

Rapid Communication

Geographic Distribution of the 20210 G to A Prothrombin Variant

F. R. Rosendaal[1 2], C. J. M Doggen[1], A. Zivelin[3], V. R. Arruda[4], M. Aiach[5], D S. Siscovick[6], A. Hillarp[7], H. H. Watzke[8], F. Bernardi[9], A. M. Cumming[10], F. E. Preston[11], P. H. Reitsma[12]

From the[1]Department of Clinical Epidemiology, Leiden University Medical Centre, Leiden The Netherlands,
[2]Department of Hematology, Leiden University Medical Centre, Leiden, The Netherlands,
[3]Institute of Thrombosis and Hemostasis, Department of Hematology, Sheba Medical Centre Tel-Hashomer Israel,
[4]Hematology-Hemotherapy Centre, State University of Campinas, Campinas, Brazil,
[5]Laboratory of Hemostasis, Broussais Hospital, Paris, France,
[6]Department of Epidemiology, Department of Internal Medicine, University of Washington, Seattle, USA,
[7]Department of Clinical Chemistry, Lund University, Malmo, Sweden,
[8]Department of Hematology and Hemostasiology, University of Vienna, Vienna, Austria,
[9]Department of Biochemistry and Molecular Biology University of Ferrara, Ferrara, Italy,
[10]Department of Haematology Manchester Royal Infirmary Manchester, United Kingdom
[11]Haemophilia and Thrombosis Centre, Royal Hallamshire Hospital, Sheffield, United Kingdom
[12]Laboratory for Experimental Internal Medicine, Academic Medical Centre Amsterdam, The Netherlands

Summary

A variant in prothrombin (clotting factor II), a G to A transition at nucleotide position 20210, has recently been shown to be associated with the prothrombin plasma levels and the risk of both venous and arterial thrombosis The purpose of this study was to investigate the prevalence of carriership of this mutation in various populations

We combined data from 11 centres in nine countries, where tests for this mutation had been performed in groups representing the general population We calculated an overall prevalence estimate, by a precision weighted method, and, since the distribution of the prevalences did not appear homogeneous, by an unweighted average of the prevalences We examined differences in the prevalences by geographical location and ethnic background as a possible explanation for the heterogeneity

Among a total of 5527 individuals who had been tested, 111 heterozygous carriers of the 20210A mutation were found The prevalence estimates varied from 0 7 to 4 0 between the centres The overall prevalence estimate was 2 0 percent (CI95 1 4-2 6%) The variation around the summary estimate appeared more than was expected by chance alone, and this heterogeneity could be explained by geographic differences In southern Europe, the prevalence was 3 0 percent (CI95 2 3 to 3 7%), nearly twice as high as the prevalence in northern Europe (1 7%, CI95 1 3 to 2 2%) The prothrombin variant appeared very rare in individuals from Asian and African descent

The 20210A prothrombin variant is a common abnormality, with a prevalence of carriership between one and four percent It is more common in southern than in northern Europe Since this distribution within Europe is very different to that of another prothrombotic mutation (factor V Leiden or factor V R506Q), founder effects are the most likely explanation for the geographical distribution of both mutations

Correspondence to Dr F R Rosendaal, Dept of Clinical Epidemiology, Bldg 1, C0-P, University Hospital Leiden P O Box 9600, NL 2300 RC Leiden, The Netherlands – Tel +31 71 5264037, FAX +31 71 5248122, e mail Rosendaal@rullf2 leidenuniv nl

Introduction

Recently, we reported a mutation in clotting factor II (prothrombin) that was associated with plasma levels of prothrombin and with the risk of deep-vein thrombosis [1] The single base substitution (G to A) at position 20210 [2] of the 3' untranslated region of the prothrombin gene was found in 18 percent of probands of thrombophilic families, six percent of unselected consecutive patients with deep-vein thrombosis and 2 percent of healthy controls [1] Thus, carriership of this mutation led to a three-fold increased risk of venous thrombosis [1]

The mutation also affects the risk of arterial disease among young women we found a four fold increased risk of myocardial infarction among carriers of the factor II 20210A allele [3], while among men the risk was increased 1 5-fold [4]

