Gene 659 (2018) 93–99

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Genome-wide association study in ethnic Russians suggests an association of the MHC class III genomic region with the risk of primary varicose veins

Alexandra Shadrina[a,b,*,1], Yakov Tsepilov[c,b,1], Ekaterina Sokolova[a,b], Mariya Smetanina[a,b], Elena Voronina[a,b], Eugene Pakhomov[b], Kseniya Sevost'ianova[a], Andrey Shevela[a], Evgeny Ilyukhin[d], Evgeny Seliverstov[e], Igor Zolotukhin[a,e], Maxim Filipenko[a,b]

[a] Institute of Chemical Biology and Fundamental Medicine, 8 Lavrentjev Avenue, Novosibirsk 630090, Russia
[b] Novosibirsk State University, 2 Pirogova street, Novosibirsk 630090, Russia
[c] Institute of Cytology and Genetics, 10 Lavrentjev Avenue, Novosibirsk 630090, Russia
[d] Private Surgery Center "Medalp", 54 Leningradskaya street, Saint Petersburg 197758, Russia
[e] Pirogov Russian National Research Medical University, 1 Ostrovitianova street, Moscow 117997, Russia

ARTICLE INFO

Keywords:
Varicose veins
Genetics
Association
Russians

ABSTRACT

Heredity is a well-known risk factor for varicose veins, but genetic basis of this condition remains poorly studied. Our aim was to conduct a large-scale genetic association study for primary varicose veins (PVVs) in the population of ethnic Russians. An initial scan using Illumina HumanExome-12 v1.0 BeadChip was performed for 273 patients with PVVs and 250 controls without a history of chronic venous disease and other venous disorders. After quality control and removal of monomorphic markers, 25,424 common and 48,232 rare variants were included in the analysis. 42 single nucleotide polymorphisms (SNPs) were genotyped in the independent replication cohort of 447 PVVs patients and 443 controls. Association of common variants with PVVs was investigated by logistic regression, and the impact of rare variants was analyzed using sequence kernel association test. No effect of low frequency alleles has been revealed in our study. Common variant analysis identified a promising signal at chromosome 6 within classical major histocompatibility complex (MHC) class III subregion. The most strongly associated SNP in a combined analysis that reached a suggestive significance level of $3.2e-05$ was polymorphism rs4151657 in the complement factor B gene. Testing for potential pleiotropy with other traits indicated that the same causal variant in this region increases the risk of rheumatoid arthritis and has a negative impact on human height. Our results provide suggestive evidence for the involvement of the MHC class III genes in the pathogenesis of PVVs. Further independent studies are needed to confirm our pilot findings.

1. Introduction

Varicose veins are the most common venous pathology of lower extremities affecting over 15% of adult males and 25% of adult females in developed countries, with incidence increasing with age (Beebe-Dimmer et al., 2005). This condition causes cosmetic problems, heaviness of the legs, pain, itching, and swelling, and can be accompanied by pigmentation, lipodermatosclerosis, and ulcers at advanced stages of disease progression. Varicose veins that are not preceded by a known identifiable condition such as episode of deep vein thrombosis are defined as primary varicose veins (PVVs). Molecular mechanisms underlying PVVs formation have still not been clearly elucidated. According to the current understanding, the development of PVVs is preceded by and associated with the pathophysiological remodeling of the venous wall (Lim and Davies, 2009; Pfisterer et al., 2014). Factors promoting varicose remodeling include changes in hemodynamic forces such as increase in wall stress and decrease in laminar shear stress. These alterations along with hypoxia are able to modulate the expression of various genes important for vein wall homeostasis maintenance, induce smooth muscle cells proliferation and phenotypic switching,

Abbreviations: CEAP, clinical severity, etiology, anatomy and pathophysiology classification system; CHARGE, cohorts for heart and aging research in genomic epidemiology consortium; GWAS, genome-wide association study; HEIDI, heterogeneity in dependent instruments method; LD, linkage disequilibrium; MAF, minor allele frequency; MHC, major histocompatibility complex; OR, odds ratio; PCR, polymerase chain reaction; PVVs, primary varicose veins; SMR, summary data-based Mendelian randomization analysis; SNPs, single nucleotide polymorphisms

* Corresponding author at: Laboratory of Pharmacogenomics, Institute of Chemical Biology and Fundamental Medicine, 8 Lavrentjev Avenue, Novosibirsk 630090, Russia.
E-mail address: weiner.alexserg@gmail.com (A. Shadrina).
[1] Authors have contributed equally to this study.

https://doi.org/10.1016/j.gene.2018.03.039
Received 7 January 2018; Received in revised form 16 February 2018; Accepted 14 March 2018
Available online 15 March 2018

