG202101 homozygous, combined defects.

| Thrombophilia defect              | Relatives, n (%) |
|-----------------------------------|------------------|
| FVL <sup>1</sup>                  | 681 (30.8)       |
| PTG20210A <sup>2</sup>            | 320 (14.5)       |
| Antithrombin deficiency           | 30 (1.4)         |
| Protein C deficiency              | 157 (7.1)        |
| Protein S deficiency              | 79 (3.6)         |
| Combined defects                  | 82 (3.7)         |
| No defect                         | 865 (39.1)       |
| Mild thrombophilia <sup>3</sup>   | 969 (43.8)       |
| Severe thrombophilia <sup>4</sup> | 380 (17.2)       |

**Table S2. VTE incidence in relatives according to the score and the thrombophilia status.**<sup>1</sup>Per 1000 person-years. HR= Hazard ratio. 95% CI= 95% confidence interval.

| Population             | Number of events | Follow-up (years) | Incidence <sup>1</sup> (95% CI) | HR (95% CI)       | p-value |
|------------------------|------------------|-------------------|---------------------------------|-------------------|---------|
| All                    | 197              | 57918             | 3.40                            |                   |         |
| Score=0-1 (n=387)      | 17               | 13116             | 1.30 (0.76-2.07)                | 1                 | <.0001  |
| Score=2 (n=629)        | 61               | 20783             | 2.94 (2.25-3.77)                | 2.23 (1.30-3.85)  |         |
| Score≥3 (n=715)        | 119              | 24019             | 4.95 (4.11-5.93)                | 4.09 (2.44-6.84)  |         |
| No defect              | 37               | 21291             | 1.74                            | -                 | -       |
| Score = 0-1 (n=150)    | 3                | 5010              | 0.60 (0.12-1.75)                | 1                 | 0.0202  |
| Score = 2 (n=252)      | 13               | 8255              | 1.57 (0.84-2.69)                | 2.57 (0.73-9.05)  |         |
| Score≥ 3 (n=245)       | 21               | 8026              | 2.62 (1.62-4.00)                | 4.26 (1.26-14.27) |         |
| Moderate thrombophilia | 98               | 26944             | 3.64                            | -                 | -       |
| Score = 0-1 (n=164)    | 5                | 5492              | 0.91(0.30-2.12)                 | 1                 | <.0001  |
| Score = 2 (n=276)      | 32               | 9214              | 3.47 (2/38-4.90)                | 4.11 (1.55-10.89) |         |
| Score ≥ 3 (n=354)      | 61               | 12238             | 4.98 (3.81-6.40)                | 7.23 (2.81-18.61) |         |
| Severe thrombophilia   | 62               | 9683              | 6.40                            | -                 | -       |
| Score=0-1 (n=73)       | 9                | 2614              | 3.44 (1.58-6.53)                | 1                 |         |
| Score=2 (n=101)        | 16               | 3314              | 4.83 (2.76-7.83)                | 1.24 (0.54-2.86)  | 0.0007  |
| Score≥3 (n=116)        | 37               | 3755              | 9.85 (6.95-13.56)               | 3.01 (1.45-6.26)  |         |

## **Article 4**

**“Genetic risk factors for venous thrombosis in women using combined oral contraceptives: update of the PILGRIM study”**

*Clinical genetics 2017*



## Short Report

# Genetic risk factors for venous thrombosis in women using combined oral contraceptives: update of the PILGRIM study

Suchon P., Al Frouh F., Ibrahim M., Sarlon G., Venton G., Alessi M.-C., Trégouët D.-A., Morange P.-E. Genetic risk factors for venous thrombosis in women using combined oral contraceptives: update of the PILGRIM study. Clin Genet 2017; 91: 131–136. © John Wiley & Sons A/S. Published by John Wiley & Sons Ltd, 2016

Identifying women at risk of venous thrombosis (VT) under combined oral contraceptives (COC) is a major public health issue. The aim of this study was to investigate in COC users the impact on disease of genetic polymorphisms recently identified to associate with VT risk in the general population. Nine polymorphisms located on *KNG1*, *F11*, *F5*, *F2*, *PROCR*, *FGG*, *TSPAN* and *SLC44A2* genes were genotyped in a sample of 766 patients and 464 controls as part of the PILGRIM (PIL Genetic Risk Monitoring) study. Cases were women who experienced an episode of documented VT during COC use, while controls were women with no history of VT using COC at the time of inclusion. Among the studied polymorphisms, only *F11* rs2289252 was significantly associated with VT. The *F11* rs2289252-A allele was associated with a 1.6-fold increased risk of VT ( $p < 0.0001$ ). Besides, the combination of the rs2289252-A allele with non-O blood group, present in 52% of the cohort, was associated with an odds ratio of 4.00 (2.49–6.47;  $p < 10^{-4}$ ). The consideration of this genetic risk factor could help to better assess the risk of VT in COC users.

### Conflict of interest

The authors declare no conflict of interest.

**P. Suchon<sup>a,b</sup>, F. Al Frouh<sup>b</sup>,  
M. Ibrahim<sup>a,b</sup>, G. Sarlon<sup>c</sup>,  
G. Venton<sup>d,e</sup>, M.-C. Alessi<sup>a,b</sup>,  
D.-A. Trégouët<sup>f,g</sup> and P.-E.  
Morange<sup>a,b</sup>**

<sup>a</sup>Aix Marseille Univ, INSERM, INRA, NORT, Marseille, France, <sup>b</sup>APHM, Hôpital de la Timone, Service d'hématologie biologique, Marseille, France, <sup>c</sup>Médecine vasculaire et Hypertension artérielle, Faculté de Médecine de Marseille, Aix-Marseille Université, Assistance Publique Hôpitaux de Marseille – Hôpital de La Timone, Marseille, France, <sup>d</sup>INSERM, UMR1090 TAGC, Aix-Marseille University, Marseille, France, <sup>e</sup>Service d'hématologie, APHM, Hôpital de la Conception, Marseille, France, <sup>f</sup>Sorbonne Universités, UPMC Univ Paris 06, INSERM UMR\_S 1166, Paris, France, and <sup>g</sup>ICAN Institute for Cardiometabolism and Nutrition, Paris, France

Key words: case control studies – genetic polymorphisms – oral contraceptives – risk assessment – venous thrombosis

### Corresponding author:

Pierre-Emmanuel Morange, Laboratory of Hematology, CHU Timone, 264, Rue Saint-Pierre, 13385 Marseille Cedex 05, France.

Tel.: +33 4 91 38 60 49;

fax: +33 4 91 94 23 32;

e-mail: pierre.morange@ap-hm.fr

Received 3 May 2016, revised and accepted for publication 11 July 2016

In women, the overall incidence of venous thrombosis (VT) is about 1.2 per 1000 person-years (1); this incidence being lower in women of reproductive age, with an estimate ranging between 0.4 and 0.6 per 1000 person-years (1). The use of combined oral contraceptive (COC) is a major established risk factor for VT in women of reproductive age. The relative risk for COC

use is about 3–6. It is well known that this risk depends on two major factors: the progestogen and the ethinyl estradiol dosage (2). It is estimated that more than 100 million women worldwide are using COC, making the impact of COC use on VT risk and public health of major importance (3). In France, 20 young women die annually of pulmonary embolism (PE) as a consequence



of COC use (4). Furthermore, young and middle-aged women with a personal history of VT have a 2.3-fold increased risk of mortality (5). Besides, among survivors, 25–50% will have lasting debilitating health problems such as post-thrombotic syndrome, severely hampering mobility and quality of life (6). Chronic thromboembolic pulmonary hypertension occurs in about 3% PE patients (7).

Oral contraceptives and inherited thrombophilic defects (i.e. factor V Leiden and prothrombin mutations, deficiencies of protein C, protein S or antithrombin) are known to interact synergistically to increase the risk of VT (8). Very recently, we have found that the ABO blood group, an increasingly recognized risk factor for VT in the general population, was also associated with the risk of VT in COC users. In the PILGRIM (Pill Genetic Risk Monitoring) study, we observed that, under COC use, women of non-O blood group displayed a 1.85-fold increased risk of VT compared with O blood group carriers (9).

Last year, the International VENous Thrombosis (INVENT) consortium presented the results of the first international meta-analysis of genome-wide association studies for VT in the general population (10). This project led to the identification of two unsuspected susceptibility loci for VT (*TSPAN15* and *SLC44A2*) and confirmed the strong associations with VT of several previously reported genes (*F5*, *F2*, *FGG*, *F11* and *PROCR*). The aim of this study was to assess the impact of the lead common single nucleotide polymorphisms (SNPs) of these VT-associated genes on VT risk in COC users.

## Materials and methods

PILGRIM is a case–control study including child-bearing age women referred to a single center [Exploration Centre of hemorrhagic and thrombotic diseases (CEHT) of the Marseille University Hospital] between 1 January 2003 and 31 December 2013. Cases were all consecutive women who had a first objectively confirmed (i.e. evidenced by imaging or when the patient had received a full-dose anticoagulant therapy for at least 3 months) episode of VT (including PE and proximal or distal deep vein thrombosis of the lower limbs) during COC use, with no history of cancer or autoimmune disease. Controls were defined as consecutive women using COC, who were referred to CEHT for thrombophilia screening as they presented a family history of thrombosis but with no personal history of the disease. Women using first generation pills were excluded. Pills were categorized according to the generation of the progestogen. The whole PILGRIM cohort comprises 968 cases and 874 controls and is extensively described by Suchon et al. (9)

Stored DNA was available in 766 (79.1%) cases and 464 (53.1%) controls who participated in this study. Although there are fewer individuals in this study, their general characteristics were comparable to those of the whole PILGRIM study. These PILGRIM participants were genotyped for nine SNPs reported by the

INVENT consortium to associate with VT risk, including the lead SNPs at two novel susceptibility genes, *SLC44A2* rs2288904 and *TSPAN15* rs78707713, and seven SNPs at known VT-associated genes that achieved strong statistical association, *F5* rs4524, *F2* rs3136516, *KNG1* rs710446, *F11* rs2289252, *F11* rs2036914, *FGG* rs2066865 and *PROCR* rs867186 (10). Genotyping was performed using real-time polymerase chain reaction.

For each studied SNP, deviation from Hardy–Weinberg model was tested in controls using a chi-squared test with one degree of freedom. Association with VT risk was tested by use of the Cochran–Armitage trend test. Allelic odds ratios (ORs) and 95% confidence intervals (CIs) were also computed using logistic regression analysis, adjusting for age, type and duration of COC use, and family history score (calculated by dividing the number of first-degree relatives with a personal history of VT by the total number of first-degree relatives) (9). A statistical threshold of  $6 \times 10^{-3}$  ( $\sim 0.05/9$ ), correcting for the number of tested polymorphisms, was used to declare statistical significance for the observed SNP association with disease risk. All statistical analyses were performed using R.

We obtained the informed consent from all participants, and the study met all institutional ethics requirements. The procedures employed were reviewed and approved by the Assistance Publique des Hopitaux de Marseille institutional review committee.

## Results

Main characteristics of the population study are shown in Table 1 and are comparable to the characteristics of the whole PILGRIM cohort previously reported. In particular, the four main risk factors for VT risk we previously identified in the whole PILGRIM study (9), which were smoking, body mass index (BMI) greater than 35 kg/m<sup>2</sup> (class II obesity and above according to the World Health Organization classification of obesity), severe inherited thrombophilia and non-O blood groups, were consistently associated with the disease in the current PILGRIM participants. The corresponding ORs for VT were 1.65 [95% CI (1.30–2.10)], 3.46 (1.81–7.03), 2.13 (1.32–3.51) and 1.98 (1.57–2.49), respectively (data not shown).

Genotype and allele distributions of the nine studied SNPs in cases and controls were provided in Table 2. All SNPs were in Hardy–Weinberg equilibrium. Only the *F11* rs2289252 showed statistical significant association with VT risk. The rs2289252-A allele was significantly more frequent in cases than in controls (49% vs 42%) and associated with an adjusted 1.60-fold (1.33–1.94) ( $p < 10^{-4}$ ) increased risk of VT. The magnitude of its effect was similar to that reported in the general population (OR = 1.35) (11). The OR for VT associated with the rs2289252-A allele was similar in smoker and non-smoker women (OR = 1.62 vs OR = 1.63;  $p = 0.96$ ) and in O and non-O carrier women (OR = 1.80 vs OR = 1.50;  $p = 0.45$ ). In addition, the allele frequency of

## Genetic risk factors for venous thrombosis in oral contraceptive users

Table 1. Main characteristics of the population study<sup>a</sup>

|   | VT patients (n = 766) | Controls (n = 464) | p value           |
|---|-----------------------|--------------------|-------------------|
| Mean (SD) age (years)                                   | 32.2 (9.5)            | 30.5 (8.8)         | 0.0008            |
| Oral contraceptive use by generation                    |                       |                    |                   |
| Second <sup>b</sup>                                     | 262 (34.2)            | 173 (37.3)         | <10 <sup>-4</sup> |
| Third <sup>c</sup>                                      | 315 (41.1)            | 135 (29.1)         |                   |
| Fourth <sup>d</sup>                                     | 96 (12.5)             | 70 (15.1)          |                   |
| Cyproterone acetate                                     | 93 (12.1)             | 86 (18.5)          |                   |
| Mean duration of COC in years (SD) <sup>e</sup>         | 8.4 (7.4)             | 7.5 (5.5)          | 0.02              |
| Duration of use ≤12 months                              | 106 (14.8)            | 20 (4.7)           | <10 <sup>-4</sup> |
| Positive first-degree family history of VT <sup>f</sup> | 213 (27.8)            | 271 (58.4)         | <10 <sup>-4</sup> |
| Mean family history score (SD)                          | 0.07 (0.13)           | 0.17 (0.17)        | <10 <sup>-4</sup> |
| Mean (SD) body mass index (BMI) <sup>g</sup>            | 23.9 (5.1)            | 22.1 (3.7)         | <10 <sup>-4</sup> |
| BMI ≥ 30 kg/m <sup>2</sup>                              | 95 (12.5)             | 18 (4.0)           | <10 <sup>-4</sup> |
| BMI ≥ 35 kg/m <sup>2</sup>                              | 36 (4.7)              | 5 (1.1)            | 0.0007            |
| Current smokers <sup>h</sup>                            | 254 (33.2)            | 103 (22.2)         | <10 <sup>-4</sup> |
| Severe thrombophilia <sup>i</sup>                       | 52 (6.8)              | 22 (4.7)           | 0.14              |
| ABO blood group   |                       |                    |                   |
| A   | 440 (57.4)            | 229 (49.4)         | <10 <sup>-4</sup> |
| B   | 93 (12.1)             | 46 (9.9)           |                   |
| AB  | 54 (7.0)              | 22 (4.7)           |                   |
| O   | 179 (23.4)            | 167 (36.0)         |                   |
| Non-O   | 587 (76.6)            | 297 (64.0)         | <10 <sup>-4</sup> |

COC, combined oral contraceptives; VT, venous thrombosis.

