Brief report

Prothrombin 20210G>A is an ancestral prothrombotic mutation that occurred in whites approximately 24 000 years ago

Ariella Zivelin, Ronit Mor-Cohen, Victoria Kovalsky, Nurit Kornbrot, Jacqueline Conard, Flora Peyvandi, Paul A. Kyrle, Rogier Bertina, Ferial Peyvandi, Joseph Emmerich, and Uri Seligsohn

Prothrombin 20210G>A and factor V Leiden are common prothrombotic mutations in whites for which founder effects have been established. In this study, we analyzed the frequencies of 5 single nucleotide polymorphisms (SNPs) and 9 microsatellites flanking the prothrombin gene *(F2)* in 88 homozygotes for 20210A and 66 homozygotes for 20210G. For estimating the age of the prothrombin 20210G>A mutation, we used the DMLE+2.0 program, which analyzed linkage disequilibria between the mutation and the multiple markers that had been assessed. This analysis yielded an age estimate of 23 720 years (95% credible set, 19 080-31 340 years). A similar analysis by the DMLE+2.0 program was performed on 5 SNPs from previously studied homozygotes for factor V Leiden and controls that yielded an age estimate of 21 340 years (95% credible set, 16 880-29 480 years). The occurrence of the 2 mutations in whites toward the end of the last glaciation and their presently wide distribution in whites suggest selective evolutionary advantages for which some evidence was reported (diminished blood loss) or is controversial (protection against infections). (Blood. 2006;107:4666-4668)

Introduction

The prothrombotic prothrombin 20210G>A mutation, first described in 1996 as an inherited thrombophilia,[1] confers an increased risk for venous thrombosis with an odds ratio of 3.8 (95% CI, 3.0-4.9).[2] This G to A change in the 3′-untranslated region of the prothrombin mRNA probably introduces a more effective polyadenylation site giving rise to increased mRNA and protein expression.[3,4] Prothrombin 20210G>A is prevalent in whites but not in Africans or Asians.[5] In a previous study of 123 white heterozygotes and 10 homozygotes for prothrombin 20210G>A, haplotype analyses established a single ancestral origin for the mutation.[6]

The aim of the present study was to estimate the age of the prothrombin mutation by calculations based on analyses of the frequency and variability of polymorphisms flanking both sides of the prothrombin gene *(F2)* in 20210A homozygotes and 20210G homozygous controls from various white populations.

Study design

Unidentifiable stored DNA samples from 88 unrelated whites who were previously diagnosed as homozygotes for prothrombin 20210A were retrieved or sent to the laboratory in Israel and verified to harbor the genotype. Nineteen subjects were from Israel, 20 from France, 17 from Italy, 16 from Austria, 6 from Germany, 4 from Spain, 3 from the Netherlands, and 3 from the United States, Norway, and Australia. Unidentifiable stored DNA samples from 66 healthy whites (45 from Israel, 18 from France, and 3 from Austria) who were homozygotes for prothrombin 20210G were also examined. The institutional review board of the Sheba Medical Center approved the study, and patients provided informed consent.

Nine multiallelic microsatellite polymorphisms (3 or 4 nucleotide repeats) and 5 single nucleotide polymorphisms (SNPs) flanking the prothrombin gene *(F2)* over distances of approximately 1 Mb 3′ and 2.4 Mb 5′ to the gene were analyzed in homozygotes for 20210A and 20210G. The microsatellites and SNPs were identified in the human genome working draft (www.genome.ucsc.edu) and NCBI SNP database (www.ncbi.nlm.nih.gov)[7], respectively. The analyses of the microsatellites were performed by specific polymerase chain reactions (PCRs) in the presence of radiolabeled dCTP, separation by polyacrylamide gel electrophoresis, and assessment by autoradiography, and the SNPs were assessed by PCRs and restriction analyses (available upon request from azivelin@sheba.health.gov.il).

