The Genetic Component of Disorders of Coagulation and Thrombosis
José Manuel Soria

Unitat de Genòmica de Malalties Complexes. Institut de Recerca de l’Hospital de la Santa Creu i Sant Pau.
Barcelona. Spain.

Thrombosis plays a crucial role in the pathogenesis of acute myocardial infarction, stroke and venous thrombosis, and is the principle factor responsible for a subsequent fatal outcome. These conditions, each of which has an annual incidence of 1 to 3 per 1000 adults, are some the principle causes of morbidity and mortality in developed countries. Consequently, considerable financial and health-care resources are being devoted to their diagnosis, treatment and prevention.

Thrombosis is a good example of a complex disease, in which each individual’s susceptibility to the disease is determined by the actions of numerous genes and their interactions with environmental factors. It has been established that genetic variations in the genes that code for coagulation factors or inhibitors are important risk factors for thromboembolism. However, 50% of patients with inherited thrombophilia do not have any of these genetic variations. Consequently, the major challenge today is to identify new genetic risk factors for thrombosis.

Key words: Thrombosis. Coagulation. Genetics. Risk factors.
The next step to revolutionise the study of thrombophilia was the evolution of the concept of the disease: it is currently thought of as a continuous risk variable rather than a discrete, dichotomous variable, i.e. healthy or sick patients. In other words, rather than determining whether or not the disease will develop, genetics determines greater or smaller chances of developing it.

Based on what is mentioned above, hereditary thrombophilia is currently considered to be both a multifactor (involving both genetic and environmental factors) and a complex disease (it does not follow a simple Mendelian inheritance pattern), in which the sum of multiple genes and how each one interacts with environmental factors determines each individual’s degree of susceptibility to thrombosis.

METHODS FOR STUDYING THROMBOPHILIA

Nevertheless, despite a large amount of effort invested in studying thrombotic disease over the last decade, our knowledge of this condition’s molecular base is scant, even more so if we consider the fact that 60% of predisposition to thrombosis is attributable to genetic factors. To illustrate this situation, we must recall that known genetic factors for thrombotic risk are only identified in 50% of families with hereditary thrombosis, in the best-case scenario and depending on the population under study. The great challenge at present is to identify the genetic factors for thrombotic risk in the remaining 50%.

Experimental designs used for identifying and locating the genes implicated in thrombotic risk include genetic linkage analysis and association studies.

Association studies

Case-control association studies analyse the correlation between phenotype and genotype. The phenotype is generally the presence (cases) or absence (controls) of the disease in unrelated individuals or in families. Genotype is determined by some type of genetic polymorphism, mainly single nucleotide variations, known as single nucleotide polymorphisms (SNP). There is association when this marker’s distribution (allele frequencies) is different in cases and controls for certain level of statistical significance. The genetic marker is a polymorphism within or close to a candidate gene, defined as a specific gene that is biologically related to the disease process. This type of design was used to identify FVL and G20210A mutation in the prothrombin gene. More recently, it has also permitted the identification of A384S mutations in the SERPINAC1 gene, which codes for AT, and R67X mutation in the SERPINC gene, which codes for protein Z inhibitor. These mutations are associated with increased risk of thrombotic events.

Additionally, the genetic linkage analysis differs from association studies in that it is based on parents’ transmission of a genetic marker and a functional genetic variant to their descendents (co-segregation). The co-segregation of a genetic marker can only be detected by observing how chromosomes are inherited from one generation to another, which necessarily requires recruiting related individuals.

Genetic linkage analyses have permitted identifying C46T mutation in the F12 gene as a thrombotic risk factor. This result has been replicated in diverse case-control association studies. In addition, although the ABO blood group had reportedly been implicated as a thrombotic risk factor, the first genetic evidence of this implication came from a genetic linkage study. Subsequently,
various association studies confirmed this result and determined that A1 is the allele that increases the risk of thrombotic events\textsuperscript{15,16}.