<sup>a</sup>Shown data are counts (%) unless otherwise indicated.

<sup>b</sup>Second generation: levonorgestrel, norgestrel.

<sup>c</sup>Third generation: desogestrel, gestodene, norgestimate.

<sup>d</sup>Fourth generation: drospirenone, dienogest, nomegestrol.

<sup>e</sup>Data available for 714 cases and 429 controls.

<sup>f</sup>Only first-degree relatives were considered.

<sup>g</sup>Data available for 758 cases and 450 controls.

<sup>h</sup>At time of VT episode for cases, at time of inclusion for controls.

<sup>i</sup>Protein C, protein S and antithrombin deficiencies, homozygous for factor V Leiden, homozygous for G20210A prothrombin mutation, or combined defects.

the rs2289252-A allele was similar according to the presence or absence of family history of VT ( $p > 0.37$ , regardless of the definition used: the existence of a first-degree family history of VT or a family history score above 0.3) (9).

As a consequence of these results, women carrying both the *FII* rs2289252-A allele and non-O blood group, representing about 52% of the women population, presented a 4.00-fold (2.49–6.47) ( $p < 10^{-4}$ ) increased risk compared with women of O blood group and rs2289252-GG genotype. We further stratified this analysis according to obesity, one of the main environmental risk factor for VT in COC users, to better assess the risk associated with those two common genetic risk factors for VT. Because of the small number of patients in the highest class of obesity, we merged the top two classes: 30–35 kg/m<sup>2</sup> and >35 kg/m<sup>2</sup>. Adjusted OR estimating the VT risk associated with different combinations of these three risk factors, rs2289252, ABO blood group and obesity, are presented in Table 3. Women with BMI over 30 kg/m<sup>2</sup> (class I obese and above) of non-O blood group that harbored the rs2289252-A allele were at 13.15-fold (5.20–38.50) ( $p < 10^{-4}$ ) increased risk compared with non-obese women with O blood group and

rs2289252-GG genotype, the former group representing 4.8% of the female population.

### Discussion

Efforts to promote a better identification of women at high-risk of COC-related VT are invaluable, as they may lead to more adequate prescription of contraception, with less or no thrombogenic influence and, perhaps, a decrease of these preventable VT events.

In this study, we found that the *FII* rs2289252 is a mild and common genetic risk factor for VT in COC users. The allelic OR associated with the polymorphism was 1.60. The risk allele was common in the population, 64% in the control group, and its impact on VT risk was independent on that of the non-O blood group, another common genetic risk factor. As a consequence, the combination of these two genetic factors that occurred together in the general population at a prevalence of ~52% was associated with an ~fourfold increased risk for VT.

Eight other established VT-susceptibility alleles were investigated in our study, but failed to reach statistical significance. A nominal association ( $p = 0.02$ ) was observed for the rs78707713 whose rare allele tend

Table 2. Association of studied SNPs with the risk of venous thrombosis in combined oral contraceptive users

|                               |    | Controls  | Cases     | p <sup>a</sup>            | OR <sup>b</sup>  | p                 |
|-------------------------------|----|-----------|-----------|---------------------------|------------------|-------------------|
| rs4524 ( <i>F5</i> )          | AA | 272 (59%) | 462 (60%) | p = 0.528                 | 1                | 0.36              |
|                               | AG | 166 (36%) | 265 (35%) |                           | 0.90 (0.73–1.12) |                   |
|                               | GG | 26 (5%)   | 39 (5%)   |                           |                  |                   |
| RAF <sup>c</sup>              |    | 0.23      | 0.22      |                           |                  |                   |
| rs2289252 ( <i>F11</i> )      | GG | 166 (36%) | 195 (25%) | p = 1.84 10 <sup>-4</sup> | 1                | <10 <sup>-4</sup> |
|                               | AG | 231 (46%) | 387 (51%) |                           | 1.60 (1.33–1.94) |                   |
|                               | AA | 85 (18%)  | 184 (24%) |                           |                  |                   |
| RAF                           |    | 0.42      | 0.49      |                           |                  |                   |
| rs2036914 ( <i>F11</i> )      | GG | 136 (29%) | 266 (35%) | p = 0.037                 | 1                | 0.02              |
|                               | GA | 230 (50%) | 363 (47%) |                           | 1.24 (1.04–1.49) |                   |
|                               | AA | 98 (21%)  | 137 (18%) |                           |                  |                   |
| RAF                           |    | 0.46      | 0.42      |                           |                  |                   |
| rs2066865 ( <i>FGG</i> )      | GG | 272 (59%) | 425 (56%) | p = 0.648                 | 1                | 0.38              |
|                               | AG | 159 (34%) | 298 (39%) |                           | 1.10 (0.89–1.35) |                   |
|                               | AA | 33 (7%)   | 43 (5%)   |                           |                  |                   |
| RAF                           |    | 0.24      | 0.25      |                           |                  |                   |
| rs867186 ( <i>PROCR</i> )     | AA | 361 (78%) | 611 (80%) | p = 0.612                 | 1                | 0.38              |
|                               | AG | 98 (21%)  | 142 (18%) |                           | 0.88 (0.67–1.17) |                   |
|                               | GG | 5 (1%)    | 13 (2%)   |                           |                  |                   |
| RAF                           |    | 0.12      | 0.11      |                           |                  |                   |
| rs3136516 ( <i>F2</i> )       | AA | 117 (25%) | 193 (25%) | p = 0.345                 | 1                | 0.43              |
|                               | AG | 219 (47%) | 392 (51%) |                           | 0.93 (0.78–1.11) |                   |
|                               | GG | 128 (28%) | 181 (24%) |                           |                  |                   |
| RAF                           |    | 0.51      | 0.49      |                           |                  |                   |
| rs710446 ( <i>KNG1</i> )      | AA | 153 (33%) | 215 (28%) | p = 0.070                 | 1                | 0.06              |
|                               | AG | 227 (49%) | 393 (51%) |                           | 1.19 (0.99–1.43) |                   |
|                               | GG | 84 (18%)  | 158 (21%) |                           |                  |                   |
| RAF                           |    | 0.43      | 0.46      |                           |                  |                   |
| rs2288904 ( <i>SLC44A2</i> )  | AA | 18 (4%)   | 25 (3%)   | p = 0.231                 | 1                | 0.08              |
|                               | CG | 160 (35%) | 243 (32%) |                           | 1.23 (0.98–1.54) |                   |
|                               | GG | 285 (62%) | 495 (65%) |                           |                  |                   |
| RAF                           |    | 0.79      | 0.81      |                           |                  |                   |
| rs78707713 ( <i>TSPAN15</i> ) | CC | 1 (<1%)   | 0 (0%)    | p = 0.022                 | 1                | 0.06              |
|                               | CT | 84 (18%)  | 104 (14%) |                           | 1.40 (0.99–1.97) |                   |
|                               | TT | 378 (82%) | 658 (86%) |                           |                  |                   |
| RAF                           |    | 0.91      | 0.93      |                           |                  |                   |

CI, confidence interval; COC, combined oral contraceptives; OR, odds ratio; SNP, single nucleotide polymorphisms.

<sup>a</sup>Cochran Armitage Trend test p value.

<sup>b</sup>Allelic odds ratio (95% CI) adjusted for age, type of COC, duration of COC use and family history score calculated by logistic regression analysis.

<sup>c</sup>Allele frequency of the reported risk allele.

to associate with increased risk of disease as initially reported (11). The lack of statistical significance was probably because of some power issues, as our sample size was not large enough to detect genetic effects characterized by moderate impact on VT risk. As an illustration, the power of the PILGRIM study to detect allele ORs as low as 1.4 was only 47% for a polymorphism with a minor allele frequency of 0.10 (which corresponds to the observed association at *TSPAN15* rs78707713) at the statistical threshold of  $6 \times 10^{-3}$ . However, we cannot exclude that the genetic associations reported in the general population (10) do not hold in specific groups of subjects such as COC users.

This study underlined the role of combination of common genetic risk factors on the risk of VT in COC users and confirmed the complexity of the disease arising from

the occurrence of multiple risk factors. Family history, besides being a recognized risk factor for VT, is poorly associated with known genetic risk factors and thus cannot be considered as a surrogate for the search of these genetic factors. Thrombophilia screening is not recommended in most of the international recommendations in part because defects are rare in the general population (12). Conversely, the combination of the non-O blood group with the rs2289252-A allele is far from being rare as it accounts for 52% of COC users. Moreover, in our cohort, the magnitude of risk provided by this combination of at-risk allele is similar to that of class II obesity (BMI higher than 35 kg/m<sup>2</sup>), which is a relative contraindication for COC use in some international recommendations. Very interestingly, the combination of one demographic (partly genetic and partly environmental)

## Genetic risk factors for venous thrombosis in oral contraceptive users

Table 3. Impact on VT risk of obesity, ABO blood group and *F11* rs2289252

| BMI (kg/m <sup>2</sup> ) | ABO blood group | <i>F11</i> rs2289252 <sup>a</sup> | Cases, <i>n</i> (%) | Controls, <i>n</i> (%) | OR (95% CI) <sup>b</sup> | <i>p</i>          |
|--------------------------|-----------------|-----------------------------------|---------------------|------------------------|--------------------------|-------------------|
| <25                      | O               | GG                                | 39 (5.5)            | 46 (10.9)              | 1                        |                   |
| 25–30                    | O               | GG                                | 4 (0.6)             | 6 (1.4)                | 0.99 (0.22–4.03)         | 0.98              |
| >30                      | O               | GG                                | 0                   | 1 (0.2)                | NA                       | NA                |
| <25                      | Non-O           | GG                                | 89 (12.6)           | 81 (19.2)              | 1.51 (0.86–2.63)         | 0.15              |
| 25–30                    | Non-O           | GG                                | 26 (3.7)            | 15 (3.6)               | 2.39 (1.07–5.52)         | 0.04              |
| >30                      | Non-O           | GG                                | 24 (3.4)            | 5 (1.2)                | 6.10 (2.17–20.31)        | 0.01              |
| <25                      | O               | AG/AA                             | 83 (11.7)           | 77 (18.3)              | 1.46 (0.83–2.57)         | 0.18              |
| 25–30                    | O               | AG/AA                             | 22 (3.1)            | 13 (3.1)               | 3.55 (1.49–8.71)         | 0.005             |
| >30                      | O               | AG/AA                             | 18 (2.5)            | 6 (1.4)                | 5.97 (2.08–19.39)        | 0.002             |
| <25                      | Non-O           | AG/AA                             | 288 (40.6)          | 148 (35.2)             | 2.99 (1.81–4.96)         | <10 <sup>-4</sup> |
| 25–30                    | Non-O           | AG/AA                             | 68 (9.6)            | 17 (4.0)               | 5.58 (2.75–11.76)        | <10 <sup>-4</sup> |
| >30                      | Non-O           | AG/AA                             | 48 (6.7)            | 6 (1.4)                | 13.15 (5.20–38.50)       | <10 <sup>-4</sup> |

BMI, body mass index; CI, confidence interval; COC, combined oral contraceptives; NA, not applicable; OR, odds ratio; VT, venous thrombosis.

<sup>a</sup>Genotype for *F11* rs2289252. Risk allele = A.