Estimating the age of the prothrombin 20210G>A mutation was done by the DMLE+2.0 software program (www.dmle.org). This program was designed for high-resolution mapping of disease mutations and for estimating their age. It is based on the observed linkage disequilibrium between a mutation and linked markers tested in DNA samples of unrelated affected individuals and controls.[8,9] The parameters used were as follows: genotypes of microsatellites and SNPs in affected (20210A) and unaffected (20210G) chromosomes; chromosome map distances taken from the human genome working draft; a population growth rate of 0.0075 in Europeans[10]; the ratio of affected chromosomes sampled and the product of the current European population—730 million (Population Reference Bureau)[11]; and 2% mean carrier frequency of prothrombin 20210A allele.[12]

The DMLE+2.0 program was also used for re-estimating accurately the age of factor V Leiden by using data on 5 SNPs examined in an earlier study in 117 homozygotes for factor V Leiden and 167 controls.[13]

From the Amalia Biron Research Institute of Thrombosis and Hemostasis, Chaim Sheba Medical Center and Tel Aviv University, Israel; Hôtel Dieu, Paris, France; University of Milan, Italy; University of Vienna, Austria; Leiden University Medical Center, The Netherlands; J. W. Goethe University, Frankfurt, Germany; and University of Paris Descartes, Institut National de la Santé et de la Recherche Médicale (INSERM) U765, Paris, France.

Submitted January 3, 2005; accepted February 7, 2006. Prepublished online as *Blood* First Edition Paper, February 21, 2006; DOI 10.1182/blood-2005-12-5158.

A.Z. and R.M.-C. contributed equally to this study.

Reprints: U. Seligsohn, The Amalia Biron Research Institute of Thrombosis and Hemostasis, Chaim Sheba Medical Center, Tel Hashomer 52621, Israel; e-mail: seligson@sheba.health.gov.il.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 U.S.C. section 1734.

Results and discussion

Figure 1A displays the frequencies of 5 SNPs flanking the prothrombin 20210G>A mutation and spanning approximately 1.6 Mb of the chromosome harboring the prothrombin gene *F2*. For 4 of the 5 SNPs, there was a significant difference in the frequency observed in 20210A and 20210G alleles and for the fifth, at a distance of 0.734 Mb 5′ from the mutation, this difference was insignificant, likely because of recombinations.

Figure 1B displays the frequencies of 9 microsatellitic nucleotide repeats in 20210A and 20210G alleles. The microsatellites spanned approximately 3.2 Mb of the chromosome bearing the prothrombin gene *F2* (ie, from 2.4 Mb 5′ to the mutation to 0.8 Mb 3′ to the mutation). The differences in the frequencies were highly significant for the 3′ microsatellites and 5′ microsatellites up to 0.7 Mb from the mutation, and they diminished at a larger 5′ distance.

We then used the DMLE+2.0 program for estimating the age of the prothrombin 20210G>A mutation by separate analyses of the data for SNPs and the data for microsatellites. Figure 2 shows that when we used the SNP data and assumed 20 years for a generation, the age of the mutation was 21 060 years (95% credible set, 15 040-27 960 years), and when we used the microsatellite data, the age was 23 280 years (95% credible set, 18 680-28 720 years). Thus, these separate analyses yielded age estimates that were only 2200 years apart. When both SNP and microsatellite data were introduced into the program, the age estimate was 23 720 years (95% credible set, 19 080-31 340 years). Re-estimation of the mutational age of factor V Leiden by the advanced DMLE+2.0 program yielded 21 340 years (95% credible set, 16 880-29 480 years), which is remarkably close to the 24 000-year age estimate for the prothrombin 20210G>A mutation (Figure 2).

