Population genetics of venous thromboembolism

A narrative review

Maurizio Margaglione1,2; Elvira Grandone2
1Genetica Medica, Università di Foggia, Foggia, Italy; 2Unità di Aterosclerosi e Trombosi, I.R.C.C.S. “Casa Sollievo della Sofferenza”, S. Giovanni Rotondo, Foggia Italy

Summary

Results from epidemiological studies are consistent with the hypothesis that disparities in venous thromboembolism (VTE) burden are attributable to differences in genetic structure among populations from different genetic backgrounds. To that end, recent genetic studies have demonstrated not only potential associations between certain alleles and VTE but also clear differences in the distribution of these alleles in patients stratified by ancestry. There are a number of notable clinical and pathophysiological questions that arise from these findings. First at all is defining the precise variant(s) that alter disease susceptibility. The comparatively lower rates of VTE recorded among Asians would imply that risk profile is devoid of many risk factors on comparison to Caucasian or African counterparts or that a putative protective factor is advocated in the former population. Identification of these variants provided specific insight into VTE disease in selected populations and also shed lights on the biology of the disease. The association observed between ancestry and VTE is likely to be multifactorial, possibly reflecting, in addition to genetic variation, also socioeconomic differences. Acknowledgment of this may provide useful information in biomedical contexts and help to identify individual risk factors for VTE.

Keywords

Epidemiological studies, familial thrombosis, gene mutations, thrombophilia, inherited coagulation disorders

Introduction

Venous thromboembolism (VTE) is the third most common cardiovascular disease after ischaemic heart disease and stroke (1). Deep-vein thrombosis (DVT), the most frequent presentation of VTE, is associated with significant morbidity and mortality. The most serious complication of DVT, pulmonary embolism (PE), is a life-threatening condition with short-term survival of less than 60% (2). VTE is a common multifactorial disease resulting from the interaction of genetic and/or environmental predisposing risk factors (3). In addition to circumstantial predisposing factors, genetic abnormalities of components of the coagulation pathway leading to hypercoagulability have been found in subjects who have suffered from thromboembolic disease (4). While both acquired and inherited factors play important roles in the pathogenesis of VTE, risk varies greatly from one individual to another, and the causes for many cases remain unidentified. The involvement of genetic factors to an increased risk for VTE was confirmed in family-based studies among Spanish and Caucasian-American families that estimated a heritability of thrombosis up to about 60% (5, 6). In addition, the overall contribution of genetic factors to VTE was further strengthened in male twin pairs. The odds ratio for the second twin to have VTE if one twin had an event was 13.5 and 3.8 for monozygotic and dizygotic twins, respectively (7). However, the incidence of VTE varies widely among different groups of population from settings with a distinct genetic background, and determinants of the different risk still remain to be fully understood (8, 9). Explanations can include differences of environmental risk factors, such as lifestyle habits and the socioeconomic status, the prevalence of obesity, psychosocial stress, and accessibility to and the quality of medical care (10). Although there are other excellent reviews of population genetics and VTE available in the literature as well (see for example [11, 12]), the purpose of this review is to analyse other genetic risk factors, not strictly belonging to the coagulation pathway, such as sickle cell trait and thalassaemia, and to discuss the VTE burden among persons from different genetic backgrounds to assess the relationship between clinical demographics and genetics. Given the different prevalence of alleles across populations, the aim is to examine available information that can help to customise genetic counselling and follow-up care for members of diverse ancestries.

