

SCIENTIFIC REPORTS



OPEN

Genome-wide association analysis for chronic venous disease identifies *EFEMP1* and *KCNH8* as susceptibility loci

Received: 08 November 2016

Accepted: 28 February 2017

Published: 04 April 2017

Eva Ellinghaus^{1,*}, David Ellinghaus^{1,*}, Petra Krusche², Aljoscha Greiner², Claudia Schreiber³, Susanna Nikolaus³, Christian Gieger⁴, Konstantin Strauch^{4,5}, Wolfgang Lieb⁶, Philip Rosenstiel¹, Norbert Frings², Andreas Fiebig¹, Stefan Schreiber^{1,3} & Andre Franke¹

Chronic venous disease (CVD) is a multifactorial condition representing one of the most common disorders among populations of Western countries. The heritability of about 17% suggests genetic risk factors in CVD etiology. However, so far the genetic causes are unknown. We undertook the hitherto first genome-wide association study (GWAS) for CVD, analyzing more than 1.93 M SNPs in 4,942 German individuals, followed by replication in two independent German data sets. The combined analysis of discovery and replication stages (2,269 cases and 7,765 controls) yielded robust associations within the two genes *EFEMP1* and *KCNH8* (rs17278665, rs727139 with $P < 5 \times 10^{-8}$), and suggestive association within gene *SKAP2* (rs2030136 with $P < 5 \times 10^{-7}$). Association signals of rs17278665 and rs727139 reside in regions of low linkage disequilibrium containing no other genes. Data from the ENCODE and Roadmap Epigenomics projects show that tissue specific marks overlap with the variants. SNPs rs17278665 and rs2030136 are known eQTLs. Our study demonstrates that GWAS are a valuable tool to study the genetic component of CVD. With our approach, we identified two novel genome-wide significant susceptibility loci for this common disease. Particularly, the extracellular matrix glycoprotein *EFEMP1* is promising for future functional studies due to its antagonistic role in vessel development and angiogenesis.

CVD is a complex venous pathology and defined as “morphological and functional abnormalities of the venous system of long duration manifested either by symptoms and/or signs indicating the need for investigation and/or care”¹. Prevalence of CVD in Germany and the United States was estimated at 31%² and 29%², respectively. The term CVD encompasses patients with primary varicose veins (PVV) and chronic venous insufficiency (CVI). According to the CEAP guidelines (Comprehensive Classification System for Chronic Venous Disorders), absence of visible or palpable signs of venous disease is classified as C0 whereas the presence of telangiectasia or reticular veins leads to C1 classification. Varicose veins are distinguished from reticular veins by a diameter of at least 3 mm so that PVV patients who have dilated subcutaneous veins (≥ 3 mm \emptyset) but no signs of edema or skin changes are classified as C2. Patients with CVI can be classified as either C3 (edema), C4 (skin changes without ulceration), C5 (healed venous ulcer) or C6 (active venous ulcer)^{3,4}. Exemplary photos for C1 to C6 are shown in Fig. 1. CVD has a major socioeconomic impact. The financial burden of CVD treatment on the health-care system is immense with an estimated cost of US\$3 billion per year in the USA and up to 2% of the total health-care budget of all Western countries^{5,6}. Despite the frequency of CVD and its considerably high impact on health-care budgets and on patients’ quality of life the underlying etiology and pathophysiology are still poorly understood.

¹Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, 24105 Kiel, Germany. ²Capio Mosel-Eifel-Clinic, 56864 Bad Bertrich, Germany. ³Department of General Medicine, University Hospital Schleswig-Holstein, Campus Kiel, 24105 Kiel, Germany. ⁴Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, 85764 Neuherberg, Germany. ⁵Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, 80539 Munich, Germany. ⁶PopGen Biobank, Institute of Epidemiology, University Hospital Schleswig-Holstein, 24105 Kiel, Germany. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to D.E. (email: d.ellinghaus@ikmb.uni-kiel.de)

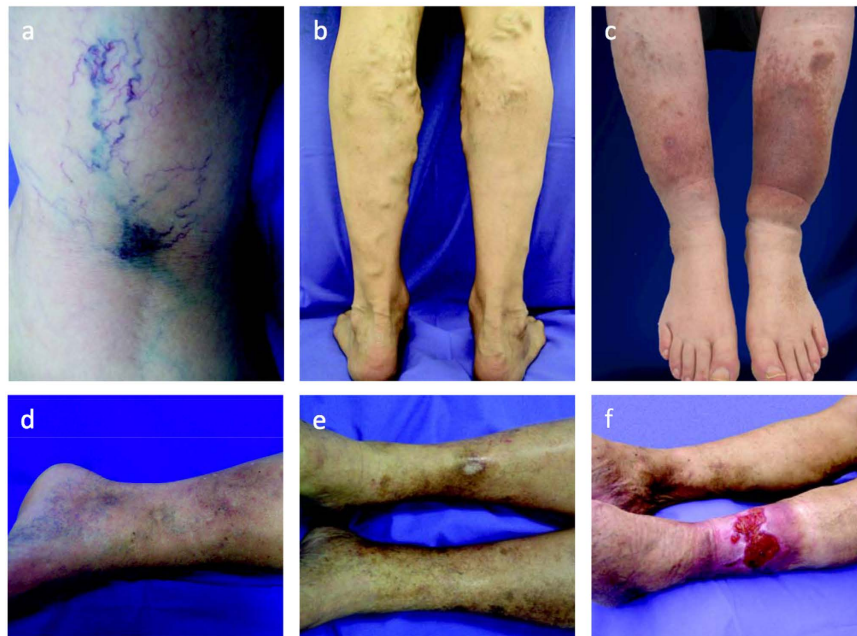


Figure 1. Exemplary photographs for a spectrum of clinical manifestations of chronic venous disease ranging from C1 to C6. (a) Spider veins, reticular veins; C1. (b) Varicose veins affecting the short saphenous vein; C2. (c) Edema due to varicose veins; C3. (d) Changes in skin and subcutaneous tissue, dermatosclerosis; C4. (e) Healed venous ulcer and Atrophie blanche; C5. (f) Active venous ulcer; C6. Photographs by courtesy of Dr. P. Krusche, Capio Mosel-Eifel-Clinic, Bad Bertrich.