### Genetic linkage studies

However, genetic linkage studies are rendering the most noteworthy results in the study of the genetic determinants of intermediary phenotypes with thromboembolic disease (table 1)\textsuperscript{58-64}. These methods are based on quantitative trait genetics analysed in families. The genetic effects are quantified in terms of heritability ($h^2$), which is the proportion of phenotype variation that is exclusively attributable to the effect of the genes. Table 2\textsuperscript{65-72} shows the heritability of several components of haemostasis. Estimating heritability is a necessary step before attempting gene localisation, since if the phenotype has no heritability, or this is very small (e.g. < 10\%), gene searching makes no sense. The basic mathematical tool for this type of investigation is variance analysis, which allows us to separate the effect due to genetic factors from the effect due to any environmental factors that affect a trait or quantitative phenotype and the complex disease under study. This methodological approximation has permitted determining that more than 60\% predisposition to thrombosis is attributable to genetic factors\textsuperscript{10}.

Once a phenotype has been shown to be heritable, the next step is to find the chromosomal sites (loci) containing genes that affect the variability of such phenotype. Each of these loci is known as a QTL (quantitative trait locus). A QTL can explain only a small proportion of the variability observed in a phenotype, while the rest is due to other QTLs and to environmental factors\textsuperscript{17}. Finding the location of QTLs is achieved through genetic linkage analysis. There are several methodological strategies for performing a genetic linkage analysis, but one of the most robust and powerful, from a statistical viewpoint, is based on variance-components linkage analysis. The idea is that relatives with more similarity in a certain phenotype would share more genetic markers near the gene that influences such phenotype. Conversely, other relatives farther removed from the phenotype under study would not bear the same alleles\textsuperscript{18}.

Different statistical tests exist for finding out if two loci are linked. The classic parameter is the LOD score or the (logarithm [base 10] of odds [OR]) between two alternative probabilities (LOD = $\log_{10}$ [probability of both loci being linked / probability of their being not linked]).

This concept is important because the most advanced techniques for searching new genes are based on the use of highly polymorphic anonymous

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Chromosome</th>
<th>Gene</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>FII-thrombosis</td>
<td>11p11</td>
<td>F2</td>
<td>58</td>
</tr>
<tr>
<td>FXII-thrombosis</td>
<td>5q35</td>
<td>F12</td>
<td>13</td>
</tr>
<tr>
<td>FXII</td>
<td>10p13</td>
<td>?</td>
<td>13</td>
</tr>
<tr>
<td>Free protein S</td>
<td>1q32</td>
<td>C4BP</td>
<td>59</td>
</tr>
<tr>
<td>APCR/FVIII-thrombosis</td>
<td>18p11</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>Von Willebrand factor</td>
<td>9p34</td>
<td>AB0</td>
<td>14</td>
</tr>
<tr>
<td>C Protein</td>
<td>16</td>
<td>NQO1</td>
<td>61</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>12</td>
<td>TCF1</td>
<td>42</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>14</td>
<td>?</td>
<td>42</td>
</tr>
<tr>
<td>FVII</td>
<td>13</td>
<td>F7</td>
<td>56</td>
</tr>
<tr>
<td>TFPI</td>
<td>2</td>
<td>TFPI</td>
<td>62</td>
</tr>
<tr>
<td>TFPI</td>
<td>13</td>
<td>CBP2</td>
<td>63</td>
</tr>
<tr>
<td>FVIII</td>
<td>5/11</td>
<td>?</td>
<td>52</td>
</tr>
<tr>
<td>Thrombosis</td>
<td>10p12</td>
<td>?</td>
<td>64</td>
</tr>
<tr>
<td>Thrombosis</td>
<td>18p11.2</td>
<td>?</td>
<td>64</td>
</tr>
<tr>
<td>Thrombosis</td>
<td>11q23</td>
<td>?</td>
<td>64</td>
</tr>
</tbody>
</table>