VTE burden

General population and hospital-based samples, involving mainly individuals from European ancestry (Caucasians), indicated that the...
incidence of first-time symptomatic VTE standardised for age varies from 71 to 132 cases per 100,000 population (13). Different exclusion and inclusion criteria for case ascertainment may well explain quite inconsistent estimates. The rate of diagnosis of DVT and PE thrombosis depends on the sensitivity and specificity of diagnostic approaches. Early studies claimed for a lower DVT incidence but a higher PE incidence among African-Americans than among Caucasians (1) and a rare occurrence of symptomatic VTE after surgery in East Asia (14–16). A common concern of any kind of study addressing VTE is to avoid the selection bias and allow all valid cases to be included. Definitely, as VTE produce few specific symptoms, and the clinical diagnosis is unreliable (17), patients will go undetected and may never be included in population surveys. In addition, misclassification of PE, particularly unsuspected deaths from PE subsequently detected at autopsy, has been shown (18). Despite the acknowledged problems of misclassification potentially affecting all studies, the definite result of a series of investigations strongly supports that the incidence and rate of VTE varies among populations of different genetic backgrounds. These studies were performed in communities where a remarkable diversity in genetic background is present (19–21). The most notable finding was a significantly lower incidence of both total and idiopathic VTE among the Asian/Pacific-Islanders and Hispanics compared to both Caucasians and African-Americans. Although not as uncommon as previously reported, VTE is less frequent in Chinese and Japanese than in Caucasians (22–25). A comprehensive look to data (Table 1) seems to indicate that persons of African ancestry carry the highest incidence of total and idiopathic VTE, the highest proportion of cases with diagnosed PE, and the highest case-fatality rate among cases with idiopathic PE. Although the incidence of VTE is higher than in Caucasians, the difference is not statistically significant. In addition, a significantly lower incidence of both total and idiopathic VTE among the Asian/Pacific-Islanders and Hispanics compared to both Caucasians and African-American’s has been found. Overall, the findings of these studies document that there is the highest incidence of VTE among individuals of African ancestry, a significantly higher incidence among Caucasians, a somewhat lower incidence among Hispanics and American Indians, and a very low incidence among individuals of Asian ancestry. Circumstantial and acquired thrombophilic risk factors may interact with each other and, at least in part, well explain different VTE incidences. Non genetic risk factors for VTE, such as immobilisation, surgery, and trauma account for 40% to 80% of VTE cases in Caucasians (26, 27). However, data from studies in Asian patients indicate a lower incidence of symptomatic VTE complicating trauma, immobilisation, surgery, and/or use of thalidomide (14, 15, 28–31), but not major orthopaedic surgery (32). On the other hand, the higher incidence of total VTE among African-Americans is mainly due to a higher incidence of VTE among the non-idiopathic or provoked cases (21), which included patients with malignancy, prior surgery or trauma, or a medical hospitalisation. This conclusion was not confirmed in a very recent study, in which a higher frequency of unprovoked VTE was found (33). In spite of these inconsistencies, ancestry seems to account for a significant portion of the VTE risk among African-Americans. Very few information is known about the rate of VTE among black populations living in Africa but tendency towards a higher prevalence of VTE seems confirmed (34). In other words, because we explain the VTE development by applying a multifactorial model, it is conceivable that the distribution of genetic risk factors may vary and play a major role in the different distribution of VTE risk among people form different genetic backgrounds.

### The multifactorial model

Since its introduction, the multifactorial threshold model has been widely applied in genetic counselling for qualitative (discrete) traits (35–37). The multifactorial model includes the influence of many genetic and environmental factors, the risk does not change

<table>
<thead>
<tr>
<th>Population</th>
<th>Incidence risk ratio for idiopathic VTE [19]</th>
<th>Rate of diagnosis of VTE per 100,000 person-years [20, 156]</th>
<th>Adjusted standardised VTE incidence rates per 100,000 person-years [21]</th>
<th>Rate/10,000 person-years [157]</th>
<th>Incidence rate per 100,000 people</th>
</tr>
</thead>
<tbody>
<tr>
<td>African-American</td>
<td>1.27</td>
<td>134–155</td>
<td>141</td>
<td>2.2</td>
<td>11 [158]</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.60</td>
<td>55</td>
<td>0.9</td>
<td>6.2</td>
<td>11.0 [24,25]</td>
</tr>
<tr>
<td>American Indians</td>
<td></td>
<td>34–71</td>
<td>33.1* [159]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>0.26</td>
<td>21</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>1.0</td>
<td>122–131</td>
<td>104</td>
<td>2.1</td>
<td>108 [160]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>124 [161]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>160 [162]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>182 [163]</td>
</tr>
</tbody>
</table>