Several recent reports, mainly in the form of meeting abstracts and presentations, have addressed the venous and arterial risk increase in other populations [5-11] One of the problems with the studies is that the relatively low prevalence of the mutation in the general population leads to large uncertainties in the risk estimates Moreover, it is possible that the prevalence depends on ethnicity and geography, as was found for the factor V Leiden mutation

Therefore, we decided to combine our data on the frequency of the factor II 20210A variant in the general population, first, to arrive at a summary allele frequency estimate, and secondly, to examine whether there is heterogeneity in its distribution

Methods

At the 11 centres from nine different countries, testing was performed for prothrombin 20210 G to A carriership in healthy individuals The origin of each study group is listed in Table 1 Most centres were in Europe and the Middle East, and ranged from Scandinavia to the Mediterranean In addition, we included a study from South-America (Campinas, Brazil) and from North

Table 1 Numbers and origin of the tested individuals per centre

centre	n	origin of tested individuals
Europe/Middle East		
Leiden [1 4]	1760	medical controls (n=646 n−291)[1] population controls (n=474) blood donors (n−249) hospital staff (n 100)
Tel Aviv [17]	1081	population controls[2]
Amsterdam [5]	400	hospital staff (n 272) population controls (n−138)
Paris [6]	398	medical controls[3]
Malmo [7]	282	blood donors (n−116) medical controls (n=166)[4]
Vienna [8]	213	population controls (n=111) newborns (n=102)
Ferrara [9]	176	population controls
Manchester [10]	164	blood donors
Sheffield [11]	150	blood donors (n−96) population controls (n−54)
America		
Brazil [18]	522	population controls (n=83) newborns (n−295) hospital staff (n=144)[5]
United States [3]	384	population controls

[1] orthopedic surgery patients (n−646) rheumatoid arthritis patients (n−291)

[2] from various ethnic groups Ashkenazi North African Middle Eastern (Iraqi and Iranian) Yemenite and Ethiopian Jews and Arabs

[3] medical checkup centre visitors excluding those with cardiovascular disease

[4] individuals who underwent orthopedic surgery (n−142) or gave blood samples for various reasons (n−24) excluding those with a history of cardiovascular disease

[5] from three regions with different ethnic background Amazonian Indians (n−83) Caucasians (n=295 Campinas south-east) and with a mainly African descent (n−144 Bahia north east)

America (Seattle, United States) These latter two centres and the Tel Aviv center also included non-whites, while all other solely or mainly comprised Caucasian individuals

The selection of individuals varied between the various studies In most cases, an effort was made to include healthy individuals, as a sample from the population These samples ranged from hospital employees, to blood donors, newborns, and various groups of individuals visiting health care providers ('medical control groups') and, finally formal control groups from the general population (Table 1) In none of the centres, individuals were selected in a way that would lead to preferential inclusion of individuals with thrombotic disease, while in most centres those with a history of cardiovascular disease were excluded Approximately equal numbers of both sexes were included

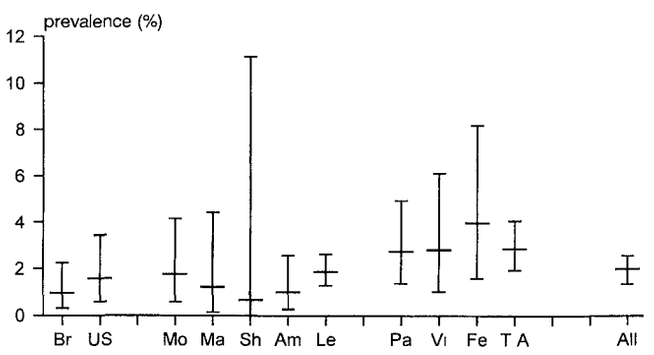

Fig 1 Prevalence of carriership (in percentage points) of prothrombin 20210A, per centre, in geographical order America US = United States (Seattle), Br = Brazil, northern Europe Mo = Malmo, Ma = Manchester, Sh = Sheffield, Am = Amsterdam, Le = Leiden, southern Europe/Middle-East Pa = Paris, Vi = Vienna, Fe = Ferrara, T A = Tel Aviv The European centres are in order of decreasing latitude On the right (All) the overall summary estimate is given (straight average) All estimates are shown with a 95% confidence interval