\textsuperscript{?}: unknown candidate gene.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Inheritability</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>aPPT</td>
<td>0.83</td>
<td>65</td>
</tr>
<tr>
<td>APCR</td>
<td>0.58-0.71</td>
<td>65, 66</td>
</tr>
<tr>
<td>Factor XII</td>
<td>0.67</td>
<td>65</td>
</tr>
<tr>
<td>Factor VII</td>
<td>0.33-0.63</td>
<td>65, 67-69</td>
</tr>
<tr>
<td>HRG</td>
<td>0.52-0.7</td>
<td>65, 70</td>
</tr>
<tr>
<td>TFPI</td>
<td>0.51</td>
<td>65</td>
</tr>
<tr>
<td>Prothrombin time</td>
<td>0.5</td>
<td>65</td>
</tr>
<tr>
<td>Protein C</td>
<td>0.41-0.5</td>
<td>65, 71</td>
</tr>
<tr>
<td>Factor II</td>
<td>0.49-0.7</td>
<td>65, 71</td>
</tr>
<tr>
<td>Antithrombin</td>
<td>0.4-0.48</td>
<td>65, 66</td>
</tr>
<tr>
<td>Free protein S</td>
<td>0.46</td>
<td>65</td>
</tr>
<tr>
<td>Functional protein S</td>
<td>0.45</td>
<td>65</td>
</tr>
<tr>
<td>Factor XI</td>
<td>0.45</td>
<td>65</td>
</tr>
<tr>
<td>Factor V</td>
<td>0.44-0.71</td>
<td>65, 71</td>
</tr>
<tr>
<td>Factor X</td>
<td>0.43</td>
<td>65</td>
</tr>
<tr>
<td>Factor VIII</td>
<td>0.4-0.61</td>
<td>65, 68</td>
</tr>
<tr>
<td>Factor IX</td>
<td>0.38-0.5</td>
<td>65, 71</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>0.33-0.51</td>
<td>65, 72</td>
</tr>
<tr>
<td>Von Willebrand factor</td>
<td>0.31-0.75</td>
<td>65, 68</td>
</tr>
<tr>
<td>PAI-1</td>
<td>0.29-0.6</td>
<td>65, 68</td>
</tr>
<tr>
<td>tPA</td>
<td>0.26-0.62</td>
<td>65, 68</td>
</tr>
<tr>
<td>Plasminogen</td>
<td>0.11-0.236</td>
<td>65</td>
</tr>
<tr>
<td>Total protein S</td>
<td>0.22-0.23</td>
<td>65, 71</td>
</tr>
<tr>
<td>Tissue factor</td>
<td>0.16</td>
<td>65</td>
</tr>
</tbody>
</table>

APCR: activated protein C resistance; HRG: Histidine-rich glycoprotein; PAI-1: plasminogen activator inhibitor 1; TFPI: tissue factor pathway inhibitor; t-PA: tissue plasminogen activator; aPPT: activated partial thromboplastin time.
genetic markers which, in the presence of linkage with complex phenotypes, permit detecting nearby presence of true functional gene.

These studies have enabled us to identify regions in the genome that determine variability of coagulation component levels or those that determine the risk of thromboembolic events occurrence (table 1).

Both the genetic linkage analysis and the association studies have their advantages and disadvantages. The former are generally good at finding new genes, while the latter are effective for analysing genes already known. Therefore, both methods may be considered complementary, and the question is not so much which of them to use, but rather when and how to apply each method. This point is currently one of the most controversial topics in the field of complex genetic diseases19.

**GENETIC BASIS FOR THROMBOSIS**

As stated previously, the anomalies identified as genetic risk factors for thrombosis are AT, PC and PS deficiencies and the FVL and G20210A PT mutations20 (table 3). However, in recent years we have improved our knowledge of the genetic basis for other thromboembolic event risk factors.

**Activated protein C resistance**

One of these advances makes reference to “activated protein C resistance” phenotype (APCR) described by Dahlback et al21 (1993), which is characterised by low anticoagulant activity in the PC system. Although the FVL mutation has been identified as the main cause of this plasma alteration4, there are still 10-20% APCR cases that do not bear the FVL mutation, thus indicating the presence of other mutations that cause the same phenotype. *FV Cambridge*, *FV Hong Kong*, and others are genetic variants that affect amino acid Arg306, and which also alter PC system activation22,23.