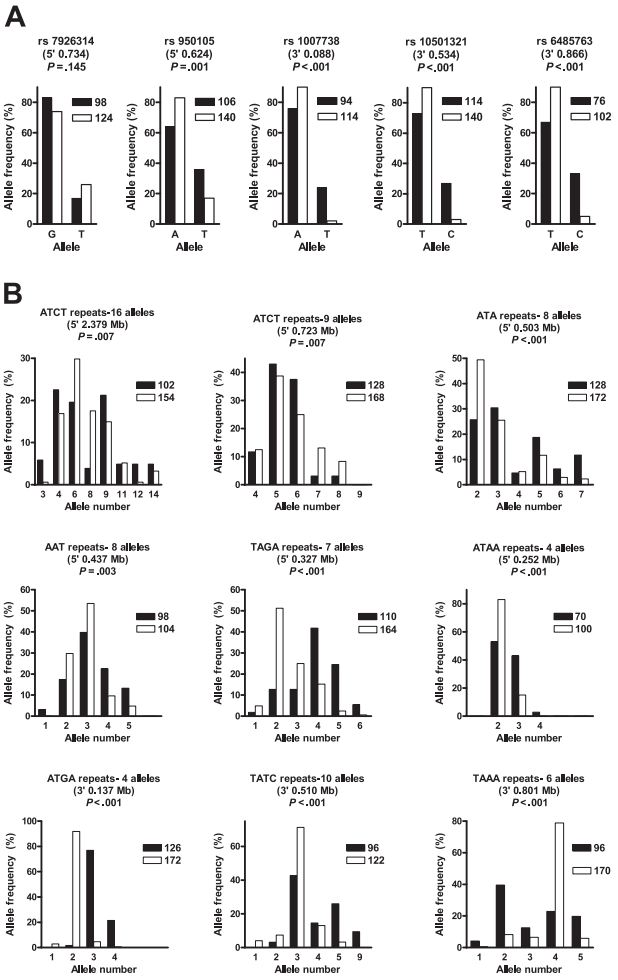

Figure 1. Frequency of extragenic polymorphisms in alleles bearing normal prothrombin 20210G (■) and mutant prothrombin 20210A (□). (A) Five diallelic polymorphisms designated by their rs number (NCBI SNP database). (B) Nine microsatellites (depicted are only the alleles with a frequency of more than 2.5% in either normal or mutant alleles). The number of chromosomes examined, distance of each polymorphism from the mutation, and the statistical significance of the difference in the frequency between the normal and mutant alleles are shown.

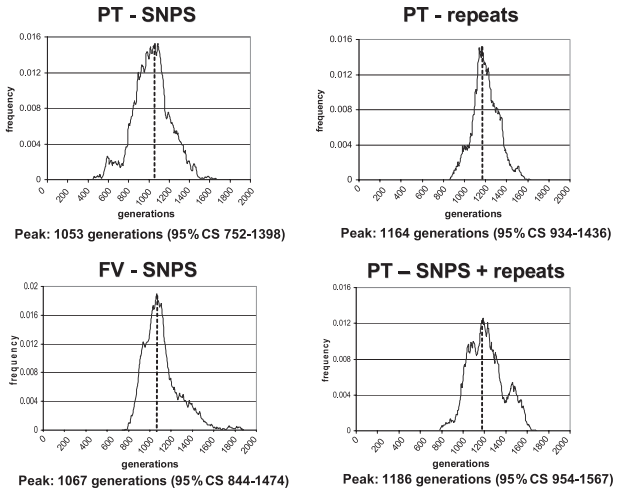

Figure 2. Age estimates of the prothrombin 20210G>A (PT) and factor V Leiden (FV) mutations by the DMLE+2.0 program. For the prothrombin mutation, separate analyses were carried out using single nucleotide polymorphisms (SNPs), microsatellites (repeats), and both. For FV Leiden only SNPs were used.

These data imply that the white founders of prothrombin 20210G>A mutation and factor V Leiden lived toward the end of the last Pleistocene glaciation. This period heralded extensive growth of populations, migration throughout the world, and development of agriculture. It is tempting to speculate that the founders of both mutations lived in the Middle East because currently the prevalence of both mutations is higher in ancient Middle Eastern populations than in other white populations. Thus, factor V Leiden is present in 12% to 14% of Arabs residing in Israel, Jordan, Lebanon, and Syria,[14,15] and prothrombin 20210G>A is present in 6.7% of Ashkenazi Jews, 5.5% of North African Jews, and 4% of Iraqi Jews,[6] whereas in European populations the prevalence of factor V Leiden (except for Southern Swedes) is approximately 4% to 5% and the prevalence of prothrombin 20210G>A is 1.7% to 4%.[12]

The widespread distribution of both mutations among all whites raises the question whether they have conferred a selective advantage during evolution. A selective disadvantage (ie, thrombosis) is unlikely because until recent centuries humans did not live long enough to manifest a meaningful incidence of thrombosis. On the other hand, augmented hemostasis conceivably conferred a selective advantage by reducing mortality from postpartum hemorrhage, menorrhagia associated with severe iron deficiency anemia, and posttraumatic bleeding. Indeed, the amount of blood lost during labor is significantly smaller in heterozygotes for factor V Leiden than in women not carrying the mutation,[16] and profuse menstrual bleeding is significantly less common in factor V heterozygotes.[17]

Another proposed selective advantage for factor V Leiden is protection against infection. In a study of the effect of recombinant activated protein C infusion or placebo in cases with severe sepsis,

it was shown that the 28-day all-cause mortality was significantly reduced in heterozygous carriers of factor V Leiden compared with noncarriers in both treatment and control arms.[18] Moreover, mortality following endotoxin administration in mice heterozygous for factor V Leiden was significantly lower than in wild-type animals or homozygotes for factor V Leiden. However, in a later single-arm study of recombinant activated protein C infusion in patients with severe sepsis, there was no significant difference in mortality in carriers and noncarriers of factor V Leiden.[19] Also, the Copenhagen city population-based study failed to discern a persistent significant difference in the vulnerability of carriers and noncarriers toward infections.[20] Thus, whether factor V Leiden confers protection against infections is controversial.