All centres except one used a strategy for the detection of the 20210 A allele similar to the one from the original report [1] Leiden, Campinas, Amsterdam, Seattle Malmo, Vienna, and Manchester used amplification primers exactly as described before [1] Tel-Aviv, Paris, and Sheffield used primers that differ from the originals but that still use the concept of introducing a Hind III recognition site with a mutagenic primer The sequences of these primers and the reaction conditions can be obtained from the respective authors Each centre reported that their local protocol was robust, which indicates that primer choice is not critical in the assay for the prothrombin mutation Ferrara used allele specific amplification (ASA) to discriminate the prothrombin alleles Two allele specific primers, 5'-CTGGGAGCATTGAAGCTC 3' (nucleotides 20227-20210) and 5' CTGGGAGCATTGAAGCTT 3' (nucleotides 20227-20210) were used as reverse primers in combination with the forward primer 5' GGGAAATATGGCTTCTACA-3' (nucleotides 20035-20054) The presence of an additional mismatch in the ASA primers increased selectivity of the ASA

We categorized the prevalence of carriers of the prothrombin 20210 A allele per centre, and pooled the estimates by two methods first, by calculating an average weighted for the number of carriers per centre This average gives most weight to the more precise estimates, since the variance, under Poisson assumptions, depends on the number of carriers This is the appropriate method when it is assumed that the 'true' allelic prevalence is the same in all study groups, and that all observed variation in prevalences is random This precision-weighted average is not the appropriate approach when the allelic prevalences are not constant We assessed homogeneity of the prevalences by Poisson modeling of the prevalence rates and comparing the model with only the constant, to the saturated model with a dummy variable for centre, with the likelihood ratio statistic As a second approach we calculated an unweighted average of the prevalence estimates (ie, the average with equal weights, which is the straight mean of the prevalence estimates) This unweighted straight average is more appropriate when homogeneity cannot be assumed Prompted by the reports on a north south difference in the prevalence of another common prothrombotic mutation, factor V Leiden [12], we stratified the analysis by geography into northern and southern European countries (the latter including the Middle East)

For an analysis by ethnicity, we grouped together African Americans from the United States and Brazil, and Ethiopian Jews from Israel as individuals with an African descent Native-Americans, i e , American Indians from north western United States and the Amazon region in Brazil, were grouped together as individuals with an Asian background

Confidence intervals for individual prevalence estimates were based on exact Poisson limits [13] and for pooled estimates on the normal approximation of the Poisson distribution

Results

In the 11 centres, a total of 5527 individuals were tested The total number of individuals per centre varied from 150 to 1760 Among these 5527 individuals, 111 heterozygous carriers of prothrombin 20210A were identified No homozygous individuals were found The number of carriers varied from 1 to 33 per centre, and the prevalence of carriership ranged from 0 7 percent to 4 0 percent (Fig 1)

The overall prevalence was 2 3 percent (CI95 1 9 to 2 6%), when a weighted average was calculated The variation of the prevalence estimates appeared to be more than would be expected by random variation around one underlying mean (LRS = 16 9, p = 0 08) Therefore, we calculated the average with equal weights, this straight average of prevalences was 2 0 percent (CI95 1 4 to 2 6%)

There was a clear relationship of the prevalence with the geographic latitude in Europe (including Israel) in the centres north of 50° N (Malmo, Manchester, Sheffield, Amsterdam, Leiden) there were 45 carriers among 2756 individuals tested, yielding a precision-weighted prevalence of 1 7 percent (CI95 1 3 to 2 2%), in the centres located

south of 50°N (Paris, Vienna, Ferrara, Tel Aviv) we found 55 carriers among a total of 1868, yielding a prevalence of 3.0 percent (CI95 2.3 to 3.7%). So, the prevalence of the prothrombin variant was nearly twice as common in the southern as in the northern regions ($\chi^2 = 9.05$, df = 1, p = 0.003). As can be seen in the figure, where the centres are shown in geographical order, in each of the four centres south of 50° N the prevalence estimate was higher than in each of the five centres north of 50° N.

