Published in final edited form as:

Curr Opin Hematol. 2015 September ; 22(5): 428–436. doi:10.1097/MOH.000000000000165.

Dissecting the Genetic Determinants of Hemostasis and Thrombosis

Karl C. Desch

Department of Pediatrics and Communicable Disease, University of Michigan, Ann Arbor, MI 48109

Abstract

Purpose of Review—New DNA genotyping and sequencing technologies have facilitated the rapid advancement in our knowledge of human genomic variation and a search for the heritable determinants of complex genetic traits. This review highlights findings from recent genetic studies of complex traits primarily related to venous thromboembolism and provides tools to understand and interpret genome-wide association studies and next generation sequencing studies.

Recent Findings—Genome-wide studies of venous thromboembolic disease and the variation of the protein components of the hemostatic system have been reported. The results of these studies have suggested that variants in a diverse set of known and new genes contribute to the heritability of these traits but that many of the genetic determinants of these traits still remain undiscovered.

Summary—Next generation sequencing studies and functional studies of the gene loci that contribute to hemostatic traits are currently underway. Future studies that explore the role of rare genetic variants, regulatory elements of the genome and gene-gene interactions will be required for a more complete understanding of the genetic control of the hemostatic system and for the application of this knowledge to the care of patients with disorders of thrombosis and hemostasis.

Keywords

Venous thromboembolism; genome-wide association study; complex genetic traits; von Willebrand Factor; Factor V Leiden

Introduction

A closed circulation requires a regulated and redundant system to respond to blood vessel injury and prevent life-threatening hemorrhage. The maintenance of hemostasis in humans is achieved through the controlled interaction of many components of the blood and vessel wall.(1, 2) Our current understanding of the numerous genetic loci contributing to the hemostatic balance has been achieved through biochemical studies and through family studies that led to the mapping and sequence identification of genes encoding components of

Corresponding Author Address: Karl C. Desch, Department of Pediatrics and Communicable Diseases, Division of Neonatal-Perinatal Medicine, University of Michigan C.S. Mott Children's Hospital, 8-621 C&W Mott Hospital, SPC 4254, 1540 E. Hospital Drive, Ann Arbor, MI 48019-4254, Ph (734) 763-4109, kdesch@umich.edu.

Conflicts of Interest: None

the clotting cascade, anticoagulant and fibrinolytic systems. Therefore, unlike many physiologic systems, most of the genes encoding critical components of the hemostatic system have been well described. But the underlying genetic risk factors for common vascular disorders, such as venous thromboembolic disease (VTE), myocardial infarction or stroke, are not the result of single gene defects. Rather, these problems are the result of a complex interaction between gene variants, environmental exposures and chance. For example, VTE is a common disorder affecting >900,000 individuals a year in the United States. Family studies suggest that > 60% of the variation in susceptibility to venous thrombosis is attributable to genetic factors (3). Yet, monogenic disorders associated with familial thrombosis or “thrombophilia”, such as protein C deficiency, antithrombin III deficiency, or protein S deficiency, are rare and cannot explain the vast majority of unprovoked (not associated with known environmental risk factors for thrombosis) VTE. Pioneering work by several labs identified common genetic risk factors for VTE in people of European ancestry such as the 3' UTR prothrombin G20210A mutation and the R506Q mutation in coagulation factor V (FV Leiden).(4–6) Although these polymorphisms are associated with an increased risk in an individual's lifetime risk of VTE, they do not appear to be useful for medical decision making.(7, 8) In Asian populations, where the FV Leiden and F2 G20210A variants are rare, mutations in *PROC* and *PROS1* may play a more dominant role.(9) Elevations in the plasma levels of several coagulation factors, including factor VIII (10, 11), factor IX and factor XI have also been associated with increased thrombosis susceptibility (2, 12), although the specific causal environmental and genetic factors resulting in these elevated levels are unknown. Importantly, investigators have been unable to stitch together a coherent landscape of thrombophilia based on the numerous single gene association studies due to the difficulty of combining independent studies of different genes. The promise of adequately powered genome-wide association studies was the ability to judge associations across the entire genome and rank order their effects in a population.

GWAS Strengths and Limitations

Since 2005, hundreds of genome-wide association studies (GWAS) of complex traits affecting human health have been reported.(13, 14) These studies capitalize on the increasing speed and lower cost of genotyping human DNA samples and are powerful ways of connecting common human genetic variation to heritable traits determined by a number of different genes. GWAS test millions of genotyped and imputed single-nucleotide polymorphisms (SNPs) in large cohorts of people to detect any association with the quantitative phenotype or dichotomous status. The principle advantage of GWAS is the ability to test common genetic variants across the entire genome as opposed to single or limited candidate gene association tests, which are more prone to false positive results. The genome-wide context of GWAS allows for a ranked interpretation of association signals and the testing of variants without a preconceived bias. The results of these GWAS studies in VTE have helped investigators understand which common genetic variants play the most dominant role in determining the VTE risk. GWAS are very powerful ways to detect SNPs with a statistically significant difference in allele frequency in a case population compared to a control population. Quantitative traits (blood pressure, height, factor VIII levels) can also

be analyzed by GWAS to identify SNP associations referred to as quantitative trait loci (QTLs). Since, at their core, GWAS report a summary of a series of thousands of independent SNP association tests, they are able to detect locus heterogeneity, where multiple genes contribute to phenotypes, without any loss of power. Likewise, to avoid false positive associations, a multiple testing correction must be employed when interpreting GWAS results. When groups of GWAS are analyzed together in meta-analyses the increase in power allows even weaker associations to be detected.