<sup>b</sup>Adjusted OR. Adjustment factors = age, type and duration of COC use, and family history score.

risk factor (obesity) and two common genetic risk factors (non-O blood groups and *F11* rs2289252) represented a major risk factor for VT. All current guidelines and recommendations consider the risk associated with COC in women harboring a mild thrombophilia unacceptable. As a consequence, COC use is contraindicated in all women with mild thrombophilia. However, van Vlijmen et al. evaluated in a recent review the additional effect of inherited thrombophilia in COC users and suggested that COC could be prescribed in women with mild thrombophilia in the absence of other risk factors (8). Indeed, most women using COC and harboring mild thrombophilia would not ever undergo VT. As a consequence, environmental and genetic modifiers that may increase the risk of VT in these women need to be identified. In this goal, other common genetic risk factors, such as ABO blood group and *F11* rs2289252, could be taken into account in the VT risk assessment. More generally, as VT is a complex disease, combination of risk factors could be more useful to consider for assessing the risk than to consider each risk factor individually. In this study, the OR associated with the combination of three common risk factors was found to be 13.15 (5.20–38.50), which is higher than the risk ratio (RR) associated with factor V Leiden or prothrombin mutation in COC users according to van Vlijmen et al. meta-analysis (RR = 5.89; 95% CI: 4.21–8.23) (8). Of note, only severe obesity (BMI > 35 kg/m<sup>2</sup>) might be considered as a contraindication for COC use, whereas a 30 kg/m<sup>2</sup> cut-off showed a high level of risk in the present study. The magnitude of this effect was highly influenced by the combination based on the two common genetic risk factors, which suggests that this combination could provide useful information for the VT risk assessment in COC users exposed to environmental risk factors such as obesity. The addition of those three risk factors could result in a state of hypercoagulability. Indeed, *F11* rs2289252 is associated with increased plasmatic levels of FXI, which is a key enzyme of the coagulation pathway as it activates factor IX (13). Moreover, thrombin generation is

increased in patients with elevated FXI (14). Similarly non-O blood groups are associated with a 30% increased level of factor VIII, which elevation has been shown to increase thrombin generation (15). Finally, obesity is also associated with increased thrombin generation (16). Altogether elevated FXI and FVIII and obesity could increase thrombin generation and VT risk. It is currently unclear whether obesity should contraindicate COC use or not. Obese women (i.e. BMI > 30 kg/m<sup>2</sup>) could be eligible for the screening of the two common genetic risk factors. Based on their prevalence in general population (non-O blood groups ≈ 50%; *F11* rs2289252 ≈ 60%) the screening would identify the combination of the two alleles in 30% of obese women whom should be considered at high risk.

This study suffers from limitations which can prevent us from generalizing the results. The first one is relative to the study design which can cause bias. The control group is not derived from the general population. Indeed, controls were recruited as they had a family history of thrombosis and thus were putatively enriched in genetic risk factors. This explains why factor V Leiden was not associated with VT while it is the most robust and strongest common genetic risk factor of VT (9). However, despite this putative genetically enriched control group, the study has been able to show a 60% increased risk of VT in women using COC who harbored the risk allele for *F11* rs2289252. Another limitation is that we used a subcohort of PILGRIM (i.e. patients with DNA available). We have checked for homogeneity between the total cohort and the subcohort. Main characteristics of patients were similar between the two groups as well as the impact of the major risk factors for VT. The former point allows us to consider that this subcohort is representative for the whole PILGRIM cohort. Finally, this study was monocentric and has included individuals only from Marseilles area. Results should thus be replicated in individuals from other countries in order to verify their generalizability.

## Suchon et al.

This study provides evidence of the difficulty in assessing individual risk of VT, which complexity arises from the multiplicity of risk factors that need to be taken into account. In that matter different predictive scores including genetic and environmental risk factors have emerged in the recent literature as useful predictors of VT of an individual (17–19). It is mandatory to validate these scores and design new ones specific for COC users. Effectiveness of these scores should be assessed and compared with current practice in management studies. Their cost-effectiveness in contraceptive counseling should also be evaluated before implementation in clinical practice.

## Acknowledgements

We are grateful to Gene Predictis SA (Switzerland) for genotyping of some SNPs.

## References

1. Martinez C, Cohen AT, Bamber L, Rietbrock S. Epidemiology of first and recurrent venous thromboembolism: a population-based cohort study in patients without active cancer. *Thromb Haemost* 2014; 112: 255–263.
2. Stegeman BH, de Bastos M, Rosendaal FR et al. Different combined oral contraceptives and the risk of venous thrombosis: systematic review and network meta-analysis. *BMJ* 2013; 347: f5298.
3. United Nations. Department of Economic and Social Affairs. Population Division. World contraceptive use 2011.
4. Tricotel A, Raguideau F, Collin C, Zureik M. Estimate of venous thromboembolism and related-deaths attributable to the use of combined oral contraceptives in France. *PLoS One* 2014; 9: e93792.
5. Ljungqvist M, Holmström M, Kieler H, Odeberg J, Lärfars G. Cardiovascular disease and mortality after a first episode of venous thromboembolism in young and middle-aged women. *Thromb Res* 2016; 138: 80–85.
6. Tick LW, Doggen CJM, Rosendaal FR et al. Predictors of the post-thrombotic syndrome with non-invasive venous examinations in patients 6 weeks after a first episode of deep vein thrombosis. *J Thromb Haemost* 2010; 8: 2685–2692.
7. Klok FA, Dzikowska-Diduch O, Kostrubiec M et al. Derivation of a clinical prediction score for chronic thromboembolic pulmonary hypertension after acute pulmonary embolism. *J Thromb Haemost* 2016; 14: 121–128.
8. van Vlijmen EF, Wiewel-Verschueren S, Monster TB, Meijer K. Combined oral contraceptives, thrombophilia and the risk of venous thromboembolism: a systematic review and meta-analysis. *J Thromb Haemost* 2016; 14: 1393–1403.
9. Suchon P, Al Frouh F, Henneuse A et al. Risk factors for venous thromboembolism in women under combined oral contraceptive. The PILL Genetic Risk Monitoring (PILGRIM) Study. *Thromb Haemost* 2015; 115: 135–142.
10. Germain M, Chasman DI, de Haan H et al. Meta-analysis of 65,734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism. *Am J Hum Genet* 2015; 96: 532–542.
11. Morange P-E, Suchon P, Trégouët D-A. Genetics of venous thrombosis: update in 2015. *Thromb Haemost* 2015; 114: 910–919.
12. Wu O, Robertson L, Twaddle S et al. Screening for thrombophilia in high-risk situations: systematic review and cost-effectiveness analysis. The Thrombosis: Risk and Economic Assessment of Thrombophilia Screening (TREATS) study. *Health Technol Assess* 2006; 10: 1–110.
13. Li Y, Bezemer ID, Rowland CM et al. Genetic variants associated with deep vein thrombosis: the F11 locus. *J Thromb Haemost* 2009; 7: 1802–1808.
14. Siegemund A, Petros S, Siegemund T, Scholz U, Seyfarth HJ, Engelmann L. The endogenous thrombin potential and high levels of coagulation factor VIII, factor IX and factor XI. *Blood Coagul Fibrinolysis* 2004; 15: 241–244.
15. Ryland JK, Lawrie AS, Mackie IJ, Machin SJ. Persistent high factor VIII activity leading to increased thrombin generation – a prospective cohort study. *Thromb Res* 2012; 129: 447–452.
16. Campello E, Zabeo E, Radu CM et al. Hypercoagulability in overweight and obese subjects who are asymptomatic for thrombotic events. *Thromb Haemost* 2015; 113: 85–96.
17. de Haan HG, Bezemer ID, Doggen CJ et al. Multiple SNP testing improves risk prediction of first venous thrombosis. *Blood* 2012; 120: 656–663.
18. Soria JM, Morange P-E, Vila J et al. Multilocus genetic risk scores for venous thromboembolism risk assessment. *J Am Heart Assoc* 2014; 3: e001060.
19. Bruzelius M, Bottai M, Sabater-Lleal M et al. Predicting venous thrombosis in women using a combination of genetic markers and clinical risk factors. *J Thromb Haemost* 2015; 13: 219–227.

# **Discussion et conclusion**

Actuellement, le bilan de thrombophilie constitutionnelle comportant la recherche de 5 anomalies biologiques est le seul reconnu par les sociétés savantes et remboursé par la sécurité sociale. Au demeurant, ce bilan ne permet pas d'identifier de manière convaincante les individus à risque de MTEV (et les individus non à risque) dans les familles avec un facteur génétique diagnostiqué. La réalisation d'une enquête familiale est donc débattue, en particulier pour les thrombophilies modérées, représentées par les mutations hétérozygotes pour le FVL ou PTG20210A.

Concernant la thrombophilie sévère, l'intérêt d'un tel bilan semble plus évident mais il se heurte à une interprétation complexe liée aux techniques de dosage utilisées, aux seuils mal définis et à la variabilité intra-individuelle des taux. Dès lors, on peut se demander si cette stratégie, reposant sur le dépistage de 5 anomalies biologiques, notamment par la réalisation de dosages plasmatiques, est la bonne.

Le meilleur exemple est probablement celui de la PS. D'un côté, seules des mutations délétères situées sur *PROS1* apparaissent associées au risque de MTEV chez des patients présentant un déficit en PS. De l'autre, dans notre étude, conduite chez plus de 10000 patients non sélectionnés par le taux de PS, le variant Heerlen était associé à la MTEV avec un risque estimé à 6,57. Ce dernier variant, généralement identifié chez des patients présentant un déficit en PS de type III (PSL diminuée et PST normale), était associé à des taux modérément diminués (mais au-dessus du seuil actuellement utilisé pour le diagnostic de déficit en PS) de PSL dans la cohorte MARTHA. Ce résultat remet en question la stratégie qui semble se dessiner consistant à utiliser un seuil de PS de plus en plus bas (bien inférieur à ceux actuellement pratiqués) pour le dépistage des mutations délétères. En effet, dans la première étude présentée, des taux inférieurs à 25% permettaient d'identifier une mutation délétère dans 94% des cas contre seulement 34% pour un seuil fixé à 55%, ce qui semblait démontrer une corrélation inverse entre les taux de PS et la découverte d'une mutation délétère. Ce paradigme est observé dans la plupart des maladies de l'hémostase comme par exemple la maladie de Willebrand. La PS Heerlen a longtemps été considérée comme un polymorphisme neutre. Des études utilisant une PS Heerlen recombinante ont en effet démontré qu'il n'existait pas d'altération fonctionnelle, en termes de liaison à la C4bBP et d'inactivation des facteurs V et VIII activés <sup>103</sup>. Peu d'études avaient à ce jour permis d'évaluer le risque de MTEV associé à la mutation Heerlen. Seule une étude avait comparé la prévalence de la mutation chez des patients présentant une histoire personnelle de MTEV et

des donneurs de sang <sup>87</sup>. Il n'existait alors pas de différence entre les deux groupes. Cette étude s'appuyait sur des effectifs modestes. Le groupe témoin, en particulier, ne comprenait que 96 donneurs. Depuis, des études s'intéressant spécifiquement à des patients présentant un déficit plasmatique en PS n'ont pas permis de relier le génotype Heerlen au phénotype thrombotique. Notre première étude confirme ces résultats puisque le variant Heerlen, identifié chez 78 patients, était associé à une augmentation non significative du risque de MTEV, comparativement au groupe de patients sans mutation : RR ajusté (sur le sexe et le taux de PS) = 1,23 (IC95% = 0,84-1,79). Pour cette étude, tous les patients séquencés ont été inclus en raison d'un déficit quantitatif en PS. Pour la seconde étude, spécifiquement dédiée au variant Heerlen, nous avons utilisé une approche complètement différente. Une étude d'association, réalisée à l'aide d'une puce exomique enrichie en variants rares, a permis d'observer une association fortement significative entre le variant Heerlen et la MTEV. Ces résultats ont ensuite été répliqués dans 3 études cas-témoins indépendantes françaises. Les taux de PS n'étaient connus que dans une étude, l'étude MARTHA qui présentait comme critère d'exclusion l'existence d'une thrombophilie sévère. Aucun patient inclus ne présentait donc de déficit en PS. Il est à noter que seulement 2 d'entre elles ont permis de confirmer ces résultats. En effet, dans l'étude FARIVE, il n'existait pas de différence significative entre les deux groupes de patients. Ce résultat négatif s'explique probablement par une prévalence étonnamment élevée de la mutation dans le groupe témoin. Alors que la prévalence de la mutation était inférieure à 0,6% chez les témoins des 3 autres études, soit une prévalence comparable à la prévalence estimée dans la population générale, elle était estimée à 1,8% chez les témoins de l'étude FARIVE. Ce résultat ne trouve pas d'explication évidente. Au demeurant, les témoins de FARIVE étaient des patients appariés de cardiologie sans antécédent personnel de MTEV. De façon intéressante, l'association entre déficit en PS et thrombose artérielle a déjà été rapportée <sup>158</sup>. Il est à noter que dans la première étude, les patients avec antécédent personnel de thrombose artérielle n'ont pas été exclus. Il serait ainsi intéressant d'estimer la prévalence de la mutation Heerlen dans ce groupe spécifique de patients. Une prévalence élevée pourrait ainsi expliquer l'absence d'association entre le variant et le phénotype thrombotique. Les différences de méthodologies employées peuvent également expliquer les résultats discordants observés. En effet, alors que dans la première étude les patients étaient sélectionnés sur la base d'un déficit en PS, le taux de PS n'était pas pris en compte dans la seconde (à l'exception de l'étude MARTHA qui excluait les patients



déficitaires en PS). Les deux études ne sont donc pas comparables et n'apportent ainsi pas les mêmes informations. Dans la première étude, tous les sujets inclus présentaient un taux diminué de PS. Parmi eux, seuls les patients présentant une mutation délétère apparaissaient à risque de MTEV, comparativement aux patients sans mutation (RR = 1,41 ; IC95% = 1,05-1,89). Seule une tendance a pu être observée pour la mutation Heerlen. Cependant, le groupe de référence ne permet pas d'exclure un risque de MTEV associé à la mutation Heerlen. En effet, il n'est pas exclu que les patients sans mutation identifiée présentent en fait une anomalie moléculaire que la technique de séquençage ne permettait pas d'explorer. De même, il n'est pas exclu que des taux bas en PS trouvent leur explication dans des mécanismes non liés à *PROS1*, génétiques et/ou environnementaux. Par exemple, une récente publication a rapporté le cas d'un déficit en AT chez un jeune homme de 18 ans présentant des antécédents de MTEV <sup>159</sup>. Un déficit en AT à 67% a été diagnostiqué sans anomalie moléculaire de *SERPINC1* (la MLPA ayant permis d'exclure une grande délétion). L'électrophorèse plasmatique et l'analyse moléculaire ont permis de diagnostiquer un trouble congénital de la glycosylation (congenital disorder of glycosylation - CDG) lié à une mutation sur *PMM2*. Il n'est donc pas exclu que des mutations non situées sur *PROS1* puissent expliquer des déficits en PS.