Three centres included non-Caucasians: Brazil, Seattle and Tel-Aviv, which allowed a separate analysis of individuals with a different ethnic background. Among individuals with an African background, 2 carriers of factor II 20210A were found among 300 individuals (0/9 in Seattle, 2/144 in Bahia, Brazil, and 0/147 among Ethiopian Jews in Israel). Taken together, this accounts for a prevalence of carriership of only 0.67 percent (CI95: 0.08 to 2.4%, comparison with Caucasians χ^2 =3.02, p = 0.08). Among 103 Native Americans, ie, individuals with an Asian background (20 from Seattle, 83 from the Amazon region in Brazil) no carrier was detected (prevalence CI95: 0 to 3.6%, comparison with Caucasians χ^2 =2.24, p = 0.13).

Discussion

We have combined the data on the prevalence of carriership of the prothrombin 20210A variant from 11 centres, including a total of 5527 healthy individuals among whom 111 carried the variant. The estimate of the overall prevalence, under the assumption of homogeneity, was 2.3 percent (CI95 1.9 to 2.6%). The straight average of the prevalences was 2 percent (CI95 1.4-2.6%). We did not find any homozygous individuals This is not surprising, since with an allele frequency of 1 percent, we would expect homozygotes to occur in only 1 in 10,000.

The assumption of homogeneity did not appear true. Within Europe, there was an association with geographical location, with higher prevalences in the southern than in the northern countries. In southern Europe/Middle East, the prevalence was 3 0 percent (CI95 2.3 to 3.7%), and in northern Europe it was 1.7 percent (CI95 1.2 to 2.2%). This distribution in Europe is different to that reported for factor V Leiden, which has a high prevalence in northern Europe, especially in southern Sweden, and a low prevalence in Italy [12]. It has been hypothesized that such geographical differences might be the result of geographical differences in other risk factors, and thus by differences in genetic fitness and selection associated with gene-environment interaction. The finding of a very different geographical distribution for the prothrombin variant renders these explanations less likely. Probably, founder effects have played a much more important role. Interestingly, both factor V Leiden and prothrombin 20210A appear common in the Middle-East [14] and in Cyprus [9,15].

We have combined data from centres that used different criteria to recruit individuals, which raises questions about the comparability. Most centres excluded individuals with a history of thrombosis, or recruited blood donors and volunteers amongst whom such a history is unlikely. Although this would lead to lower estimates than obtained from a theoretically ideal study group of consecutive newborns, this difference is immaterial. Selection bias leading to overestimation would have occurred if for some reason individuals with thrombotic disease had been preferentially included. This bias did not occur, since the medical controls either comprised patients with disorders that are not associated with prothrombotic mutations (rheumatoid arthritis, orthopedic trauma), or, individuals without a history of thrombosis.

Factor V Leiden is extremely rare among non-Caucasians, which is in accordance with the estimated age of the mutation of 21,000 to 34,000 years [16], ie, occurring after the divergence of Africans from non-Africans and Caucasians from Asians [16]. Although the number of non-Caucasians in our collaborative study was relatively small, we found only 0.67 percent prevalence of carriers among individuals with an African background, and no carriers among Native Americans. In another recent report, no carriers were found among 40 Somalians [9]. These findings suggests a low prevalence of prothrombin 20210A among Africans and Asians, as has been reported also for factor V Leiden [15,16].

We conclude that prothrombin 20210A is a common prothrombotic genetic variant, that has a prevalence of carriership between one and four percent. There are geographic differences in this prevalence, with a higher frequency in southern than in northern Europe. In areas with a high prevalence this variant may contribute significantly to the occurrence of thrombotic disease in the population.