Despite their power, multiple limitations to even large GWAS must be considered. The power to detect an association is determined by both the allele frequency and the magnitude of difference in allele distributions in case and control populations. For quantitative traits, the effect direction and the strength of a SNP association can be expressed by the beta value (the average change in a quantitative trait per allele count of a SNP). The principle limitation of GWAS involves the reduced power to detect associations with rare variants, mutations that may have a large effect size in an individual but have a low population frequency. These rare variants are also more likely to produce false positives and are often removed from the genotyping set before performing a GWAS. Even studies in large cohorts have a limited power to detect any association with variants that have allele frequencies below 1–5%. The second important limitation is determined by the SNPs used on the genotyping chip. It is often difficult or impossible to directly identify the functional mutations driving the association from the tested SNPs in a GWAS. SNPs mark a stretch of DNA sequence, in linkage disequilibrium, meaning that they are statistically likely to be inherited together. So the significant SNP associations mark a block of DNA rather than a specific mutation influencing a phenotype. This problem of association without identification of functional genomic variants is somewhat alleviated by the ever expanding catalog of SNP variants and whole-genome sequencing data that can be used to impute the non-genotyped variants in an individual and the use of gene-centric SNP chips that include more non-synonymous variants.(15) However, imputation has a very limited ability to shed light on the influence of low frequency variants because these rare variants are usually not in LD with common SNPs, even by chance.(16)

GWAS of VTE

Recent genetic studies of VTE are an excellent example of the power and limitations of GWAS and are summarized in Table 1. Although previous studies had linked elevated Factor VIII, ABO blood group, FV Leiden, prothrombin G20210A mutation and several other gene variants with the risk for VTE,(17) the first truly genome-wide study for VTE was published in 2009 by Tregouet et al.(18) Here, the authors studied 453 cases and 1327 controls and demonstrated that among common genetic variants, only SNPs tagging variants in *ABO* and *F5* were associated with VTE, Table 1. Despite numerous previously published single gene association studies and meta-analyses purporting to associate common variants in genes like *SERPINE1* 4G/5G polymorphism or *MTHFR* C677T polymorphism with VTE, (19) the GWAS uncovered these previous associations as either statistical false positives or too weakly associated with VTE to be clinically relevant. This study was relatively modest in size compared to many GWAS of the time, but it still had >80% power to detect associations of SNPs with allele frequencies greater than 5% and odds ratios greater than

2.0, (*MTHFR*, rs1801133, minor allele frequency of 31% in CEU Hapmap). On the other hand, the prothrombin G20210A mutation, with an allele frequency of ~1.6% in European populations (CEU Hapmap), was not included on the genotyping chip used in the study and was not in LD with any of the genotyped SNPs. Therefore, the association could not be tested. However, given the low allele frequency of this variant and the modest size of this study, it is likely that no association would have been detected. Subsequent GWAS studies for VTE have addressed the power limitations of this initial study, Table 1.(20–22) In fact, a recent meta-analysis of VTE examined 7,507 VTE cases and 52,632 controls and performed replication in an additional 3,009 cases and 2,586 controls.(23) In this study, six loci previously associated with VTE were confirmed (*ABO*, *F2*, *F5*, *F11*, *FGG* and *PROCR*) and two new loci were identified and confirmed in replication in SNPs with high allele frequencies tagging the *TSPAN15* and the *SLC44A2* loci. Odds ratios for the associated risk alleles ranged from high of 3.25 (2.91–3.64) for FV Leiden to a low of 1.15 (1.10–1.21) for a common variant in *PROCR*, rs6087685, allele frequency ~15% in Europeans (CEU-Hapmap). As expected, no associations with variants in *SERPINE1* or *MTHFR* were identified in this large meta-analysis. So while larger GWAS of VTE have identified several genes where variants play a role altering the risk for VTE, no new “smoking guns” with effect sizes comparable to FV Leiden have been identified and a significant proportion of the heritable risk for VTE remains unexplained.

Examining the biochemical components of VTE

Since genetic studies are very dependent on the quality of the phenotyping, many investigators have pursued more specific biochemical phenotypes in order to further dissect the genetic determinants of hemostasis. While VTE itself may be a relatively straightforward clinical diagnosis, on the biologic level, VTE may have multiple underlying etiologies that cause an increase in the heterogeneity between cases and decrease the power of an association study. By concentrating on more specific quantitative phenotypes, such as the level of a protein component of the hemostatic system, investigators have been able to use unselected populations (not case-control cohorts) that are much easier to assemble into very large studies. Through more specific quantitative traits and larger cohorts, investigators have been able to find genetic associations with biomarkers that may further inform important underlying genetic determinants of VTE and other clinical disorders of hemostasis.