La difficulté à déterminer un taux permettant de diagnostiquer un déficit constitutionnel en PS et les résultats obtenus au cours de la seconde étude soulèvent la question de la réalisation d'un génotypage systématique du variant Heerlen. Ainsi, le génotypage du variant Heerlen apporterait une information indépendante du dosage plasmatique de la PS, dont l'interprétation serait comparable à celle du génotypage des deux mutations ponctuelles du bilan de thrombophilie, le FVL et PTG20210A. Il serait alors envisageable de réaliser l'analyse simultanée des trois mutations (PCR multiplex) afin de maîtriser l'augmentation des coûts liée à l'extension du bilan de thrombophilie. A titre de comparaison, le génotypage seul d'une mutation parmi le FVL et PTG20210A est coté B100 (correspondant à 27 euros) alors que la prescription simultanée des deux analyses est cotée B150 (soit 40 euros environ). La seule étude présentée dans ce manuscrit traitant spécifiquement du variant Heerlen ne permet pas de répondre à la question du génotypage systématique. En effet, bien que l'analyse ait porté sur plusieurs études, permettant ainsi de conclure à la robustesse du résultat publié, une étude ne retrouvait pas d'association entre le variant et le risque de premier épisode de MTEV (cette absence d'association étant

probablement expliquée par la prévalence anormalement élevée de la mutation dans le groupe témoin). De même, seules des données sur la survenue d'un premier épisode thrombotique sont disponibles. Il apparaît donc nécessaire d'évaluer le risque de récurrence associé au variant Heerlen. Les résultats d'une telle étude, en complément des résultats déjà publiés, permettraient de mieux évaluer l'intérêt du génotypage du variant Heerlen dans le bilan de thrombophilie.

De façon plus générale, les difficultés à obtenir une corrélation entre phénotype biologique (taux plasmatiques) et génotype pour les inhibiteurs de la coagulation, doivent faire envisager le séquençage systématique des gènes impliqués : *SERPINC1*, *PROC* et *PROS1*. Outre la complexité du diagnostic de déficit constitutionnel en PS évoquée plus haut, les déficits en AT et PC ne sont pas exempts de complications. Par exemple, une mutation de *SERPINC1* n'est identifiée que dans 83% des déficits en AT<sup>160</sup>. A contrario, le variant Dublin situé sur *SERPINC1* est associé à un risque de MTEV évalué à 2,9 malgré des taux normaux d'AT<sup>161</sup>. Un séquençage systématique, rendu possible à moindre coût grâce au développement des techniques de NGS, chez des patients sélectionnés (patients jeunes avec une histoire personnelle et familiale de MTEV...), pourrait permettre d'améliorer les performances du bilan de thrombophilie en termes de prédiction du risque thrombotique.

L'amélioration de la prédiction de la survenue d'un épisode de MTEV passe probablement par la prise en compte d'autres facteurs de risque génétiques, comme le suggère l'héritabilité manquante de la pathologie, et des facteurs de risque environnementaux, dont l'interaction avec les premiers facteurs permet d'expliquer la physiopathologie de la MTEV. A ce jour, de nombreux polymorphismes ont été identifiés comme associés à la MTEV en population générale, avec des niveaux de preuve variables. Leur impact dans des populations spécifiques de familles dites thrombophiles est mal évalué. De même, l'impact de facteurs de risque environnementaux communs tels que le tabagisme et l'obésité est méconnu. Nous avons conduit une étude chez 2214 apparentés issus de 651 familles avec histoire thrombotique et diagnostic d'une anomalie du bilan de thrombophilie. Nous avons ainsi identifié 5 facteurs de risque de MTEV fréquents à effet faible permettant d'améliorer l'évaluation du risque individuel : l'obésité, le tabagisme, le groupe sanguin, le polymorphisme rs2036914 situé sur *F11* et le polymorphisme rs2066865 situé sur *FGG*.

Parmi les 11 polymorphismes sélectionnés pour la robustesse de leur association en population générale, seuls 2 étaient donc associés au risque de MTEV dans une population

spécifique d'apparentés issus de familles présentant un facteur de risque biologique majeur. L'absence d'association pour 9 polymorphismes peut s'expliquer par un manque de puissance. L'effet serait alors identique à celui observé en population générale mais le manque de puissance ne permettrait pas d'identifier un effet statistique. En effet, seuls 246 sujets avec antécédent personnel de MTEV ont pu être analysés. Cet effectif peut sembler faible comparativement aux études qui ont permis l'identification des polymorphismes sélectionnés, en général plusieurs centaines, voire milliers de patients. Ces résultats peuvent cependant signifier que le signal identifié en population générale ne s'observe pas dans une population spécifique, enrichie en facteurs génétiques.

En revanche, notre étude confirme les résultats publiés concernant les polymorphismes rs2036914 et rs2066865 respectivement situés sur *F11* et *FGG*.

Deux études permettent notamment d'estimer le risque associé au polymorphisme rs2036914 en population générale. La première rapportait un HR estimé à 1,20 (IC95% = 1,11-1,31)<sup>129</sup>, et la seconde un RR estimé à 1,30 (1,19-1,42)<sup>127</sup>. Les HR observés dans notre étude semblent sensiblement plus élevés que dans la première étude (HR = 1.55 (IC95% = 1.06-2.26)). Le polymorphisme pourrait ainsi constituer un facteur modulateur important du risque dans ces familles.

Le facteur XI est un acteur essentiel de la cascade de la coagulation et plus particulièrement de la génération de thrombine et de la protection contre la fibrinolyse. En pathologie, le déficit en facteur XI entraîne des manifestations hémorragiques dont la sévérité n'est pas corrélée au taux<sup>162</sup>. A l'inverse, des taux élevés de facteur XI sont associés à une augmentation du risque de MTEV<sup>136</sup>. Des taux élevés de facteur XI amplifieraient la génération de thrombine et diminueraient la fibrinolyse de façon plus soutenue dans le temps qu'en présence de taux normaux, ce qui favoriserait la constitution du thrombus veineux. Des données de cinétique enzymatique viennent étayer cette hypothèse<sup>163</sup>. Le polymorphisme rs2036914 est associé à des taux élevés de facteur XI<sup>129</sup> qui augmentent la génération de thrombine<sup>164</sup>. Ce résultat explique probablement une partie de l'association entre le polymorphisme et la MTEV. En effet, cette dernière persiste après ajustement pour les taux de facteur XI<sup>129</sup>. Ce dernier résultat laisse ainsi penser que l'implication du polymorphisme rs2036914 dans la survenue de la MTEV fait appel à des mécanismes moins intuitifs. Il s'agit d'un polymorphisme intronique dont la caractérisation fonctionnelle est par définition délicate à entreprendre.

Le polymorphisme rs2289252, significativement associé à la MTEV en analyse univariée, ne l'était plus en analyse multivariée, incluant notamment l'autre polymorphisme de *F11* précédemment discuté. Les deux polymorphismes sont ainsi probablement en déséquilibre de liaison et les études haplotypiques ont permis de déterminer que les 2 polymorphismes sont organisés en haplotypes <sup>129</sup>. Il serait donc intéressant de compléter l'analyse dans les familles en prenant en compte cette donnée.

En ce qui concerne le polymorphisme rs2066865 situé sur *FGG*, chez des patients non apparentés, le génotype TT était associé à une multiplication par 2 du risque de MTEV chez les individus d'origine allemande <sup>118</sup> et d'origine australienne <sup>165</sup> et par 1,20 dans la population afro-américaine <sup>119</sup>. Le fibrinogène est un acteur fondamental de la formation du thrombus. Il est clivé par la thrombine pour former la fibrine, qui est ensuite stabilisée par le facteur XIII. Le mécanisme par lequel ce polymorphisme est responsable d'une susceptibilité à la MTEV n'est que partiellement connu. Il se traduit par une substitution d'une cytosine par une thymine en position 10034 de *FGG*, dans la partie 3' non codante du gène, au sein d'une séquence nécessaire pour un clivage efficace et la polyadénylation du pré-ARNm. Des modifications post-transcriptionnelles liées au polymorphisme entraîneraient une synthèse favorisée de la chaîne  $\gamma$ A comparativement à la chaîne  $\gamma'$  <sup>118</sup>. La chaîne  $\gamma'$  résulte de l'épissage alternatif de l'ARNm de la chaîne  $\gamma$  du fibrinogène. Elle diffère de la chaîne  $\gamma$ A par la substitution des 4 acides aminés C-terminaux par un polypeptide de 20 acides aminés <sup>166</sup>. La chaîne  $\gamma'$  porte un site de haute affinité pour la thrombine. Ainsi la liaison inhiberait la thrombine et expliquerait l'association entre des taux bas de fibrinogène  $\gamma'$  et l'augmentation de la MTEV <sup>118</sup>.

Trois autres facteurs semblaient améliorer l'évaluation du risque de MTEV à l'échelle individuelle dans les familles thrombophiles : le groupe sanguin (non-O), l'obésité et le tabac. Le groupe sanguin est un facteur de risque robuste de MTEV, bien connu depuis les années 60 <sup>167</sup>. En effet, une méta-analyse réunissant de nombreuses études a permis d'évaluer son impact sur le risque de MTEV. Celui-ci est estimé à 1,79 (IC95% = 1,56-2,05) <sup>168</sup>. Sa détermination ne fait toutefois pas partie du bilan de thrombophilie. Les résultats obtenus dans notre étude semblaient confirmer les données de la littérature obtenues en population générale puisque les HR étaient estimés à 1,37 et 2,51 respectivement pour les groupes A/B et AB, comparativement au groupe O. De même, l'obésité était associée à un risque de 1,78 (IC95% = 1,24-2,55), résultat comparable aux données précédemment

publiées <sup>25</sup>. En revanche, alors que des données récentes ont permis d'évaluer que le tabagisme actif est associé à une augmentation modérée du risque de MTEV (OR voisins de 1,25), mettant ainsi fin au débat ayant fait rage pendant de nombreuses années <sup>34,35</sup>, le tabagisme actif apparaissait un facteur de risque fréquent majeur chez les apparentés issus de familles thrombophiles. En effet, parmi les facteurs de risque communs étudiés, le tabagisme était, avec le groupe sanguin AB, le facteur entraînant le plus fort risque de MTEV : HR = 2,08 (IC95% = 1,51-2,86). Alors que le tabagisme actif constitue un facteur de risque très modéré de MTEV en population générale, il semble constituer un facteur de risque important dans des familles sélectionnées, avec un HR comparable à un facteur de risque génétique reconnu comme le groupe sanguin. Les apparentés sélectionnés pour l'étude appartenaient à des familles qui concentrent des facteurs de risque génétiques (majeurs ou modérés), comme l'atteste, par exemple, la prévalence observée des groupes sanguins non-O (77,2 et 65,2% des apparentés avec et sans antécédent personnel de MTEV respectivement, alors que la prévalence en population générale est estimée à 58% d'après les données de l'Etablissement Français du Sang). Dans une population enrichie en facteurs de risque génétiques, l'effet du tabagisme semble donc plus marqué. Des études d'interaction entre ce facteur de risque environnemental et les différents facteurs génétiques pourraient étayer cette hypothèse. Au demeurant, dans notre étude, les interactions entre les polymorphismes génotypés et le tabagisme ont été testées mais aucune n'était significative.