References

1 Poort SR, Rosendaal FR, Reitsma PH, Bertina RM A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis Blood 1996, 88 3698-703
2 Degen SJ, Davie EW Nucleotide sequence of the gene for human prothrombin Biochemistry 1987, 26 6165-77
3 Rosendaal FR, Siscovick DS, Schwartz SM, Psaty BM, Raghunathan TE, Vos HL A common prothrombin variant (20210 G to A) increases the risk of myocardial infarction in young women Blood 1997, 90 1747-50
4 Doggen CJM, Manger Cats V, Bertina RM, Rosendaal FR Interaction of coagulation defects and cardiovascular risk factors increased risk of myocardial infarction associated with factor V Leiden or prothrombin 20210A Circulation 1998 (in press)
5 Franco RF, Trip MD, ten Cate H, Prins MH, Kastelein JJP, Reitsma PH The prevalence of the 20210 G>A mutation in the 3'-untranslated region of the prothrombin gene in patients with premature coronary artery disease (abstract) Thromb Haemost 1997, Suppl 769-70
6 Alhenc-Gelas M, Le Cam Duchez V, Emmerich J, Frebourg T, Fiessinger JN, Borg JY, Aiach M The 20210A allele of the prothrombin gene is not frequently associated with the factor V Arg 506 to Gln mutation in thrombophilic families Blood 1997, 90 1711
7 Hillarp A, Zoller B, Svensson PJ, Dahlback B The 20210 A allele of the prothrombin gene is a common risk factor among Swedish outpatients with verified deep venous thrombosis Thromb Haemost 1997, 78 990-2
8 Watzke HH, Schuttrumpf J, Graf S, Huber K, Panzer S Increased prevalence of a polymorphism in the gene coding for human prothrombin in patients with coronary heart disease Thromb Res 1997, 87 521-6
9 Ferraresi P, Marchetti G, Legnani C, Cavallari E, Castoldi E, Mascoli F, Ardissino D, Palareti G, Bernardi F The heterozygous 20210 G/A prothrombin genotype is associated with early venous thrombosis in inherited thrombophilia and is not increased in frequency in arterial disease Arterioscler Thromb Vasc Biol 1997 (in press)
10 Cumming AM, Keeney S, Salden A, Bhavnani M, Shwe KH, Hay CRM The prothrombin gene G20210A variant prevalence in a U K anticoagulant clinic population Br J Haematol 1997, 98 353-5
11 Makris M, Preston FE, Beauchamp NJ, Hampton KK, Daly ME, Cooper P, Bayliss P, Peake IR Co-inheritance of the 20210A allele of the prothrombin gene increases the thrombotic risk in subjects with familial thrombophilia (abstract) Thromb Haemost 1997, Suppl 165
12 Cox MJ, Rees DC, Martinson JJ, Clegg JB Evidence for a single origin of factor V Leiden Br J Haematol 1996, 92 1022-5
13 CIBA-GEIGY AG Wissentschaftliche Tabellen Geigy Teilband Statistik Basel, CIBA-GEIGY, 1980
14 Seligsohn U, Zivelin A Thrombophilia as a multigenic disorder Thromb Haemost 1997, 78 297-301
15 Rees DC, Cox M, Clegg JB World distribution of factor V Leiden Lancet 1995, 346 1133-4
16 Zivelin A, Griffin JH, Xu X, Pabinger I, Samama M, Conard J, Brenner B, Eldor A, Seligsohn U A single genetic origin for a common Caucasian risk factor for venous thrombosis Blood 1997, 2 397-402
17 Zivelin A, Rosenberg N, Dardik R, Amit Y, Kenet G, Kornbrot N, Fridman A, Seligsohn U A low frequency of two genetic risk factors for thrombosis in Yemenite-Jews who manifest a decreased incidence of coronary artery disease (abstract) Thromb Haemost 1997, Suppl 222-3
18 Arruda VR, Von Zuben PM, Chiaparini LC, Annichino-Bizzacchi JM, Costa FF The mutation Ala677>Val in the methylene tetrahydrofolate reductase gene a risk factor for arterial disease and venous thrombosis Thromb Haemost 1997,77 818-21

Received November 4, 1997 Accepted December 18, 1997