*hospitalisation rate; *VTE rate.
as drastically with the presence or the absence of one of them. This implies that the susceptibility to VTE is largely accounted for by the clustering of several, possibly inherited, risk factors. Hyperhomocystinaemia is a mild cause of VTE and provides an excellent example of the interaction among genetic (MTHFR 677TT genotype) and acquired factors (vitamins intake) affecting an intermediate phenotype (38). As the cause for VTE, a multifactorial non-Mendelian inheritance model was recently formally proposed, while a model postulating a purely environmental cause was rejected (6). Usually, geneticists admit that in a multifactorial disease several, but not an unlimited number, loci are involved in the expression of the trait; there is no classical (Mendelian) dominance or recessivity at each of these loci; the loci act in concert in an additive fashion, each adding or detracting a small amount from the phenotype; and the environment interacts with the genotype to produce the final phenotype. None of them is neither necessary nor sufficient for the disease to occur but makes it more likely that one will become ill. This means that liability to develop the disease is continuously distributed in the population because of the additive effects of genetic and environmental factors. Our current thinking is that only the individuals whose liability exceeds a certain threshold will manifest the disorder. However, because the phenotype alone may not be able to distinguish different subgroups, genetic heterogeneity may affect the ability to identify genetic risk factors associated with the disease.

Genetic heterogeneity of VTE

Among the three major causes of thrombosis, postulated by Virchow in 1856 (39); blood hypercoagulability plays a major role in VTE development. Reports on families in which a high number of VTE events suggested a familial hypercoagulability emerged as early as 1905 (40, 41). Inherited VTE risk factors, resulting from single-gene defects, have been known since 1965 (42–45). Then, a large array of gene mutations was suggested to modulate such risk (46). Usually, genetic heterogeneity refers to the presence of a variety of different genetic mechanisms, due to different mutations in the same gene or occurring in different unrelated loci, which lead to the same clinical disease, a finding common to many human diseases including VTE. Because most of these gene mutations give rise to a hypercoagulability state, one can refer to them as “thrombophilias”. Thus, genetic thrombophilia can be defined as the tendency to develop blood clots as a result of a genetic/ inherited molecular defect (47). From a clinical point of view, multifactorial diseases as VTE are the result of a large collection of a number, even high, of common and rare, private, genetic and acquired risk factors.

Clinical usefulness of genetic information

Rare (sporadic) and common (polymorphic) mutations in gene loci involved in the coagulation pathway are known to affect the risk for VTE occurring not only at lower extremities but also, in Caucasians, at upper extremities and in the portal vein district (48, 49). Identification of genetic risk factors, genetic profiling, may be of help in order to predict VTE risk and can guide decisions on prophylaxis in or avoidance of high-risk situations. A series of easy genetic tests has set-up to investigate the presence of sporadic mutations, using functional tests (i.e. for deficiency of a natural coagulation inhibitor), or polymorphisms, by means of genetic techniques (as for gain-of-function gene variants). This has led, mainly in Western countries, to a marked increase in the screening for these abnormalities in clinical laboratories (50, 51). However, since human populations are outbred, and each individual is heterogeneous for several million polymorphisms, the impact in an individual of each gene variant depends on the potential contribution of other genetic and acquired risk factors. Regardless of circumstantial risk factors, caution must be used in applying the same panel of genetic tests for thrombophilia to members of different ancestries.

Deficiencies of natural coagulation inhibitors

Deficiencies of the endogenous anticoagulant proteins antithrombin (AT), protein C, and protein S were the first identified genetic causes of VTE. Albeit rather uncommon risk factors (4), until several years ago they accounted for most of the few cases in which an inherited cause of thrombosis could be identified. Rare severe alleles have been implicated in all forms of inherited deficiency of natural coagulation inhibitors. Hundreds of different loss-of-function mutations, which result in very low or absent gene product activity, have been detected in gene loci encoding for AT, protein C, and protein S. All of these mutations are individually rare and, because half of normal levels do not ensure the required anticoagulant activity, each independently confers a high risk for VTE. These loss-of-function gene mutations result in a dominant pattern of inheritance, as for the biochemical (or intermediate) phenotype is concerned, and persons carrying two mutant alleles show a severe, often fatal, disease (52, 53). In Caucasian unselected patients presenting with DVT the aggregate prevalence of these three conditions is approximately 5–7% (4, 46, 47). At variance with this, deficiencies in AT, protein C, and protein S were consistently reported to be significantly higher in Chinese, Thai, and Japanese patients with VTE (29, 54–59) (Table 2). The most frequent defect was the occurrence of a protein S deficiency, followed by protein C and AT deficiencies. In keeping with this, the prevalence of AT, protein C, and protein S deficiencies in general population settings from China, Japan, and Thailand has also been reported (58, 60, 61) to be higher than that found in Caucasians (6, 46, 47). These findings indicate that an overall higher prevalence of this class of genetic risk factors may exist in general Asian population and has to be expected in patients with VTE from the same genetic background. In healthy African subjects, significantly lower levels of protein S, and slightly of protein C but not of AT, were observed as compared with those recorded in Caucasian controls (62). These findings are partially in agreement with previous data obtained in...
an asymptomatic population (63). However, in African VTE patients the prevalence of deficiencies of natural coagulation inhibitors (7.7%) was comparable to that expected in Caucasians (64).