Au total, 5 facteurs fréquents semblent améliorer l'évaluation du risque de premier épisode de MTEV chez les apparentés. Ces résultats laissent envisager de nouvelles stratégies d'exploration. A ce jour, l'évaluation individuelle du risque de MTEV dans le cadre d'une enquête familiale de thrombophilie repose principalement sur la recherche de l'anomalie familiale. Les résultats observés dans l'étude présentée suggèrent que cette stratégie n'est pas satisfaisante. Le score reposant sur 5 critères permet d'ajouter un niveau d'interprétation au bilan de thrombophilie. De façon schématique, 3 groupes d'apparentés se dessinent après dépistage de l'anomalie familiale : les apparentés ne présentant pas l'anomalie, les apparentés présentant une thrombophilie modérée et les apparentés présentant une thrombophilie sévère. Dans le premier groupe, les individus avec un score élevé semblent présenter un risque plus élevé que la population générale (incidence évaluée à 2,62/1000/an). L'application du score pourrait ainsi être bénéfique dans ce groupe. Dans le

second groupe, l'utilisation du score semble extrêmement intéressante. En effet, le score permet d'individualiser deux catégories d'apparentés : des individus avec un score faible et une incidence comparable à la population générale, et des individus avec un score élevé et un risque comparable aux apparentés porteurs d'une thrombophilie sévère. Enfin, le troisième groupe constitué des apparentés présentant une thrombophilie sévère ne semble pas bénéficier du score puisque ce dernier ne permet pas d'identifier des individus à faible risque (malgré la grande disparité des incidences observées en fonction du score). Dans ce dernier groupe, le résultat du dépistage de l'anomalie familiale, en cas de positivité, constitue l'information essentielle. La place du score dans la stratégie d'évaluation du risque individuel de MTEV dans les familles est donc variable. S'il apparaît essentiel chez les apparentés présentant une thrombophilie modérée, et potentiellement intéressant chez les patients ne présentant pas l'anomalie familiale, il est probablement moins important en cas de thrombophilie sévère. Il reste alors à définir une stratégie prenant en compte cette interprétation ainsi que l'impact économique d'un tel score. Ce dernier point mérite d'être précisé. Parmi les 5 paramètres intégrés dans le score, 2 sont obtenus lors de la consultation (IMC et statut tabagique), un pourrait l'être également (le groupe sanguin peut être obtenu lors de l'interrogatoire ou devra être déterminé), et 2 autres nécessitent une analyse de biologie moléculaire. Le coût de cette analyse doit être déterminé. La comparaison avec la recherche du FVL et de PTG20210A précédemment exposée permet d'envisager le génotypage simultané des 4 mutations (FVL, PTG20210A et les deux polymorphismes intégrés dans le score), favorisant ainsi la maîtrise des surcoûts occasionnés par l'ajout de deux facteurs biologiques de risque. La prise en compte du paramètre économique doit alors conduire à mieux définir les indications de la prescription du score : quels patients devront être testés ? A quel moment par rapport à la recherche de l'anomalie familiale ? Il pourrait être intéressant de définir des stratégies différentes en fonction de l'anomalie familiale. Le score ne permettant pas d'identifier des individus à faible risque chez les apparentés porteurs d'une thrombophilie sévère, il serait envisageable de ne le réaliser qu'en seconde intention chez les apparentés non porteurs de l'anomalie familiale. Au contraire, afin de réduire les coûts de biologie moléculaire, il serait probablement préférable de considérer le score comme une analyse de première intention dans le cadre de l'enquête familiale d'une thrombophilie modérée. Enfin, il serait intéressant d'estimer l'utilisation d'un score simplifié, ne prenant en compte que des paramètres dont le recueil est possible au cours de la

consultation (statut tabagique, IMC et groupe sanguin). Une telle stratégie permettrait alors de s'affranchir du facteur économique.

Les résultats intéressants obtenus dans une population spécifique d'apparentés a permis de confirmer l'hypothèse que la prise en compte de facteurs de risque fréquents à effet faible pourrait permettre de mieux évaluer le risque individuel. Cette dernière hypothèse étant sous-tendue par la physiopathologie même de la MTEV, pathologie complexe. D'autres populations spécifiques pourraient justifier d'une telle démarche. Nous avons ainsi choisi d'étudier une population spécifique de femmes sous contraceptifs oraux combinés. Les femmes en âge de procréer constituent en effet le groupe principalement ciblé par le dépistage d'une anomalie constitutionnelle de l'hémostase prédisposant à la MTEV. Et la contraception orale est une situation extrêmement fréquente qui se traduit par une augmentation du risque de MTEV, dont le niveau varie en fonction de la combinaison de 2 caractéristiques : le progestatif et la concentration en éthinyl-estradiol <sup>169</sup>. Plus de 1200 utilisatrices de contraceptifs oraux combinés ont été incluses dans l'étude PILGRIM présentée dans ce manuscrit. Les cas étaient définis comme étant des femmes avec antécédent personnel de MTEV documentée au cours de l'utilisation d'une pilule oestroprogestative. Les témoins étaient des femmes sous contraceptifs oraux combinés au moment de leur consultation au Centre d'Exploration des pathologies Hémostatiques et Thrombotiques de Marseille. Neuf polymorphismes, tous compris dans l'étude précédemment citée, ont été génotypés. Seul le polymorphisme rs2289252 situé sur *F11* était associé à la MTEV avec un risque évalué à 1,6. Il faut ici noter la discordance de ce résultat avec l'étude précédente. En effet, les deux mêmes polymorphismes de *F11* ont été sélectionnés et génotypés dans les deux études (MARFAST et PILGRIM) présentées avec des résultats différents. En effet, dans la première étude conduite chez des apparentés, c'est le polymorphisme rs2036914 qui était associé à la MTEV en analyse multivariée, alors que c'est le polymorphisme rs2289252 qui restait associé au risque en analyse multivariée. Il faut noter que de façon similaire dans les deux études, les 2 polymorphismes étaient associés à la MTEV en analyse univariée. Ce résultat semble indiquer, à nouveau, que l'analyse haplotypique pourrait être plus pertinente. Au demeurant, il n'est pas exclu que des mécanismes physiopathologiques différents soient impliqués dans les différentes populations (familles ou femmes sous contraceptifs oraux combinés). Parmi les autres facteurs étudiés, l'obésité (en particulier un IMC supérieur à 35 kg.m<sup>-2</sup>), les groupes sanguins

non-O et le tabagisme étaient associés à des OR respectifs de 3,46 (IC95% = 1,81-7,03), 1,98 (IC95% = 1,57-2,49) et 1,65 (IC95% = 1,30-2,10). En dehors des polymorphismes, les résultats observés dans MARFAST étaient comparables dans l'étude PILGRIM. Des stratégies similaires ont ainsi été employées dans les 2 études pour mieux caractériser le risque individuel. Dans l'étude MARFAST, un score prédictif réunissant 5 paramètres a été construit : le groupe sanguin, l'obésité, le tabagisme et les polymorphismes situés sur *F11* et *FGG*. De façon simple, un point était attribué à chacun des 5 éléments lorsqu'il était présent, à l'exception du groupe sanguin, dont la complexité des phénotypes a conduit à une cotation spécifique (1 point pour les groupes A et B et 2 points pour le groupe AB). Trois niveaux de risque ont ensuite été définis permettant une prédiction plus fine du risque de MTEV que le seul bilan de thrombophilie. En effet, alors que les incidences étaient de 1,74, 3,64 et 6,40 pour 1000 personnes-années respectivement chez les apparentés sans anomalie du bilan de thrombophilie, avec une thrombophilie modérée et une thrombophilie sévère, les incidences étaient en fait comprises entre 0,60 et 9,85 pour 1000 personnes-années en appliquant le score. De façon intéressante, le score permettait d'identifier des individus à risque faible (comparable à la population générale) chez les apparentés présentant une thrombophilie modérée. Dans ce même groupe, le score permettait d'identifier des individus dont le risque était comparable aux apparentés qui présentaient une thrombophilie sévère et un score moyen (incidences respectives de 4,98 et 4,83 pour 1000 personnes-années). La combinaison des facteurs de risque permettait également d'identifier des individus à haut risque chez les utilisatrices de contraceptifs oraux combinés. En effet, la présence simultanée de 3 facteurs de risque (obésité, groupe sanguin non-O et allèle à risque pour le polymorphisme rs2289252 du *F11*) était associée à un OR estimé de 13,15 (IC95% = 5,20-38,50). Ainsi, la prise en compte de facteurs de risque fréquents à effet faible, génétiques et environnementaux, semblait améliorer l'évaluation du risque individuel de MTEV dans des populations spécifiques d'apparentés et de femmes sous contraceptifs oraux combinés. Cette stratégie a déjà été rapportée dans une population spécifique de femmes. La combinaison de 7 facteurs de risque fréquents à effet faible/modéré (FVL, PTG20210A, ABO rs514659, *FGG*\_rs2066865, *F11*\_rs2289252, *PROC*\_rs1799810 et *KNG1*\_rs710446) et leurs interactions, en association avec l'histoire familiale de MTEV, permettait d'obtenir une bonne performance pour la prédiction du risque thrombotique, évaluée par l'aire sous la



courbe (AUC) (AUC = 0,84). L'utilisation de cette stratégie dans des populations spécifiques (par exemple la femme enceinte) pourrait permettre d'obtenir des outils prédictifs efficaces.

Les études d'association pangénomiques ont permis d'identifier de nombreux polymorphismes associés à la MTEV dont l'utilisation en pratique clinique reste à définir. Au demeurant, la liste des facteurs génétiques permettant d'expliquer la MTEV est probablement encore longue. En effet, tous les facteurs génétiques à ce jour identifiés ne permettent d'expliquer qu'une faible part de l'hérédité de la pathologie. De nouvelles stratégies restent donc probablement à mettre en œuvre pour identifier cette héritabilité manquante. Avant de lister ces stratégies, notons qu'il est envisageable d'identifier de nouveaux variants génétiques en utilisant des stratégies classiques. L'augmentation toujours plus importante des effectifs des études d'association pangénomiques devrait permettre d'identifier des polymorphismes avec des effets de plus en plus faibles<sup>170</sup>. Mais la pertinence clinique de tels polymorphismes reste questionnable. La course à l'effectif, nécessaire pour obtenir une puissance suffisante pour observer des effets faibles, peut se traduire par une dilution de certaines associations. En effet, les populations sélectionnées pour réaliser de telles études sont généralement très hétérogènes (critères d'inclusion peu contraignants pour garantir le recrutement). La sélection de plus petits échantillons pourrait au contraire avoir certaines vertus. En effet, des critères de sélection plus restrictifs pourraient permettre de sélectionner des échantillons homogènes en termes de phénotype et donc potentiellement de physiopathologie. Les mécanismes à l'œuvre dans la MTEV ne sont pas nécessairement universels et l'approche agnostique appliquée à diverses populations de patients pourrait identifier de nouveaux polymorphismes et ainsi potentiellement de nouvelles voies physiopathologiques.

Les études d'association pangénomiques se sont focalisées sur des variants fréquents. Au demeurant, les puces exomiques actuellement commercialisées permettent de génotyper des variants rares (fréquence allélique inférieure à 1%). De telles puces, enrichies en variants rares, ont ainsi permis d'étudier le variant Heerlen, que les puces classiquement utilisées n'avaient pu identifier jusqu'alors.

Le développement et la démocratisation des techniques de biologie moléculaire ouvre de nouvelles perspectives de recherche. Plutôt que de restreindre la recherche de facteurs de risque génétiques au seul bilan de thrombophilie, il est désormais possible de séquencer les

gènes impliqués dans la MTEV connus, en particulier les gènes de l'AT, la PC et la PS <sup>152</sup>. Cette stratégie se justifie par le fait que le dosage plasmatique des inhibiteurs ne permet pas toujours l'identification de la mutation causale. C'est notamment le cas du variant Heerlen et de l'AT Dublin <sup>161</sup>. Le principe peut même être étendu à tout l'exome, voire le génome. De telles stratégies soulèvent bien entendu de nouvelles questions et problématiques : comment analyser l'information ? Comment stocker l'information ? Quels patients explorer ? A quel coût ?

Enfin, la recherche de nouveaux facteurs de risque de MTEV s'est principalement concentrée sur les facteurs génétiques. Cependant, l'expression des facteurs génétiques est soumise aux mécanismes épigénétiques. L'expression de nombreuses protéines de la coagulation est, par exemple, en partie modulée par les mécanismes de méthylation de l'ADN <sup>171,172</sup>. Les polymorphismes mis en évidence par les GWAS ont des conséquences très diverses sur l'expression protéique. Certains polymorphismes semblent neutres (le polymorphisme identifié comme associé au phénotype étudié est alors probablement en déséquilibre de liaison avec le polymorphisme causal), d'autres conduisent à une dysfonction protéique, et d'autres, enfin, sont responsables d'une variation de l'expression protéique (production diminuée ou augmentée). D'autres paramètres peuvent moduler cette dernière expression. Outre les mécanismes épigénétiques précédemment abordés, l'environnement peut expliquer cette variabilité. Dès lors, l'analyse des seuls mécanismes génétiques ne permet d'obtenir qu'une information partielle. Et le dosage des protéines exprimées pourrait permettre d'obtenir une information plus globale, résultant des différents mécanismes à l'œuvre dans la synthèse protéique. Une récente étude a identifié une nouvelle protéine potentiellement impliquée dans la physiopathologie de la MTEV en comparant les taux plasmatiques de 408 protéines candidates chez 88 patients avec antécédent de MTEV et 85 sujets sains. Les taux de PDGFB (platelet-derived growth factor beta) étaient associés à la MTEV (170). La même approche, transposée aux métabolites, a permis d'identifier les acylcarnitines comme marqueurs potentiels du risque de MTEV (171). De nouvelles stratégies restent donc à mettre en œuvre pour identifier de nouveaux marqueurs du risque de MTEV.