Many factors may increase risk for CVD, including age, female gender, pregnancy, hormonal changes, obesity and standing occupation and predispose individuals to the dilatation, elongation and tortuosity of the saphenous vein and its tributaries⁷. In addition, a genetic component has been proposed for many years, supported by reports on familial clustering and twin studies^{8–10}. The narrow-sense heritability (proportion of the phenotypic variance that is explained by additive genetic variance) of PVV and CVD has recently been estimated to equal 18.5% and 17.3%, respectively, in a large sample of affected nuclear families from Germany⁴. These heritability estimates suggest the existence of susceptibility genes, necessitating large systematic association studies. So far, no disease association has been confirmed. To further our understanding of the genetic etiology of CVD, we undertook the hitherto first genome-wide association study (GWAS) for this multifactorial condition comprising 10,034 individuals in total.

Results

To identify common variants that confer susceptibility to CVD, we performed genome-wide SNP genotyping of 323 unrelated German C2, C3 and C4 CVD cases and 4,619 healthy German control individuals employing Affymetrix SNP arrays (panel A, Supplementary Table 1). Principal component analysis supported European ancestry for all cases and controls (Methods, Supplementary Figure 1). To facilitate the analysis of the four data sets genotyped on different SNP chips and to increase the genomic coverage of our study by predicting untyped markers, SNP genotype imputation was performed. For imputation and subsequent statistical analysis, we used data sets that passed stringent quality control filters (Methods). In total, 1,934,349 quality-controlled autosomal-imputed SNP markers were available for association analysis (Supplementary Figure 2). Assuming a frequency of the disease-associated allele of at least 20% in controls, our screening panel had 79% power to detect a variant with an odds ratio of 1.3 or higher at the 0.001 significance level, which was the value used to define regions for attempting replication in a larger sample set. For stage one replication, we selected 70 SNPs based on their *P*-value rankings in the discovery stage and support by other SNPs in linkage disequilibrium (Methods, Supplementary Table 2). The replication sample of stage one consisted of 1,258 independent CVD cases and 1,925 controls from Germany (panel B, Supplementary Table 1). For stage two replication analysis, we selected three SNPs that replicated at a significance level of $P < 0.05/70 = 7.14 \times 10^{-4}$ (based on Bonferroni correction) in replication stage one for TaqMan[®] genotyping (Supplementary Table 3) in an additional German panel comprising 688 CVD cases and 1,221 control individuals (panel C, Supplementary Table 1). We used a significance level of 7.14×10^{-4} for the statistical association analysis in the replication panels (panel B and C; Supplementary Table 1) based on Bonferroni correction (Methods). All three SNPs were significant in a meta-analysis of the combined replication data (panel B and C) after adjusting for multiple testing ($P_{\text{Repl1}\&2} = 2.59 \times 10^{-5}$ for rs17278665; $P_{\text{Repl1}\&2} = 3.39 \times 10^{-8}$ for rs727139; $P_{\text{Repl1}\&2} = 6.84 \times 10^{-5}$ for rs2030136; Table 1). In a combined meta-analysis of the discovery and replication samples (2,269 CVD cases and 7,765 controls) SNPs rs17278665 ($P_{\text{combined}} = 1.74 \times 10^{-8}$) and rs727139 ($P_{\text{combined}} = 5.42 \times 10^{-11}$) reached genome-wide significance ($P < 5 \times 10^{-8}$ Table 1). To assess the robustness of our results, the association analyses were also carried

Chr	Association boundaries (kb)	dbSNP id ^a	A1	A2	AF _{ca}	AF _{co}	Key gene (N additional in locus)	Discover GWAS (323/4,619)		Replication stage 1&2 (1,946/3,146)		Discovery&Replication (2,269/7,765)		Associations ($P < 5 \times 10^{-8}$) to other traits
								P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	
2p16.1	55979–56163	rs17278665	G	C	0.17	0.15	EFEMP1 (0)	5.18×10^{-5}	1.59 (1.28–1.98)	2.59×10^{-5}	1.26 (1.13–1.41)	1.74×10^{-8}	1.32 (1.19–1.45)	H, O, FVC
3p24.3	19121–19548	rs727139	G	A	0.16	0.21	KCNH8 (0)	1.69×10^{-4}	0.63 (0.50–0.80)	3.39×10^{-8}	0.75 (0.68–0.83)	5.42×10^{-11}	0.73 (0.66–0.80)	—
7p15.2	26657–27190	rs2030136	C	T	0.26	0.24	SKAP2 (9)	3.93×10^{-4}	1.48 (1.20–1.83)	6.84×10^{-5}	1.21 (1.10–1.33)	3.39×10^{-7}	1.25 (1.15–1.36)	T1D*, CD*

Table 1. Loci with strongest evidence of association with CVD in combined sample. We used a significance level of $0.05/70 = 7.14 \times 10^{-4}$ for the statistical association analysis in replication panels (panel B and C; Supplementary Table 1) based on Bonferroni correction. All three SNPs replicated. Loci 2p16.1 and 3p24.3 are genome-wide significant risk loci ($P < 5 \times 10^{-8}$) for CVD. **Chr:** chromosome; **Association boundaries (kb):** association boundaries for each index SNP in kb (see Methods) with genomic positions retrieved from NCBI's dbSNP build v138 (genome build hg19); **dbSNP id:** rs ID of index SNP; **A1:** minor allele; **A2:** major allele; **AF_{ca}/AF_{co}:** allele frequency of A1 estimated from replication stage1&2 (panel B-C) in cases and controls, respectively; **Key gene:** candidate gene in the region; for 7p15.2, 10 genes are covered by association boundaries: *C7orf71*, *HOTAIRM1*, *HOXA1*, *HOXA2*, *HOXA3*, *HOXA4*, *HOXA5*, *HOXA6*, *HOXA-AS3* and *SKAP2*. **P** and **OR (95% CI):** P-value and corresponding odds ratio and 95% confidence interval with respect to minor allele. For each panel, numbers of CVD cases/controls are displayed in parentheses. **Associations ($P < 5 \times 10^{-8}$) to other traits:** Overlaps with other disease phenotypes (listed if anywhere within association boundaries, * = known genome-wide significant risk SNP in perfect LD ($r^2 = |D'| = 1$) with CVD hit SNP, see also Methods); CD = Crohn's disease (rs10486483), FVC = forced vital capacity (rs1430193), H = Height (rs1367226, rs3791679, rs3791675), O = optic disc morphology (rs1346786), T1D = Type 1 diabetes (rs7804356). ^aSNPs attaining genome-wide significance and their notable nearby genes are indicated in bold.

out with adjustments made for sex, age at ascertainment and body mass index (BMI). No notably different results were obtained in the sex-, age- and BMI-adjusted analyses (Supplementary Table 4).