**FV:Q506 (FV Leiden)**

Since it was reported (65), the inherited resistance to activated protein C (APC resistance) has been shown to be the most frequent inherited defect associated with VTE accounting for most of cases of unselected patients with DVT (4). The point mutation G1691A in the factor (F)V gene (FV:Q506), later called FV Leiden, is a common mutation, gene polymorphism, which confers an improvement in the activity of the gene product by slowing the inactivation rate of FVa by APC (66). FV Leiden is responsible for the vast majority of APC resistance cases and, in turn, is the leading cause of genetic thrombophilia in VTE patients. From a genetic point of view, FV Leiden is a gain-of-function mutation and shows a dominant pattern of inheritance of the biochemical phenotype. This common gene variant, although it is a major contributor, has a low penetrance of disease by conferring only a small degree of the VTE risk. Population studies in asymptomatic individuals established the FV Leiden mutation as one of the most common susceptibility alleles in Caucasians, ranging from 2 to 7% (67, 68). However, when analysing allele frequency distributions it is particularly important to recognise that definite conclusions about their geographical distribution require assuming that the test population is randomly mating, has no substructure, has no recent immigration, loci are neutral (or allelic frequencies are maintained by selection behaviorally) and that sampling is representative of the population. In addition, it needs to assess the Hardy-Weinberg equilibrium. Departures from it, if not due to chance, point to violation of these assumptions and data should be rejected or used only with caution. Therefore, methodological limitations may represent major pitfalls that must be kept in mind when examining the literature on population genetics, mainly when gene frequencies are assessed within a small geographical setting.

Notwithstanding, all data confirm a widespread presence among Caucasians and suggest a general South to North Europe positive gradient (69, 70). Surveys in populations without a substantial Caucasian admixture found that the FV Leiden allele frequency is very low, or even rare (Table 2) (68–73). This population-based distribution fits with the time for the origin of the mutation, estimated arising in a single Caucasian ancestor approximately 21,000 years ago (74).

**FII G20210A**

A polymorphism, a G to A nucleotide transition at nt 20210 in the 3’ untranslated region of the factor (F)II gene represents the second most frequent genetic risk factor for VTE in humans (75). The site of mutation is located in close proximity to the cleavage site of the mRNA precursor to which poly A is added (76). Also the FII A20210 behaves as a gain-of-function mutation, enhancing the gene expression and increasing protein activity, and the transmission of the intermediate phenotype is dominant (77). Carriership of this mutation is associated with an estimated three-fold increased risk of VTE (4). As for the FV Leiden, the FII A20210 distribution is restricted quite exclusively in Caucasians with a South to North gradient in Europe and a prevalence of 3.0 and 1.7%, respectively (78). Actually, the mutation is rare in African, Asian and indigenous populations of America and Australasia (Table 2) (79–83). Among African-Americans and Asian-Americans, the allelic prevalence is intermediate, likely a result of population admixture. The high prevalence of FII A20210 allele in Caucasians reflects a founder effect dating back approximately 24,000 years ago (84). This estimate is very close to that calculated for the emergence of the FV Leiden. In addition, both common alleles share similar features: a gain-of-function mutation with a low penetrance of the phenotype; and a similar geographical distribution. The latter may explain, at least in part, the higher VTE incidence in Caucasians than in settings with a different genetic background. Asians and American Indians share a favourable genetic profile, representing one explanation for the lower VTE incidence recorded in these two geographical groups.