Acknowledgment

We are extremely grateful to Dr Jeff Reeve from the Department of Medical Genetics, University of Alberta, Edmonton, for guiding us in using the DMLE+2.0 program.

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;88:3698-3703.
2. Emmerich J, Rosendaal FR, Cattaneo M, et al. Combined effect of factor V Leiden and prothrombin 20210A on the risk of venous thromboembolism: pooled analysis of 8 case-control studies including 2310 cases and 3204 controls: Study Group for Pooled-Analysis in Venous Thromboembolism. Thromb Haemost. 2001;86:809-816.
3. Gehring NH, Frede U, Neu-Yilik G, et al. Increased efficiency of mRNA 3′ end formation: a new genetic mechanism contributing to hereditary thrombophilia. Nat Genet. 2001;28:389-392.
4. Ceelie H, Spaargaren-van Riel CC, Bertina RM, Vos HL. G20210A is a functional mutation in the prothrombin gene: effect on protein levels and 3′-end formation. J Thromb Haemost. 2004;2:119-127.
5. Rees DC, Chapman NH, Webster MT, Guerreiro JF, Rochette J, Clegg JB. Born to clot: the European burden. Br J Haematol. 1999;105:564-566.
6. Zivelin A, Rosenberg N, Faier S, et al. A single genetic origin for the common prothrombotic G20210A polymorphism in the prothrombin gene. Blood. 1998;92:1119-1124.
7. National Center for Biotechnology Information. NCBI SNP database. http://www.ncbi.nlm.nih.gov. Accessed April 14, 2006.
8. Rannala B, Reeve JP. High-resolution multipoint linkage-disequilibrium mapping in the context of a human genome sequence. Am J Hum Genet. 2001;69:159-178.
9. Reeve JP, Rannala B. DMLE+: Bayesian linkage disequilibrium gene mapping. Bioinformatics. 2002;18:894-895.
10. Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. Mol Biol Evol. 1999;16:1791-1798.
11. Population Reference Bureau. 2005 World Population Data Sheet of the Population Reference Bureau. http://www.prb.org/pdf05/05WorldDataSheet_Eng.pdf. Accessed April 14, 2006.
12. Rosendaal FR, Doggen CJ, Zivelin A, et al. Geographic distribution of the 20210 G to A prothrombin variant. Thromb Haemost. 1998;79:706-708.
13. Zivelin A, Griffin JH, Xu X, et al. A single genetic origin for a common Caucasian risk factor for venous thrombosis. Blood. 1997;89:397-402.
14. Seligsohn U, Zivelin A. Thrombophilia as a multigenic disorder. Thromb Haemost. 1997;78:297-301.
15. Irani-Hakime N, Tamim H, Kreidy R, Almawi WY. The prevalence of factor V R506Q mutation-Leiden among apparently healthy Lebanese. Am J Hematol. 2000;65:45-49.
16. Lindqvist PG, Svensson PJ, Dahlback B, Marsal K. Factor V Q506 mutation (activated protein C resistance) associated with reduced intrapartum blood loss: a possible evolutionary selection mechanism. Thromb Haemost. 1998;79:69-73.
17. Lindqvist PG, Zoller B, Dahlback B. Improved hemoglobin status and reduced menstrual blood loss among female carriers of factor V Leiden: an evolutionary advantage? Thromb Haemost. 2001;86:1122-1123.
18. Kerlin BA, Yan SB, Isermann BH, et al. Survival advantage associated with heterozygous factor V Leiden mutation in patients with severe sepsis and in mouse endotoxemia. Blood. 2003;102:3085-3092.
19. Yan SB, Nelson DR. Effect of factor V Leiden polymorphism in severe sepsis and on treatment with recombinant human activated protein C. Crit Care Med. 2004;32:S239-S246.
20. Benfield TL, Dahl M, Nordestgaard BG, Tybjaerg-Hansen A. Influence of the factor V leiden mutation on infectious disease susceptibility and outcome: a population-based study. J Infect Dis. 2005;192:1851-1857.