For example, multiple investigators have employed genomic studies to further determine the genetic determinants of von Willebrand Factor (VWF) levels.(24–26) VWF facilitates the interaction of platelets with sites of vascular injury and is a critical mediator of the initial platelet plug in flowing blood. Additionally, as a carrier protein for coagulation Factor VIII, levels of VWF are highly correlated with FVIII activity. VWF levels are highly heritable (~65%) and vary 5-fold in healthy populations. Individuals with elevated VWF have increased risk for VTE while individuals with low VWF levels are at higher risk for type I von Willebrand Disease, a common genetic disorder of bleeding.(27) Previous studies made the connection between variation in *VWF* and ABO blood group serotypes.(28) Populations with ABO type O blood have lower VWF levels (and FVIII activity) than people with type A or B serotypes. In fact, ~20% of the variation in VWF levels can be explained by ABO blood types. It is likely that the *ABO* signal generated in the multiple VTE GWAS, Table 1,

is driven by the influence of *ABO* variants on VWF and Factor VIII levels. In 2010, Smith et al reported the first genome-wide association meta-analysis for VWF in a study of ~17,596 people in several cardiovascular disease related cohorts, the CHARGE consortium.(26) This study confirmed the influence of common *ABO* haplotypes on VWF levels and identified the *VWF* locus as well as six new loci (*STXBP5*, *SCARA5*, *STAB2*, *STX2*, *TC2N*, *CLEC4M*) as genetic determinants of VWF levels, Table 2. In a follow up study, Smith et al examined the top SNPs from the VWF/Factor VIII meta-analysis in a VTE case-control cohort. In addition to the known associations with the *ABO* haplotypes, this study uncovered associations with incident VTE and variants in *VWF* and *STXBP5* that were presumably driven by altered VWF/FVIII levels.(29) In a subsequent functional study, Zhu et al examined the interaction of *STXBP5* and VWF and demonstrated increased VWF secretion from human endothelial cells with reduced *STXBP5* and found that mice deficient in *STXBP5* had higher VWF levels than littermate controls.(30) In 2013, Rydz et al reported that similar to *ABO* type A and B serotypes, specific *CLEC4M* polymorphisms associated with increased levels of VWF were under-represented in patients with Type 1 VWD compared to their unaffected family members.(31) Conversely, some VWF GWAS variants that are associated with decreased VWF levels (*ABO*, *CLEC4M*, *STXBP5*) seem to have higher allele frequencies in Type 1 VWD populations.(32)

Do Linkage Studies still have a role in Complex Traits?

As discussed above, even well powered GWAS cannot detect associations with rare variants. In 2012, our group reported the results of a smaller meta-analysis of VWF levels in two young healthy cohorts, N= 3,462. We analyzed a sibling cohort in order to perform both GWAS and linkage studies. Although much less powerful than association studies, linkage studies are not blind to rare variants clustered into loci with high allelic heterogeneity. We hypothesized that these types of loci, where single nucleotide variants are individually rare but in aggregate common, would accrue linkage signals across affected families and may explain a portion of the missing heritability for complex traits.(25) In our GWAS, common variants in *ABO* and *VWF* explained about 18% of the variation in VWF levels. Linkage analysis detected a strong signal near the centromere of chromosome 2 that collectively explained about 19% of the variance in VWF levels. The chromosome 2 linkage interval had not been detected in a previous larger pedigree based study,(33) but was identified in a linkage study for FVIII levels in a French-Canadian VTE enriched pedigree.(20) This linkage interval was too large (>20 megabases) to usefully pick out candidate genes although the interval included several genes involved in post-translational modifications that could affect clearance of VWF and SNARE complex proteins that may modify the secretion rate of VWF. Follow up linkage studies to narrow the linkage interval and DNA sequencing of top candidate genes will be required to better understand the functional mutations driving this linkage signal and their connection to VWF and VTE phenotypes.

A Survey of GWAS studies in Thrombosis and Hemostasis

Table 2 lists several other GWAS of biochemical traits related to VTE. Groups have tested genome-wide associations with more global measures of the procoagulant cascade, such as the activated partial thromboplastin time (aPTT).(34, 35) An initial study in 1477

individuals found associations with variants in *F12* (factor 12), *KNG1* (kininogen 1), and *HRG* (histidine rich glycoprotein), all proteins with previously described functions in the coagulation cascade.(34) A subsequent larger study in 9240 individuals found additional signals in *ABO*, *F11* (factor 11), *F5* (factor 5) and *C6orf10* (chromosome 6 open reading frame 10), all but *C6orf10* with known functions in the coagulation cascade.(35)