### Table 2: Estimated prevalences of carriers of genetic abnormalities of components of the coagulation pathway.

<table>
<thead>
<tr>
<th>Region</th>
<th>AT (%)</th>
<th>PC (%)</th>
<th>PS (%)</th>
<th>FV Leiden (%)</th>
<th>FII A20210 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>African</strong></td>
<td>1–6 [64, 164]</td>
<td>4–6 [64, 164]</td>
<td>2–3 [64, 164]</td>
<td>&lt;0.1</td>
<td>0.1* [72]</td>
</tr>
<tr>
<td><strong>American</strong></td>
<td>0 [165]</td>
<td>5 [165]</td>
<td>2 [165]</td>
<td>0–1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td><strong>Asian</strong></td>
<td>2–5 [54–61]</td>
<td>8–19 [54–61]</td>
<td>8–30 [54–61]</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td><strong>European</strong></td>
<td>1</td>
<td>3</td>
<td>1–2</td>
<td>3–7</td>
<td>2–4</td>
</tr>
<tr>
<td><strong>Oceania</strong></td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
</tr>
</tbody>
</table>
Geographic distribution of FV Leiden and FII 20210A prevalence suggests that both arose roughly at the same time in the western branch of the Indo-European group and spread toward Europe following migration of Neolithic farmers (85). However, whether Caucasians represent a founding population with a high prevalence of FV Leiden and FII 20210A mutations, why they are so common today? There is no simple reason, and one could reasonably provide two possible explanations: the balanced polymorphic alleles, FV Leiden and FII 20210A, may be the consequence of the change in the gene pool that occurs in a small population due to chance (genetic drift), but also reflect a selective advantage in heterozygotes. It is difficult to distinguish effects of population history from those of natural selection. The former hypothesis implies that both alleles have been largely neutral for most of human history, and that this distribution is the result of the founder effect and genetic drift. The latter implies that adaptive evolution by acting through positive or negative natural selection forces shapes the human genome. FV Leiden and FII 20210A mutations have been suggested to modulate the clinical phenotype of haemophilia (86), to play a role in reducing blood loss related to menstruation and post-partum (87) or trauma and transfusion after cardiac surgery (88), to be associated with a decreased risk of intracranial haemorrhage (89), and to reduce mortality due to severe infections (90). Finally, the FII locus has recently been found to show evidence of positive selection (i.e. advantageous mutations may be occurred) (91). However, these suggestions are far to be conclusive: the selective pressure could have been too recent or too weak to leave a trace in sequence variation data.

Nevertheless, these mutations are almost absent in African populations despite they show a VTE risk and a prevalence of VTE thrombosis (ranging from 1.8 to 2.5– fold) associated with the non-O blood group and a recent meta-analysis confirmed this linkage (93–105). ABO blood groups vary extensively around the world and not all population settings carry the same frequencies. Selective pressure on frequency distributions has been suggested (106). For example, cholera and other diarrhoeal diseases that may be substantial causes of death in sub-Saharan Africa and Asia are more common or more severe in group O individuals (107, 108). On the other hand, ABO blood groups affect host susceptibility to severe Plasmodium falciparum infection, O alleles being associated with a lower risk of having severe malaria phenotypes (109). Asian people carry the relatively high frequency of the B blood group and Africans show a greater proportion than Caucasians (Table 3). The frequency of blood group O in Europe increases going from Southern to Northern countries and is the most common among Americans and Hispanics. Interestingly, as compared to Caucasians and African-Americans, American Indians and Hispanics have a higher proportion, well above 50%, of O blood group and a lower VTE rate (110).