**C46T mutation in the F12 gene**

The C46T mutation in the F12 gene is much more relevant. Factor XII (FXII) is essential for initiating the blood coagulation process. The plasma concentration of this protein has a significant positive genetic correlation ($r = 0.351$) with thromboembolic disease24. A genome-wide scan was used to demonstrate the involvement of the F12 structural gene in determining plasma variability of factor XII (LOD score = 4.76; $p = 1.5 \times 10^{-6}$) and susceptibility to thrombotic events24. Further case-control association studies confirmed that C46T mutation of F12 is a risk factor for venous thrombosis25,26 or arterial thrombosis27-29. Specifically, homozygotic carriers of T allele have five times more risk of thromboembolic events than non-carriers. Recently, a similar study was published30 that points to this polymorphism as a thrombosis risk factor during first pregnancy in a population of 32,463 previously asymptomatic women.

**ABO blood group**

Meanwhile, we have known since the late 1960s that individuals with a non-O blood type have two to four times more risk of suffering thrombotic events. This correlation was established based on the associations observed in the ABO group and the levels of factor VIII31 as well as the levels of the Von Willebrand factor (vWF)32, since concentrations of these proteins are clearly related to cardiovascular risk33. Further understanding of the blood group as a cardiovascular risk factor came when the A1 allele was shown to be one with the highest risk of thrombosis34. These results attest to the implication of the ABO blood group in thromboembolic risk and its synergistic role with other genetic factors involved in this disease.
R67X mutation in the SERPINA 10 gene

Recently, a multi-centre study in Spain\textsuperscript{19} identified a new mutation (R67X) in the SERPINA10 gene (which codes for the protein Z inhibitor) as an important thromboembolic risk factor. Protein Z inhibitor is a protein in the serpin family that is considered to be a new member of the haemostatic system because its anticoagulant activity inhibits activated factors X and XI. Carriers of this mutation (R67X) have 3.3 times more risk of suffering a thrombotic event than non-carriers\textsuperscript{8}, which is similar to the risk for carriers of the FVL or G20210A mutations in the F2 gene. Furthermore, this mutation presents a strong association with a family history of thrombosis (p < 0.001).

A384S mutation in the SERPINC1 gene

The same Spanish multi-centre study\textsuperscript{18} also identified another mutation, A384S, in the SERPINC1 gene (which codes for AT) as an important thromboembolic risk factor. Carriers of this mutation have ten times more risk of suffering a thrombotic event than non-carriers. From a phenotype standpoint, A384S causes a very peculiar deficiency in AT since it shows normal antigenic levels and normal anti-FXa activity, but decreased anti-IIa activity in the presence of heparin\textsuperscript{15}. Because of these characteristics, the effect of the A384S mutation in AT cannot be detected with the plasma testing methods that are normally used in clinical laboratories. Additionally, it is important to highlight the clinical relevance of identifying the patients carrying this mutation in order to administer anticoagulant treatment correctly. This is because unfractioned heparins may be an inefficient anticoagulant treatment due to the defect that this mutation causes in the AT protein. Therefore, detection of this genetic alteration is essential and it constitutes a significant advance in diagnosing AT deficiencies.

V34L mutation in the F13 gene

Lastly, there is evidence that the V34L mutation in the F13 gene is involved in thromboembolic disease. The step whereby the activated factor XIII (aFXIII) stabilises fibrin molecules is essential to the clot formation process. This process has a positive feedback mechanism by which the fibrin activates FXIII. This activation takes place more quickly when the amino acid 34 in FXIII is Leu rather than Val\textsuperscript{16}. As a result, this alteration changes the structure of the fibrin which then polymerises by forming a finer mesh with smaller pores in the clot; this changes the clot’s permeability characteristics\textsuperscript{36}. Various studies have indicated that allele Leu34 has a protective effect against the risk of thromboembolism\textsuperscript{37,38} and this has been confirmed in a meta-analysis\textsuperscript{39}. Since there is no plasma-based method for detecting the functional effect of the Leu variant\textsuperscript{34} on clot formation, detecting this genetic alteration is an important advance for diagnosing thromboembolic disease.