When checking for potential coding SNPs being highly correlated ($r^2 > 0.8$ in 1000 Genomes European samples) with the three lead SNPs (Methods), we identified only one further intronic SNP at the *EFEMP1* and *KCNH8* loci, respectively. At the *SKAP2* locus 163 intronic variants, one 5' UTR variant, five variants upstream and one variant downstream of *SKAP2* correlate with the lead SNP. However, no coding variants in high LD were detected. 31 of these non-coding LD SNPs, all located at the *SKAP2* locus on chromosome 7, were predicted to be deleterious by the fathmm¹¹ algorithm (with values above 0.5), and 23 variants with a more stringent cutoff value above 0.8, respectively (Supplementary Table 5). The CADD¹² method assigned scaled C scores of 10 or greater for 27 variants at the *SKAP2* locus, indicating that these variants are among the highest 10% of all scores, whereas five variants were assigned a scaled C score above 15, a suggested cutoff value on the likelihood of deleteriousness to identify potentially pathogenic variants (Supplementary Table 5). Of these five variants, four variants (rs10266759, rs3801828, rs3801808, rs3801814) had a fathmm score greater 0.8, strongly confirming their potential deleterious effect (Supplementary Table 5).

The intronic SNP rs17278665 at 2p16.1 maps to the last intron of the gene *EFEMP1* (epidermal growth factor-containing, fibulin-like extracellular matrix protein 1; also known as *FBNL*), which encodes the extracellular matrix protein fibulin-3. The association signal resides in a region of low linkage disequilibrium containing no other genes (Fig. 2a), pointing out the possibility of a direct involvement of *EFEMP1* in disease development. The odds ratio for carriership of the minor risk allele G of rs17278665, as estimated from replication samples, was 1.26 (95% CI = 1.13–1.41). The variant is a known local eQTL for *EFEMP1* ($P_{eQTL} = 3.45 \times 10^{-7}$ in lung tissue and $P_{eQTL} = 2.25 \times 10^{-6}$ in skin)^{13,14} (Methods, Supplementary Table 6) with the minor allele being associated with decreased expression¹⁴ (Supplementary Figure 3a). Although classified as benign, the Mutation Taster¹⁵ prediction tool reports the minor allele to increase a donor splice site and as a consequence, the potential loss of the protein region that mediates interaction with TIMP3, located partially downstream of the altered splice site (Supplementary Table 5). *EFEMP1* is highly expressed in endothelial cells like Telomerase-immortalized microvascular endothelial cells (RNA: 1866 FPKM, protein: 8931)¹⁶, human umbilical vein endothelial cells (HUVECs; 7680 RPKM), human dermal blood endothelial cells (5896 RPKM) and human brain microvascular endothelial cells (4650 RPKM)¹⁷ (Supplementary Table 7). Data from the ENCODE¹⁷ and Roadmap Epigenomics¹⁸ projects imply that in HUVECs and mesenchymal tissue the variant overlaps with an HMM-predicted enhancer that is classified as a genic enhancer (15-state core model) and transcribed 3' enhancer (25-state model), respectively (Supplementary Table 8).

The other genome-wide significantly associated SNP rs727139 at 3p24.3 ($OR_{Rep1\&2} = 0.75$, 95% CI = 0.68–0.83) is located within the third intron of the *KCNH8* gene without any other genes contained between the association boundaries (Fig. 2b). *KCNH8* encodes a potassium voltage-gated channel. The Mutation Taster¹⁵ tool reports the minor allele of the variant to cause the generation of a new donor splice site, and the potential loss of several protein domains downstream of the altered splice site (Supplementary Table 5). According to the ENCODE¹⁷ data, *KCNH8* is almost exclusively expressed in CD20⁺ B cells (13.1 RPKM) and embryonic stem cells (6.6 RPKM)