The ability to change the H-determinant to obtain the different antigens of the ABO blood group has been correlated with the capacity to modify N-linked glycosylation of vonWillebrand factor (VWF). Actually, the occurrence of blood group A, B, and H(O) structures in the N-linked oligosaccharides of vWF has been shown (111). In individuals with the ABO Bombay phenotype, a rare condition in which the H determinant is not formed, VWF:Ag levels and susceptibility to ADAMTS13 proteolysis were significantly lower than in non-O and also lower than in group O individuals (112). Thus, changes in VWF glycan composition driven by different ABO allele products not only influence plasma VWF antigen levels, but may also have significant differential effects on VWF biologic activity (113), determining the rate of hepatic clearance and of VWF clearance rate of proteolysis by ADAMTS13 (114). ABO blood groups account for about 20% of the total variance in VWF:Ag (115, 116), and VWF has been observed to be significantly higher in African(-Americans) than in Caucasians (116–118) and Indians (119). In keeping with this, black patients are more likely to have elevated FVIII:C levels. An increase of VWF:Ag, and FVIII:C activity, have been suggested as the major reason for the high risk of VTE in persons with non-O blood groups (120). Finally, genetic background contributes for an additional 7% of the total variance in VWF:Ag, independently of the ABO blood group system (116).

### ABO blood group

Most studies indicate an increased risk of VTE thrombosis (ranging from 1.8 to 2.5– fold) associated with the non-O blood group and a recent meta-analysis confirmed this linkage (93–105). ABO blood groups vary extensively around the world and not all population settings carry the same frequencies. Selective pressure on frequency distributions has been suggested (106). For example, cholera and other diarrhoeal diseases that may be substantial causes of death in sub-Saharan Africa and Asia are more common or more severe in group O individuals (107, 108). On the other hand, ABO blood groups affect host susceptibility to severe Plasmodium falciparum infection, O alleles being associated with a lower risk of having severe malaria phenotypes (109). Asian people carry the relatively high frequency of the B blood group and Africans show a greater proportion than Caucasians (Table 3). The frequency of blood group O in Europe increases going from Southern to Northern countries and is the most common among Americans and Hispanics. Interestingly, as compared to Caucasians and African-Americans, American Indians and Hispanics have a higher proportion, well above 50%, of O blood group and a lower VTE rate (110).

### Sickle cell trait

Hereditary haemolytic anemias are the most common monogenic diseases, with approximately 7% of the world’s population carrying a haemoglobinopathy, thalassemia and sickle cell disease (SCD) representing the most common forms (Table 4) (121). SCD refers

<table>
<thead>
<tr>
<th>Region</th>
<th>O (%)</th>
<th>A (%)</th>
<th>B (%)</th>
<th>AB (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>African</td>
<td>49</td>
<td>27</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>African-American</td>
<td>51</td>
<td>26</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>American</td>
<td>79</td>
<td>16</td>
<td>4</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Hispanic</td>
<td>57</td>
<td>31</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Asian</td>
<td>40</td>
<td>28</td>
<td>25</td>
<td>7</td>
</tr>
<tr>
<td>European</td>
<td>44</td>
<td>43</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Oceania</td>
<td>43</td>
<td>57</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
</tr>
</tbody>
</table>

to a collection of autosomal recessive genetic disorders characterised by the inheritance of two abnormal \( \beta \)-globin alleles, both carrying a single amino acid mutation (\( \beta \) Glu-to-Val) leading to sickle haemoglobin (HbS), or a HbS allele and another carrying an abnormal \( \alpha \)-globin allele, such as HbC or \( \beta \)-thalassemia. There is increasing evidence that SCD and \( \beta \)-thalassemia are characterised by a hypercoagulable state (122, 123). On the other hand, persons with a normal \( \alpha \) globin allele and a HbS allele (sickle cell trait) have been viewed as carrying a benign condition, having no apparent effect on life expectancy (124) partially protective against malaria. It is estimated that 300 million people worldwide carry the sickle cell trait, with the highest concentration in Africa, where the prevalence of sickle cell trait varies from 5% to 40%, and in the Mediterranean region (125). In communities with a remarkable diversity in genetic background, i.e. US, the prevalence of sickle cell trait is estimated at approximately 7% in African-Americans and 0.3% in European-Americans (126). However, a series of case reports and uncontrolled observational studies suggested that sickle cell trait is not completely benign (127–129). As compared with controls, persons with sickle cell trait showed a two-fold increased risk for idiopathic and provoked VTE as well as for first and recurrent VTE (130). Assuming this estimate, the proportion of VTE among individuals of African ancestry attributable to sickle cell trait is approximately 7%. Moreover, the risk is further increased in women with a sickle cell trait assuming oral contraceptives (131).