# Références bibliographiques

1. Silverstein MD, Heit JA, Mohr DN, et al. Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. *Arch. Intern. Med.* 1998;158(6):585–593.
2. Naess IA, Christiansen SC, Romundstad P, et al. Incidence and mortality of venous thrombosis: a population-based study. *J. Thromb. Haemost. JTH.* 2007;5(4):692–699.
3. White RH. The Epidemiology of Venous Thromboembolism. *Circulation.* 2003;107(90231):4I–8.
4. Oger E. Incidence of venous thromboembolism: a community-based study in Western France. EPI-GETBP Study Group. Groupe d’Etude de la Thrombose de Bretagne Occidentale. *Thromb. Haemost.* 2000;83(5):657–660.
5. White RH, Zhou H, Romano PS. Incidence of idiopathic deep venous thrombosis and secondary thromboembolism among ethnic groups in California. *Ann. Intern. Med.* 1998;128(9):737–740.
6. Ridker PM, Miletich JP, Hennekens CH, Buring JE. Ethnic distribution of factor V Leiden in 4047 men and women. Implications for venous thromboembolism screening. *JAMA.* 1997;277(16):1305–1307.
7. Morange P-E, Trégouët D-A. Current knowledge on the genetics of incident venous thrombosis. *J. Thromb. Haemost.* 2013;11:111–121.
8. Schulman S, Rhedin A-S, Lindmarker P, et al. A comparison of six weeks with six months of oral anticoagulant therapy after a first episode of venous thromboembolism. *N. Engl. J. Med.* 1995;332(25):1661–1665.
9. Kearon C, Gent M, Hirsh J, et al. A comparison of three months of anticoagulation with extended anticoagulation for a first episode of idiopathic venous thromboembolism. *N. Engl. J. Med.* 1999;340(12):901–907.
10. Agnelli G, Prandoni P, Santamaria MG, et al. Three months versus one year of oral anticoagulant therapy for idiopathic deep venous thrombosis. *N. Engl. J. Med.* 2001;345(3):165–169.
11. Pinede L, Ninet J, Duhaut P, et al. Comparison of 3 and 6 months of oral anticoagulant therapy after a first episode of proximal deep vein thrombosis or pulmonary embolism and comparison of 6 and 12 weeks of therapy after isolated calf deep vein thrombosis. *Circulation.* 2001;103(20):2453–2460.

12. Furie B, Furie BC. Role of platelet P-selectin and microparticle PSGL-1 in thrombus formation. *Trends Mol. Med.* 2004;10(4):171–178.
13. Sevitt S. Pathology and pathogenesis of deep vein thrombi. *Proc. R. Soc. Med.* 1975;68(4):261.
14. Palabrica T, Lobb R, Furie BC, et al. Leukocyte accumulation promoting fibrin deposition is mediated in vivo by P-selectin on adherent platelets. *Nature.* 1992;359(6398):848–851.
15. Kamikubo Y, Mendolicchio GL, Zampolli A, et al. Selective factor VIII activation by the tissue factor-factor VIIa-factor Xa complex. *Blood.* 2017;130(14):1661–1670.
16. Fredenburgh JC, Gross PL, Weitz JI. Emerging anticoagulant strategies. *Blood.* 2017;129(2):147–154.
17. Kahn SR, Comerota AJ, Cushman M, et al. The Postthrombotic Syndrome: Evidence-Based Prevention, Diagnosis, and Treatment Strategies: A Scientific Statement From the American Heart Association. *Circulation.* 2014;130(18):1636–1661.
18. Sandler DA, Martin JF. Autopsy proven pulmonary embolism in hospital patients: are we detecting enough deep vein thrombosis? *J. R. Soc. Med.* 1989;82(4):203–205.
19. Rosendaal FR. Venous thrombosis: a multicausal disease. *Lancet.* 1999;353(9159):1167–1173.
20. Kearon C, Ageno W, Cannegieter SC, et al. Categorization of patients as having provoked or unprovoked VTE: guidance from the SSC of ISTH. *J. Thromb. Haemost.* 2016;
21. Timp JF, Braekkan SK, Versteeg HH, Cannegieter SC. Epidemiology of cancer-associated venous thrombosis. *Blood.* 2013;122(10):1712–1723.
22. Tsai AW, Cushman M, Rosamond WD, et al. Cardiovascular risk factors and venous thromboembolism incidence: the longitudinal investigation of thromboembolism etiology. *Arch. Intern. Med.* 2002;162(10):1182–1189.
23. Samama MM. An epidemiologic study of risk factors for deep vein thrombosis in medical outpatients: the Sirius study. *Arch. Intern. Med.* 2000;160(22):3415–3420.
24. Abdollahi M, Cushman M, Rosendaal FR. Obesity: risk of venous thrombosis and the interaction with coagulation factor levels and oral contraceptive use. *Thromb. Haemost.* 2003;89(3):493–498.

25. Stein PD, Beemath A, Olson RE. Obesity as a risk factor in venous thromboembolism. *Am. J. Med.* 2005;118(9):978–980.
26. Goldhaber SZ, Grodstein F, Stampfer MJ, et al. A prospective study of risk factors for pulmonary embolism in women. *JAMA.* 1997;277(8):642–645.
27. Klatsky AL, Armstrong MA, Poggi J. Risk of pulmonary embolism and/or deep venous thrombosis in Asian-Americans. *Am. J. Cardiol.* 2000;85(11):1334–1337.
28. Lindqvist PG, Epstein E, Olsson H. The relationship between lifestyle factors and venous thromboembolism among women: a report from the MISS study. *Br. J. Haematol.* 2009;144(2):234–240.
29. Hippisley-Cox J, Coupland C. Development and validation of risk prediction algorithm (QThrombosis) to estimate future risk of venous thromboembolism: prospective cohort study. *BMJ.* 2011;343:d4656.
30. Wattanakit K, Lutsey PL, Bell EJ, et al. Association between cardiovascular disease risk factors and occurrence of venous thromboembolism. A time-dependent analysis. *Thromb. Haemost.* 2012;108(3):508–515.
31. Sweetland S, Parkin L, Balkwill A, et al. Smoking, surgery, and venous thromboembolism risk in women: United Kingdom cohort study. *Circulation.* 2013;127(12):1276–1282.
32. Ray JG, Langman LJ, Vermeulen MJ, et al. Genetics University of Toronto Thrombophilia Study in Women (GUTTSI): genetic and other risk factors for venous thromboembolism in women. *Curr. Control. Trials Cardiovasc. Med.* 2001;2(3):141–149.
33. Yamada N, Ota S, Liu Y, et al. Risk factors for nonfatal pulmonary embolism in a Japanese population: A hospital-based case-control study. *Angiology.* 2010;61(3):269–274.
34. Cheng Y-J, Liu Z-H, Yao F-J, et al. Current and former smoking and risk for venous thromboembolism: a systematic review and meta-analysis. *PLoS Med.* 2013;10(9):e1001515.
35. Mahmoodi BK, Cushman M, Anne Næss I, et al. Association of Traditional Cardiovascular Risk Factors With Venous Thromboembolism: An Individual Participant Data Meta-Analysis of Prospective Studies. *Circulation.* 2017;135(1):7–16.
36. Di Nisio M, van Es N, Büller HR. Deep vein thrombosis and pulmonary embolism. *The Lancet.* 2016;

37. Heit JA, Phelps MA, Ward SA, et al. Familial segregation of venous thromboembolism. *J. Thromb. Haemost. JTH*. 2004;2(5):731–736.
38. Souto JC, Almasy L, Borrell M, et al. Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study. *Am. J. Hum. Genet.* 2000;67(6):1452–1459.
39. Larsen TB, Sørensen HT, Skytthe A, et al. Major genetic susceptibility for venous thromboembolism in men: a study of Danish twins. *Epidemiology*. 2003;14(3):328–332.
40. Tosetto A, Frezzato M, Rodeghiero F. Prevalence and risk factors of non-fatal venous thromboembolism in the active population of the VITA Project. *J. Thromb. Haemost.* 2003;1(8):1724–1729.
41. Noboa S, Le Gal G, Lacut K, et al. Family history as a risk factor for venous thromboembolism. *Thromb. Res.* 2008;122(5):624–629.
42. Bezemer ID, van der Meer FJ, Eikenboom JC, Rosendaal FR, Doggen CJ. The value of family history as a risk indicator for venous thrombosis. *Arch. Intern. Med.* 2009;169(6):610–615.
43. Zöller B, Li X, Sundquist J, Sundquist K. Parental history and venous thromboembolism: a nationwide study of age-specific and sex-specific familial risks in Sweden: Parental history and venous thromboembolism. *J. Thromb. Haemost.* 2011;9(1):64–70.
44. Sørensen HT, Riis AH, Diaz LJ, et al. Familial risk of venous thromboembolism: a nationwide cohort study: Familial risk and venous thromboembolism. *J. Thromb. Haemost.* 2011;9(2):320–324.
45. Zöller B, Ohlsson H, Sundquist J, Sundquist K. Familial risk of venous thromboembolism in first-, second- and third-degree relatives: a nationwide family study in Sweden: *Thromb. Haemost.* 2013;109(3):458–463.
46. Pernod G, Biron-Andreani C, Morange P-E, et al. Recommandations pour la recherche de facteurs biologiques de risque dans le cadre de la maladie thromboembolique veineuse. Recommandations GEHT–SFMV2008. *J. Mal. Vasc.* 2009;34(2):90.
47. DiScipio RG, Davie EW. Characterization of protein S, a gamma-carboxyglutamic acid containing protein from bovine and human plasma. *Biochemistry (Mosc.)*. 1979;18(5):899–904.

48. Denis CV, Roberts SJ, Hackeng TM, Lenting PJ. In vivo clearance of human protein S in a mouse model: influence of C4b-binding protein and the Heerlen polymorphism. *Arterioscler. Thromb. Vasc. Biol.* 2005;25(10):2209–2215.
49. Dahlbäck B. Protein S and C4b-binding protein: components involved in the regulation of the protein C anticoagulant system. *Thromb. Haemost.* 1991;66(1):49–61.
50. Stenberg Y, Linse S, Drakenberg T, Stenflo J. The high affinity calcium-binding sites in the epidermal growth factor module region of vitamin K-dependent protein S. *J. Biol. Chem.* 1997;272(37):23255–23260.
51. Villoutreix BO, García de Frutos P, Lövenklev M, et al. SHBG region of the anticoagulant cofactor protein S: secondary structure prediction, circular dichroism spectroscopy, and analysis of naturally occurring mutations. *Proteins.* 1997;29(4):478–491.
52. Esmon CT. The protein C pathway. *Chest.* 2003;124(3 Suppl):26S–32S.
53. Yegneswaran S, Smirnov MD, Safa O, et al. Relocating the active site of activated protein C eliminates the need for its protein S cofactor. A fluorescence resonance energy transfer study. *J. Biol. Chem.* 1999;274(9):5462–5468.
54. Nicolaes GAF, Dahlbäck B. Factor V and thrombotic disease: description of a janus-faced protein. *Arterioscler. Thromb. Vasc. Biol.* 2002;22(4):530–538.
55. Rosing J, Tans G. Coagulation factor V: an old star shines again. *Thromb. Haemost.* 1997;78(1):427–433.
56. He X, Shen L, Villoutreix BO, Dahlbäck B. Amino acid residues in thrombin-sensitive region and first epidermal growth factor domain of vitamin K-dependent protein S determining specificity of the activated protein C cofactor function. *J. Biol. Chem.* 1998;273(42):27449–27458.
57. Long GL, Lu D, Xie RL, Kalafatis M. Human protein S cleavage and inactivation by coagulation factor Xa. *J. Biol. Chem.* 1998;273(19):11521–11526.
58. Dahlbäck B, Hildebrand B, Malm J. Characterization of functionally important domains in human vitamin K-dependent protein S using monoclonal antibodies. *J. Biol. Chem.* 1990;265(14):8127–8135.
59. Van Wijnen M, Stam JG, Chang GT, et al. Characterization of mini-protein S, a recombinant variant of protein S that lacks the sex hormone binding globulin-like domain. *Biochem. J.* 1998;330 ( Pt 1):389–396.



60. Heeb MJ, Mesters RM, Tans G, Rosing J, Griffin JH. Binding of protein S to factor Va associated with inhibition of prothrombinase that is independent of activated protein C. *J. Biol. Chem.* 1993;268(4):2872–2877.
61. Hackeng TM, van 't Veer C, Meijers JC, Bouma BN. Human protein S inhibits prothrombinase complex activity on endothelial cells and platelets via direct interactions with factors Va and Xa. *J. Biol. Chem.* 1994;269(33):21051–21058.
62. Heeb MJ, Rosing J, Bakker HM, et al. Protein S binds to and inhibits factor Xa. *Proc. Natl. Acad. Sci. U. S. A.* 1994;91(7):2728–2732.
63. Castoldi E, Hackeng TM. Regulation of coagulation by protein S. *Curr. Opin. Hematol.* 2008;15(5):529–536.
64. Heeb MJ. Role of the PROS1 gene in thrombosis: lessons and controversies. *Expert Rev. Hematol.* 2008;1(1):9–12.
65. Heeb MJ, Kojima Y, Rosing J, Tans G, Griffin JH. C-terminal residues 621-635 of protein S are essential for binding to factor Va. *J. Biol. Chem.* 1999;274(51):36187–36192.
66. García de Frutos P, Fuentes-Prior P, Hurtado B, Sala N. Molecular basis of protein S deficiency. *Thromb. Haemost.* 2007;98(3):543–556.
67. Mahasandana C, Veerakul G, Tanphaichitr VS, et al. Homozygous protein S deficiency: 7-year follow-up. *Thromb. Haemost.* 1996;76(6):1122.
68. Miyata T, Kimura R, Kokubo Y, Sakata T. Genetic risk factors for deep vein thrombosis among Japanese: importance of protein S K196E mutation. *Int. J. Hematol.* 2006;83(3):217–223.
69. Suehisa E, Nomura T, Kawasaki T, Kanakura Y. Frequency of natural coagulation inhibitor (antithrombin III, protein C and protein S) deficiencies in Japanese patients with spontaneous deep vein thrombosis. *Blood Coagul. Fibrinolysis Int. J. Haemost. Thromb.* 2001;12(2):95–99.
70. Biguzzi E, Razzari C, Lane DA, et al. Molecular diversity and thrombotic risk in protein S deficiency: the PROSIT study. *Hum. Mutat.* 2005;25(3):259–269.
71. Pabinger I, Schneider B. Thrombotic risk in hereditary antithrombin III, protein C, or protein S deficiency. A cooperative, retrospective study. Gesellschaft für Thrombose- und Hamostaseforschung (GTH) Study Group on Natural Inhibitors. *Arterioscler. Thromb. Vasc. Biol.* 1996;16(6):742–748.