Several groups have investigated the genetic determinants of proteins in the fibrinolytic system. Huang et al. have reported GWAS studies for both tissue plasminogen activator (tPA) and plasminogen activator inhibitor 1 (PAI-1) in ~20,000 individuals.(36, 37) Variants in two SNARE complex proteins, *STX2* and *STXBP5* were associated with t-PA levels. These two loci were also associated with VWF levels consistent with genetic regulation of a shared secretory pathway. The most recent fibrinogen meta-analysis in ~90,000 individuals reported significant associations with 23 loci.(38) The strongest association was in the *FGB* (fibrinogen beta chain) locus itself. However the total amount of variance explained by these 23 loci was only 3.7% suggesting lower overall fibrinogen heritability, difficulty with accurate fibrinogen measurement and variance due to unmeasured factors. Our group reported an analysis of plasminogen levels in ~3,000 people finding significant associations with variants in the *PLG* (plasminogen) locus as well as *LPA* (Apolipoprotein (a)) and *SIGLEC14* (sialic acid-binding Ig-like lectin 14) that explained a total of 6.8% of the variation in plasminogen levels.(39) In the anticoagulant pathway, several groups have reported GWAS of protein C levels.(40, 41) Interestingly, in a study of ~3000 individuals, the strongest association signal for protein C level was in *PROCR* (protein C receptor) explaining ~11% of the variance in protein C levels. In a larger study of ~9,000 people, variants in *PROC* itself were also detected explaining ~2% of the variance. The later study also made associations with variants in *EDEM2* (ER degradation enhancer, mannosidase alpha-like), *GCKR* (glucokinase regulatory protein) and *BAZ1B* (bromodomain adjacent to zinc finger domain 1B), genes whose function in protein C regulation has yet to be described. More recently, a study of the African American subset of the ARIC cohort (N = 2701) reported associations in *PROCR*, *PROC* and an intergenic region near *PROC* that collectively explained ~14% of the variance in protein C levels.(42)

Rare Variants and Beyond

In order to move closer to the goal of using an individual's genetic information to inform disease risk, individualize care and create targeted therapies, the genetic determinants of complex disease traits need to be more thoroughly defined. Towards this end, studies are underway employing next-generation sequencing technology to examine the DNA sequence from the 2% of the genome that encodes proteins (exome) as well as the whole genome to understand the role of rare mutations and regulatory variants in complex disease.(43) Aggregate tests of mutation burden are being performed to identify genes with significant differences in mutation load between cases and controls.(44) In these tests, the mutations in a specific gene are, in a way, added together and analyzed in aggregate to increase their apparent frequency so that associations can be made with rare or extremely rare variants if they are accumulating in a single locus. More complex computational methods and complementary functional studies will be required to improve the power of mutation burden

tests and differentiate mutations causing loss of function from mutations causing gain of function in the same locus.

A large number of GWAS have discovered strong signals in intergenic areas. These areas may harbor regulatory variants such as enhancers that change disease risk through altered gene expression patterns. Understanding the genetic signature of these regulatory elements remains a research priority for groups studying complex genetic traits.(45) Likewise, a better understanding of gene-gene interactions (epistasis) may shed light on how specific gene variants function in a gene network. For example, most individuals with FV Leiden mutation never suffer from VTE. This could be due to the presence of other protective genetic variants in these individuals or the requirement of specific genetic variants in people with FV Leiden who develop VTE. To detect these interactions, a large case-control cohort of individuals with FV Leiden would be required. Unfortunately, the search for epistatic interactions is computationally intensive and will likely require larger studies in order to detect significant interactions in VTE.(46)

Conclusion

More complete knowledge of the genetic determinants of the hemostatic and thrombotic systems will require further studies into the common, rare and regulatory variants in the genome. These studies will provide a context for the impact of a given variant in a population and aid in the discovery of new genes playing a role in thrombosis and hemostasis. Ultimately, the results from these studies should inform clinicians about an individual's risk for thrombotic disease and help guide therapeutic decisions.

Acknowledgments

K.D. receives funding from the National Institute of Health

Financial Support and Sponsorship: K.D. receives funding from the National Institute of Health.

References

1. Ginsburg D. Identifying novel genetic determinants of hemostatic balance. *J Thromb Haemost*. 2005 Aug; 3(8):1561–8. eng. [PubMed: 16102020]
2. Ginsburg, D. Hemophilia and Other Disorders of Hemostasis. In: Rimoïn, DLCL.; Pyeritz, RE., editors. *Emery and Rimion's Principles and Practice of Medical Genetics*. Vol. II. New York: Churchill Livingstone; 2002. p. 1926-58.
3. Souto JC, Almasy L, Borrell M, Blanco-Vaca F, Mateo J, Soria JM, et al. Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study. *Genetic Analysis of Idiopathic Thrombophilia*. *American journal of human genetics*. 2000 Dec; 67(6):1452–9. eng. [PubMed: 11038326]
4. Dahlback 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 U S A*. 1993 Feb 1; 90(3):1004–8. [PubMed: 8430067]
5. Bertina RM, Koeleman 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 May 5; 369(6475):64–7. [PubMed: 8164741]