In persons with sickle cell trait, haemostatic abnormalities of laboratory markers suggesting an activation of coagulation, such as TAT complexes, D-dimers, and prothrombin fragment F1–2, have been observed (132). In SCD, repeated cycles of sickling and unsickling that result from polymerisation and depolymerisation of HbS lead to the disruption of membrane phospholipid asymmetry (133–136). The resulting phosphatidylserine exposure in the outer monolayer of sickle erythrocytes is thought to contribute to the haemostatic perturbations, providing a binding site for both factor X and prothrombinase complexes (137). Subclinical sickling of red cells probably occurs in most persons with sickle cell trait and may give an explanation for activation of the coagulation (138).

### Compound heterozygous haemoglobinopathies

A combination of differently mutated \( \beta \)-globin genes may be inherited in the same individual. These individuals (compound heterozygotes) produce a mixture of variant haemoglobins and belong to a clinically and genotypically very heterogeneous group of patients, ranging in severity from asymptomatic to a severe phenotype. Individuals with HbSC and HbS/\( \beta \)^+thalassaemia tend to have a more benign course of the disease than homozygous SCD patients (139). However, individuals with HbSC have increased risk for thromboembolic complications when compared with homozygous SCD patients (140).

Thalassaemia intermedia encompasses a wide spectrum of clinically and genotypically heterogeneous thalassaemia-like disorders, ranging in severity from the asymptomatic carrier state to the severe transfusion-dependent type. Patients with thalassaemia intermedia are at greater risk of VTE events (141–143), the overall prevalence ranging from 4% to 8%, and even higher in splenectomised patients (144). These patients may be more at risk because they have increased platelet counts and more circulating damaged erythrocytes. Likewise, thromboembolic events have been recorded in compound heterozygotes of \( \beta \)-thalassaemia and HbE (145). The high VTE incidence suggested all these form of com-

<table>
<thead>
<tr>
<th>Region</th>
<th>HbC (%)</th>
<th>HbS (%)</th>
<th>( \beta )-thalassaemia (%)</th>
<th>HbE (%)</th>
<th>Aggregate carrier (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>African</td>
<td>2.531</td>
<td>15.861</td>
<td>0.64</td>
<td>0.001</td>
<td>18.2</td>
</tr>
<tr>
<td></td>
<td>up to 50</td>
<td>up to 40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American</td>
<td>0.631</td>
<td>1.699</td>
<td>0.585</td>
<td>0.036</td>
<td>3.0</td>
</tr>
<tr>
<td>Eastern Mediterranean</td>
<td>0.047</td>
<td>0.744</td>
<td>2.85</td>
<td>0.044</td>
<td>4.4</td>
</tr>
<tr>
<td>European</td>
<td>0.02</td>
<td>0.202</td>
<td>0.967</td>
<td>0.021</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>up to 61</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South East Asian</td>
<td>0</td>
<td>1.345</td>
<td>2.07</td>
<td>2.595</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>up to 25</td>
<td>up to 50**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Western Pacific</td>
<td>0</td>
<td>0.001</td>
<td>1.141</td>
<td>1.726</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>up to 60**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* \( \alpha \)-carriers included; \( \beta \)Ivory Coast; \( \gamma \)Tanzania; \( \delta \)Greece, Italy; \( \gamma \)India; \( \beta \)++Thailand; ++Vietnam.
Pound heterozygous haemoglobinopathies as a hypercoagulable state, particularly if in the presence of splenectomy (144, 146). The distribution of the haemoglobinopathies widely differs through the world (147, 148). HbS is distributed widely at high frequencies (20%) throughout sub-Saharan Africa, the Middle East and parts of the Indian sub-continent, and HbC is found only in African populations. HbE occurs in the Eastern half of the Indian sub-continent and throughout South-East Asia, where carrier rates may exceed 60% of the population, but is absent elsewhere. β-Thalassaemias are widespread extending from the Mediterranean area and parts of Africa, throughout the Middle East, the Indian sub-continent, and South-East Asia, mainly regions where malaria was endemic. The carrier frequency for β-thalassaemia in these areas ranges from 1% to 20%. Thus, the possibility for an individual to be a compound heterozygote is high in persons from Africa, Middle East, the Indian sub-continent, and South-East Asia (Table 4). However, as a result of changing demographics, this condition is becoming more common in many other regions.