72. Simmonds RE, Ireland H, Lane DA, et al. Clarification of the risk for venous thrombosis associated with hereditary protein S deficiency by investigation of a large kindred with a characterized gene defect. *Ann. Intern. Med.* 1998;128(1):8–14.
73. Faioni EM, Valsecchi C, Palla A, et al. Free protein S deficiency is a risk factor for venous thrombosis. *Thromb. Haemost.* 1997;78(5):1343–1346.
74. Persson KEM, Dahlbäck B, Hillarp A. Diagnosing protein S deficiency: analytical considerations. *Clin. Lab.* 2003;49(3–4):103–110.
75. Meijer P, Haverkate F, Kluft C. Performance goals for the laboratory testing of antithrombin, protein C and protein S. *Thromb. Haemost.* 2006;96(5):584–589.
76. Makris M, Leach M, Beauchamp NJ, et al. Genetic analysis, phenotypic diagnosis, and risk of venous thrombosis in families with inherited deficiencies of protein S. *Blood.* 2000;95(6):1935–1941.
77. Henkens CM, Bom VJ, Van der Schaaf W, et al. Plasma levels of protein S, protein C, and factor X: effects of sex, hormonal state and age. *Thromb. Haemost.* 1995;74(5):1271–1275.
78. Aiach M, Borgel D, Gaussem P, et al. Protein C and protein S deficiencies. *Semin. Hematol.* 1997;34(3):205–216.
79. Woodward M, Lowe GD, Rumley A, et al. Epidemiology of coagulation factors, inhibitors and activation markers: The Third Glasgow MONICA Survey. II. Relationships to cardiovascular risk factors and prevalent cardiovascular disease. *Br. J. Haematol.* 1997;97(4):785–797.
80. Dykes AC, Walker ID, McMahon AD, Islam SI, Tait RC. A study of Protein S antigen levels in 3788 healthy volunteers: influence of age, sex and hormone use, and estimate for prevalence of deficiency state. *Br. J. Haematol.* 2001;113(3):636–641.
81. Sakata T, Okamoto A, Mannami T, Tomoike H, Miyata T. Prevalence of protein S deficiency in the Japanese general population: the Suita Study. *J. Thromb. Haemost. JTH.* 2004;2(6):1012–1013.
82. Alhenc-Gelas M, Canonico M, Morange PE, Emmerich J, Geht Genetic Thrombophilia Group. Protein S inherited qualitative deficiency: novel mutations and phenotypic influence. *J. Thromb. Haemost. JTH.* 2010;8(12):2718–2726.
83. Brouwer J-LP, Veeger NJGM, van der Schaaf W, Kluin-Nelemans HC, van der Meer J. Difference in absolute risk of venous and arterial thrombosis between familial protein S

- deficiency type I and type III. Results from a family cohort study to assess the clinical impact of a laboratory test-based classification. *Br. J. Haematol.* 2005;128(5):703–710.
84. Libourel EJ, Bank I, Veeger NJGM, et al. Protein S type III deficiency is no risk factor for venous and arterial thromboembolism in 168 thrombophilic families: a retrospective study. *Blood Coagul. Fibrinolysis Int. J. Haemost. Thromb.* 2005;16(2):135–140.
  85. Espinosa-Parrilla Y, Morell M, Souto JC, et al. Absence of linkage between type III protein S deficiency and the PROS1 and C4BP genes in families carrying the protein S Heerlen allele. *Blood.* 1997;89(8):2799–2806.
  86. Brunet D, Barthet MC, Morange PE, et al. Protein S deficiency: different biological phenotypes according to the assays used. *Thromb. Haemost.* 1998;79(2):446–447.
  87. Bertina RM, Ploos van Amstel HK, van Wijngaarden A, et al. Heerlen polymorphism of protein S, an immunologic polymorphism due to dimorphism of residue 460. *Blood.* 1990;76(3):538–548.
  88. Beauchamp NJ, Dykes AC, Parikh N, Campbell Tait R, Daly ME. The prevalence of, and molecular defects underlying, inherited protein S deficiency in the general population. *Br. J. Haematol.* 2004;125(5):647–654.
  89. Pintao MC, Ribeiro DD, Bezemer ID, et al. Protein S levels and the risk of venous thrombosis: results from the MEGA case-control study. *Blood.* 2013;122(18):3210–3219.
  90. Lijfering WM, Mulder R, ten Kate MK, et al. Clinical relevance of decreased free protein S levels: results from a retrospective family cohort study involving 1143 relatives. *Blood.* 2009;113(6):1225–1230.
  91. Schmidel DK, Tatro AV, Phelps LG, Tomczak JA, Long GL. Organization of the human protein S genes. *Biochemistry (Mosc.).* 1990;29(34):7845–7852.
  92. Ploos van Amstel HK, Reitsma PH, van der Logt CP, Bertina RM. Intron-exon organization of the active human protein S gene PS alpha and its pseudogene PS beta: duplication and silencing during primate evolution. *Biochemistry (Mosc.).* 1990;29(34):7853–7861.
  93. Gandrille S, Borgel D, Sala N, et al. Protein S deficiency: a database of mutations--summary of the first update. *Thromb. Haemost.* 2000;84(5):918.
  94. Kimura R, Honda S, Kawasaki T, et al. Protein S-K196E mutation as a genetic risk factor for deep vein thrombosis in Japanese patients. *Blood.* 2006;107(4):1737–1738.

95. Liu W, Yin T, Okuda H, et al. Protein S K196E mutation, a genetic risk factor for venous thromboembolism, is limited to Japanese. *Thromb. Res.* 2013;132(2):314–315.
96. Espinosa-Parrilla Y, Morell M, Souto JC, et al. Protein S gene analysis reveals the presence of a cosegregating mutation in most pedigrees with type I but not type III PS deficiency. *Hum. Mutat.* 1999;14(1):30–39.
97. Johansson AM, Hillarp A, Säll T, et al. Large deletions of the PROS1 gene in a large fraction of mutation-negative patients with protein S deficiency. *Thromb. Haemost.* 2005;94(5):951–957.
98. Pintao MC, Garcia AA, Borgel D, et al. Gross deletions/duplications in PROS1 are relatively common in point mutation-negative hereditary protein S deficiency. *Hum. Genet.* 2009;126(3):449–456.
99. Duchemin J, Gandrille S, Borgel D, et al. The Ser 460 to Pro substitution of the protein S alpha (PROS1) gene is a frequent mutation associated with free protein S (type IIa) deficiency. *Blood.* 1995;86(9):3436–3443.
100. Espinosa-Parrilla Y, Navarro G, Morell M, et al. Homozygosity for the protein S Heerlen allele is associated with type I PS deficiency in a thrombophilic pedigree with multiple risk factors. *Thromb. Haemost.* 2000;83(1):102–106.
101. Wypasek E, Potaczek DP, Płonka J, Alhenc-Gelas M, Undas A. Protein S deficiency and Heerlen polymorphism in a Polish patient with acute myocardial infarction and previous venous thromboembolism. *Thromb. Res.* 2013;132(6):776–777.
102. Morboeuf O, Borgel D, Aiach M, et al. Expression and characterization of recombinant protein S with the Ser 460 Pro mutation. *Thromb. Res.* 2000;100(1):81–88.
103. Koenen RR, Gomes L, Tans G, Rosing J, Hackeng TM. The Ser460Pro mutation in recombinant protein S Heerlen does not affect its APC-cofactor and APC-independent anticoagulant activities. *Thromb. Haemost.* 2004;91(6):1105–1114.
104. Germain M, Saut N, Greliche N, et al. Genetics of Venous Thrombosis: Insights from a New Genome Wide Association Study. *PLoS ONE.* 2011;6(9):e25581.
105. Sudmant PH, Rausch T, Gardner EJ, et al. An integrated map of structural variation in 2,504 human genomes. *Nature.* 2015;526(7571):75–81.
106. Ireland H, Konstantoulas CJ, Cooper JA, et al. EPCR Ser219Gly: elevated sEPCR, prothrombin F1+2, risk for coronary heart disease, and increased sEPCR shedding in vitro. *Atherosclerosis.* 2005;183(2):283–292.

107. Reiner AP, Carty CL, Jenny NS, et al. *PROC*, *PROCR* and *PROS1* polymorphisms, plasma anticoagulant phenotypes, and risk of cardiovascular disease and mortality in older adults: the Cardiovascular Health Study. *J. Thromb. Haemost.* 2008;6(10):1625–1632.
108. Saposnik B, Reny J-L, Gaussem P, 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):1311–1318.
109. Uitte de Willige S, Van Marion V, Rosendaal FR, et al. Haplotypes of the EPCR gene, plasma sEPCR levels and the risk of deep venous thrombosis. *J. Thromb. Haemost.* 2004;2(8):1305–1310.
110. Medina P, Navarro S, Estellés A, et al. Contribution of polymorphisms in the endothelial protein C receptor gene to soluble endothelial protein C receptor, circulating activated protein C levels and thrombotic risk. *Thromb. Haemost.* 2004;
111. Tang W, Basu S, Kong X, et al. Genome-wide association study identifies novel loci for plasma levels of protein C: the ARIC study. *Blood.* 2010;116(23):5032–5036.
112. Ghosh S, Pendurthi UR, Steinoe A, Esmon CT, Rao LVM. Endothelial Cell Protein C Receptor Acts as a Cellular Receptor for Factor VIIa on Endothelium. *J. Biol. Chem.* 2007;282(16):11849–11857.
113. Ireland HA, Cooper JA, Drenos F, et al. FVII, FVIIa, and downstream markers of extrinsic pathway activation differ by EPCR Ser219Gly variant in healthy men. *Arterioscler. Thromb. Vasc. Biol.* 2009;29(11):1968–1974.
114. Dennis J, Johnson CY, Adediran AS, et al. The endothelial protein C receptor (PROCR) Ser219Gly variant and risk of common thrombotic disorders: a HuGE review and meta-analysis of evidence from observational studies. *Blood.* 2012;119(10):2392–2400.
115. Saposnik B. A haplotype of the EPCR gene is associated with increased plasma levels of sEPCR and is a candidate risk factor for thrombosis. *Blood.* 2003;103(4):1311–1318.
116. van Hylckama Vlieg A, Rosendaal FR. High levels of fibrinogen are associated with the risk of deep venous thrombosis mainly in the elderly. *J. Thromb. Haemost. JTH.* 2003;1(12):2677–2678.
117. Koster T, Rosendaal FR, Reitsma PH, et al. Factor VII and fibrinogen levels as risk factors for venous thrombosis. A case-control study of plasma levels and DNA polymorphisms--the Leiden Thrombophilia Study (LETS). *Thromb. Haemost.* 1994;71(6):719–722.

118. Uitte de Willige S. Genetic variation in the fibrinogen gamma gene increases the risk for deep venous thrombosis by reducing plasma fibrinogen ' levels. *Blood*. 2005;106(13):4176–4183.
119. de Willige SU, Pyle ME, Vos HL, et al. Fibrinogen gamma gene 3'-end polymorphisms and risk of venous thromboembolism in the African-American and Caucasian population. *Thromb. Haemost.* 2009;
120. Pérez-Ceballos E, Corral J, Alberca I, et al. Prothrombin A19911G and G20210A polymorphisms' role in thrombosis. *Br. J. Haematol.* 2002;118(2):610–614.
121. Martinelli I, Battaglioli T, Tosetto A, et al. Prothrombin A19911G polymorphism and the risk of venous thromboembolism. *J. Thromb. Haemost.* 2006;4(12):2582–2586.
122. Ceelie H, Bertina RM, van Hylckama Vlieg A, Rosendaal FR, Vos HL. Polymorphisms in the prothrombin gene and their association with plasma prothrombin levels. *Thromb. Haemost.* 2001;85(6):1066–1070.
123. Smith NL, Hindorff LA, Heckbert SR, et al. Association of genetic variations with nonfatal venous thrombosis in postmenopausal women. *JAMA*. 2007;297(5):489–498.
124. Bezemer ID. Updated Analysis of Gene Variants Associated With Deep Vein Thrombosis. *JAMA*. 2010;303(5):421.
125. Germain M, Chasman DI, de Haan H, et al. Meta-analysis of 65,734 Individuals Identifies TSPAN15 and SLC44A2 as Two Susceptibility Loci for Venous Thromboembolism. *Am. J. Hum. Genet.* 2015;
126. Daly M, O'Meara A, Hallinan FM. Identification and characterization of a new antithrombin III familial variant (AT Dublin) with possible increased frequency in children with cancer. *Br. J. Haematol.* 1987;65(4):457–462.
127. Bezemer ID, Bare LA, Doggen CJM, et al. Gene variants associated with deep vein thrombosis. *JAMA*. 2008;299(11):1306–1314.
128. Antón AI, Teruel R, Corral J, et al. Functional consequences of the prothrombotic SERPINC1 rs2227589 polymorphism on antithrombin levels. *Haematologica*. 2009;94(4):589–592.
129. Li Y, Bezemer ID, Rowland CM, et al. Genetic variants associated with deep vein thrombosis: the *F11* locus. *J. Thromb. Haemost.* 2009;7(11):1802–1808.