6. 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 Nov 15; 88(10):3698–703. [PubMed: 8916933]
7. Kearon C. Influence of hereditary or acquired thrombophilias on the treatment of venous thromboembolism. *Curr Opin Hematol*. 2012 Sep; 19(5):363–70. [PubMed: 22759630]
8. Kearon C, Julian JA, Kovacs MJ, Anderson DR, Wells P, Mackinnon B, et al. Influence of thrombophilia on risk of recurrent venous thromboembolism while on warfarin: results from a randomized trial. *Blood*. 2008 Dec 1; 112(12):4432–6. [PubMed: 18791166]
9. Yin T, Miyata T. Dysfunction of protein C anticoagulant system, main genetic risk factor for venous thromboembolism in northeast Asians. *J Thromb Thrombolysis*. 2014 Jan; 37(1):56–65. [PubMed: 24233386]
10. Rosendaal FR. High levels of factor VIII and venous thrombosis. *Thromb Haemost*. 2000 Jan; 83(1):1–2. [PubMed: 10669143]
11. Kraaijenhagen RA, in't Anker PS, Koopman MM, Reitsma PH, Prins MH, van den Ende A, et al. High plasma concentration of factor VIIIc is a major risk factor for venous thromboembolism. *Thromb Haemost*. 2000 Jan; 83(1):5–9. [PubMed: 10669145]
12. Crowther MA, Kelton JG. Congenital thrombophilic states associated with venous thrombosis: a qualitative overview and proposed classification system. *Ann Intern Med*. 2003 Jan 21; 138(2):128–34. [PubMed: 12529095]
13. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005 Apr 15; 308(5720):385–9. [PubMed: 15761122]
- 14*. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res*. 2014 Jan; 42(Database issue):D1001–6. This study provides links to the NIH GWAS catalog, a searchable database of GWAS studies. [PubMed: 24316577]
15. Pasaniuc B, Rohland N, McLaren PJ, Garimella K, Zaitlen N, Li H, et al. Extremely low-coverage sequencing and imputation increases power for genome-wide association studies. *Nat Genet*. 2012 Jun; 44(6):631–5. [PubMed: 22610117]
16. Wray NR, Purcell SM, Visscher PM. Synthetic associations created by rare variants do not explain most GWAS results. *PLoS Biol*. 2011; 9(1):e1000579. [PubMed: 21267061]
17. Bezemer ID, Bare LA, Doggen CJ, Arellano AR, Tong C, Rowland CM, et al. Gene variants associated with deep vein thrombosis. *JAMA*. 2008 Mar 19; 299(11):1306–14. [PubMed: 18349091]
18. Tregouet DA, Heath S, Saut N, Biron-Andreani C, Schved JF, Pernod G, 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 May 21; 113(21):5298–303. [PubMed: 19278955]
19. 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 Aug; 102(2):360–70. [PubMed: 19652888]
20. Antoni G, Morange PE, Luo Y, Saut N, Burgos G, Heath S, et al. A multi-stage multi-design strategy provides strong evidence that the BAI3 locus is associated with early-onset venous thromboembolism. *J Thromb Haemost*. 2010 Dec; 8(12):2671–9. Epub 2010/10/16. eng. [PubMed: 20946148]
21. Heit JA, Armasu SM, Asmann YW, Cunningham JM, Matsumoto ME, Petterson TM, et al. A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. *J Thromb Haemost*. 2012 Aug; 10(8):1521–31. [PubMed: 22672568]
22. Tang W, Teichert M, Chasman DI, Heit JA, Morange PE, Li G, 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 Jul; 37(5):512–21. [PubMed: 23650146]
- 23*. Germain M, Chasman DI, de Haan H, Tang W, Lindstrom S, Weng LC, et al. Meta-analysis of 65,734 Individuals Identifies TSPAN15 and SLC44A2 as Two Susceptibility Loci for Venous