Population genetics of VTE

VTE is a complex disorder that is affected by many interacting genetic and non-genetic factors. Albeit investigations on VTE risk factors have been mostly carried out in Western countries and as expected mainly in Caucasians, there is compelling evidence that persons from a different genetic background often carry a significant different VTE risk, suggesting that geographically prevalent susceptibility alleles play a pivotal role. Because humans form a genetically homogeneous group, genetic variations are expected to be shared widely among populations. At variance with this, as a result of processes occurred during much of their history, such as mutation, migration, genetic drift, and selection, genetic variation is geographically structured. Genetic studies of groups from different ancestries produced diverse results because of the variations in their genomes that have evolved over generations to cope with environmental pressures specific to their region (149). There is little doubt that some mutations offer a better protection to malaria, whereas some suppositions have been formulated to explain other geographically prevalent susceptibility alleles. Depending on the particular combination of mutations present, different populations carry their own risk for VTE. Well-known thrombophilic risk factors, mainly FV Leiden and FII A20210 alleles, confer a susceptibility to disease in Caucasians. In persons from African ancestry, the main source of the genetic risk appears to be sickle cell trait and compound heterozygous haemoglobinopathies. The situation is more puzzling in Asia where distribution of the haemoglobinopathies is high in the Middle East (sickle cell trait), parts of the Indian sub-continent (sickle cell trait and HbE) and throughout South-East Asia (HbE). On the contrary, they are relatively rare in northeast Asia and Japan, where the VTE incidence appear to be low. In the latter geographical regions, the most represented genetic risk factor for VTE is a deficiency of natural coagulation inhibitors. Epidemiological data suggest that American Indians and Hispanics have an intermediate VTE risk.

American Indians have an Asian origin (150), although current non-European-derived populations in the Americas show higher levels of African and European genetic admixture (151). Hispanics possess a complex genetic structure reflecting recent admixture with American Indian and African populations (152). As a consequence, they carry a high frequency of O blood group and intermediate frequencies of sickle cell trait. Therefore, knowledge of ancestry can be clinically important providing information on specific risk factors (Fig. 1) as other personal features, such as gender and age. In addition to genetic structure, ancestry may capture relevant cultural habits and environmental determinants shared by specific population groups. However, genetic diversity among populations is lower than that existing within populations (153, 154). It has to be carefully taken into account that understanding of the human genome does not support the existence of discrete boundaries between populations: each person represents a unique combination of alleles (155). Keeping this information in mind, knowledge about individual genetic background may be useful to identify genetic risk factors, which can interact with other genes or the environment during the life course and modify the VTE risk.

© Schattauer 2011

Thrombosis and Haemostasis 105.2/2011
A brief note of terminology

From a genetic point of view, as a measure of the genetic distance or variation terms as “race” and “ethnicity” are biologically meaningless. Actually, populations do not present genetic boundaries and genetic diversity among them is lower than that existing within the vast majority being found among individuals from Africa. The term “ancestry” is used in the present work to capture information correlated with genetic relatedness. However, the definition is arbitrary because ancestry refers also to shared cultural and environmental characteristics.

Search strategy

All relevant published studies of human medicine were searched for. In MEDLINE, a combination of several Medical Subject Heading (MeSH) terms was used in the following way: venous thrombosis/thromboembolism AND (ancestry OR ethnic OR inherited OR race OR Africa OR America OR Asia) AND (epidemiology OR prevalence OR gene OR risk factor). Titles and abstracts to be reviewed were listed. On the basis of abstract reading, full papers were selected, reviewed and those that matched the selection criteria were retained. Then, abstracts of the related articles of this first series of papers were explored, reviewed and eventually the full paper was read if deemed appropriate. All references of the retained papers were also examined. This process was performed iteratively until no new suitable study could be found.

References

13. White RH. The epidemiology of venous thromboembolism; Circulation 2003; 107 (23 Suppl 1); 4–8.