130. Anton AI, Teruel R, Corral J, et al. Functional consequences of the prothrombotic SERPINC1 rs2227589 polymorphism on antithrombin levels. *Haematologica*. 2009;94(4):589–592.
131. Segers O, Simioni P, Tormene D, Castoldi E. Influence of single nucleotide polymorphisms on thrombin generation in factor V Leiden heterozygotes: *Thromb. Haemost.* 2013;111(3):438–446.
132. Austin H, De Staercke C, Lally C, et al. New gene variants associated with venous thrombosis: a replication study in White and Black Americans: Gene variants and venous thrombosis: a replication study. *J. Thromb. Haemost.* 2011;9(3):489–495.
133. Massberg S, Gawaz M, Grüner S, et al. A Crucial Role of Glycoprotein VI for Platelet Recruitment to the Injured Arterial Wall In Vivo. *J. Exp. Med.* 2003;197(1):41–49.
134. Snoep JD, Gaussem P, Eikenboom JCJ, 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: Functional and clinical effects of GP6. *J. Thromb. Haemost.* 2010;8(11):2377–2384.
135. Trégouët D-A, Heath S, Saut N, 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):5298–5303.
136. Meijers JC, Tekelenburg WL, Bouma BN, Bertina RM, Rosendaal FR. High levels of coagulation factor XI as a risk factor for venous thrombosis. *N. Engl. J. Med.* 2000;342(10):696–701.
137. Germain M, Saut N, Greliche N, et al. Genetics of Venous Thrombosis: Insights from a New Genome Wide Association Study. *PLoS ONE*. 2011;6(9):e25581.
138. Middelburg RA, van Stein D, Briët E, van der Bom JG. The role of donor antibodies in the pathogenesis of transfusion-related acute lung injury: a systematic review. *Transfusion (Paris)*. 2008;48(10):2167–2176.
139. Prox J, Willenbrock M, Weber S, et al. Tetraspanin15 regulates cellular trafficking and activity of the ectodomain sheddase ADAM10. *Cell. Mol. Life Sci. CMLS*. 2012;69(17):2919–2932.
140. Facey A, Pinar I, Arthur JF, et al. A-Disintegrin-And-Metalloproteinase (ADAM) 10 Activity on Resting and Activated Platelets. *Biochemistry (Mosc.)*. 2016;55(8):1187–1194.

141. Smith NL, Chen M-H, Dehghan A, et al. Novel Associations of Multiple Genetic Loci With Plasma Levels of Factor VII, Factor VIII, and von Willebrand Factor: The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. *Circulation*. 2010;121(12):1382–1392.
142. Smith NL, Rice KM, Bovill EG, et al. Genetic variation associated with plasma von Willebrand factor levels and the risk of incident venous thrombosis. *Blood*. 2011;117(22):6007–6011.
143. van Loon JE, Sanders YV, de Wee EM, et al. Effect of Genetic Variation in STXBP5 and STX2 on von Willebrand Factor and Bleeding Phenotype in Type 1 von Willebrand Disease Patients. *PLoS ONE*. 2012;7(7):e40624.
144. Huang J, Huffman JE, Yamakuchi M, et al. Genome-Wide Association Study for Circulating Tissue Plasminogen Activator Levels and Functional Follow-Up Implicates Endothelial STXBP5 and STX2. *Arterioscler. Thromb. Vasc. Biol*. 2014;34(5):1093–1101.
145. Ye S, Huang Y, Joshi S, et al. Platelet secretion and hemostasis require syntaxin-binding protein STXBP5. *J. Clin. Invest*. 2014;124(10):4517–4528.
146. Zhu Q, Yamakuchi M, Ture S, et al. Syntaxin-binding protein STXBP5 inhibits endothelial exocytosis and promotes platelet secretion. *J. Clin. Invest*. 2014;124(10):4503–4516.
147. Lowe GD, Haverkate F, Thompson SG, et al. Prediction of deep vein thrombosis after elective hip replacement surgery by preoperative clinical and haemostatic variables: the ECAT DVT Study. European Concerted Action on Thrombosis. *Thromb. Haemost.* 1999;81(6):879–886.
148. Tripodi A. A shortened activated partial thromboplastin time is associated with the risk of venous thromboembolism. *Blood*. 2004;104(12):3631–3634.
149. Houlihan LM, Davies G, Tenesa A, et al. Common Variants of Large Effect in F12, KNG1, and HRG Are Associated with Activated Partial Thromboplastin Time. *Am. J. Hum. Genet*. 2010;86(4):626–631.
150. Morange P-E, Oudot-Mellakh T, Cohen W, et al. KNG1 Ile581Thr and susceptibility to venous thrombosis. *Blood*. 2011;117(13):3692–3694.
151. Gailani D, Renné T. The intrinsic pathway of coagulation: a target for treating thromboembolic disease? *J. Thromb. Haemost.* 2007;5(6):1106–1112.



152. Trégouët D-A, Morange P-E. What is currently known about the genetics of venous thromboembolism at the dawn of next generation sequencing technologies. *Br. J. Haematol.* 2017;
153. Baglin T, Luddington R, Brown K, Baglin C. Incidence of recurrent venous thromboembolism in relation to clinical and thrombophilic risk factors: prospective cohort study. *Lancet Lond. Engl.* 2003;362(9383):523–526.
154. Coppens M, Reijnders JH, Middeldorp S, Doggen CJM, Rosendaal FR. Testing for inherited thrombophilia does not reduce the recurrence of venous thrombosis. *J. Thromb. Haemost. JTH.* 2008;6(9):1474–1477.
155. Connors JM. Thrombophilia Testing and Venous Thrombosis. *N. Engl. J. Med.* 2017;377(12):1177–1187.
156. Middeldorp S. Inherited thrombophilia: a double-edged sword. *Hematol. Am. Soc. Hematol. Educ. Program.* 2016;2016(1):1–9.
157. van Vlijmen EFW, Wiewel-Verschueren S, Monster TBM, Meijer K. Combined oral contraceptives, thrombophilia and the risk of venous thromboembolism: a systematic review and meta-analysis. *J. Thromb. Haemost. JTH.* 2016;14(7):1393–1403.
158. Mahmoodi BK, Brouwer J-LP, Veeger NJGM, van der Meer J. Hereditary deficiency of protein C or protein S confers increased risk of arterial thromboembolic events at a young age: results from a large family cohort study. *Circulation.* 2008;118(16):1659–1667.
159. de la Morena-Barrio ME, Sevivas TS, Martinez-Martinez I, et al. Congenital disorder of glycosylation (PMM2-CDG) in a patient with antithrombin deficiency and severe thrombophilia. *J. Thromb. Haemost. JTH.* 2012;10(12):2625–2627.
160. Caspers M, Pavlova A, Driesen J, et al. Deficiencies of antithrombin, protein C and protein S - practical experience in genetic analysis of a large patient cohort. *Thromb. Haemost.* 2012;108(2):247–257.
161. Navarro-Fernández J, de la Morena-Barrio ME, Padilla J, et al. Antithrombin Dublin (p.Val30Glu): a relatively common variant with moderate thrombosis risk of causing transient antithrombin deficiency. *Thromb. Haemost.* 2016;116(1):146–154.
162. Seligsohn U. Factor XI deficiency. *Thromb. Haemost.* 1993;70(1):68–71.
163. Gailani D, Broze GJ. Factor XI activation in a revised model of blood coagulation. *Science.* 1991;253(5022):909–912.

164. Von dem Borne PA, Bajzar L, Meijers JC, Nesheim ME, Bouma BN. Thrombin-mediated activation of factor XI results in a thrombin-activatable fibrinolysis inhibitor-dependent inhibition of fibrinolysis. *J. Clin. Invest.* 1997;99(10):2323–2327.
165. Grünbacher G, Weger W, Marx-Neuhold E, et al. The fibrinogen gamma (FGG) 10034C>T polymorphism is associated with venous thrombosis. *Thromb. Res.* 2007;121(1):33–36.
166. Chung DW, Davie EW. gamma and gamma' chains of human fibrinogen are produced by alternative mRNA processing. *Biochemistry (Mosc.)*. 1984;23(18):4232–4236.
167. Jick H, Slone D, Westerholm B, et al. Venous thromboembolic disease and ABO blood type. A cooperative study. *Lancet Lond. Engl.* 1969;1(7594):539–542.
168. Wu O, Bayoumi N, Vickers MA, Clark P. ABO(H) blood groups and vascular disease: a systematic review and meta-analysis. *J. Thromb. Haemost. JTH.* 2008;6(1):62–69.
169. Lidegaard Ø, Nielsen LH, Skovlund CW, Skjeldstad FE, Løkkegaard E. Risk of venous thromboembolism from use of oral contraceptives containing different progestogens and oestrogen doses: Danish cohort study, 2001–9. *BMJ.* 2011;343:d6423.
170. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747–753.
171. Friso S, Lotto V, Choi S-W, et al. Promoter methylation in coagulation F7 gene influences plasma FVII concentrations and relates to coronary artery disease. *J. Med. Genet.* 2012;49(3):192–199.
172. El-Maarri O, Becker T, Junen J, et al. Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. *Hum. Genet.* 2007;122(5):505–514.

# **Liste des publications relatives au sujet de la thèse**

Cohen W, Castelli C, **Suchon P**, Bouvet S, Aillaud MF, Brunet D, Barthet MC, Alessi MC, Trégouët DA, Morange PE. Risk assessment of venous thrombosis in families with known hereditary thrombophilia: the MARseilles-Nîmes prediction model. *J Thromb Haemost.* 2014;12:138-146.

Rocañín-Arjón A, Dennis J, **Suchon P**, Aïssi D, Truong V, Trégouët DA, Gagnon F, Morange PE. Thrombin generation potential and whole-blood DNA methylation. *Thromb Res.* 2015;135:561-4.

Germain M, Chasman DI, de Haan H, Tang W, Lindström S, Weng LC, de Andrade M, de Visser MC, Wiggins KL, **Suchon P**, Saut N, Smadja DM, Le Gal G, van Hylckama Vlieg A, Di Narzo A, Hao K, Nelson CP, Rocanin-Arjo A, Folkersen L, Monajemi R, Rose LM, Brody JA, Slagboom E, Aïssi D, Gagnon F, Deleuze JF, Deloukas P, Tzourio C, Dartigues JF, Berr C, Taylor KD, Civelek M, Eriksson P; Cardiogenics Consortium, Psaty BM, Houwing-Duitermaat J, Goodall AH, Cambien F, Kraft P, Amouyel P, Samani NJ, Basu S, Ridker PM, Rosendaal FR, Kabrhele C, Folsom AR, Heit J, Reitsma PH, Trégouët DA, Smith NL, Morange PE. Meta-analysis of 65,734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism. *Am J Hum Genet.* 2015;96:532-42.

Morange PE, **Suchon P**, Trégouët DA. Genetics of Venous Thrombosis: update in 2015. *Thromb Haemost.* 2015;114:910-9.

**Suchon P**, Al Frouh F, Henneuse A, Ibrahim M, Brunet D, Barthet MC, Aillaud MF, Venton G, Alessi MC, Trégouët DA, Morange PE. Risk factors for venous thromboembolism in women under combined oral contraceptive. The PILI Genetic Risk Monitoring (PILGRIM) Study. *Thromb Haemost.* 2016;115:135-42.

Alhenc-Gelas M, Plu-Bureau G, Horellou MH, Rauch A, **Suchon P**; GEHT genetic thrombophilia group. PROS1 genotype phenotype relationships in a large cohort of adults with suspicion of inherited quantitative protein S deficiency. *Thromb Haemost.* 2016;115:570-9.

**Suchon P**, Al Frouh F, Ibrahim M, Sarlon G, Venton G, Alessi MC, Trégouët DA, Morange PE. Genetic risk factors for venous thrombosis in women using combined oral contraceptives: update of the PILGRIM study. *Clin Genet.* 2017;91:131-136.

Sennblad B, Basu S, Mazur J, **Suchon P**, Martinez-Perez A, van Hylckama Vlieg A, Truong V, Li Y, Gådin JR, Tang W, Grossman V, de Haan HG, Handin N, Silveira A, Souto JC, Franco-Cereceda A, Morange PE, Gagnon F, Soria JM, Eriksson P, Hamsten A, Maegdefessel L, Rosendaal FR, Wild P, Folsom AR, Trégouët DA, Sabater-Lleal M. Genome-wide association study with additional genetic and post-transcriptional analyses reveals novel regulators of plasma factor XI levels. *Hum Mol Genet.* 2017;26:637-649.

**P. Suchon**, M. Germain, A. Delluc, D. Smadja, X. Jouven, B. Gyorgy, N. Saut, M. Ibrahim, J. F. Deleuze, M. C. Alessi, P. E. Morange, & D. A. Trégouët. Protein S Heerlen mutation heterozygosity is associated with venous thrombosis risk. *Scientific reports.* 2017. 2017;7:45507

McDaid A, Logette E, Buchillier V, Muriset M, **Suchon P**, Pache TD, Tanackovic G, Kutalik Z, Michaud J. Risk prediction of developing venous thrombosis in combined oral contraceptive users. *PLoS One.* 2017;12:e0182041.