- Thromboembolism. *Am J Hum Genet.* 2015 Mar 11. A large meta-analysis of VTE GWAS. With the increased power, investigators identify 9 loci associated with VTE.
24. Antoni G, Oudot-Mellakh T, Dimitromanolakis A, Germain M, Cohen W, Wells P, et al. Combined analysis of three genome-wide association studies on vWF and FVIII plasma levels. *BMC Med Genet.* 2011; 12:102. Epub 2011/08/04. eng. [PubMed: 21810271]
 25. Desch KC, Ozel AB, Siemieniak D, Kalish Y, Shavit JA, Thornburg CD, et al. Linkage analysis identifies a locus for plasma von Willebrand factor undetected by genome-wide association. *Proc Natl Acad Sci U S A.* 2013 Jan 8; 110(2):588–93. Epub 2012/12/26. eng. [PubMed: 23267103]
 26. Smith NL, Chen MH, Dehghan A, Strachan DP, Basu S, Soranzo N, 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 Mar 30; 121(12):1382–92. Epub 2010/03/17. eng. [PubMed: 20231535]
 27. Sadler JE. von Willebrand factor: two sides of a coin. *J Thromb Haemost.* 2005 Aug; 3(8):1702–9. Epub 2005/08/17. eng. [PubMed: 16102036]
 28. Orstavik KH, Magnus P, Reisner H, Berg K, Graham JB, Nance W. Factor VIII and factor IX in a twin population. Evidence for a major effect of ABO locus on factor VIII level. *Am J Hum Genet.* 1985 Jan; 37(1):89–101. Epub 1985/01/01. eng. [PubMed: 3919575]
 29. Smith NL, Rice KM, Bovill EG, Cushman M, Bis JC, McKnight B, et al. Genetic variation associated with plasma von Willebrand factor levels and the risk of incident venous thrombosis. *Blood.* 2011 Jun 2; 117(22):6007–11. Epub 2010/12/18. eng. [PubMed: 21163921]
 - 30*. Zhu Q, Yamakuchi M, Ture S, de la Luz Garcia-Hernandez M, Ko KA, Modjeski KL, et al. Syntaxin-binding protein STXBP5 inhibits endothelial exocytosis and promotes platelet secretion. *J Clin Invest.* 2014 Oct; 124(10):4503–16. This study offers a mechanistic link between variants in STXBP5 and VWF levels. [PubMed: 25244095]
 31. Rydz N, Swystun LL, Notley C, Paterson AD, Riches JJ, Sponagle K, et al. The C-type lectin receptor CLEC4M binds, internalizes, and clears von Willebrand factor and contributes to the variation in plasma von Willebrand factor levels. *Blood.* 2013 Jun 27; 121(26):5228–37. [PubMed: 23529928]
 32. Sanders YV, van der Bom JG, Isaacs A, Cnossen MH, de Maat MP, Laros-van Gorkom BA, et al. CLEC4M and STXBP5 gene variation contribute to von Willebrand factor level variation in von Willebrand disease. *J Thromb Haemost.* 2015 Mar 31.
 33. Souto JC, Almasy L, Soria JM, Buil A, Stone W, Lathrop M, et al. Genome-wide linkage analysis of von Willebrand factor plasma levels: results from the GAIT project. *Thromb Haemost.* 2003 Mar; 89(3):468–74. Epub 2003/03/08. eng. [PubMed: 12624629]
 34. Houlihan LM, Davies G, Tenesa A, Harris SE, Luciano M, Gow AJ, et al. Common variants of large effect in F12, KNG1, and HRG are associated with activated partial thromboplastin time. *Am J Hum Genet.* 2010 Apr 9; 86(4):626–31. [PubMed: 20303064]
 35. Tang W, Schwienbacher C, Lopez LM, Ben-Shlomo Y, Oudot-Mellakh T, Johnson AD, et al. Genetic Associations for Activated Partial Thromboplastin Time and Prothrombin Time, their Gene Expression Profiles, and Risk of Coronary Artery Disease. *Am J Hum Genet.* 2012 Jul 13; 91(1):152–62. Epub 2012/06/19. eng. [PubMed: 22703881]
 36. Huang J, Huffman JE, Yamkauchi M, Trompet S, Asselbergs FW, Sabater-Lleal 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 May; 34(5):1093–101. [PubMed: 24578379]
 37. Huang J, Sabater-Lleal M, Asselbergs FW, Tregouet D, Shin S-Y, Ding J, et al. Genome-wide association study for circulating levels of PAI-1 provides novel insights into its regulation. *Blood.* 2012 Dec 6; 120(24):4873–81. [PubMed: 22990020]
 38. Sabater-Lleal M, Huang J, Chasman D, Naitza S, Dehghan A, Johnson AD, 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 Sep 17; 128(12):1310–24. [PubMed: 23969696]
 - 39*. Ma Q, Ozel AB, Ramdas S, McGee B, Khoriaty R, Siemieniak D, et al. Genetic variants in PLG, LPA, and SIGLEC 14 as well as smoking contribute to plasma plasminogen levels. *Blood.* 2014

- Nov 13; 124(20):3155–64. This study reports both linkage and GWAS of plasminogen levels in a healthy sibling cohort. [PubMed: 25208887]
40. Oudot-Mellakh T, Cohen W, Germain M, Saut N, Kallel C, Zelenika D, 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 Apr; 157(2):230–9. [PubMed: 22443383]
41. Tang W, Basu S, Kong X, Pankow JS, Aleksic N, Tan A, et al. Genome-wide association study identifies novel loci for plasma levels of protein C: the ARIC study. *Blood*. 2010 Dec 2; 116(23):5032–6. Epub 2010/08/31. eng. [PubMed: 20802025]
- 42*. Munir MS, Weng LC, Tang W, Basu S, Pankow JS, Matijevic N, et al. Genetic markers associated with plasma protein C level in African Americans: the atherosclerosis risk in communities (ARIC) study. *Genet Epidemiol*. 2014 Dec; 38(8):709–13. This study highlights the importance of population structure in GWAS. [PubMed: 25376901]
43. Cirulli ET, Lasseigne BN, Petrovski S, Sapp PC, Dion PA, Leblond CS, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science*. 2015 Mar 27; 347(6229):1436–41. [PubMed: 25700176]
- 44*. Lange LA, Hu Y, Zhang H, Xue C, Schmidt EM, Tang ZZ, et al. Whole-exome sequencing identifies rare and low-frequency coding variants associated with LDL cholesterol. *Am J Hum Genet*. 2014 Feb 6; 94(2):233–45. This study provides a framework for the discovery of rare variants influencing complex genetic traits. [PubMed: 24507775]
45. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. *Nature reviews Genetics*. 2015 Apr; 16(4):197–212.
46. Greliche N, Germain M, Lambert JC, Cohen W, Bertrand M, Dupuis AM, et al. A genome-wide search for common SNP x SNP interactions on the risk of venous thrombosis. *BMC Med Genet*. 2013; 14:36. [PubMed: 23509962]

Key Points

- Venous thromboembolism is heritable complex genetic trait determined by multiple gene variants, gene-environmental interactions and altered gene regulation.
- GWAS studies of VTE and biochemical markers related to VTE have identified common genetic variants altering the risk of disease but the majority of the heritability remains unexplained.
- Future studies will concentrate on the identification of rare mutations, regulatory variants and gene-gene interactions to further understand individual risk for VTE.