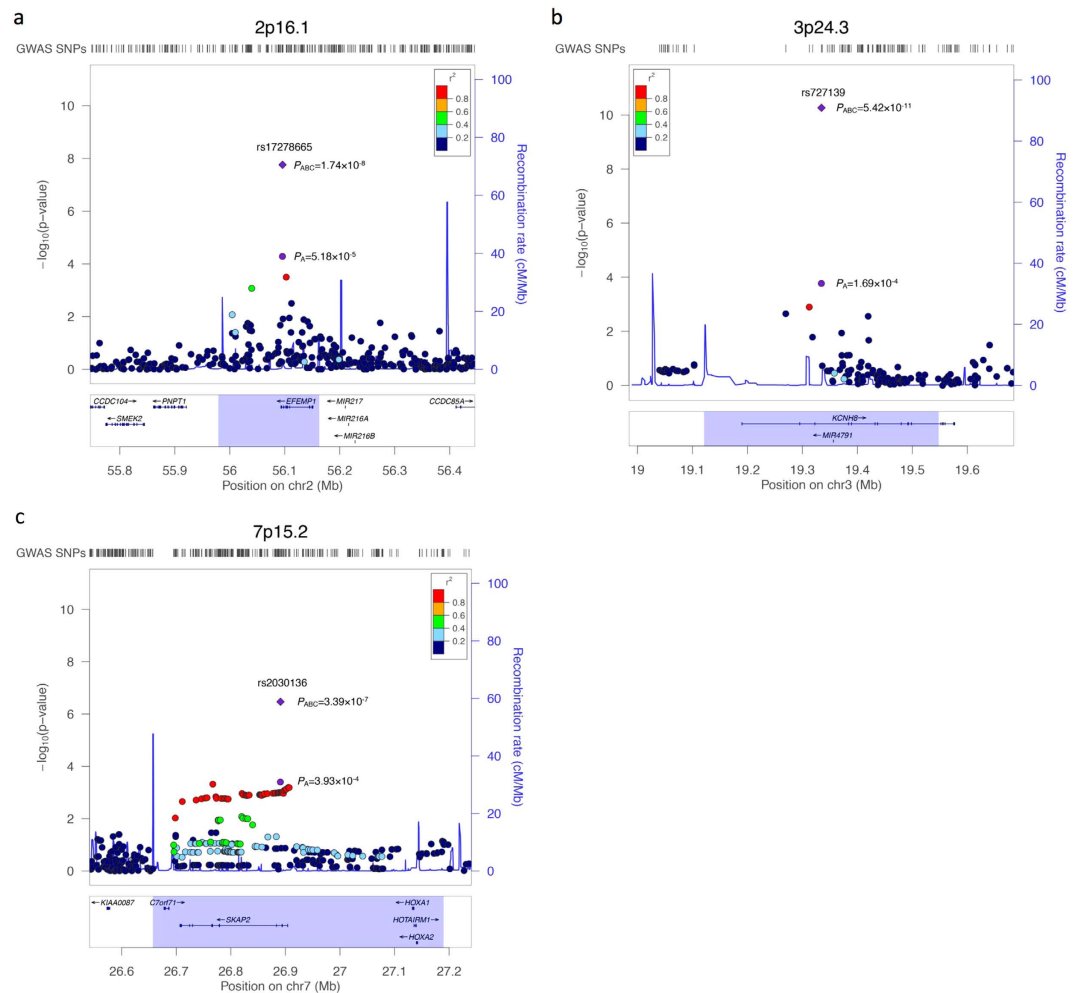


Figure 2. Regional association plots of loci with strongest evidence of association with CVD in combined sample. Shown are the negative \log_{10} (genomic control adjusted) P -values from the discovery analysis of 323 unrelated German CVD cases and 4,619 healthy control individuals (panel A, Supplementary Table 1) with regard to the physical location of markers. The intronic SNPs (a) rs17278665, (b) rs727139, and (c) rs2030136 were genotyped in the two replication panels and therefore the combined P -value of panels A, B and C (P_{ABC}) is indicated for these three lead SNPs (purple diamond) in addition to the P -value of panel A (P_A), whereas for the surrounding SNPs in each plot only the respective P -values of panel A are available. Loci 2p16.1 and 3p24.3 are genome-wide significant risk loci ($P < 5 \times 10^{-8}$) in the combined analysis of discovery and replication stages (2,269 CVD cases and 7,765 controls). **Purple circle:** index SNP with P -value of panel A; **Purple diamond:** index SNP with combined P -value of panel A, B and C; **filled circles:** analyzed SNPs with the fill color corresponding to the strength of linkage disequilibrium (r^2) with the lead SNP (for color coding see legend in the upper right corner of each plot); **blue line:** recombination intensity (cM/Mb); **highlighted region:** association boundaries for each index SNP (see Methods and Table 1). Positions and gene annotations are according to NCBI's build 37 (hg19).

(Supplementary Table 7). In the Roadmap Epigenomics¹⁸ data, there is a cluster of promoter and enhancer activity in embryonic stem cells and induced pluripotent stem cells (iPSCs) overlapping with the variant. The 15-state core model classified it as enhancer and flanking active transcription start site whereas the 25-state core model classified it as active enhancer. In addition, a DNase hypersensitivity data peak overlaps with rs727139 in iPSCs (Supplementary Table 8).

SNP rs2030136 is located within the third intron of the *SKAP2* gene (src kinase associated phosphoprotein 2; 7p15.2) and showed highly significant association in the combined analysis ($P_{combined} = 3.39 \times 10^{-7}$) (Fig. 2c). The gene product of *SKAP2* is potentially involved in leukocyte adhesion processes and couples the B-cell receptor in B-cells to integrin activation¹⁹. According to the 1000 Genomes Project EUR reference²⁰, SNP rs2030136 is in perfect LD ($r^2 = 1.0$) with SNPs rs7804356 and rs10486483 that were previously reported to be associated with type I diabetes²¹ and Crohn's disease²², respectively. SNP rs2030136 is a known local eQTL of *SKAP2* ($P_{eQTL} = 4.99 \times 10^{-141}$)²³ and *HOXA5* ($P_{eQTL} = 4.69 \times 10^{-7}$)¹⁴ (Supplementary Table 6), a gene that blocks angiogenesis, with the minor allele being associated with increased expression (Supplementary Figure 3b). In addition, this SNP is in high LD ($r^2 > 0.8$) with 168 other known eQTL SNPs for *SKAP2* and *HOXA5*

(Supplementary Table 6). Although *SKAP2* is expressed ubiquitously in human tissues²⁴, it is expressed stronger in human renal cortical epithelial cells (31 RPKM), lymphoblastoid cells (18 RPKM) and human dermal blood endothelial cells (17 RPKM) according to the ENCODE data (Supplementary Table 7).

Discussion

We performed to our knowledge the first genome-wide association study for CVD and identified three novel CVD susceptibility loci. The analysis of our combined sample of 2,296 CVD cases and 7,765 controls implicates genetic variants in the genes *EFEMP1*, *KCNH8* and *SKAP2* in CVD susceptibility for the first time. The gene product of *EFEMP1*, fibulin-3, is an extracellular matrix glycoprotein, and belongs to the six-member fibulin family. The protein family is phylogenetically conserved and shows a high conservation of orthologues among different species. A single missense mutation in the *EFEMP1* gene causes the autosomal dominant maculopathy Malattia Leventinese/Doyle honeycomb retinal dystrophy in all patients and families so far reported²⁵. *Efemp1*^{-/-} mice showed reduced fertility, premature aging, had a reduced life span and other phenotypes. Histologic analysis of these mice revealed a marked reduction of elastic fibers in fascia throughout the body²⁶ whereas mice that lack fibulin-1 die at birth from defects associated with endothelial cell abnormalities²⁷. Fibulin-3 is most homologous to fibulin-4 and -5, but also shows a strong homology to members of the fibrillin family²⁸. Mutations in fibulin-5 have been identified in patients with cutis laxa^{29,30} and mutations in fibrillin I cause Marfan syndrome, a disorder of connective tissue with manifestations involving the skeletal, ocular and cardiovascular systems³¹. There is evidence that fibulin-3 and -5 antagonize vessel development *in vivo* and fibulin-3 and -5 both repress the expression of the matrix metalloproteinases MMP-2 and MMP-3 while they induce the expression of the tissue inhibitors of metalloproteinases TIMP-1, TIMP-3 and TSP-1 in endothelial cells³². In varicose veins the expression of MMP1, -2, -3, -7, -9 and -13, and TIMP-1 and -3 is increased^{33–35}. Both groups of enzymes (MMPs and TIMPs) regulate homeostasis of the extracellular matrix with MMPs being involved in degradation of the extracellular matrix and TIMPs influencing vascular remodeling^{34,36}. An imbalance between these proteins may lead to changes in the vessel wall and, eventually, to CVD. Because *EFEMP1* is expressed in disease-relevant tissue (Supplementary Table 7), we hypothesize that genetic variation in *EFEMP1* could influence the expression of certain MMPs and TIMPs which might result in defects in cellular and extracellular matrix components, causing an altered vein elasticity. However, fine-mapping studies are needed to identify the actual causal variant(s) and its/their functional consequence(s).