activate endothelial cells, stimulate pro-inflammatory response, and change extracellular matrix composition (Lim et al., 2011; Pfisterer et al., 2014). Consequent vein wall weakening and altered tone can provoke further stasis and blood flow disturbance, leading to disease progression. In most cases, PVVs initiation does not seem to be caused by a single factor triggering this vicious cycle. It is postulated now that PVVs is a multifactorial disorder, and a combination of multiple lifestyle, environmental, physiological and genetic factors contribute to its development (Lim and Davies, 2009; Pfisterer et al., 2014). The important role of genetics in the etiology of PVVs is supported by a strong body of evidence (Grant et al., 2017), and the narrow-sense heritability was estimated to be nearly 17% (Fiebig et al., 2010). However, current knowledge of the genetic background of this condition is scarce, and until recently only few genetic association studies have been published investigating the effect of single nucleotide polymorphisms (SNPs) on PVVs risk. Genome-wide association study (GWAS) is a powerful tool for exploring genes implicated in disease susceptibility. The first GWAS for varicose veins was conducted by "23andMe" biotechnology company (Mountain View, California, USA) and presented at the 64th Annual Meeting of The American Society of Human Genetics (Bell et al., 2014). This study was performed on European individuals from the customer base of "23andMe" and revealed 12 SNPs associated with varicose veins at a genome-wide significance level, with the strongest signal being shown for the *ABO* gene. The second GWAS was conducted by Ellinghaus et al. and comprised German individuals (Ellinghaus et al., 2017). Two robust associations were identified for SNPs within the *EFEMP1* and *KCNH8* genes, and one suggestive association was shown for polymorphism within the *SKAP2* gene. Notably, no overlapping association signals have been observed in these two studies. Differences in local genomic structure between populations of diverse ethnic origin as well as influence of non-genetic factors modulating the effect of SNPs could underlie discrepancies in the results obtained by different research groups. In our study, we aimed to perform a GWAS on the sample of ethnic Russian individuals in order to identify novel genetic loci affecting the risk of PVVs and investigate the contribution of both common and rare genetic variants to the disease development. Rare variants are hypothesized to be an important component of complex traits genetics and a potential source of unexplained heritability, although their role often remains unexplored (Auer and Lettre, 2015). Studying the impact of rare genetic variation would provide deeper insights into the understanding of PVVs genetics.

2. Materials and methods

2.1. Ethics statement

The study was approved by the Ethics Committee of Institute of Chemical Biology and Fundamental Medicine (protocol No. 15, 13 September 2013) and the Ethics Committee of Pirogov Russian National Research Medical University (protocol No.123, 21 January 2013). All the individuals enrolled in this study gave signed informed consent. All clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki.

2.2. Study sample

Overall, 720 patients with primary varicose veins of lower extremities and 693 control individuals were recruited for the study. 273 cases and 250 controls were used as a discovery set, and after stringent quality control (see below) 210 case patients and 240 control individuals were left. Initial selection of samples for discovery set was performed basing on the quality and quantity of isolated DNA. Samples with the best characteristics were chosen. For a replication stage, 447 PVVs patients and 443 controls were used. A descriptive characteristic of the studied groups is presented in Table 1. All cases and controls identified themselves and their parents as ethnic Russians. Blood samples were collected in Novosibirsk State Regional Clinical Hospital (Novosibirsk, Russia), Center of New Medical Technologies (Novosibirsk, Russia), Pirogov City Clinical Hospital No. 1 (Moscow, Russia), and Private Surgery Center "Medalp" (Saint Petersburg, Russia).

All the case patients had visible varicose veins and reflux in the great saphenous vein or the small saphenous vein on duplex ultrasound examination and had objectively confirmed indications to varicose veins surgery. Duplex ultrasound was performed with the patient in standing position. The reflux in superficial venous system was evaluated using the distal compression manoeuvre, and the reflux duration of > 0.5 s was considered pathologic. The CEAP (clinical severity, etiology, anatomy and pathophysiology) classification system was used to describe the patients (Eklöf et al., 2004). Patients with reticular veins and/or telangiectasis, venous malformations, postthrombotic syndrome as well as individuals with a history of prior deep venous thrombosis were excluded from the study.

Control subjects were recruited at the same clinics as PVVs patients, had no visible signs of varicose veins and declared no history of venous disorders.

Genomic DNA was isolated from leucocytes in venous blood by proteinase K digestion followed by phenol/chloroform extraction and ethanol precipitation.

2.3. Genotyping and quality control

2.3.1. Discovery stage

Genotyping was performed using HumanExome-12 v1.0 BeadChip (Illumina, San Diego, California, USA) querying 247,869 common and rare genetic markers. Samples were processed with Illumina GenomeStudio Software 2011.1 using a cluster file provided by Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium (Grove et al., 2013). 8994 variants were excluded from the study according to CHARGE recommendations. Data were then converted into PLINK format (Purcell et al., 2007). A further quality control procedure was performed using the GenABEL statistical package for the R language (Aulchenko et al., 2007). In order to control genotyping quality, 6 samples were analyzed in duplicates (one sample on the same chip and 5 samples on different chips), and 1 sample was analyzed in 5 replicates on different chips. Markers different for replicates were excluded from the analysis. Replication error rate was 0.0056. Next, we removed SNPs with a call rate $< 95\%$ and deviation from Hardy-Weinberg equilibrium at P-value $< 1.0e-06$. Filtering of study participants was performed according to the following criteria: genotyping call rate $< 95\%$ and deviation from the autosomal heterozygosity at FDR $< 1\%$. Additionally, we estimated kinship matrix in order to assess cryptic relatedness of study subjects, and performed multidimensional scaling for detection of possible outliers. After applying the above quality control parameters and removal of related individuals, 450 out of 523 study samples and 238,312 out of 247,869 markers remained. Finally, we removed all the monomorphic variants, that yielded 73,656 SNPs for analysis, from which 25,424 SNPs had minor allele frequency (MAF) $\geq 5\%$ and 48,232 were rare variants with MAF $< 5\%$.

Population structure was evaluated using autosomal markers with MAF $> 5\%$. Genomic control inflation factor λ was evaluated to be 1.046 that corresponds to a low degree of population stratification. A quantile-quantile plot for observed vs. expected distribution of P-values is presented in Supplementary Fig. 1.

2.3.2. Replication stage

Genotyping of the top common SNPs most significantly associated with PVVs in our discovery study sample was carried out by real-time polymerase chain reaction (PCR) allelic discrimination with TaqMan probes. Primers and probes were designed using sequences obtained from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). Oligonucleotides' structures as well as PCR cycling