Table 1

GWAS of Venous Thromboembolism

Trait	Year	Study	N:	SNPs Tested	Locus	SNP (Minor Allele)	Functional Class	Odds Ratio (95% CI)	P-Value
VTE	2009	Tregouet DA, et al.	453 Cases 1327 Controls	291,872	<i>ABO</i>	rs505922(C)	intronic	1.91 (1.53–2.39)	1.48×10^{-14}
					<i>ABO</i>	rs657152 (A)	intronic	1.89 (1.51–2.36)	2.22×10^{-13}
					<i>F5</i>	rs2420371 (G)	intronic	2.27 (1.62–3.18)	8.08×10^{-10}
VTE	2012	Heit JA, et al.	Discovery: 1503 Cases 1459 Controls Replication: 1407 Cases 1418 Controls	557,112 2500K imputed	<i>F5</i> (Leiden)	rs6025 (T)	missense	3.57 (2.76–4.60)	1.68×10^{-22}
					<i>ABO</i>	rs8176719 (G)	indel	1.47 (1.32–1.64)	5.68×10^{-12}
					<i>ABO</i>	rs2519093 (T)	intronic	1.69 (1.48–1.91)	8.08×10^{-16}
					<i>F2</i> (20210)	rs1799963 (A)	3'UTR	2.46	1.70×10^{-6}
VTE	2013	Tang W et al.	Discovery: 1618 Cases 42881 Controls Replication: 3231 Cases 3536 Controls Meta-Analysis performed	2500K imputed	<i>F5</i>	rs6427196 (G) *	3' UTR	2.07 (1.89–2.28)	4.47×10^{-51}
					<i>ABO</i>	rs687621 (G)	intronic	1.55 (1.47–1.64)	1.55×10^{-52}
					<i>F11</i>	rs4253399 (G)	intronic	1.24 (1.17–1.31)	2.78×10^{-14}
					Near <i>FGG</i>	rs6536024	intergenic	0.80 (0.76–0.85)	1.75×10^{-13}
VTE	2015	Germain M et al	Discovery: 7507 Cases 52632 Controls Replication: 3009 Cases 2586 Controls	6.751K imputed	<i>F5</i> (Leiden)	rs6025 (T)	missense	3.25 (2.91–3.64)	1.10×10^{-96}
					<i>F5</i>	rs4524 (T) **	missense	1.20 (1.14–1.26)	2.65×10^{-11}
					<i>ABO</i>	rs529565 (C)	intronic	1.55 (1.48–1.63)	4.23×10^{-75}
					<i>F11</i>	rs4253417 (C)	intronic	1.27 (1.22–1.34)	1.21×10^{-23}
					<i>FGG</i>	rs2066865 (A)	3' UTR	1.24 (1.18–1.31)	1.03×10^{-16}
					<i>F2</i>	rs1799963 (A) ***	3' UTR	2.29 (1.75–2.99)	1.73×10^{-9}
					<i>PROCR</i>	rs6087685 (C)	intronic	1.15 (1.10–1.21)	1.15×10^{-8}
					<i>TSPAN15</i>	rs78707713 (T)	intronic	1.28 (1.19–1.39)	5.74×10^{-11}
					<i>SLC44A2</i>	rs2288904 (G)	missense	1.19(1.12–1.26)	1.07×10^{-9}