Other thrombotic risk factors

Other coagulation or fibrinolytic system factors may be implicated in predisposition to thrombosis. It has been stated that high fibrinogen concentrations are associated with increased thrombotic risk\textsuperscript{40}; however, only 5-9% of the variability of these concentrations is explained by polymorphisms arising in the beta chain gene, while only 4.2% is determined by polymorphisms in the alpha chain gene\textsuperscript{41}. These results imply that other genetic factors cause most of the quantitative variation for this phenotype. Recently, two genome regions related to the plasma fibrinogen concentration were located (table 1), one in chromosome 12 and the other in chromosome 14\textsuperscript{42}.

From a genetic point of view, there are many polymorphisms in different candidate genes (the majority of those that code for the proteins listed above) which are being examined in numerous association studies to determine whether or not they are implicated in thromboembolic disease. We can cite several polymorphisms in the endothelial PC receptor gene ($EPCR$)\textsuperscript{43-45}, polymorphisms in genes that code for platelet receptors\textsuperscript{46,47}, and polymorphisms in the TFPI gene which could modulate the risk of thrombosis\textsuperscript{48} by determining the plasma TFPI concentration. In the best possible scenario, the actual participation of these polymorphisms in thromboembolic disease is yet to be seen or is obscured by the appearance of contradictory results\textsuperscript{49,50}.

Apart from the previously listed blood coagulation factors and inhibitors, there are other haemostasis factors that may be implicated in predisposition to thrombosis. These parameters represent a good example of a complex quantitative phenotype in which plasma concentrations are the result of the combined influence of genetic and environmental factors. But genetic factors are the ones that determine most of the haemostasis factors; in other words, there is a very sizeable genetic basis for the larger part of these phenotypes (table 2).

Among these phenotypes we can highlight the FVIII levels. In case-control association studies, this phenotype has been related to an increased risk of thrombosis and a relative risk that is more than
four times higher in patients whose plasma FVIII concentration is at or above 150 U/dl. Presently, as explained in the previous section, the only genetic determining factor to clearly be involved in determining FVIII plasma values is the ABO blood group. However, we now have scientific data that point to the presence of a (still unidentified) gene in chromosome 5, which could be implicated in determining FVIII levels (table 1).

It has also been demonstrated that figures above the 90th percentile for both FXI and FIX are associated with a thrombotic risk that is 2.2 and 2-3 times higher respectively (table 3).

Genetic linkage studies in large families have shown a marked genetic correlation between susceptibility to thrombotic events and various haemostasis phenotypes. In particular, the levels of APCR, FVIII, FIX, FXI, FXII, vWF, tissue plasminogen activator (t-PA), thrombin production, FVII, serum folate and homocysteine are determined by genes which in turn increase the risk of thrombotic events.

Given that coagulation factor levels are clearly involved in raising the risk of thrombosis, identifying the genetic factors that determine these levels constitutes one of the main research lines in the field of thromboembolic disease. At present, thanks to genome-wide studies, we have information on what regions of the genome contain genes that determine the variability of numerous coagulation components (table 1). For some cases, we already know what gene and which of its variant alleles cause this variability, and how it may be involved with thromboembolic disease. For other cases, we still have a long way to go before we are able to identify these genetic factors.

The future perspective will involve compiling a list of all genetic factors that contribute to thrombotic events. This knowledge will help us to design treatment and prevention strategies to fit an individual's genetic profile.

Due to the automation of genetic marker genotyping methods and DNA sequencing, the information provided by the Human Genome Project and most of all, to the huge strides made in genetic statistics and the calculation power provided by modern computers, we currently possess the necessary tools and are rightly positioned to successfully overcome the challenge posed by studying the genetic base of complex traits such as thromboembolic disease and intermediary phenotypes that affect the risk of developing this disease. The identification of the new genetic factors and the study of their physiopathological mechanisms is a fundamental step towards understanding the molecular basis of thrombosis. Basic research is necessary in order to develop more effective diagnostic, prophylactic and therapeutic methods. Furthermore, this knowledge is accompanied by very clear advantages for the health profession, such as improving the preventive and therapeutic strategies for patients faced with future risk situations and thrombotic events, and identifying affected family members, most of whom asymptomatic, who otherwise would not benefit from these prevention strategies.

**Conflict of interests**

The author declares he has no conflict of interests.

**REFERENCES**

Rev Esp Cardiol Supl. 2009;9:58B-65B