KCNH8 is a member of the human Elk K⁺ channel gene family. Voltage-gated potassium channels represent the most complex class of voltage-gated ion channels from both functional and structural standpoints. Their diverse known functions include regulating neurotransmitter release, heart rate, insulin secretion, neuronal excitability, epithelial electrolyte transport and smooth muscle contraction. Interestingly, several studies have demonstrated the involvement of ion channels in venous dilatation and varicose vein formation³⁷. For example, Raffetto *et al.* described MMP-2 induced venous dilatation via hyperpolarization and activation of K⁺ channels. Among several possible mechanisms by which MMP-2 could activate K⁺ channels, they suggested that as a result of collagen degradation by the protease activity of MMP-2, soluble Arg-Gly-Asp (RGD) tripeptide can bind to smooth muscle $\alpha_v\beta_3$ integrin receptors and lead to activation of K⁺ channels, venous hyperpolarization and relaxation³⁸. This is further supported by an observed partial inhibition of vasodilatation with inhibitors of voltage-gated, ATP-sensitive and inward rectifying K⁺ channels³⁹. Whether the herein identified marker rs727139 is tagging causative variant(s) at the *KCNH8* locus and what role the protein itself might play in disease etiology needs to be elucidated in the future.

The gene product of *SKAP2* is an adaptor protein that is thought to be involved in the Src signaling pathway and to regulate proper activation of the immune system. Among its related super-pathways is “cell junction organization”. Variants at the *SKAP2* gene locus have previously been implicated in type 1 diabetes⁴⁰ and in Crohn’s disease²². *SKAP2* is expressed broadly in lymphoid and myeloid cells and plays an adhesive/migratory role in B cells and dendritic cells. It interacts with the immune inhibitory receptor protein SIRP α (Signal Regulatory Protein- α) and drives the transduction of integrin-evoked signals which cause cytoskeletal reorganization required for macrophage migration⁴¹. *Skap2*- and *Sirp α* -deficient mice are resistant to experimental autoimmune encephalomyelitis and collagen-induced arthritis, two models for the human diseases multiple sclerosis and rheumatoid arthritis that have integrin-dependent components. These findings imply an important role for *SKAP2* in inflammation, particularly in integrin-dependent and macrophage-centric processes like atherosclerosis and inflammatory bowel disease⁴¹. A key aspect of CVD is inflammation, particularly caused by macrophages and monocytes, resulting from persistently increased venous pressure which is sensed by the endothelial cells and followed by recruitment of leukocytes, endothelial attachment, and finally, initiation of an inflammatory cascade⁴². Accordingly, *SKAP2* is an interesting candidate gene for functional studies in CVD development. The minor allele of SNP rs2030136 is strongly associated with increased expression of *SKAP2* and *HOXA5*. Furthermore, there are 168 other known eQTL SNPs for *SKAP2* and *HOXA5* in high LD. *HOXA5* is abundantly expressed in quiescent blood vessels, whereas its expression is diminished or absent in active angiogenic endothelial cells⁴³. The *HOXA5* gene blocks angiogenesis by up-regulating expression of anti-angiogenic genes like Thrombospondin-2 and down-regulating many pro-angiogenic genes such as vascular endothelial growth factor receptor 2 (*VEGFR2*) and hypoxia inducible factor-1 alpha (*HIF1 α*)⁴³. *VEGFR-2* mediates most of the known cellular responses to VEGF, which plays a central role in maintaining vascular integrity and reactivity. The proangiogenic cytokine VEGF and its upstream transcription factor HIF-1 α are important for new vessel formation such as vasa vasora^{44,45}.

Identification of genetic susceptibility factors in the etiology of chronic venous disease is fundamental for future translational and clinical research in order to understand disease mechanisms and to be able to translate knowledge of involved genes into better prevention and/or treatment of patients. The three risk variants newly identified here collectively account for 2.0% of the total variance in disease liability (heritability) for CVD, indicating that more powerful GWAS studies including larger sample sets are required not only to confirm the herein

reported loci but to detect potential additional causal variants. Fine-mapping and resequencing efforts together with extensive functional studies may help to specify the contribution of the respective genes to overall disease susceptibility and to study the mechanisms that are altered through the risk variants.

Methods

Study population. To narrow down the wide spectrum of clinical severity of CVD, we included only patients with primary varicose veins with and without signs of edema and/or skin changes (C2-C4) but without healed or active venous ulcer (C5/C6). The prevalence of the different CVD grades (C2-C4) in the overall study population is presented in Supplementary Table 9. Personal characteristics of the CVD patients and healthy control individuals can be found in Supplementary Tables 10 and 11, respectively. In all patients, the primary varicose veins affected the superficial great saphenous vein (vena saphena magna), frequently accompanied by additional varicosities of the short saphenous vein (vena saphena parva). Patients with any signs of acute deep vein thrombosis were excluded. The samples were organized in panels that corresponded to the successive steps of the present study. All individual panels (A-C in Supplementary Table 1) were independent from each other. All patients were from Germany, recruited at the Capio Mosel-Eifel-Clinic (Bad Bertrich, Germany). Ascertainment criteria for CVD are based on the CEAP classification system, which was developed in 1994 and subsequently revised³. Its key elements are clinical severity, etiology or cause, anatomy and pathophysiology.

4,374 German healthy control individuals of panel A-C were obtained from the PopGen biobank⁴⁶. 3,391 German healthy controls of panel A were selected from the KORA S3 and S4 survey, an independent population-based sample from the general population living in the region of Augsburg in southern Germany⁴⁷.

Genomic DNA was prepared from blood samples using the Invisorb Blood Giga Kit (Invitex, Berlin, Germany). All DNA samples were quality-checked on agarose gels.

Written, informed consent was obtained from all study participants and the Ethics Committee of the Medical Faculty of the CAU Kiel approved all protocols. All methods carried out were in accordance with the approved study protocols.

SNP genotyping for discovery sample and quality control. Genotyping of GWAS cases and PopGen controls - which was part of the German GWAS initiative funded by the National Genome Research Network (NGFN) - was performed as a service project by Affymetrix. All experimental steps were carried out according to standard protocols. SNPs that had > 5% missing data in cases or controls, a minor allele frequency < 1% or deviated from Hardy-Weinberg equilibrium (exact $P < 10^{-4}$ in controls) were excluded using the PLINK software version 1.07⁴⁸. We excluded individuals with more than 5% missing genotypes (i.e. call rate < 95%), individuals from each pair of unexpected duplicates or relatives, as well as outlier individuals with average marker heterozygosities of ± 5 s.d. away from the sample mean. Pair-wise percentage identity-by-state (IBS) values were computed for all individuals and all QCed markers using PLINK. 'Extreme outliers' were detected by comparing the distribution of the pair-wise IBS values for each individual with the combined IBS distribution of the entire study population. Two types of "outliers" were detected: 1) Individuals less related to the entire study population than expected were defined as those for whom > 60% of the IBS values were smaller than the median minus three times the interquartile range ($3 \times$ IQR) of the population distribution. In this case, the individual was removed from the population. 2) Individuals with a close cognate (relative) in the population were defined as those who had at least one observed IBS value above the median plus $3 \times$ IQR. In this case, the member of the cognate pair with the lower call rate was removed from the population. 'Extreme outliers' were calculated using R 2.15.2⁴⁹.

These quality control measures left genotype data of a set of 302,457 autosomal overlapping SNPs in 323 CVD samples and 4,619 control samples for inclusion in the discovery phase (panel A, Supplementary Table 1). These samples were tested for population stratification using the principal components stratification method, as implemented in EIGENSTRAT⁵⁰. Principal component analysis revealed no population outliers (Supplementary Figure 1).

Genome-wide genotype imputation and association analysis. As input for the imputation only genotyped SNPs that passed quality control were used. SNP genotype imputation was performed using the MACH version 1.10.16⁵¹ and 690 HapMap3 reference haplotypes from the CEU, TSI, MXL and GIH⁵² as well as 120 HapMap II CEU phased haplotypes cohorts to predict missing autosomal genotypes *in-silico*⁵³. Including samples from diverse populations such as Mexican (MXL) and Indian (GIH) can provide significant improvement in imputation for low-frequency variants (MAF < 5%)⁵⁴. For downstream analyses, we included only SNPs that were imputed with high confidence (estimated r^2 between imputed and true genotypes > 0.8) and had a MAF $\geq 1\%$. We further excluded markers deviating from HWE in controls after imputation ($P < 10^{-6}$). In total, 1,934,349 quality-controlled autosomal genotyped or imputed SNP markers were available for association analysis (Supplementary Figure 2). To take imputation uncertainty into account, phenotypic association was tested for allele dosage data using MACH2DAT's logistic regression framework for dosage data⁵¹. The allelic dosage is the weighted sum of the genotype class probabilities. To control potentially confounding effects due to population stratification, we adjusted for the top five eigenvectors (indicating statistical significance by Tracy-Widom statistics) from EIGENSTRAT in the regression analysis. In addition, genomic control correction was applied to results by correction of χ^2 test statistics using inflation factor $\lambda = 1.156$ and multiplying SNP standard errors by the square root of the inflation factor in order to minimize artificial differences in allele frequencies due to different genotyping platforms (see quantile-quantile plot in Supplementary Figure 4).

Replication genotyping. Since only a few variants showed a substantial deviation from the expected distribution of genome-wide P -values in the tail of the distribution, we genotyped in stage one the 70 most strongly

associated SNPs with $P < 10^{-3}$ from each associated locus by means of PLINK's clumping procedure (using settings $P_1 < 0.001$, $P_2 < 0.05$, $r^2 \geq 0.8$, $\text{kb} = 250$) in an independent German panel of 1,258 CVD cases and 1,925 control individuals (panel B; Supplementary Table 1), using the Sequenom® iPLEX platform and TaqMan® technology from Life Technologies. Adjustment for multiple testing was performed by Bonferroni correction for 70 SNPs (association results of 70 SNPs; Supplementary Table 2). In stage two, we selected 3 SNPs that replicated at $P < 0.05/70 = 7.14 \times 10^{-4}$ in stage one for TaqMan® genotyping in an additional German panel comprising 688 CVD cases and 1,221 control individuals (panel C; Supplementary Table 1, association results of 3 SNPs; Supplementary Table 3). Individual samples with $>5\%$ missing data were excluded from the downstream analyses. SNPs that had $>3\%$ missing data or deviated from Hardy-Weinberg equilibrium (exact $P < 10^{-4}$ in controls) were excluded. P -values for allele-based tests were calculated using PLINK. The fixed-effects meta-analysis function of PLINK was then used to obtain P -values for the combined replication data set (panel B-C) and the combined discovery-replication data set (panel A-C). We used the commonly accepted threshold of 5×10^{-8} for joint P -values to define genome-wide significance.

Annotation of association boundaries & regional association plots. Linkage disequilibrium regions (association boundaries) around lead SNPs from Table 1 were defined by extending in both directions a distance of 0.1 centimorgans (cM). For each locus, candidate genes within regions are listed in the column 'Key gene' and more details are listed in the legend of Table 1. Regional association plots of Fig. 2 were generated using LocusZoom⁵⁵.

Annotation of associations to other phenotypes. Overlaps with other phenotypes were annotated with the NHGRI GWAS-catalog⁵⁶ (<http://www.genome.gov/gwastudies/>, date of access Dec-1-2015). All known associations with $P < 5 \times 10^{-8}$ with any disease or primary phenotype were included. For each susceptibility locus with association boundaries defined in Table 1, we annotated all phenotypes that had at least one associated SNP within the region. We also checked whether the hit SNP in the NHGRI GWAS catalogue was the same as, or in high LD with ($r^2 > 0.8$) the CVD lead SNP.

Functional annotation of associated variants. All SNPs in high LD ($r^2 > 0.8$) with the three lead variants at CVD associated loci were identified using the SNP Annotation and Proxy Search⁵⁷ (SNAP) online tool (version 2.2; <https://www.broadinstitute.org/mpg/snap/index.php>). The three non-coding lead SNPs and all variants in high LD (171 non-coding SNPs) were annotated using the following online tools that aim to predict the functional impact of non-coding sequence variants by obtaining information about sequence conservation (conservation metrics like GERP, phastCons and phyloP), chromatin state and protein binding annotation from the Roadmap Epigenomics¹⁸ and ENCODE¹⁷ projects, the effect of SNPs on regulatory motifs, on expression from eQTL studies and on splice sites (Supplementary Tables 5, 6, 8): Combined Annotation-Dependent Depletion (CADD)¹² v1.3 (GRCh37/hg19; <http://cadd.gs.washington.edu/home>), fathmm¹¹ v2.3 (<http://fathmm.biocompute.org.uk/index.html>), MutationTaster¹⁵ (NCBI build 37, Ensemble 69; <http://www.mutationtaster.org>) and HaploReg⁵⁸ v4.1 (<http://www.broadinstitute.org/mammals/haploreg/haploreg.php>). The tool CADD uses the Ensembl Variant Effect Predictor (VEP⁵⁹), data from the ENCODE¹⁷ project and information from UCSC genome browser tracks for annotation. The output of MutationTaster includes 'regulatory features' like histone modification sites, open chromatin or transcription factor binding sites from the Ensembl Regulation database (<http://www.ensembl.org/info/genome/funcgen/index.html>) and 'splice sites' to show possible changes in splice sites using the locally installed third party splice site prediction program NNSplice⁶⁰. For the 'splice site' results, MutationTaster determines the position of the splice site change (increased/decreased or gained/lost splice site) relative to intron/exon borders. Only a loss/decrease of a splice site that occurs at an intron/exon (or reverse) border is shown, whereas such a change distant from any intron/exon border is ignored. For a significant finding the confidence score for the gain of a new splice site has to be greater than 0.3 and for an increase in an existing splice site greater than 10%. The online tool HaploReg annotates variants by their effect on regulatory motifs by using a library of position weight matrices collected from TRANSFAC, JASPAR, protein-binding microarray and ENCODE ChIP-seq experiments⁶¹ and extracts functional annotations from the ENCODE¹⁷ and Roadmap Epigenomics¹⁸ projects, including a chromatin 15-state and 25-state core model learned on five and 12 core chromatin marks, respectively. The chromatin state learning was performed to capture significant combinatorial interactions between different chromatin marks in their spatial context with a tool based on a multivariate Hidden Markov Model (HMM)¹⁸.

Expression quantitative trait loci (eQTL). eQTL data was obtained from the genotype-tissue expression (GTEx) project¹⁴ (<http://www.gtexportal.org/>), the GEUVADIS⁶² data browser (<http://www.ebi.ac.uk/Tools/geuvaldis-das/>) and 11 other studies using HaploReg⁵⁸ v4.1 (see Supplementary Table 6).

Gene expression. Information about *EFEMP1*, *KCNH8* and *SKAP2* gene expression (Supplementary Table 7) was obtained from the publicly available database *THE HUMAN PROTEIN ATLAS*¹⁶ (<http://www.proteinatlas.org/>) and the ENCODE¹⁷ expression data (http://promoter.bx.psu.edu/ENCODE/search_human.php). *THE HUMAN PROTEIN ATLAS*¹⁶ provides quantitative data on gene expression levels (RNA expression in fragments per kilobase of exon model per million mapped reads (FPKM) and protein expression as antibody staining intensity with values from 0 to 9000) for 46 different human cell lines and 44 human tissues. The ENCODE¹⁷ project provides gene expression data of 58 human cell types consisting of 108 datasets (RNA expression in reads per kilobase per million mapped reads (RPKM)).

Variance explained and heritability of single SNPs. The proportion of variance explained by each association signal per population was calculated using a liability threshold model⁶³ assuming a disease prevalence of 0.313 for CVD and log-additive disease risk from replication stage1&2 samples (panel B-C).

References

- Eklof, B., Perrin, M., Delis, K. T., Rutherford, R. B. & Glocviczki, P. Updated terminology of chronic venous disorders: the VEIN-TERM transatlantic interdisciplinary consensus document. *J Vasc Surg* **49**, 498–501 (2009).
- Rabe, E. *et al.* Bonner Venenstudie der Deutschen Gesellschaft für Phlebologie - Epidemiologische Untersuchung zur Frage der Häufigkeit und Ausprägung von chronischen Venenkrankheiten in der städtischen und ländlichen Wohnbevölkerung. *Phlebologie* **32**, 1–14 (2003).
- Eklof, B. *et al.* Revision of the CEAP classification for chronic venous disorders: consensus statement. *J Vasc Surg* **40**, 1248–52 (2004).
- Fiebig, A. *et al.* Heritability of chronic venous disease. *Hum Genet* **127**, 669–74 (2010).
- Milic, D. J. Prevalence and socioeconomic data in chronic venous disease: how useful are they in planning appropriate management? *Medicographia* **33**, 253–258 (2011).
- McGuckin, M. *et al.* Validation of venous leg ulcer guidelines in the United States and United Kingdom. *Am J Surg* **183**, 132–7 (2002).
- Meissner, M. H. *et al.* Primary chronic venous disorders. *J Vasc Surg* **46 Suppl S**, 54S–67S (2007).
- Brinsuk, M., Tank, J., Luft, F. C., Busjahn, A. & Jordan, J. Heritability of venous function in humans. *Arterioscler Thromb Vasc Biol* **24**, 207–11 (2004).
- Matousek, V. & Prerovsky, I. A contribution to the problem of the inheritance of primary varicose veins. *Hum Hered* **24**, 225–35 (1974).
- Cornu-Thenard, A., Boivin, P., Baud, J. M., De Vincenzi, I. & Carpentier, P. H. Importance of the familial factor in varicose disease. Clinical study of 134 families. *J Dermatol Surg Oncol* **20**, 318–26 (1994).
- Shihab, H. A. *et al.* An integrative approach to predicting the functional effects of non-coding and coding sequence variation. *Bioinformatics* **31**, 1536–43 (2015).
- Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* **46**, 310–5 (2014).
- Hao, K. *et al.* Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS Genet* **8**, e1003029 (2012).
- GTeX-Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–60 (2015).
- Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* **11**, 361–2 (2014).
- Uhlen, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
- Birney, E. *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799–816 (2007).
- Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–30 (2015).
- Togni, M. *et al.* Regulation of *in vitro* and *in vivo* immune functions by the cytosolic adaptor protein SKAP-HOM. *Mol Cell Biol* **25**, 8052–63 (2005).
- The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–73 (2010).
- Barrett, J. C. *et al.* Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* **41**, 703–7 (2009).
- Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–24 (2012).
- Westra, H. J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* **45**, 1238–43 (2013).
- Zhou, L. *et al.* SKAP2, a novel target of HSF4b, associates with NCK2/F-actin at membrane ruffles and regulates actin reorganization in lens cell. *J Cell Mol Med* **15**, 783–95 (2011).
- Stone, E. M. *et al.* A single EFEMP1 mutation associated with both Malattia Leventinese and Doyme honeycomb retinal dystrophy. *Nat Genet* **22**, 199–202 (1999).
- McLaughlin, P. J. *et al.* Lack of fibulin-3 causes early aging and herniation, but not macular degeneration in mice. *Hum Mol Genet* **16**, 3059–70 (2007).
- Kostka, G. *et al.* Perinatal lethality and endothelial cell abnormalities in several vessel compartments of fibulin-1-deficient mice. *Mol Cell Biol* **21**, 7025–34 (2001).
- Marmorstein, L. Association of EFEMP1 with malattia leventinese and age-related macular degeneration: a mini-review. *Ophthalmic Genet* **25**, 219–26 (2004).
- Markova, D. *et al.* Genetic heterogeneity of cutis laxa: a heterozygous tandem duplication within the fibulin-5 (FBLN5) gene. *Am J Hum Genet* **72**, 998–1004 (2003).
- Loeys, B. *et al.* Homozygosity for a missense mutation in fibulin-5 (FBLN5) results in a severe form of cutis laxa. *Hum Mol Genet* **11**, 2113–8 (2002).
- Collod-Beroud, G. & Boileau, C. Marfan syndrome in the third Millennium. *Eur J Hum Genet* **10**, 673–81 (2002).
- Albig, A. R., Neil, J. R. & Schiemann, W. P. Fibulins 3 and 5 antagonize tumor angiogenesis *in vivo*. *Cancer Res* **66**, 2621–9 (2006).
- Gillespie, D. L. *et al.* Varicose veins possess greater quantities of MMP-1 than normal veins and demonstrate regional variation in MMP-1 and MMP-13. *J Surg Res* **106**, 233–8 (2002).
- Lim, C. S. & Davies, A. H. Pathogenesis of primary varicose veins. *Br J Surg* **96**, 1231–42 (2009).
- Kowalewski, R., Sobolewski, K., Wolanska, M. & Gacko, M. Matrix metalloproteinases in the vein wall. *Int Angiol* **23**, 164–9 (2004).
- Bharath, V., Kahn, S. R. & Lazo-Langner, A. Genetic polymorphisms of vein wall remodeling in chronic venous disease: a narrative and systematic review. *Blood* **124**, 1242–50 (2014).
- Atta, H. M. Varicose veins: role of mechanotransduction of venous hypertension. *Int J Vasc Med* **2012**, 538627 (2012).
- Raffetto, J. D., Ross, R. L. & Khalil, R. A. Matrix metalloproteinase 2-induced venous dilation via hyperpolarization and activation of K⁺ channels: relevance to varicose vein formation. *J Vasc Surg* **45**, 373–80 (2007).
- Platts, S. H., Mogford, J. E., Davis, M. J. & Meiningner, G. A. Role of K⁺ channels in arteriolar vasodilation mediated by integrin interaction with RGD-containing peptide. *Am J Physiol* **275**, H1449–54 (1998).
- Reddy, M. V. *et al.* Association between type 1 diabetes and GWAS SNPs in the southeast US Caucasian population. *Genes Immun* **12**, 208–12 (2011).
- Alenghat, F. J. *et al.* Macrophages require Skap2 and Sirpalpha for integrin-stimulated cytoskeletal rearrangement. *J Cell Sci* **125**, 5535–45 (2012).
- Raffetto, J. D. & Mannello, F. Pathophysiology of chronic venous disease. *Int Angiol* **33**, 212–21 (2014).
- Rhoads, K. *et al.* A role for Hox A5 in regulating angiogenesis and vascular patterning. *Lymphat Res Biol* **3**, 240–52 (2005).
- Kuwahara, F. *et al.* Hypoxia-inducible factor-1alpha/vascular endothelial growth factor pathway for adventitial vasa vasorum formation in hypertensive rat aorta. *Hypertension* **39**, 46–50 (2002).

45. Bayer, I. M., Caniggia, I., Adamson, S. L. & Langille, B. L. Experimental angiogenesis of arterial vasa vasorum. *Cell Tissue Res* **307**, 303–13 (2002).
46. Krawczak, M. *et al.* PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet* **9**, 55–61 (2006).
47. Wichmann, H. E., Gieger, C. & Illig, T. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67** Suppl 1, S26–30 (2005).
48. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–75 (2007).
49. Team, R. D. C. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria* (2007).
50. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904–9 (2006).
51. Li, Y., Willer, C. J., Ding, J., Scheet, P. & Abecasis, G. R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* **34**, 816–34 (2010).
52. The International HapMap Consortium. Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–8 (2010).
53. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–61 (2007).
54. Jostins, L., Morley, K. I. & Barrett, J. C. Imputation of low-frequency variants using the HapMap3 benefits from large, diverse reference sets. *Eur J Hum Genet* **19**, 662–6 (2011).
55. Pruim, R. J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–7 (2010).
56. Hindorf, L. A. *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* **106**, 9362–7 (2009).
57. Johnson, A. D. *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24**, 2938–9 (2008).
58. Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40**, D930–4 (2012).
59. McLaren, W. *et al.* Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26**, 2069–70 (2010).
60. Reese, M. G., Eckman, F. H., Kulp, D. & Haussler, D. Improved splice site detection in Genie. *J Comput Biol* **4**, 311–23 (1997).
61. Kheradpour, P. & Kellis, M. Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Res* **42**, 2976–87 (2014).
62. Lappalainen, T. *et al.* Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* **501**, 506–11 (2013).
63. So, H. C., Gui, A. H., Cherny, S. S. & Sham, P. C. Evaluating the heritability explained by known susceptibility variants: a survey of ten complex diseases. *Genet Epidemiol* **35**, 310–7 (2011).

Acknowledgements

We thank all patients with chronic venous disease and healthy individuals who contributed samples as well as clinicians and nursing staff for their participation in this study. This work was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (SysInflame grant 01ZX1306 A). This project received infrastructure support from the DFG Excellence Cluster No. 306 “Inflammation at Interfaces” and the PopGen Biobank (Kiel, Germany). The PopGen 2.0 network is supported by a grant from the German Federal Ministry of Education and Research (01EY1103). Andre Franke receives an endowment professorship (Peter Hans Hofschneider Professorship) by the “Stiftung Experimentelle Biomedizin” (Zuerich, Switzerland). The KORA research platform (KORA, Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München - German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

Author Contributions

E.E. and D.E. contributed equally to the manuscript. A.Fi and A.Fr conceived the study. E.E. and D.E. carried out statistical analyses and prepared the manuscript. P.K., A.G., N.F., C.S., S.N., P.R., A.Fi and S.S. were involved in patient recruitment and assembling of phenotypic data. C.G., K.S. and W.L. were responsible for the control panel. All authors read, edited and approved the final manuscript before submission.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing Interests: The authors declare no competing financial interests.

How to cite this article: Ellinghaus, E. *et al.* Genome-wide association analysis for chronic venous disease identifies *EFEMP1* and *KCNH8* as susceptibility loci. *Sci. Rep.* **7**, 45652; doi: 10.1038/srep45652 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017