K = × 1000

* In high linkage disequilibrium with rs6025, FV Leiden

** Independent association from rs6025, FV Leiden

*** Prothrombin G20210A

Table 2

GWAS of biochemical traits in thrombosis and hemostasis

Trait	Year	Study	N:	SNPs Tested	Locus	SNP (Tested Allele)	Beta Coeff. (95% CI)	P-Value	Functional Class
VWF	2010	Smith NL et al.	23608	2,600K imputed	<i>ABO</i> <i>VWF</i> <i>STXBP5</i> <i>SCARA5</i> <i>STX2</i> <i>TC2N</i> <i>STAB2</i> <i>CLEC4M</i>	rs687621 (C)*	24.1 (21.4–26.8)	5.0×10^{-324}	intronic
						rs1063857 (C)	6.0 (3.2 – 8.8)	1.7×10^{-32}	synonymous
						rs9390459 (A)	–4.8 (–7.5 – –2.1)	1.2×10^{-22}	synonymous
						rs27857224 (T)	4.5 (1.5 – 7.5)	1.3×10^{-16}	intronic
						rs7978987 (A)	3.4 (0.6 – 6.2)	3.8×10^{-11}	intronic
						rs10133762 (T)	3.1 (0.4 – 5.8)	2.3×10^{-10}	intronic
						rs4981022 (C)	–3.6 (–6.8 – –0.4)	7.3×10^{-10}	intronic
VWF	2013	Desch KC et al.	3462	723716	<i>ABO</i> <i>VWF</i> <i>2p11.2 (linkage)</i>	rs687289 (T)*	0.36 ***	1.3×10^{-128}	intronic
						rs1063856(A)	–0.095	4.9×10^{-16}	missense
						N/A	N/A	N/A	N/A
aPTT	2010	Houlihan LM et al.	1477	542050	<i>F12</i> <i>KNG1</i> <i>HRG</i>	rs2731672 (A)	0.45	2.16×10^{-30}	intronic
						rs710446 (C)	–0.36	9.52×10^{-22}	missense
						rs9898 (T)	–0.26	1.34×10^{-11}	missense
aPTT	2012	Tang W et al.	9240, 869-3467 replication	2500K imputed	<i>KNG1</i> <i>HRG</i> <i>ABO</i> <i>F12</i> <i>F11</i> <i>ABO</i> <i>C6orf10</i> <i>F5</i>	rs710446 (C)	–1.19	1.06×10^{-185}	missense
						rs9898 (T)	–1.01	1.65×10^{-116}	missense
						rs687621 (G)*	–0.82	3.64×10^{-81}	intronic
						rs2545801 (G)	1.55	4.30×10^{-60}	5' upstream
						rs2289252 (T)	–0.48	4.30×10^{-30}	intronic
						rs8176704 (T)**	0.89	4.26×10^{-24}	intronic
						rs2050190 (G)	–0.25	1.29×10^{-8}	intronic
						rs9332701 (G)	0.54	3.70×10^{-8}	missense
						rs4252129 (C)	–0.16 ***	1.9×10^{-27}	missense
						rs1084651 (G)	–0.038	1.6×10^{-15}	intronic
Plasminogen	2014	Ma Q et al.	3244	741807 4,500K imputed	<i>PLG</i> <i>LPA</i>	rs4252129 (C)	–0.16 ***	1.9×10^{-27}	missense
						rs1084651 (G)	–0.038	1.6×10^{-15}	intronic

Trait	Year	Study	N:	SNPs Tested	Locus	SNP (Tested Allele)	Beta Coeff. (95% CI)	P-Value	Functional Class
					<i>SIGLEC14</i>	rs10412972 (G)	0.026	1.1×10^{-8}	5' upstream
Fibrinogen	2013	Sabater-Lleal M et al.	91323	2500K imputed	<i>FGB</i>	rs1800789 (G)	0.031 [*]	1.68×10^{-127}	5' upstream
					<i>C5orf56</i>	rs2106854 (C)	-0.019	1.72×10^{-48}	intronic
					<i>IL6R</i>	rs4129267 (C)	-0.011	5.97×10^{-27}	intronic
					<i>NLRP3</i>	rs10157379 (C)	0.010	1.15×10^{-19}	intronic
					<i>IL1F10</i>	rs6734238 (G)	-0.009	5.77×10^{-19}	3' downstream
					<i>MSL2</i>	rs1154988 (T)	-0.010	9.64×10^{-17}	3' upstream
					<i>LEPR + 16 other</i>	rs1938492 (C)	0.008	5.28×10^{-14}	3' downstream
PAI-1	2012	Huang J et al.	19599, 10796 replication	2500K imputed	<i>SERPINE1</i>	rs2227631 (G)	0.073	3.2×10^{-24}	5' upstream
					<i>ACHE</i>	rs6976053 (C)	0.048	5.8×10^{-13}	3' downstream
					<i>ARNTL</i>	rs6486122 (C)	0.046	1.7×10^{-10}	intronic
					<i>PPARG</i>	rs11128603 (G)	0.066	9.4×10^{-8}	intronic
t-PA	2014	Huang J et al.	26929	2500K imputed	<i>STXBP5</i>	rs9399599 (A)	0.032 ^{***}	2.9×10^{-14}	intronic
					<i>STX2</i>	rs7301826 (T)	0.027	1.0×10^{-9}	intronic
					<i>POLB</i>	rs3136739 (G)	0.063	1.3×10^{-9}	intronic
Protein C	2012	Oudot-Mellakh T et al.	951 1921	472173	<i>PROCR</i>	rs8119351 (A)	20.6	1.11×10^{-31}	5' upstream
Protein C	2010	Tang W et al.	8048, 1376 replication	2500K imputed	<i>PROCR</i>	rs8119351 (A)	0.480 ^{***}	2.68×10^{-203}	5' upstream
					<i>EDEM2</i>	rs6120849 (T)	-0.141	7.19×10^{-37}	intronic
					<i>PROC</i>	rs1158867 (C)	-0.123	3.77×10^{-36}	intronic
					<i>GCKR</i>	rs1260326 (T)	0.082	2.04×10^{-17}	missense
					<i>BAZ1B</i>	rs17145713 (T)	-0.062	2.50×10^{-7}	intronic
Protein C	2014	Munir MS et al.	2,701	2,600K	<i>PROCR</i>	rs867186 (T)	0.49 ^{***}	9.84×10^{-65}	missense
					Intergenic	rs7580658 (G)	0.15	1.70×10^{-12}	intergenic
					<i>PROC</i>	rs1799808 (C)	0.15	2.03×10^{-12}	5' upstream

K = × 1000

Desch

Beta values reported on log transformed levels

Tags the A2 allele of ABO

Tags the O allele of ABO

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript