Repository of the Max Delbrück Center for Molecular Medicine (MDC) in the Helmholtz Association

Associations of clonal hematopoiesis with recurrent vascular events and death in patients with incident ischemic stroke

Arends C.M., Liman T.G., Strzelecka P.M., Kufner A., Löwe P., Huo S., Stein C.M., Piper S.K., Tilgner M., Sperber P.S., Dimitriou S., Heuschmann P., Hablesreiter R., Harms C., Bullinger L., Weber J.E., Endres M., Damm F.

MYELOID NEOPLASIA

CME Article

Associations of clonal hematopoiesis with recurrent vascular events and death in patients with incident ischemic stroke

Christopher M. Arends,[1,2,*] Thomas G. Liman,[2-7,*] Paulina M. Strzelecka,[1,*] Anna Kufner,[2,3] Pelle Löwe,[1] Shufan Huo,[3] Catarina M. Stein,[1] Sophie K. Piper,[3,8,9] Marlon Tilgner,[1] Pia S. Sperber,[4,10,11] Savvina Dimitriou,[1] Peter U. Heuschmann,[12-14] Raphael Hablesreiter,[1] Christoph Harms,[2-4,6] Lars Bullinger,[1,2,15,16] Joachim E. Weber,[2,3] Matthias Endres,[2-6,†] and Frederik Damm[1,2,15,16,†]

[1]Department of Hematology, Oncology, and Cancer Immunology, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [2]Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany; [3]Department of Neurology and [4]Center for Stroke Research Berlin, Department of Experimental Neurology, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [5]German Center for Neurodegenerative Diseases (Deutsches Zentrum für Neurodegenerative Erkrankungen), Partner Site, Berlin, Germany; [6]German Center for Cardiovascular Research (Deutsches Zentrum für Herz-Kreislauferkrankungen), Partner Site, Berlin, Germany; [7]Department of Neurology, Evangelical Hospital Oldenburg, Carl von Ossietzky-University, Oldenburg, Germany; [8]Institute of Biometry and Clinical Epidemiology and [9]Institute of Medical Informatics, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [10]Experimental and Clinical Research Center, a Cooperation Between the Max Delbrück Center for Molecular Medicine in the Helmholtz Association and Charité Universitätsmedizin Berlin, Berlin, Germany; [11]NeuroCure Clinical Research Center, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [12]Institute of Clinical Epidemiology and Biometry and [13]Comprehensive Heart Failure Center Würzburg, University of Würzburg, Würzburg, Germany; [14]Clinical Trial Center Würzburg, University Hospital Würzburg, Würzburg, Germany; [15]German Cancer Consortium (Deutsches Konsortium für Translationale Krebsforschung, DKTK), Partner Site, Berlin, Germany; and [16]German Cancer Research Center (Deutsches Krebsforschungszentrum), Heidelberg, Germany

KEY POINTS

- CH is associated with large-artery atherosclerosis, white matter lesion load, and a proinflammatory profile in patients with ischemic stroke.

- CH and, in particular, mutations in *TET2* and *PPM1D* are associated with higher risk for second vascular events and death after ischemic stroke.

Clonal hematopoiesis (CH) is common among older people and is associated with an increased risk of atherosclerosis, inflammation, and shorter overall survival. Age and inflammation are major risk factors for ischemic stroke, yet the association of CH with risk of secondary vascular events and death is unknown. We investigated CH in peripheral blood DNA from 581 patients with first-ever ischemic stroke from the Prospective Cohort With Incident Stroke–Berlin study using error-corrected targeted sequencing. The primary composite end point (CEP) consisted of recurrent stroke, myocardial infarction, and all-cause mortality. A total of 348 somatic mutations with a variant allele frequency ≥1% were identified in 236 of 581 patients (41%). CH was associated with large-artery atherosclerosis stroke (P = .01) and white matter lesion (P < .001). CH-positive patients showed increased levels of proinflammatory cytokines, such as interleukin-6 (IL-6), interferon gamma, high-sensitivity C-reactive protein, and vascular cell adhesion molecule 1. CH-positive patients had a higher risk for the primary CEP (hazard ratio [HR], 1.55; 95% confidence interval [CI], 1.04-2.31; P = .03), which was more pronounced in patients with larger clones. CH clone size remained an independent risk factor (HR, 1.30; 95% CI, 1.04-1.62; P = .022) in multivariable Cox regression. Although our data show that, in particular, larger and *TET2*- or *PPM1D*-mutated clones are associated with increased risk of recurrent vascular events and death, this risk is partially mitigated by a common germline variant of the IL-6 receptor (*IL-6R* p.D358A). The CH mutation profile is accompanied by a proinflammatory profile, opening new avenues for preventive precision medicine approaches to resolve the self-perpetuating cycle of inflammation and clonal expansion.

Introduction

With aging, the risk of cardiovascular disease and cancer is increasing. Clonal hematopoiesis (CH), defined by the acquisition of somatic mutations in hematopoietic stem cells (HSCs), has been identified as a commonality between these 2 age-related conditions. CH occurs in 20% to 30% of individuals aged >60 years and most frequently involves mutations in epigenetic regulatory genes (eg, *DNMT3A*, *TET2*, and *ASXL1*).[1-7] It is associated with a higher all-cause mortality, an increased risk for cardiovascular events, and an approximately 10-fold risk of developing hematologic malignancies.[6,8] A causal relationship between CH and cardiovascular disease is best known for *TET2*, for which accelerated development of atherosclerosis driven by an altered function of the nucleotide-binding domain (NOD)-like receptor protein 3 (NLRP3)/interleukin-1β (IL-1β) inflammasome of mutated monocytes/macrophages was reported in preclinical models.[8-10] Moreover, in patients with ischemic and nonischemic heart failure, CH has been reported to be associated with rapid progression and unfavorable overall survival.[11,12] These data point toward multifaceted effects of mutated clones in CH-positive individuals, not only affecting self-renewal and differentiation but also inflammatory signaling of mature blood cells.[13,14] Inflammation plays a crucial role in the pathogenesis of ischemic stroke and its functional consequences after brain injury.[15-17] Compared with the rapidly increasing number of reports in the field of myocardial infarction (MI) and heart failure, little is known with respect to CH and ischemic stroke. In their original 2014 article, Jaiswal and colleagues reported an increased risk of ischemic stroke in individuals with CH, which has recently been confirmed in large patient series.[6,18] To fill the knowledge gap concerning the role of CH in patients with ischemic stroke, we conducted a thorough genetic study investigating secondary cardiovascular risk of patients with ischemic stroke and CH in a large prospective cohort.[19]

Patients and methods

Patients

The Prospective Cohort With Incident Stroke–Berlin (PROSCIS-B; ClinicalTrials.gov identifier: NCT01363856) is a prospective, observational, hospital-based cohort study of patients enrolled after first-ever stroke. For further details, notice the previously published study protocol.[19] In short, patients with ischemic stroke were recruited at 1 of 3 stroke units of Charité–Universitätsmedizin Berlin. Within 7 days of stroke onset, patients received interviews, extensive clinical examinations, and blood draws for laboratory analysis. During 3 years of follow-up, annual telephone-based interviews assessed patients' vital status, incidence of cardiovascular diseases, and functional outcome. Vital

status was additionally obtained from the local registry office, even if patients were lost to follow-up. Patients aged ≥18 years were included after first-ever stroke, as defined by World Health Organization criteria.[20] Exclusion criteria were previous stroke (not counting transitory ischemic attacks), brain tumor or metastases, and participation in any intervention study. A Consolidated Standards of Reporting Trials (CONSORT) flow diagram for the inclusion and exclusion of patients is shown in supplemental Figure 1, available on the *Blood* website. In addition, serial blood samples of 24 patients with ischemic stroke from the Berlin Longterm Observation of Vascular Events Study (BeLOVE; German Clinical Trials Registry DRKS00016852) were investigated. BeLOVE is an ongoing observational clinical cohort study of patients with existing cardiovascular disease. The study was approved by the ethics committee (internal review board) of the Charité–Universitätsmedizin Berlin (EA1/218/09) and was conducted in accordance with ethical principles described in the Declaration of Helsinki.

Patient characteristics

Additional information at baseline was collected, including the following: sociodemographic parameters; etiologic subtype of stroke, according to the Trial of ORG 10172 in Acute Stroke Treatment classification[21]; stroke severity, according to National Institutes of Health Stroke Scale (NIHSS); cardiovascular risk factors (current smoking, alcohol consumption, hypertension, peripheral artery disease, prevalent atrial fibrillation, prevalent diabetes mellitus, and history of MI); and use of statin or antithrombotic/anticoagulant medication.

Outcome measures

Primary end point of the PROSCIS-B study is a composite end point (CEP) of recurrent stroke, MI, and all-cause death within 3 years, based on a structured interview with the patient or his/her relatives. Moreover, medical records were screened for any unreported end points. Reported recurrent stroke events or MIs were validated according to World Health Organization criteria[22] using medical records from the responsible hospital and/or the treating physician by 2 senior vascular neurologists who were not involved in the PROSCIS-B study and were blinded to CH status. Survival status was obtained from the Berlin local registration office. Only confirmed end points were used in the analysis.

Targeted sequencing

Peripheral blood (PB) samples were obtained within 7 days after the incident stroke and stored at −80°C. DNA was extracted from PB samples and subjected to an error-corrected targeted sequencing workflow, as published previously.[23,24] Briefly, sequencing libraries were prepared using a commercially available library preparation kit and a customized targeted sequencing panel (Twist BioScience) covering 45 genes recurrently mutated in CH (supplemental Table 1). Bioinformatic error correction was implemented via 9-bp unique molecular identifiers (xGen UDI-UMI adapters by Integrated DNA Technologies). Libraries were sequenced in paired-end mode on Illumina's NovaSeq 6000 sequencing platform. Somatic variants with a variant allele frequency (VAF) ≥1% were identified using our in-house variant calling pipeline (see supplemental Patients and methods).[24] Selected variants with VAF <2% were validated with an independent targeted ultradeep sequencing approach,

as previously described.[25,26] The genotype of the interleukin-6 receptor (*IL-6R*) single-nucleotide polymorphism p.D358A (*rs2228145* C>A) was determined via Sanger sequencing (supplemental Patients and methods).

Biomarker analysis

Several inflammatory biomarkers (ie, high-sensitivity C-reactive protein [hsCRP], IL-1β, IL-6, IL-18, interferon gamma [IFN-γ], and tumor necrosis factor-α [TNF-α]) and markers of endothelial dysfunction (ie, vascular cell adhesion molecule 1 [VCAM-1] and soluble P-selectin) were measured in baseline serum samples at a commercial laboratory (Synevo, Clinical Trials Service Laboratory, Berlin, Germany). hsCRP was determined using the High Sensitivity C-Reactive Protein Assay (Siemens Healthcare, Germany) with a limit of detection of 0.3 mg/L and the IMMULITE 1000 system. IFN-γ, IL-6, IL-18, IL-1β, and TNF-α were determined using the MILLIPLEX Human Cytokine/Chemokine Magnetic Bead Panel multiplex assay (Merck, Germany) with a limit of detection of 3.20 pg/mL and the Luminex 200 System. VCAM-1 and soluble P-selectin were determined using the MILLIPLEX Human Cardiovascular Disease Magnetic Bead Panel 2 multiplex assay (Merck) with a limit of detection of 0.122 and 0.244 ng/mL, respectively, and the Luminex 200 System.

MRI data analysis

Magnetic resonance imaging (MRI) images were taken from clinical routine data using different 1.5- and 3-T scanners at 3 tertiary sites. Fluid-attenuated inversion recovery or T2-weighted images were used to assess white matter lesion severity, according to the Wahlund classification system.[27] The Wahlund visual rating scale, or age-related white matter changes scale, is a scale ranging from a score of 0 to 30 that takes into account both sides of the brain (right and left) and prespecified brain regions (frontal, parieto-occipital, temporal, infratentorial/cerebellum, and basal ganglia). Herein, white matter hyperintensities are assigned a score of 0 to 3 in each region on both sides of the brain. The final score is the sum of all regions. Rating was performed independently by 2 raters (neurologist [A.K.] and senior neuroradiologist).

Statistical analysis

Statistical analysis was performed in R version 4.0.1 using the packages "ggplot2," "stats," "tableone," "survival," "powerSurvEpi," and "forestmodel." The primary end point was assessed using univariable and multivariable Cox regression models to obtain crude and adjusted hazard ratios (HRs). A sample size calculation was performed in advance using the "powerSurvEpi" package in R: Presuming a CH prevalence of 25%, a 2-tailed type I error α of 5%, and an HR of order 2.0 (based on the exploratory analysis for ischemic stroke and cardiovascular disease by Jaiswal et al[6]), 87 events were needed to obtain a power of 80%. Event-free survival curves were estimated using Kaplan-Meier analysis. To account for potential confounders, adjustment was made for baseline risk factors found to be imbalanced (*P* < .2 in Fisher test or Wilcoxon rank-sum test) between the groups of CH-positive and CH-negative individuals: age, sex, arterial hypertension, diabetes, obesity (defined as body mass index >30 kg/m^2), smoking status, stroke severity (as measured by the NIHSS), peripheral artery disease, atrial fibrillation, physical activity before stroke incidence (low vs high), and use of anticoagulants. The Cox proportional hazards

assumption was tested for all covariates using a score test of time-weighted residuals.[28] The primary end point was assessed separately for the variables CH (all somatic mutations with VAF ≥1%), CH of indeterminate potential[29] (CHIP; all somatic mutations with VAF ≥2%), CH with large clones (all somatic mutations with VAF ≥10%), as well as VAF as a continuous measure of clone size. In patients with >1 mutation, the largest VAF was used for clone size calculation. In analyses on single gene level, all patients carrying a mutation in the respective gene X were classified as X mutated, regardless of the VAF of the mutation or other comutations. Pairwise comparisons of variables for exploratory purposes were performed using Wilcoxon rank-sum tests or Fisher exact tests. The 2-sided level of significance was set at a $P < .05$ without adjustment for multiple testing, if not stated otherwise. In all analyses, we considered only complete cases.

Results

Sequencing data analysis

PB samples were available for 581 patients with incident ischemic stroke (supplemental Figure 1). We identified 348 somatic mutations with VAF ≥1% in 236 patients, representing 41% of our cohort. Although 155 patients harbored a single mutation, 81 patients harbored multiple (up to 5) mutations (Figure 1A; supplemental Figure 2). The VAF ranged from 1% to 48% (Figure 1B), with a median of 2.7%. *ASXL1* mutations had a higher median VAF than non-*ASXL1* mutations (5.6% vs 2.6%; $P = .03$; supplemental Figure 3). The most frequently mutated genes were *DNMT3A*, *TET2*, and *ASXL1* (Figure 1C), affecting 81% of all CH-positive patients, consistent with the mutational spectrum described in other nonmalignant cohorts.[6,26,30-32] The most frequent comutation pair was *DNMT3A/TET2* in 19 patients (Figure 1D). *TET2*-mutated patients more frequently harbored a second mutation in another gene than *DNMT3A*-mutated patients (35/73 vs 37/119; $P = .02$). *CBL*-mutated patients frequently harbored second mutations (8/11), in particular *DNMT3A* mutations (Figure 1D; supplemental Figure 2).

Clonal dynamics after ischemic stroke

To investigate the dynamics of CH clones after ischemic stroke, we analyzed 24 CH-positive patients with ischemic stroke with available follow-up blood sample 2 years after stroke from the BeLOVE cohort. We identified 35 CH clones with VAF ≥1% in at least 1 time point. Median VAF increased from 1.7% to 2.2% (maximum increase, 3%; maximum decrease, 3%). To quantify clonal dynamics, we calculated the clonal fitness index following Robertson et al[33] and classified clones as increasing, decreasing, or stable (supplemental Patients and methods). Median fitness was s = 0.09 per year, corresponding to a doubling time of ~8 years for a clone of VAF 1%. Interestingly, we found that 13 of 35 clones showed positive selection and only 2 of 35 clones showed negative selection (supplemental Figure 4). On individual gene level, *SRSF2*-mutated clones showed the highest expansion properties (median fitness, s = 0.25 per year, corresponding to a doubling time of 2.8 years; supplemental Figure 5).

Associations of CH with baseline clinical characteristics

Baseline characteristics of the 581 patients by CH status are shown in Table 1. As previously reported, the presence of CH mutations correlated strongly with patient age, ranging from 0% in patients aged <40 years to >60% in patients aged >80 years (Figure 1E-F). In terms of the different etiological stroke subtypes according to the Trial of ORG 10172 in Acute Stroke Treatment criteria,[21] the highest prevalence of CH was found in patients with evidence of large-artery atherosclerosis (49.7% vs 37.4% in patients with all other stroke subtypes; $P = .01$; Figure 1G). This association remained significant after correction for age, sex, and cardiovascular risk factors in a logistic regression (odds ratio, 1.67; 95% confidence interval [CI], 1.1-2.54; $P = .016$; supplemental Table 2). In terms of laboratory measures, CH-positive patients had lower baseline hemoglobin values (14.1 vs 14.4 g/dL; $P = .04$) and red blood cell counts (4.58/μL vs 4.72/μL; $P = .007$), as well as a lower estimated glomerular filtration rate (74.3 vs 83.3 ml/min per 1.73 m²; $P < .001$; supplemental Figure 6), as recently reported in patients with chronic kidney disease.[34]

Association of CH with inflammatory biomarkers

We measured levels of inflammatory biomarkers that have previously been related to CH (hsCRP, IL-1β, IL-6, TNF-α, IFN-γ, and IL-18) and markers of endothelial activation (VCAM-1 and P-selectin) at baseline in 562 patients. Patients with CH showed higher median values of hsCRP (6.1 [2.1-17.4] vs 4.0 [1.7-10.8] mg/mL; $P = .006$), IL-6 (4.9 [3.2-11.6] vs 3.7 [3.2-7.1] pg/mL; $P = .002$), and VCAM-1 (503 [412-605] vs 459.5 [370-571] ng/mL; $P = .001$) compared with patients without, at a Bonferroni-corrected significance level of $P < .00625$ (Figure 2A; supplemental Figure 7). The fraction of patients with elevated levels of IFN-γ was higher in patients with CH (21.7% vs 10.1%; $P < .001$; Figure 2A). In line with the previously reported association of *TET2* deficiency with elevated IL-1β,[8,9] we found higher IL-1β levels in patients with a *TET2* mutation compared with CH-negative patients (3.20 [3.20-10.22] vs 3.20 [3.20-4.26] pg/mL; $P = .015$; Figure 2B).

Association of CH with white matter lesion load

White matter lesions (WMLs) are areas of abnormal myelination in the brain that reflect a mixture of inflammation, swelling, and damage to the myelin. There is increasing evidence that WMLs may be an early component of neurodegenerative conditions, such as Alzheimer disease and stroke.[35] They can be detected in T2-weighted or fluid-attenuated inversion recovery sequences on MRI, are typically quantified by the Wahlund visual rating scale,[27] and are considered a marker for small-vessel disease.[35,36] Among 398 patients with available MRI data, the Wahlund score significantly differed between CH-positive and CH-negative patients (median, 7 [interquartile range, 4-11] vs 4 [interquartile range, 2-8]; $P < .001$; Figure 2C). As expected, WML load correlated with patient age (supplemental Figure 8A). However, we also found a statistically significant correlation of WML load with clone size (supplemental Figure 8B). In a linear regression with age, sex, arterial hypertension, diabetes mellitus, and clone size as independent variables, age ($P < .0001$), sex ($P = .04$), arterial hypertension ($P = .007$), and clone size ($P = .06$) were predictive for WML load (supplemental Figure 8C).

Associations of CH and clone size with recurrent vascular events and death

With a median follow-up time of 36.2 months, primary end point data were available for all 581 patients with a total of 97 events (50 deaths, 42 strokes, and 5 MIs). In a univariable time-to-event analysis, CH was associated with a higher risk of the
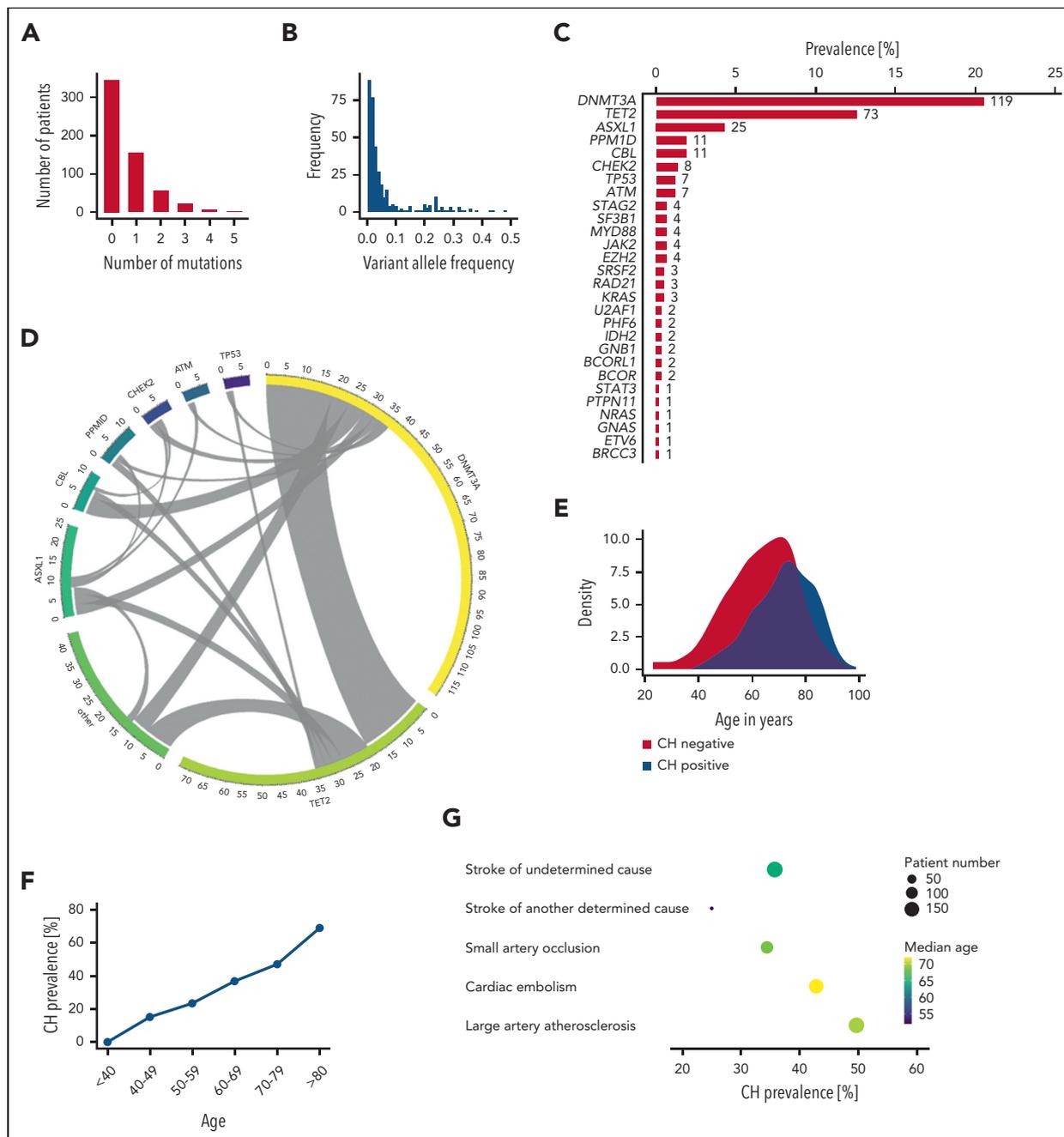
Figure 1. Sequencing analysis and demographic and clinical characteristics of 581 patients with ischemic stroke from the PROSCIS-B study with respect to CH status. (A) Number of patients stratified by the number of detected mutations. (B) Histogram of VAF. (C) Gene-specific somatic mutation prevalence in 581 patients of the PROSCIS-B cohort. (D) CIRCOS plot (www.circos.ca) visualizing the comutational spectrum of the PROSCIS-B cohort. Segment length depicts number of patients with mutation(s) in the respective gene. Multiple mutations in the same gene are not considered. Ribbons depict the frequency of co-occurrence of 2 gene mutations in the same patient. (E) Age distribution of the PROSCIS-B cohort of CH-positive and CH-negative patients. (F) CH prevalence according to patient age. (G) CH prevalence according to etiological cause of stroke, as defined by the Trial of ORG 10172 in Acute Stroke Treatment criteria. Point size visualizes patient number, and color visualizes median age of the group.

occurrence of the CEP (HR, 1.55; 95% CI, 1.04-2.31; *P* = .03; Figure 3A). The risk was higher for patients with multiple mutations (HR, 1.94; 95% CI, 1.16-3.28; *P* = .01; supplemental Figure 9) and patients with large clone CH (defined by VAF ≥ 10%: HR, 2.45; 95% CI, 1.39-4.31; *P* = .002; Figure 3B; supplemental Figure 10). In fact, in a univariable Cox regression with the group of CH-negative patients as reference group, HR increases steadily with clone size (Figure 3C). This led us to investigate the effect of clone size as measured by the VAF as a

continuous parameter. In a univariable Cox regression, clone size was associated with a 51% higher risk per 10% increase in VAF (HR, 1.51; 95% CI, 1.24-1.84; *P* < .001). Because of its strong association with CH, age represents a major source of confounding in our analyses. However, in a multivariable analysis with age, sex, NIHSS score, cardiovascular risk factors, atrial fibrillation, peripheral artery disease, and anticoagulant therapy as covariates, CH clone size remained an independent risk factor (HR, 1.30; 95% CI, 1.04-1.62; *P* = .022; Table 2). Of note,

Table 1. Demographic and clinical characteristics of the PROSCIS-B cohort

Characteristic	CH negative (n = 345)	CH positive (n = 236)	Total (n = 581)	P value*
Age, median (IQR), y	64 (55-73)	73.5 (66.75-81)	68 (59-76)	<.001
Sex, no. (%)				.012
Female	227 (65.8)	130 (55.1)	357 (61.4)	
Male	118 (34.2)	106 (44.9)	224 (38.6)	
TOAST classification, no. (%)				.037
Large-artery atherosclerosis	77 (22.3)	76 (32.2)	153 (26.3)	
Cardiac embolism	80 (23.2)	60 (25.4)	140 (24.1)	
Small-artery occlusion	61 (17.7)	32 (13.6)	93 (16.0)	
Other determined cause	12 (3.5)	4 (1.7)	16 (2.8)	
Undetermined cause	115 (33.3)	64 (27.1)	179 (30.8)	
Modified ranking scale, no. (%)				1
≤1	224 (64.9)	153 (64.8)	377 (64.9)	
≥2	121 (35.1)	83 (35.2)	204 (35.1)	
NIHSS score, no. (%)				.08
≤4	268 (77.7)	168 (71.2)	436 (75.0)	
>4	77 (22.3)	68 (28.8)	145 (25.0)	
Smoking status, no. (%)				.002
Never smoker	114 (33.4)	108 (46.8)	222 (38.8)	
Current or former smoker	227 (66.5)	123 (53.2)	350 (61.2)	
Arterial hypertension, no. (%)				.01
No	136 (39.4)	68 (28.8)	204 (35.1)	
Yes	209 (60.6)	168 (71.2)	377 (64.9)	
Diabetes mellitus, no. (%)				.066
No	279 (80.9)	175 (74.2)	454 (78.1)	
Yes	66 (19.1)	61 (25.8)	127 (21.9)	
Dyslipidemia, no. (%)				.754
No	272 (79.8)	183 (78.5)	455 (79.3)	
Yes	69 (20.2)	50 (21.5)	119 (20.7)	
Obesity, no. (%)				.022
No	250 (72.7)	186 (80.9)	436 (76.0)	
Yes	94 (27.3)	44 (19.1)	138 (24.0)	
Atrial fibrillation, no. (%)				.01
No	283 (82.0)	172 (72.9)	455 (78.3)	
Yes	62 (18.0)	64 (27.1)	126 (21.7)	
Coronary heart disease, no. (%)				.91
No	287 (83.2)	198 (83.9)	485 (83.5)	
Yes	58 (16.8)	38 (16.1)	96 (16.5)	
Peripheral artery disease, no. (%)				.043
No	328 (95.1)	214 (90.7)	542 (93.3)	
Yes	17 (4.9)	22 (9.3)	39 (6.7)	

Analyses were restricted to patients without missing values in the respective category.

IQR, interquartile range; TOAST, Trial of Org 10172 in Acute Stroke Treatment.

*P value from Wilcoxon rank sum test or Fisher exact test.

†High-dose statin therapy is defined as a dose equivalent to atorvastatin ≥ 40 mg/d.

ARENDS et al

Table 1 (continued)

Characteristic	CH negative (n = 345)	CH positive (n = 236)	Total (n = 581)	P value*
History of cancer, no. (%)				.30
No	313 (92.3)	210 (89.7)	523 (91.3)	
Yes	26 (7.7)	24 (10.3)	50 (8.7)	
Statin therapy, no. (%)				.41
None	50 (14.7)	26 (11.1)	76 (13.3)	
Low dose	243 (71.7)	178 (76.1)	421 (73.5)	
High dose†	46 (13.6)	30 (12.8)	76 (13.3)	
Antithrombotics, no. (%)				.19
None	0 (0.0)	1 (0.4)	1 (0.2)	
Antiplatelet	271 (79.9)	174 (74.4)	445 (77.7)	
Anticoagulation	61 (18.0)	50 (21.4)	111 (19.4)	
Both	7 (2.1)	9 (3.8)	16 (2.8)	
IL-6R p.D358A, no. (%)				.069
D/D	148 (42.9)	91 (38.6)	239 (41.1)	
D/A	146 (42.3)	121 (51.3)	267 (46.0)	
A/A	51 (14.8)	24 (10.2)	75 (12.9)	

Analyses were restricted to patients without missing values in the respective category.

IQR, interquartile range; TOAST, Trial of Org 10172 in Acute Stroke Treatment.

*P value from Wilcoxon rank sum test or Fisher exact test.

†High-dose statin therapy is defined as a dose equivalent to atorvastatin ≥ 40 mg/d.

adding inflammatory biomarker levels associated with CH to the multivariable model did not substantially change the magnitude of this association (supplemental Table 3). On gene group level, in particular the group of non-*DNMT3A* (defined by the presence of a CH mutation[s] in other gene[s] than *DNMT3A*) was associated with higher risk of CEP occurrence in univariable analysis (Figure 3D), mainly driven by *TET2* and *PPM1D* mutations (Figure 3E). In multivariable analysis, clone size of non-*DNMT3A* mutations (HR, 1.44 per 10% VAF; 95% CI, 1.16-1.79; *P* = .001), *TET2* mutations (HR, 1.40 per 10% VAF; 95% CI, 1.01-1.91; *P* = .04), and *PPM1D* mutations (HR, 2.30 per 10% VAF; 95% CI, 1.37-3.87; *P* = .002) were significantly associated with the CEP (supplemental Table 4). Referring to the proposed definition of CH with indeterminate potential[29] (CHIP), defined by CH with a VAF ≥2%, *TET2*-CHIP (HR, 2.25; 95% CI, 1.13-4.50; *P* = .021) and *PPM1D*-CHIP (HR, 6.75; 95% CI, 2.15-21.19; *P* = .001) were associated with a higher risk of CEP occurrence in multivariable analysis (supplemental Table 4).

To understand whether the risk for the CEP was concordant across vascular events and mortality, we separately analyzed for the end points recurrent stroke and death using Kaplan-Meier analysis. Interestingly, the increased risk for the CEP in CH-positive patients seems to be primarily driven by mortality rather than recurrent vascular events (supplemental Figure 11).

Interplay of CH, systemic inflammation, and *IL-6R* genotype

A common hypomorphic variant of the IL-6 receptor (*IL-6R* p.D358A; *rs2228145* A>C) functionally impairs IL-6R signaling,[37] is associated with reduced cardiovascular risk in large patient cohorts,[38] and has been suggested as a proxy phenotype for IL-6 inhibition.[39] Recently, Bick et al demonstrated that the risk attenuation attributable to the *IL-6R* p.D358A variant is more pronounced in the presence of CH.[40] To explore the interplay of CH, systemic inflammation, and impaired IL-6R signaling in the context of ischemic stroke, we genotyped all 581 patients for the presence of the *rs2228145* variant. The minor allele (*C*) frequency in our patient cohort was 35.9%. Although IL-6 levels did not significantly differ across different genotypes, individuals with A/D or A/A variants had significantly lower hsCRP values (Figure 4A-B). The *IL-6R* p.D358A variant was not associated with the CEP (HR, 0.93; 95% CI, 0.62-1.41; *P* = .76; supplemental Figure 12). When further stratifying by CH mutation status, however, our data indicate differences in the risk for CEP occurrence depending on CH/*TET2* status (Figure 4C-F). The presence of A/D or A/A variants seems to partially attenuate the risk in the group of CH-positive and *TET2*-mutated patients, whereas it does not in CH/*TET2*-negative patients.

Discussion

In our large prospective cohort of patients with ischemic stroke, we found that CH is associated with a higher risk for recurrent vascular events and death. Associations with adverse outcome have likewise been documented for cardiovascular diseases, such as heart failure or aortic valve stenosis.[41,42]

Although the association of CH with secondary vascular risk is substantially driven by age, several arguments are in favor of an age-independent biologic effect of CH in this context. First, the risk for secondary vascular events in our cohort increased with
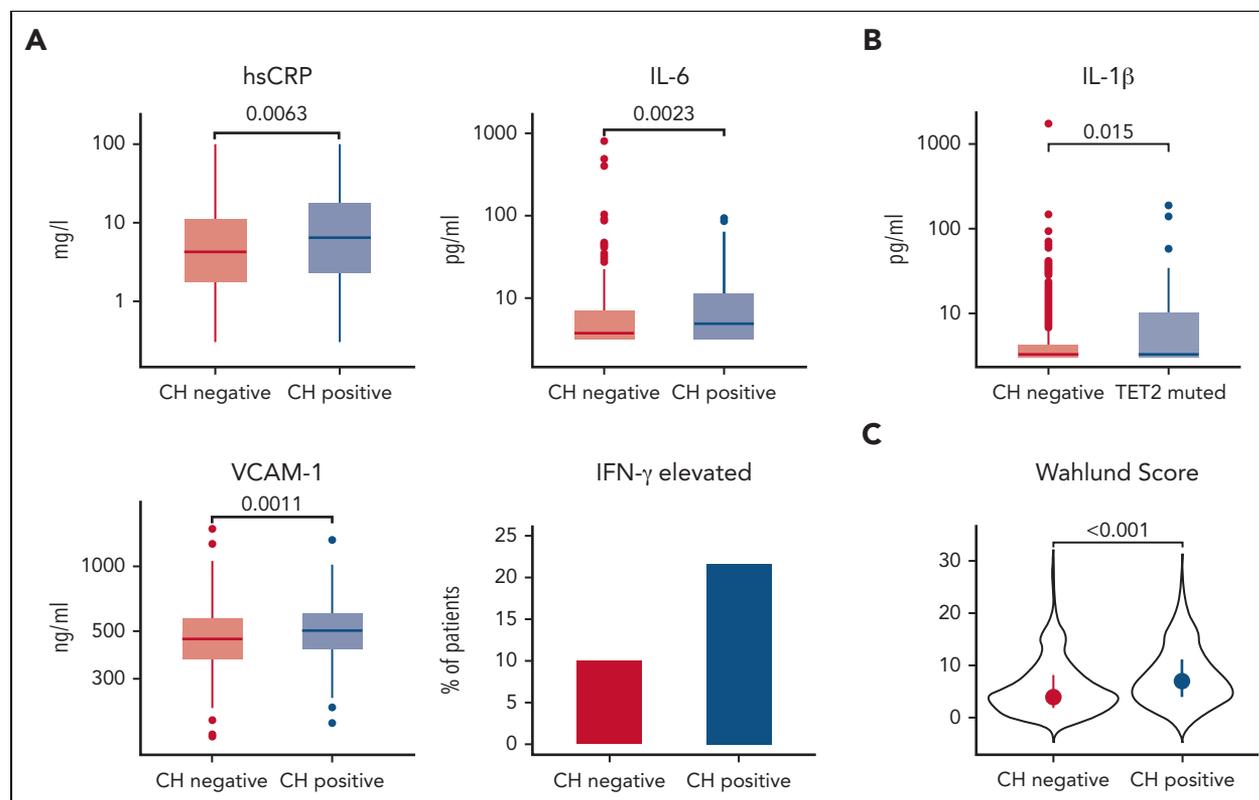
Figure 2. Analysis of inflammatory biomarkers and white matter lesion load. (A) Boxplots of levels of inflammatory biomarkers at baseline of 562 patients with respect to CH mutation status as well as fraction of patients with IFN-γ values ≥ detection limit (3.20 pg/mL). (B) Boxplots of IL-1β levels in patients with a single *TET2* mutation compared with CH-negative patients. (C) White matter lesion load in 398 patients with available MRI data, measured by the Wahlund score according to CH status. *P* values above brackets from Wilcoxon rank-sum test.

clone size and number of mutations in the sense of a dose-response relation, remaining statistically significant after correction for age and other confounders. Clone size itself has been shown to be primarily a function of the mutation-specific clonal fitness rather than patient age (supplemental Figure 13).[43,44] Similar effects of clone size have been reported in the risk prediction for hematologic malignancies.[1,6,32]

Second, we found statistically significant differences between the different mutated genes with respect to secondary vascular risk. Although mutations in the most frequently altered gene, *DNMT3A*, seem to confer a more benign phenotype, there were in contrast strong associations for mutations in genes other than *DNMT3A*, in particular *TET2* and *PPM1D*, with recurrent vascular events and death. For these 2 genes, the proinflammatory dysregulation/pathway leading to accelerated development of atherosclerosis or heart failure is best understood.[8-10,45]

Third, we found stroke subtype-specific differences in the prevalence of CH. Although CH was in general highly prevalent in our cohort of patients with ischemic stroke, it was specifically associated with large-artery occlusion, supporting a potential role of CH in the pathogenesis of this stroke subtype. Compared with the gene mutation prevalence in a cohort of 237 patients with colorectal cancer of similar median age from the FIRE-3 study that was investigated with the exact same sequencing technique,[24] we observed a significantly higher prevalence of *TET2* mutations in patients with ischemic stroke (12.7% vs 5.9%;

P = .004), whereas no significant difference was found for *DNMT3A* mutations (20.5% vs 18.6%; *P* = .56). This enrichment of *TET2* mutations in patients with ischemic stroke further suggests a central role of *TET2*-driven CH in stroke pathogenesis.[24]

Recently, Bhattacharya et al demonstrated that CH is associated with higher stroke risk in individuals without prior stroke in large population-based cohorts.[18] In contrast to our study, CH was associated with hemorrhagic stroke and small-vessel ischemic stroke in an exploratory analysis in one cohort sample of postmenopausal women.[18] Although we cannot investigate the association with hemorrhagic stroke in the PROSCIS-B cohort, because recruitment resulted in predominantly patients with ischemic stroke being included, the reason for the differing findings concerning ischemic stroke subtypes presently remains unclear. Differences in study design (population based vs incident stroke) and sequencing sensitivity (whole exome sequencing vs error-corrected targeted sequencing) might, in part, explain diverging results. However, further studies, including mechanistic studies, will be necessary to determine cause-effect relations beyond the epidemiologic association level.

Following the current concept that CH-inducing mutations in HSCs increase the inflammatory properties of myeloid cells that differentiate from them,[46] CH-positive patients showed elevated hsCRP and cytokine levels of IL-6 and IFN-γ at baseline after ischemic stroke. These findings hint toward a stronger systemic proinflammatory response as a reaction to ischemic brain
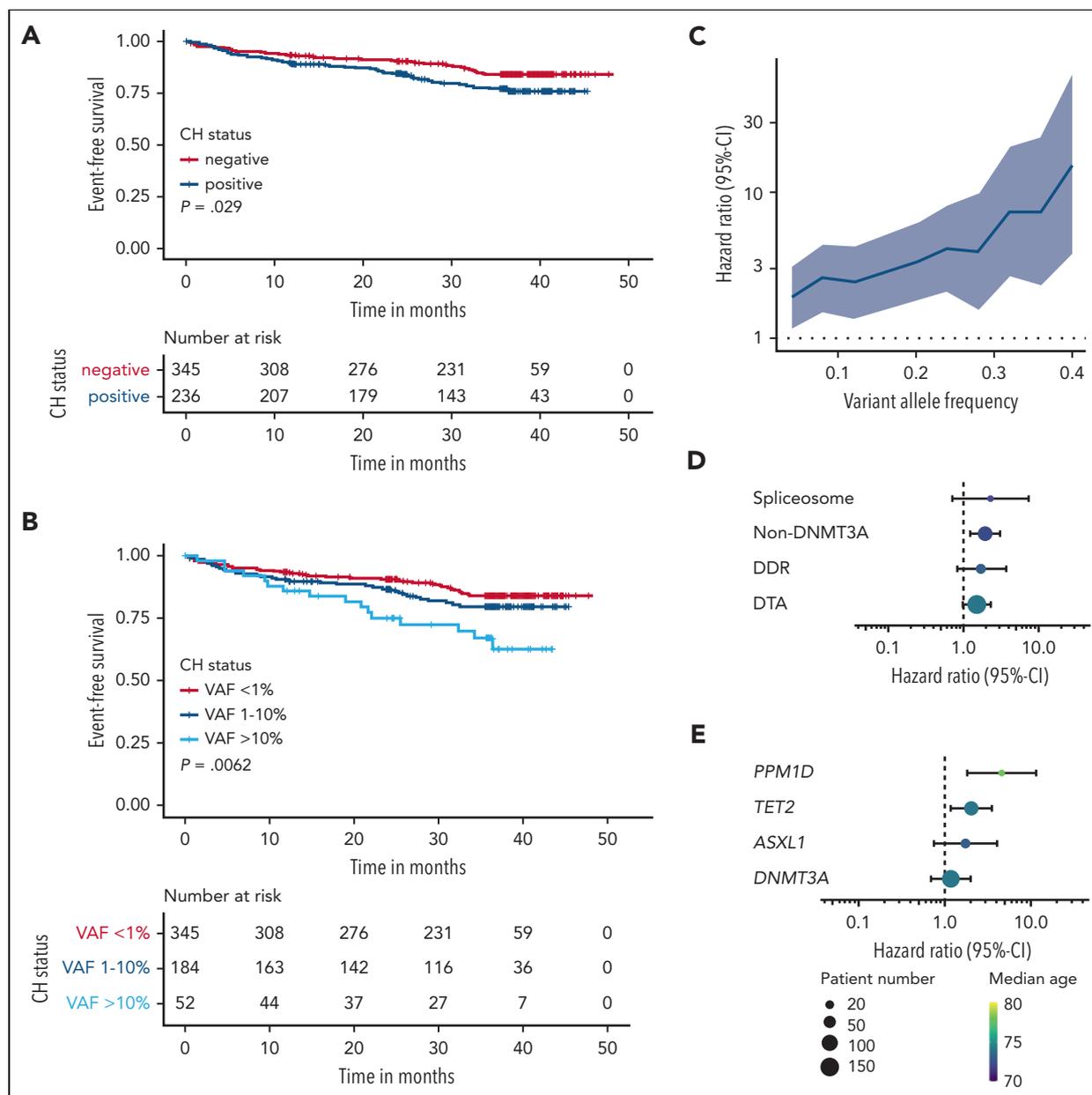
Figure 3. Analysis of secondary vascular risk and mortality according to CH status. (A) Kaplan-Meier analysis of the CEP (recurrent stroke, myocardial infarction, and all-cause death) stratified by CH status. (B) Kaplan-Meier analysis of the CEP, stratified by CH status and clone size. (C) HR for the CEP from univariable Cox regression with varying VAF cutoff and CH-negative patients as reference group. The shaded area visualizes the 95% CI. (D and E) HRs for the CEP from univariable Cox regression for different gene groups (D) and different genes (E). Point size visualizes patient number, color visualizes median age of the group, and bars visualize the 95% CI. DDR, mutations in DNA damage response pathway genes (*TP53*, *PPM1D*, *CHEK2*, *ATM*, and *RAD21*); DTA, mutations affecting the genes *DNMT3A*, *TET2*, and *ASXL1*; non-DNMT3A, mutation(s) in gene(s) other than *DNMT3A*; spliceosome, *SF3B1*, *SRSF2*, and *U2AF1*.

damage. In addition, it is conceivable that resolution of acute inflammation in the brain after primary stroke might be affected in CH-positive patients as the monocyte/granulocyte axis represents both: a key regulator of inflammation resolution and a main target of mutated CH clones.[47] Interestingly, our analysis of the *IL-6R* p.D358A single-nucleotide polymorphism indicates differences in cardiovascular risk modulation dependent on CH/*TET2* status, as previously reported in large population studies without prevalent cardiovascular disease.[40] These results warrant further studies in larger cohorts with potential implications for the applicability of anti-inflammatory treatment. Furthermore, we showed a

previously unappreciated association between CH and elevated VCAM-1 levels. VCAM-1 mediates the adhesion of leukocytes to vascular endothelium and plays a critical role in the development of atherosclerosis.[48] VCAM-1 is also an essential protein for homing and immune tolerance of HSCs.[49]

Although we comprehensively investigated CH in the by far largest cohort of patients with incident ischemic stroke to date, our study has several limitations. When dissecting the CEP into mortality and recurrent strokes, we find that the association is primarily driven by mortality. The cause of death

Table 2. Multivariable Cox regression model for the CEP

Variable		N	Hazard ratio (95% CI)	Hazard ratio (95% CI) value	P
VAF		550		1.30 (1.04, 1.62)	.022
Age		550		1.41 (1.11, 1.80)	.005
Gender	male	342		Reference	
	female	208		0.69 (0.43, 1.10)	.121
Physical activity	low	374		Reference	
	high	176		0.96 (0.58, 1.60)	.886
Stroke severity	NIHSS = 0-4	413		Reference	
	NIHSS > 4	137		1.57 (0.99, 2.49)	.057
Smoking	not active	393		Reference	
	active	157		0.80 (0.44, 1.43)	.446
Hypertension	no	192		Reference	
	yes	358		0.79 (0.48, 1.29)	.345
Diabetes	no	430		Reference	
	yes	120		1.50 (0.93, 2.44)	.098
Obesity	no	417		Reference	
	yes	133		1.31 (0.78, 2.22)	.307
Atrial fibrillation	no	434		Reference	
	yes	116		1.66 (0.78, 3.52)	.186
Peripheral artery disease	no	516		Reference	
	yes	34		1.92 (0.90, 4.06)	.090
Anticoagulant therapy	no	428		Reference	
	yes	122		0.68 (0.32, 1.43)	.304

(x-axis: 0.5 1 2)

Multivariable Cox regression analysis of clone size (measured by the VAF) with potential confounders as covariates. For the variables VAF and age, HRs are given per 10% VAF change and per age decade, respectively.

is difficult to define in individuals after stroke because of significant comorbidities. Population-based studies suggest that approximately two-thirds of deaths are caused by cardiovascular events in patients after ischemic stroke.[50] However, CH is also associated with a higher incidence of hematologic cancer. In the absence of a systematic documentation of cause of death, cancer incidence, or incidence of other vascular events in the PROSCIS-B study, it eventually remains elusive to which extent the reported association between CH and the CEP reflects secondary vascular risk. Moreover, our analyses rely on sequencing data of a single time point. As we exemplified in 24 CH-positive patients, assessing somatic mutations at 2 (or more) time points enables investigations of longitudinal clonal dynamics, which could provide additional insights into a potential relation between clonal fitness and clinical outcomes.

In summary, our results provide novel insights into the interplay of CH, systemic inflammation, and cardiovascular risk and set the stage for further investigations, potentially improving personalized risk stratification and secondary prevention strategies for patients experiencing ischemic stroke.

Acknowledgments

Figure 4. Interplay of clonal hematopoiesis, systemic inflammation, and IL-6R p.D358A. (A) IL-6 levels according to the *IL-6R* p.D358 variant. (B) hsCRP levels according to the *IL-6R* p.D358 variant. (C) Kaplan-Meier analysis of the CEP, stratified by CH status and *IL-6R* p.D358 variant. (D) Kaplan-Meier analysis of the CEP, stratified by *TET2* mutation (mut.) status and *IL-6R* p.D358 variant. (E) HRs and 95% CI for the CEP from univariable Cox regression for CH-positive (pos.) patients, according to the *IL-6R* p.D358 variant compared with CH-negative (neg.) patients. (F) HRs and 95% CI for the CEP from univariable Cox regression for *TET2*-mutated patients, according to the *IL-6R* p.D358 variant compared with CH-negative patients. WT, wild type.

Authorship

Contribution: C.M.A., T.G.L., M.E., and F.D. conceptualized the study; C.M.A., T.G.L., P.M.S., M.E., and F.D. performed methods; C.M.A. and R.H. performed software analysis; C.M.A., T.G.L., and R.H. performed formal analysis; C.M.A., T.G.L., P.M.S., A.K., P.L., S.H., C.M.S., S.K.P., M.T., P.S.S., S.D., P.U.H., C.H., M.E., and F.D. performed investigations; T.G.L., C.H., L.B., J.E.W., M.E., and F.D. obtained resources; C.M.A., P.M.S., and R.H. performed data curation; C.M.A. and F.D. performed writing (original draft preparation); C.M.A, P.M.S., T.G.L., M.E., and F.D. performed writing (review and editing); C.M.A. performed visualization; M.E. and F.D. supervised the study; M.E. and F.D. performed project administration; and M.E. and F.D. acquired funding; and all authors have read and agreed to the published version of the article.

ORCID profiles: C.M.A., 0000-0003-4417-4750; P.M.S., 0000-0003-0226-6788; P.L., 0000-0002-6222-5323; S.H., 0000-0002-0119-7517; C.M.S., 0000-0002-9015-7671; S.K.P., 0000-0002-0147-8992; M.T., 0000-0003-3196-6214; P.U.H., 0000-0002-2681-3515; C.H., 0000-0002-2063-2860; J.E.W., 0000-0002-1666-6021; M.E., 0000-0001-6520-3720; F.D., 0000-0001-5553-1173.

Correspondence: Matthias Endres, Department of Neurology, Neurologie, Charité–Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany; email: matthias.endres@charite.de; and Frederik Damm, Department of Hematology, Oncology, and Cancer Immunology, Charité–Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany; email: frederik.damm@charite.de.

Footnotes

REFERENCES

1. Abelson S, Collord G, Ng SWK, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature*. 2018;559(7714):400-404.

2. Acuna-Hidalgo R, Sengul H, Steehouwer M, et al. Ultra-sensitive sequencing identifies high prevalence of clonal hematopoiesis-associated mutations throughout adult life. *Am J Hum Genet*. 2017;101(1):50-64.

3. Buscarlet M, Provost S, Zada YF, et al. DNMT3A and TET2 dominate clonal hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. *Blood*. 2017;130(6):753-762.

4. Desai P, Mencia-Trinchant N, Savenkov O, et al. Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nat Med*. 2018;24(7):1015-1023.

5. Genovese G, Kahler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014;371(26):2477-2487.

6. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488-2498.

7. Xie M, Lu C, Wang J, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med*. 2014;20(12):1472-1478.

8. Jaiswal S, Natarajan P, Silver AJ, et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *N Engl J Med*. 2017;377(2):111-121.

9. Fuster JJ, MacLauchlan S, Zuriaga MA, et al. Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science*. 2017;355(6327):842-847.

10. Sano S, Oshima K, Wang Y, et al. Tet2-mediated clonal hematopoiesis accelerates heart failure through a mechanism involving the IL-1beta/NLRP3 inflammasome. *J Am Coll Cardiol*. 2018;71(8):875-886.

11. Assmus B, Cremer S, Kirschbaum K, et al. Clonal haematopoiesis in chronic ischaemic heart failure: prognostic role of clone size for DNMT3A- and TET2-driver gene mutations. *Eur Heart J*. 2021;42(3):257-265.

12. Pascual-Figal DA, Bayes-Genis A, Diez-Diez M, et al. Clonal hematopoiesis and risk of progression of heart failure with reduced left ventricular ejection fraction. *J Am Coll Cardiol*. 2021;77(14):1747-1759.

13. Christen F, Hablesreiter R, Hoyer K, et al. Modeling clonal hematopoiesis in umbilical cord blood cells by CRISPR/Cas9. *Leukemia*. 2022;36(4):1102-1110.

14. Tall AR, Levine RL. Cardiovascular disease: commonality with cancer. *Nature*. 2017;543(7643):45-47.

15. Endres M, Moro MA, Nolte CH, Dames C, Buckwalter MS, Meisel A. Immune pathways in etiology, acute phase, and chronic sequelae of ischemic stroke. *Circ Res*. 2022;130(8):1167-1186.

16. Elkind MS, Ramakrishnan P, Moon YP, et al. Infectious burden and risk of stroke: the northern Manhattan study. *Arch Neurol*. 2010;67(1):33-38.

17. Elkind MSV, Boehme AK, Smith CJ, Meisel A, Buckwalter MS. Infection as a stroke risk factor and determinant of outcome after stroke. *Stroke*. 2020;51(10):3156-3168.

18. Bhattacharya R, Zekavat SM, Haessler J, et al. Clonal hematopoiesis is associated with higher risk of stroke. *Stroke*. 2022;53(3):788-797.

19. Liman TG, Zietemann V, Wiedmann S, et al. Prediction of vascular risk after stroke - protocol and pilot data of the Prospective Cohort With Incident Stroke (PROSCIS). *Int J Stroke*. 2013;8(6):484-490.

20. Hatano S. Experience from a multicentre stroke register: a preliminary report. *Bull World Health Organ*. 1976;54(5):541-553.

21. Adams HP Jr, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke: definitions for use in a multicenter clinical trial: TOAST: Trial of Org 10172 in Acute Stroke Treatment. *Stroke*. 1993;24(1):35-41.

22. Aho K, Harmsen P, Hatano S, Marquardsen J, Smirnov VE, Strasser T. Cerebrovascular disease in the community: results of a WHO collaborative study. *Bull World Health Organ*. 1980;58(1):113-130.

23. Frick M, Chan W, Arends CM, et al. Role of donor clonal hematopoiesis in allogeneic hematopoietic stem-cell transplantation. *J Clin Oncol*. 2019;37(5):375-385.

24. Arends CM, Dimitriou S, Stahler A, et al. Clonal hematopoiesis is associated with improved survival in patients with metastatic colorectal cancer from the FIRE-3 trial. *Blood*. 2022;139(10):1593-1597.

25. Mylonas E, Yoshida K, Frick M, et al. Single-cell analysis based dissection of clonality in myelofibrosis. *Nat Commun*. 2020;11(1):73.

26. Arends CM, Galan-Sousa J, Hoyer K, et al. Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis. *Leukemia*. 2018;32(9):1908-1919.

27. Wahlund LO, Barkhof F, Fazekas F, et al. A new rating scale for age-related white matter changes applicable to MRI and CT. *Stroke*. 2001;32(6):1318-1322.

28. Grambsch P, Therneau T. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994;81(3):515-526.

29. Steensma DP, Bejar R, Jaiswal S, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2015;126(1):9-16.

30. Zink F, Stacey SN, Norddahl GL, et al. Clonal hematopoiesis, with and without candidate

driver mutations, is common in the elderly. *Blood.* 2017;130(6):742-752.

31. Bick AG, Weinstock JS, Nandakumar SK, et al. Inherited causes of clonal haematopoiesis in 97,691 whole genomes. *Nature.* 2020;586(7831):763-768.

32. Saiki R, Momozawa Y, Nannya Y, et al. Combined landscape of single-nucleotide variants and copy number alterations in clonal hematopoiesis. *Nat Med.* 2021;27(7): 1239-1249.

33. Robertson NA, Latorre-Crespo E, Terradas-Terradas M, et al. Longitudinal dynamics of clonal hematopoiesis identifies gene-specific fitness effects. *Nat Med.* 2022;28(7):1439-1446.

34. Vlasschaert C, McNaughton AJM, Chong M, et al. Association of clonal hematopoiesis of indeterminate potential with worse kidney function and anemia in two cohorts of patients with advanced chronic kidney disease. *J Am Soc Nephrol.* 2022;33(5):985-995.

35. Wardlaw JM, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol.* 2013;12(8):822-838.

36. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. *Lancet Neurol.* 2010; 9(7):689-701.

37. Ferreira RC, Freitag DF, Cutler AJ, et al. Functional IL6R 358Ala allele impairs classical IL-6 receptor signaling and influences risk of diverse inflammatory diseases. *PLoS Genet.* 2013;9(4):e1003444.

38. Swerdlow DI, Holmes MV, Kuchenbaecker KB, et al. Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6 MR) Consortium, The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. *Lancet.* 2012;379(9822):1214-1224.

39. Cai T, Zhang Y, Ho YL, et al. Association of interleukin 6 receptor variant with cardiovascular disease effects of interleukin 6 receptor blocking therapy: a phenome-wide association study. *JAMA Cardiol.* 2018;3(9): 849-857.

40. Bick AG, Pirruccello JP, Griffin GK, et al. Genetic interleukin 6 signaling deficiency attenuates cardiovascular risk in clonal hematopoiesis. *Circulation.* 2020;141(2):124-131.

41. Dorsheimer L, Assmus B, Rasper T, et al. Association of mutations contributing to clonal hematopoiesis with prognosis in chronic ischemic heart failure. *JAMA Cardiol.* 2019;4(1):25-33.

42. Mas-Peiro S, Hoffmann J, Fichtlscherer S, et al. Clonal haematopoiesis in patients with degenerative aortic valve stenosis undergoing transcatheter aortic valve implantation. *Eur Heart J.* 2020;41(8):933-939.

43. Fabre MA, de Almeida JG, Fiorillo E, et al. The longitudinal dynamics and natural history of clonal haematopoiesis. *Nature.* 2022; 606(7913):335-342.

44. Watson CJ, Papula AL, Poon GYP, et al. The evolutionary dynamics and fitness landscape of clonal hematopoiesis. *Science.* 2020; 367(6485):1449-1454.

45. Yura Y, Miura-Yura E, Katanasaka Y, et al. The cancer therapy-related clonal hematopoiesis driver gene Ppm1d promotes inflammation and non-ischemic heart failure in mice. *Circ Res.* 2021;129(6):684-698.

46. Speck NA. A pernicious cycle affecting premalignant stem cells. *N Engl J Med.* 2022; 386(6):596-598.

47. Schwab JM, Chiang N, Arita M, Serhan CN. Resolvin E1 and protectin D1 activate inflammation-resolution programmes. *Nature.* 2007;447(7146):869-874.

48. Cybulsky MI, Iiyama K, Li H, et al. A major role for VCAM-1, but not ICAM-1, in early atherosclerosis. *J Clin Invest.* 2001;107(10): 1255-1262.

49. Jaiswal S, Jamieson CH, Pang WW, et al. CD47 is upregulated on circulating hematopoietic stem cells and leukemia cells to avoid phagocytosis. *Cell.* 2009;138(2): 271-285.

50. Bronnum-Hansen H, Davidsen M, Thorvaldsen P, Danish MSG. Long-term survival and causes of death after stroke. *Stroke.* 2001;32(9):2131-2136.

Supplementary Material

Associations of clonal hematopoiesis with recurrent vascular events and death in patients with incident ischemic stroke

Christopher M. Arends[1*], Thomas G. Liman[2,3,4,5,6*], Paulina M. Strzelecka[1*], Anna Kufner[2], Pelle Löwe[1], Shufan Huo[2], Catarina M. Stein[1], Sophie K. Piper[2,9,10], Marlon Tilgner[1], Pia S. Sperber[3,7,8], Savvina Dimitriou[1], Peter U. Heuschmann[11,12,13], Raphael Hablesreiter[1], Christoph Harms[2,3,5], Lars Bullinger[1,14,15], Joachim E. Weber[2], Matthias Endres[2,3,4,5§] and Frederik Damm[1,14,15§]

1. Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Department of Hematology, Oncology, and Cancer Immunology, Berlin, Germany

2. Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health (BIH), Department of Neurology, Berlin, Germany.

3. Charité - Universitätsmedizin Berlin, Center for Stroke Research Berlin (CSB) with Department of Experimental Neurology, Berlin, Germany.

4. German Center for Neurodegenerative Diseases (Deutsches Zentrum für Neurodegenerative Erkrankungen, DZNE), partner site Berlin, Germany.

5. German Center for Cardiovascular Research (Deutsches Zentrum für Herz-Kreislauferkrankungen, DZHK), partner site Berlin, Germany.

6. Department of Neurology, Evangelical Hospital Oldenburg, Carl von Ossietzky-University, Oldenburg, Germany.

7. Experimental and Clinical Research Center, a cooperation between the Max Delbrück Center for Molecular Medicine in the Helmholtz Association and Charité Universitätsmedizin Berlin, Germany

8. NeuroCure Clinical Research Center, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Germany

9. Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health (BIH), Institute of Biometry and Clinical Epidemiology, Berlin, Germany.

10. Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health (BIH), Institute of Medical Informatics, Berlin, Germany.

11. Institute of Clinical Epidemiology and Biometry, University of Würzburg, Würzburg, Germany.

12. Comprehensive Heart Failure Center Würzburg, University of Würzburg, Würzburg, Germany.

13. Clinical Trial Center Würzburg, University Hospital Würzburg, Würzburg, Germany.

14. German Cancer Consortium (DKTK), partner site Berlin, Berlin, Germany

15. German Cancer Research Center (DKFZ), Heidelberg, Germany

* CMA, TGL, PMS contributed equally to this work as first authors.
§ These authors share senior authorship.

Supplementary Methods
Bioinformatic analysis
Paired-end reads (148bp+17bp+8bp+148bp) were sequenced on an Illumina NovaSeq 6000 and processed using our in-house Snakemake[1] pipeline. UMIs were extracted and FASTQs were generated using picard ExtractIlluminaBarcodes, IlluminaBasecallsToSam and SamToFastq subsequently.[2] Raw reads were aligned to GRCh38[3] using bwa mem[4] and UMI information was added using picard MergeBamAlignment.[2] Consensus reads were generated using fgbio GroupReadsByUmi with -s adjacency and fgbio CallMolecularConsensusReads with -M 3.[5] Consensus reads were aligned to GRCh38 using bwa mem and picard MergeBamAlignment. Fgbio FilterConsensusReadsQuality with a minimum of 3 consensus reads and default parameters was used for quality filtering of aligned consensus reads. Local realignment was performed using GATK3[6] RealignerTargetCreator and IndelRealigner.[7] Variants were called using VarDict[8] in single-mode with a minimum allele frequency of 0.0001. Variant calls were annotated using annovar[9] with following databases: refGen, cytoBand, clinvar_20200316, dbnsfp35c, gnomad30_genome, avsnp150, cosmic92_coding, cosmic92_noncoding, revel, nci60.

Filtering of somatic variants
The list of variants called by the above variant calling pipeline was further processed using an R-based filtering script with the following exclusion criteria:
1. Functional criteria
   a. synonymous variants
   b. intronic variants
2. Quality Criteria
   a. Strandbalance = 1
   b. Strandbalance = 0
3. Read count criteria
   a. Coverage < 50
   b. Variant supporting reads < 10
   c. Variant allele frequency < 0.01
4. Cohort/Population-based frequency criteria
   a. Allele frequency in the general population > 10% according to the gnomad30_genome database
   b. Variant frequency in this cohort > 20%
5. Germline/SNP Criteria
   a. 0.45 < VAF < 0.55 or VAF > 0.95 and allele frequency > 0.1% in the gnomad30_genome database or reported in the dbSNP database. Truncating variants at 0.45 < VAF < 0.55 were rescued, if not reported in the gnomad30_genome database and not reported in the dbSNP

Here, Strandbalance is defined as the ratio of variant reads on plus strand to minus strand. Hotspot variants such as *DNMT3A* R882C/H*, GNB1* K57E*, JAK2* V617F*, SF3B1* K666N and K700E*, SFRS2* P95L*, U2AF1* S34F*,* and Q157P/R were rescued. Variants passing these filters were manually evaluated in the Integrative Genome Viewer (Broad Institute, Cambridge, USA).

Validation of variants
Selected variant candidates were validated using targeted-deep sequencing method. Briefly, 120-220 bp amplicons covering the identified mutation region were generated with Touchdown PCR and subsequently pooled for libraries construction. Each pool contained non-overlapping amplicons. The final sequencing libraries were generated using NEBNext library preparation kit (New England Biolabs) and sequenced on Illumina's MiSeq platform with the following

sequencing mode: paired-end, 150 bp (Read 1) + 8 bp (i7 index) + 8 bp (i5 index) + 150 bp (Read 2). Alignment was performed as described above. Read counts and VAFs for previously identified mutation positions were extracted and manually verified in IGV.

IL-6R p.Asp358Ala (rs2228145) genotyping
SNP genotyping was performed using Sanger sequencing. Briefly, 20 μl PCR mixture was made containing 10 μl of KAPA HiFi HotStart ReadyMix (Roche), 1 μl of forward primer (5'-GGGGTTGGAGGGGAAGGTTCCT-3') at 10 μM, 1 μl of reverse primer (5'-GCAATGCAGAGGAGCGTTCCGA-3') at 10 μM, 20 ng of gDNA and remaining volume with nuclease-free water. The following thermal cycling protocol was used: initial denaturation at 95°C for 3 min, 35 cycles of denaturing at 98°C for 20 s, annealing at 65°C for 15 s, extension at 72°C for 10 s, and a final extension at 72°C for 10 s. PCR products were purified using AMPure beads (Beckman Coulter). Samples were sequenced with Sanger sequencing method by Microsynth and rs2228145 variant was determined by visual inspection.

Clonal fitness analysis
24 patients from the BeLOVE cohort were classified as CH-positive based on the error-corrected targeted sequencing of blood DNA collected within 7 days of the stroke onset. For those patients the follow-up samples at a second timepoint ~ 2 years after stroke were available. Sequencing and variant calling was performed analogously to the PROSCIS-B cohort. All variants with VAF ≥ 1% in at least one timepoint were taken into account. The clonal fitness index s was calculated along the lines of Robertson *et al.*[10], where the authors derived from a birth-death model of hematopoiesis, that the time evolution of the VAF $v(t)$ of a mutation can be expressed as

$$v(t) = \frac{e^{s(t-t_0)}}{2(N_W + e^{s(t-t_0)})},$$

Where N_W is the number of wildtype stem cells and t_0 the time the mutation occurred. Given two VAFs v_1 and v_2 at two timepoints t_1 and t_2, the fitness index s can be calculated as

$$s = \frac{1}{\Delta t} \log\left(\frac{v_2}{v_1} \frac{(1-2v_1)}{(1-2v_2)}\right),$$

where $\Delta t = t_2 - t_1$, provided N_W is constant over time. For the VAF, we can assume a binomial error, approximated by the Wald interval

$$\sigma_v = \sqrt{\frac{v(1-v)}{d}},$$

where d is the sequencing depth. The error σ_s of s can be approximated using Gaussian error propagation.

$$\sigma_s = \sqrt{\left(\frac{\partial s}{\partial v_1}\right)^2 \sigma_{v_1}^2 + \left(\frac{\partial s}{\partial v_2}\right)^2 \sigma_{v_2}^2} = \frac{1}{\Delta t}\sqrt{\left(h(v_1)\sigma_{v_1}\right)^2 + \left(h(v_2)\sigma_{v_2}\right)^2},$$

with

$$h(x) = \frac{1}{x} + \frac{2}{1-2x}.$$

A clone was classified as increasing, if $s - 1.96\sigma_s > 0$, as decreasing if $s + 1.96\sigma_s < 0$, and as stable if it was neither increasing nor decreasing.

Supplementary Tables

Table S1

Table S1: List of genes covered by the custom sequencing panel (Twist Bioscience).

No.	Gene	Region	No.	Gene	Region	No.	Gene	Region
1	DNMT3A	Full	16	RAD21	Full	31	ETV6	Full
2	TET2	Full	17	STAG2	Full	32	FLT3	Full
3	JAK2	Full	18	CHEK2	Full	33	GATA1	Full
4	ASXL1	Full	19	GNAS	Full	34	GATA2	Exon 2
5	SF3B1	Full	20	GNB1	Full	35	KIT	Exon 8-19, 17
6	SRSF2	Full	21	ATM	Full	36	MPL	Exon 10
7	TP53	Full	22	KRAS	Full	37	NPM1	Exon 11
8	U2AF1	Full	23	NRAS	Full	38	PTPN11	Full
9	PPM1D	Full	24	WT1	Full	39	RUNX1	Full
10	CBL	Full	25	MYD88	Full	40	SETBP1	Exon 4-9
11	IDH1	Full	26	STAT3	Full	41	NF1	Exon 28-38
12	IDH2	Full	27	BRCC3	Full	42	PHF6	Exon 3-5, 7-9
13	BCOR	Full	28	CALR	Exon 8-9	43	BRAF	Exon 15
14	BCORL1	Full	29	CEBPA	Full	44	NOTCH1	Exon 26, 27, 34
15	EZH2	Full	30	CSF3R	Exon 14,17	45	XPO1	Exon 14

Table S2

Table S2: Multivariable logistic regression for incident large artery atherosclerotic stroke (dependent variable) with CH, age, sex, arterial hypertension, diabetes mellitus, smoking status, obesity and dyslipidemia as independent variables.

Variable	Univariate		
	OR	95%-CI	p-value
CH	1.63	1.08 – 2.48	**0.020**
Age	1.01	1.00 – 1.03	0.154
Male sex	1.07	0.71 – 1.62	0.738
Arterial hypertension	1.09	0.70 – 1.73	0.695
Diabetes mellitus	1.04	0.63 – 1.66	0.889
Smoking status	1.87	1.22 – 2.90	**0.005**
Obesity	1.13	0.70 – 1.80	0.614
Dyslipidemia	1.73	1.09 – 2.74	**0.020**

Table S3

Table S3: Multivariable Cox regression for the CEP. Multivariable models include age, gender, physical activity, stroke severity, smoking status, arterial hypertension, diabetes, obesity, atrial fibrillation, peripheral artery disease and anticoagulant therapy as covariates, as well as one additional biomarker associated with CH (hsCRP, IL-6, IFN-γ, or VCAM-1). 95%-CI = 95% confidence interval.

hsCRP				
Variable	**Level**	**Hazard ratio**	**95%-CI**	**p**
VAF		1.27	1.01-1.59	0.044
Age		1.41	1.11-1.80	0.005
Gender	female	0.67	0.42-1.07	0.093
Physical activity	high	1.00	0.60-1.68	0.987
Stroke severity	NIHSS > 4	1.57	0.99-2.50	0.058
Smoking	active	0.80	0.45-1.43	0.453
Arterial hypertension	yes	0.78	0.47-1.28	0.318
Diabetes	yes	1.45	0.89-2.36	0.135
Obesity	yes	1.30	0.76-2.21	0.335
Atrial fibrillation	yes	1.47	0.68-3.19	0.329
Peripheral artery disease	yes	1.82	0.84-3.90	0.127
Anticoagulant therapy	yes	0.72	0.34-1.54	0.401
$\log_{10}$(hsCRP)		1.17	0.80,1.72	0.417
IL-6				
Variable	**Level**	**Hazard ratio**	**95%-CI**	**p**
VAF		1.31	1.05-1.64	0.016
Age		1.48	1.15-1.90	0.003
Gender	female	0.67	0.42-1.08	0.099
Physical activity	high	0.89	0.53-1.52	0.682
Stroke severity	NIHSS > 4	1.64	1.02-2.64	0.040
Smoking	active	0.80	0.44-1.46	0.472
Arterial hypertension	yes	0.76	0.45-1.26	0.285
Diabetes	yes	1.46	0.89-2.39	0.137
Obesity	yes	1.40	0.80-2.43	0.236
Atrial fibrillation	yes	1.72	0.81-3.67	0.160
Peripheral artery disease	yes	1.73	0.79-3.81	0.170
Anticoagulant therapy	yes	0.58	0.27-1.23	0.153
$\log_{10}$(IL-6)		1.45	0.82-2.58	0.202
IFNγ				
Variable	**Level**	**Hazard ratio**	**95%-CI**	**p**
VAF		1.32	1.05-1.65	0.016
Age		1.49	1.15-1.91	0.002
Gender	female	0.68	0.42-1.10	0.115
Physical activity	high	0.88	0.52-1.50	0.648
Stroke severity	NIHSS > 4	1.66	1.03-2.66	0.036
Smoking	active	0.81	0.45-1.48	0.502
Arterial hypertension	yes	0.76	0.46-1.27	0.299
Diabetes	yes	1.48	0.90-2.43	0.120
Obesity	yes	1.36	0.78-2.36	0.274
Atrial fibrillation	yes	1.84	0.86-3.93	0.117

Variable	Level	Hazard ratio	95%-CI	p
Peripheral artery disease	yes	2.00	0.94-4.26	0.074
Anticoagulant therapy	yes	0.57	0.27-1.22	0.145
log$_{10}$(IFNγ)		1.22	0.60,2.47	0.577
VCAM-1				
Variable	**Level**	**Hazard ratio**	**95%-CI**	**p**
VAF		1.28	1.02-1.61	0.036
Age		1.32	1.03-1.70	0.030
Gender	female	0.71	0.44-1.14	0.153
Physical activity	high	1.03	0.62-1.71	0.913
Stroke severity	NIHSS > 4	1.69	1.06-2.69	0.028
Smoking	active	0.81	0.45-1.46	0.489
Arterial hypertension	yes	0.77	0.47-1.28	0.317
Diabetes	yes	1.39	0.85-2.26	0.186
Obesity	yes	1.28	0.76-2.18	0.354
Atrial fibrillation	yes	1.48	0.68-3.21	0.320
Peripheral artery disease	yes	2.01	0.94-4.29	0.072
Anticoagulant therapy	yes	0.74	0.35-1.59	0.447
log$_{10}$(VCAM-1)		4.80	0.97-23.82	0.055

Table S4

Table S4: Univariable and multivariable hazard ratios (HR) for the CEP from Cox regression for different (sub-) groups with CH negative patients as reference group. Multivariable analysis includes age, gender, physical activity, stroke severity, smoking status, arterial hypertension, diabetes, obesity, atrial fibrillation, peripheral artery disease and anticoagulant therapy as covariates. Only complete cases entered the analysis. CH = clonal haematopoiesis, defined by somatic mutation with variant allele frequency (VAF) ≥ 1%, CHIP = clonal haematopoiesis of indeterminate potential defined by CH with VAF ≥ 2%, LCCH = large clone CH, defined by VAF ≥ 10%, CH VAF = VAF as a continuous parameter, non-DNMT3A CHIP = CHIP defined by the presence of one or more mutation(s) in other gene(s) than DNMT3A, non-DNMT3A clone size = VAF of non-DNMT3A mutations as a continuous parameter, TET2 clone size = VAF of TET2 mutations as a continuous parameter, TET2 CHIP = somatic mutation in TET2 with VAF ≥ 2%, PPM1D clone size = VAF of PPM1D mutations as a continuous parameter, PPM1D CHIP = somatic mutation in PPM1D with VAF ≥ 2%, 95%-CI = 95% confidence interval.

Subgroup	Univariable			Multivariable		
	HR	95%-CI	p-value	HR	95%-CI	p-value
CH	1.55	1.04 - 2.31	0.031	1.14	0.73 - 1.80	0.56
CHIP	1.69	1.10 - 2.61	0.017	1.19	0.72 – 1.98	0.48
LCCH	2.45	1.39 - 4.31	0.002	1.67	0.87 - 3.20	0.14
CH clone size	1.51	1.24 - 1.84	< 0.001	1.30	1.04 - 1.62	0.022
non-*DNMT3A* clone size	1.60	1.32 - 1.95	< 0.001	1.44	1.16 - 1.79	0.001
non-*DNMT3A* CHIP	2.18	1.37 - 3.47	0.001	1.71	1.01 - 2.91	0.048
TET2 clone size	1.53	1.14 – 2.05	0.005	1.40	1.01 - 1.93	0.04
TET2 CHIP	2.72	1.47 – 5.01	0.001	2.25	1.13 - 4.50	0.021
PPM1D clone size	2.16	1.32 - 3.52	0.002	2.30	1.37 - 3.87	0.002
PPM1D CHIP	7.65	2.75 - 21.22	< 0.001	6.75	2.15 - 21.19	0.001

Supplementary Figures

Figure S1


```
┌─────────────────────────────┐          ┌─────────────────────────────┐
│ 690 patients included       │─────────▶│ 21 withdrew consent         │
└─────────────────────────────┘          └─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐          ┌─────────────────────────────┐
│ 669 patients in PROSCIS- B  │─────────▶│ 37 intracerebral hemorrhage │
└─────────────────────────────┘          │ 5 sinus venous thrombosis   │
                                         └─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐          ┌─────────────────────────────┐
│ 627 ischemic stroke         │─────────▶│ 5 no consent for blood -    │
└─────────────────────────────┘          │ based analysis              │
                                         └─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐          ┌─────────────────────────────┐
│ 622 consent for blood -     │─────────▶│ 37 no material available/   │
│ based analysis              │          │ technical issues preventing │
└─────────────────────────────┘          │ genetic analysis            │
                                         └─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐          ┌─────────────────────────────┐
│ 585 patients with           │─────────▶│ 4 with manifest hematologic │
│ sequencing data             │          │ disease at study entry      │
└─────────────────────────────┘          └─────────────────────────────┘
              │
              ▼
┌─────────────────────────────────────────────────────────────────────┐
│ 581 patients for final analysis                                      │
└─────────────────────────────────────────────────────────────────────┘
```
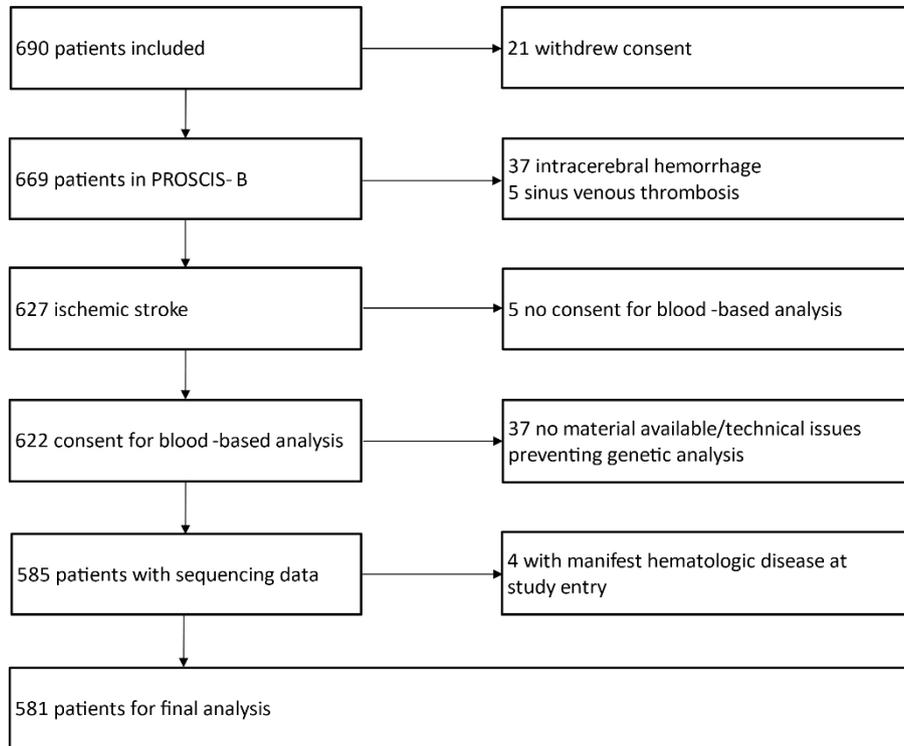
Figure S1: CONSORT flow diagram of inclusion and exclusion of patients from the Prospective Cohort With Incident Stroke Berlin.
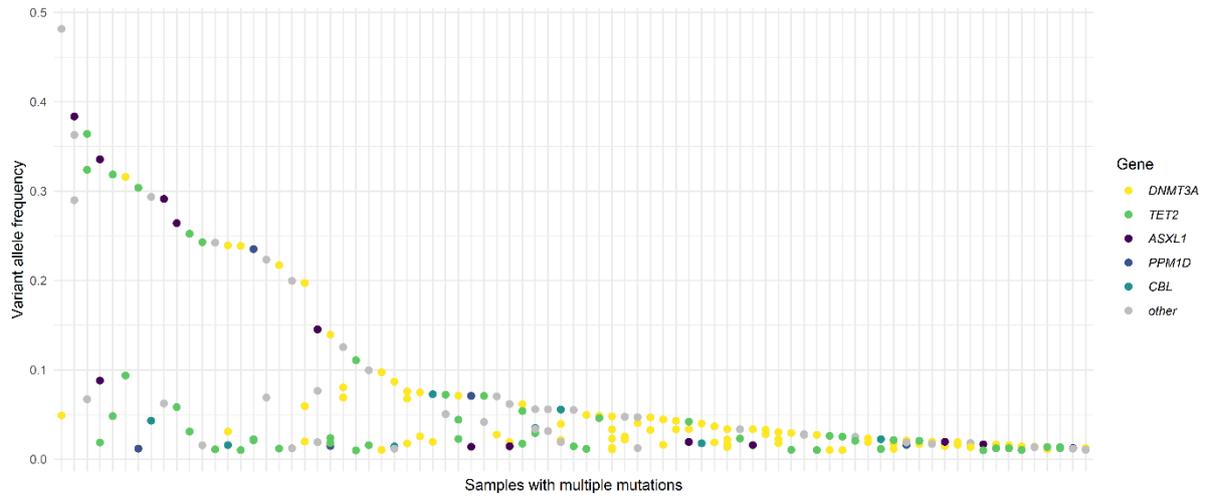
Figure S2

Figure S2: Variant allele frequencies of somatic mutations for patients with multiple mutations, ordered by the largest VAF in the sample.
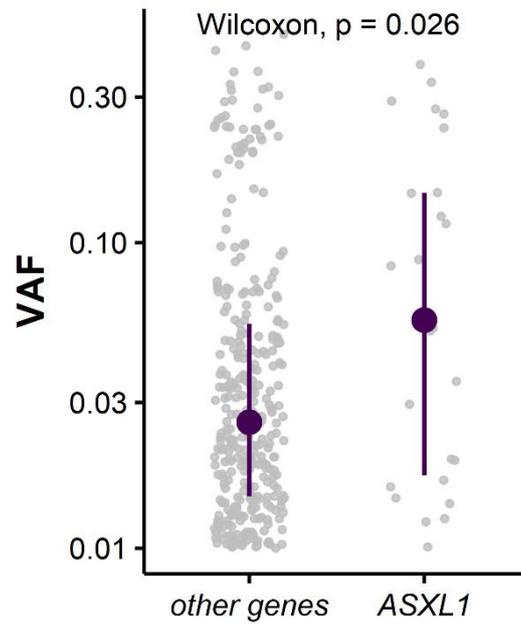
Figure S3

Figure S3: Variant allele frequencies (VAF) of ASXL1 mutations (n = 26) compared to other mutations (n = 322). P-value from Wilcoxon rank-sum test.
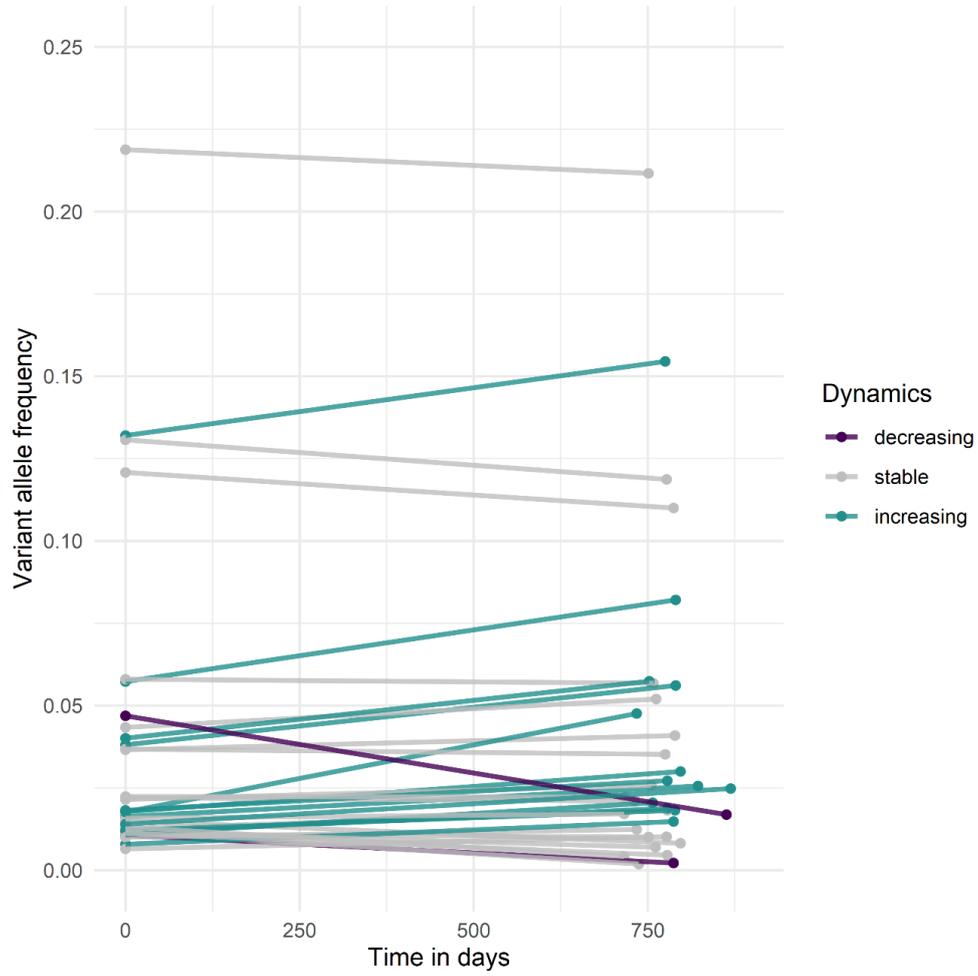
Figure S4

Figure S4: Dynamics of 35 CH clones in 24 patients with ischemic stroke from the BeLOVE cohort. Clones are classified as increasing, stable or decreasing, according to their fitness index s (see Supplementary methods).

Figure S5
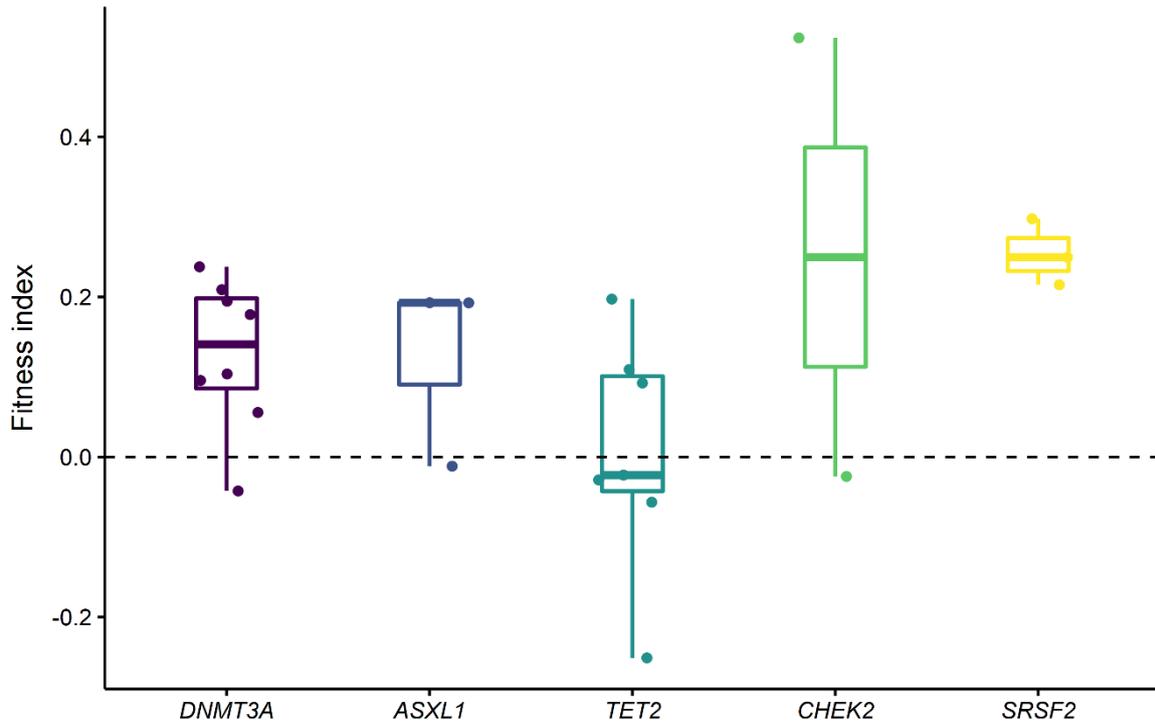
Figure S5: Fitness indices for the five most frequently mutated genes in 24 patients from the BeLOVE cohort.

Figure S6
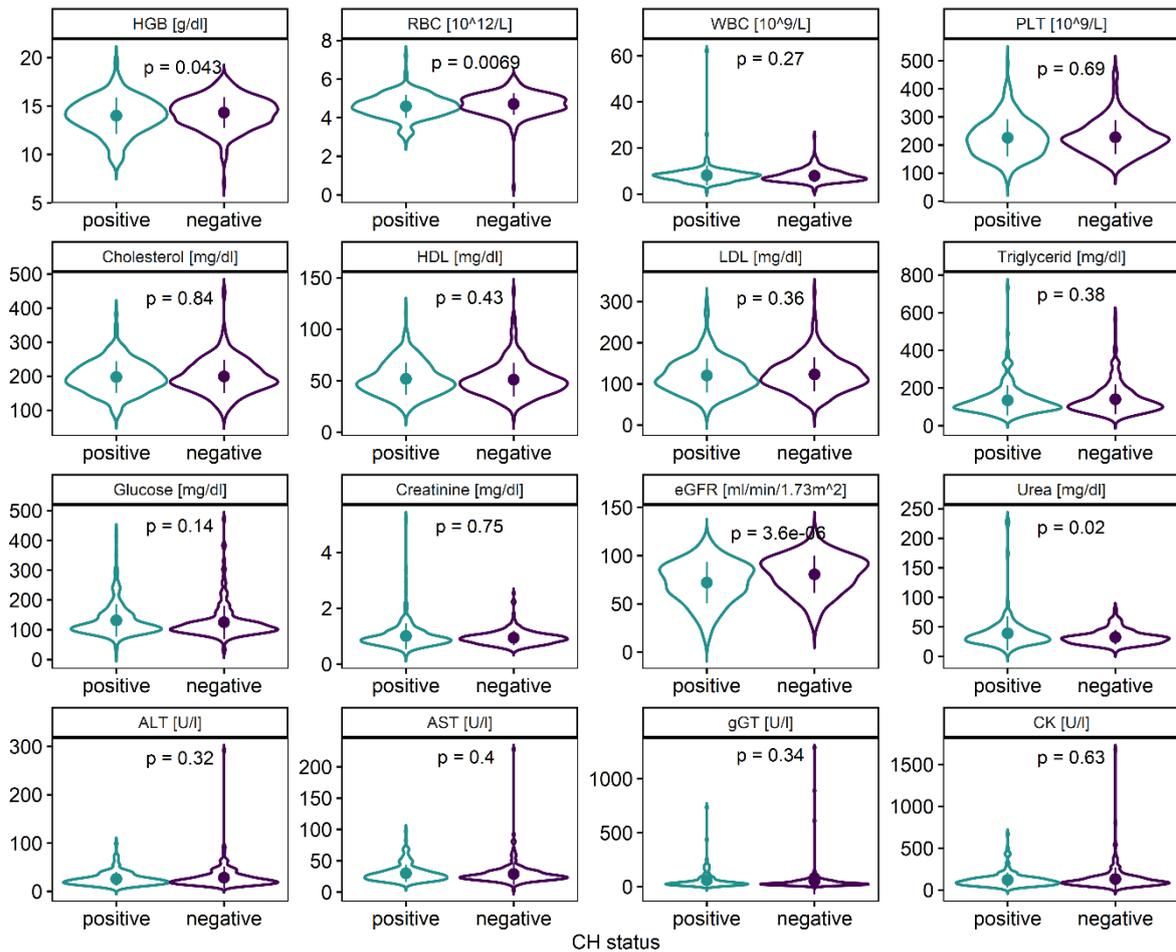
Figure S6: Violin plots of baseline laboratory measures in CH positive and CH negative patients. HGB = hemoglobin, RBC = red blood cell count, WBC = white blood cell count, PLT = platelet count, HDL = high-density lipoprotein, LDL = low-density lipoprotein, eGFR = estimated glomerular filtration rate, ALT = alanine transaminase, AST = aspartate transaminase, gGT = Gamma-glutamyltransferase, CK = creatinkinase. P-values from Wilcoxon rank-sum test without correction for multiple testing.

Figure S7
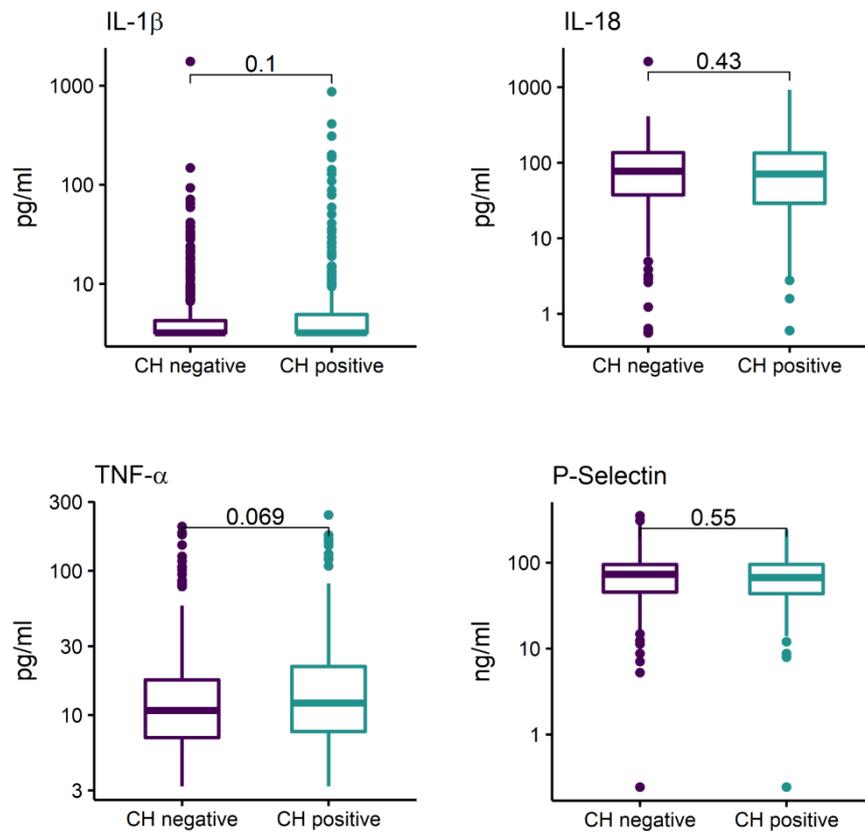
Figure S7: Inflammatory biomarkers that did not significantly differ between CH positive and CH negative patients at a Bonferroni corrected significance level of P < 0.0065. P-values above brackets from Wilcoxon rank sum test.
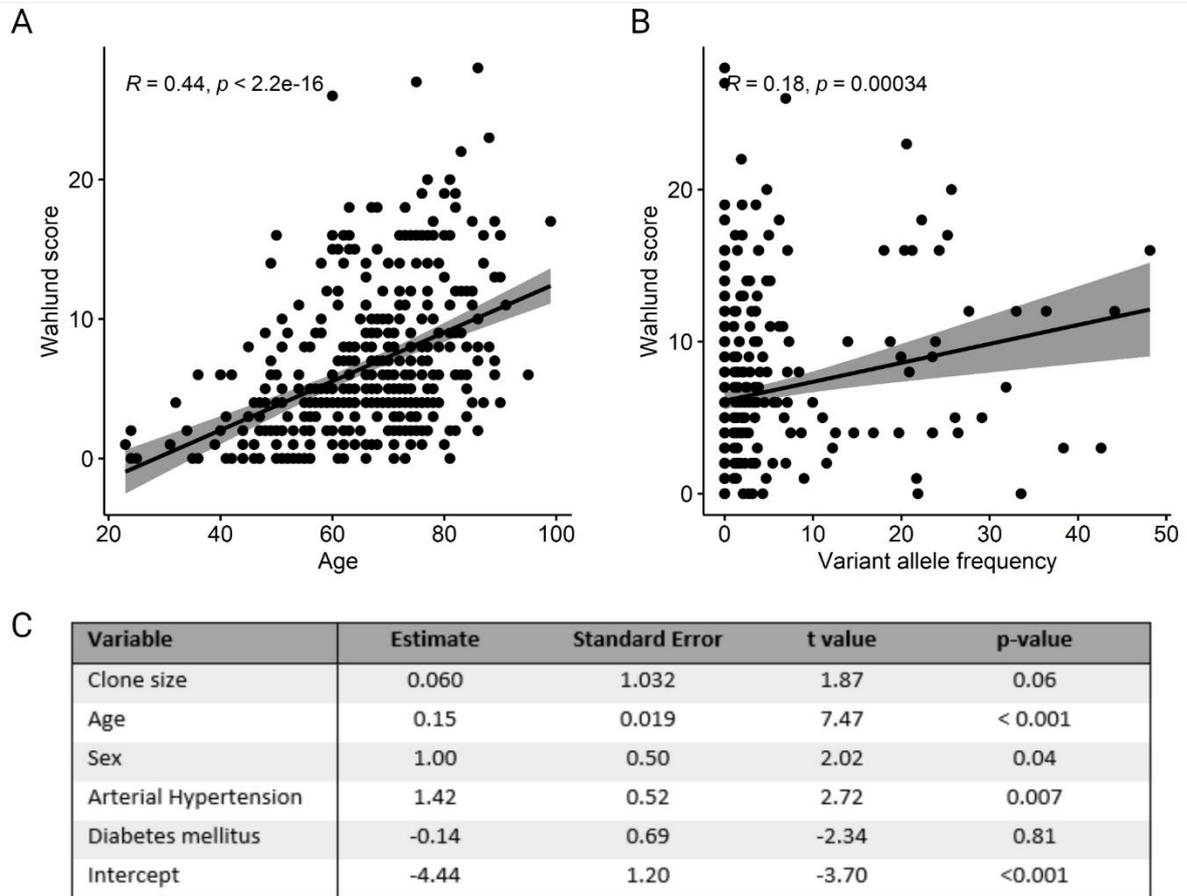
Figure S8

A

B

C

Variable	Estimate	Standard Error	t value	p-value
Clone size	0.060	1.032	1.87	0.06
Age	0.15	0.019	7.47	< 0.001
Sex	1.00	0.50	2.02	0.04
Arterial Hypertension	1.42	0.52	2.72	0.007
Diabetes mellitus	-0.14	0.69	-2.34	0.81
Intercept	-4.44	1.20	-3.70	<0.001

Figure S8: A) Scatterplot of white matter lesion load as measured by the Wahlund score and age. B) Scatterplot of the Wahlund score and clone size given by the variant allele frequency. R represents the Pearson correlation coefficient. C) Linear regression with Wahlund score as dependent variable and clone size, age, sex, arterial hypertension and diabetes mellitus as independent variables.
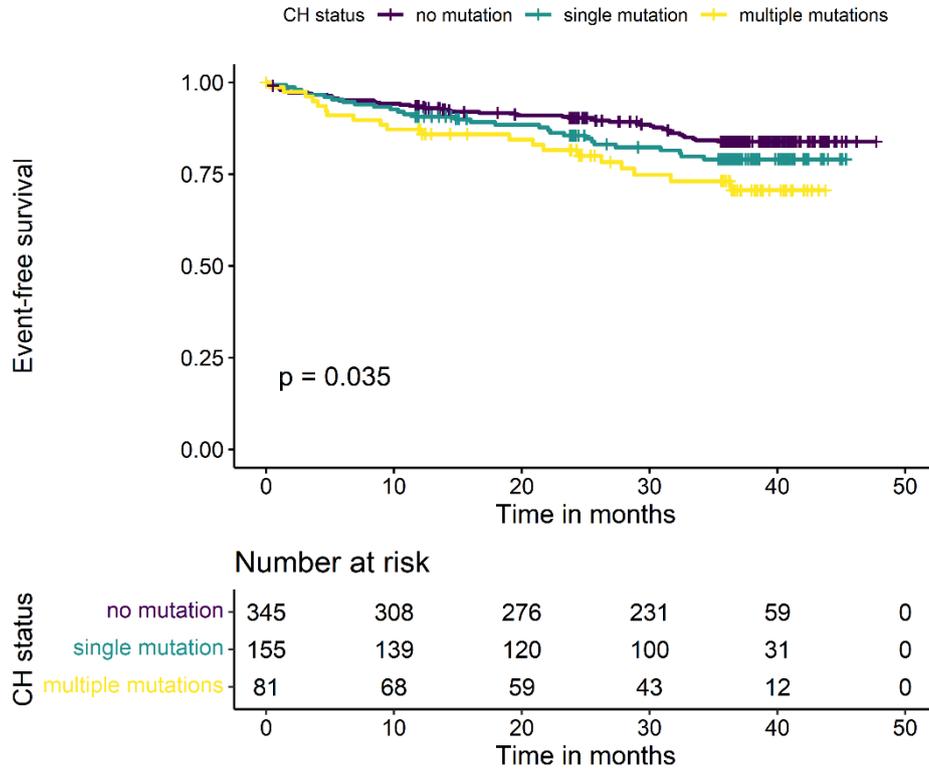
Figure S9

Figure S9: Kaplan-Meier analysis of the composite endpoint stratified by number of CH mutations.
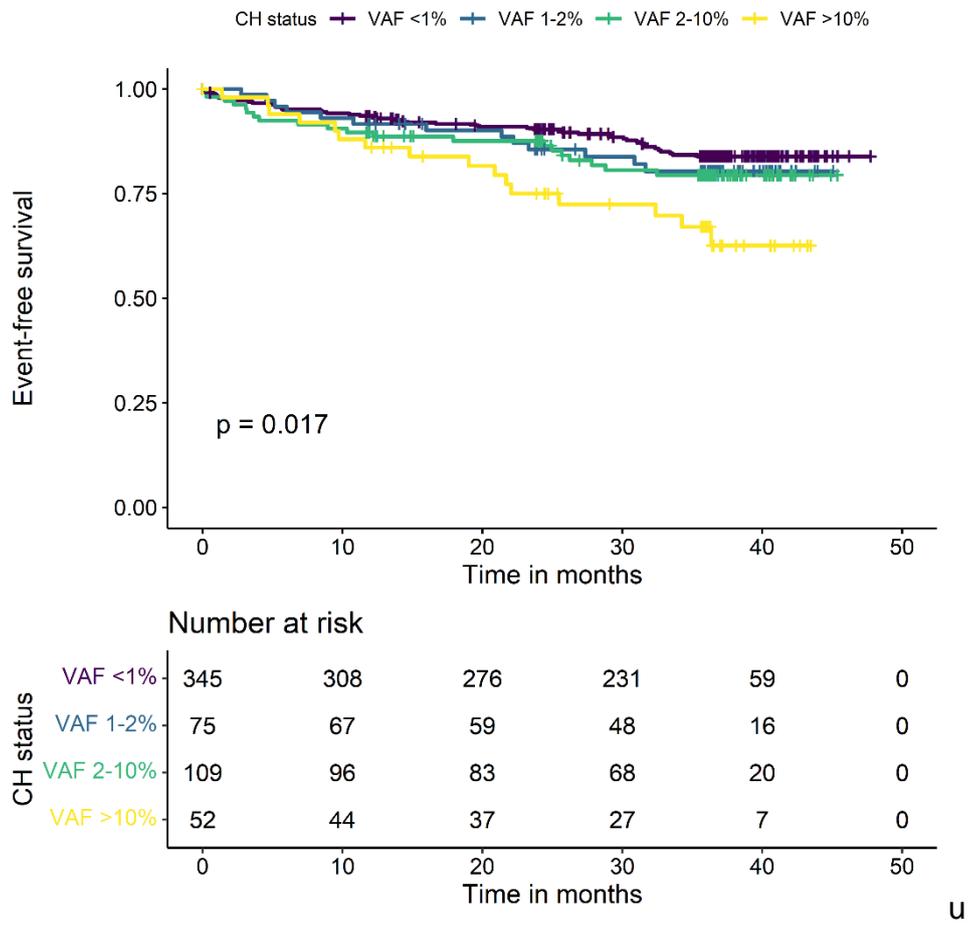
Figure S10

Figure S10: Kaplan-Meier analysis of the composite endpoint (recurrent stroke, myocardial infarction and all-cause death) stratified by four classes of CH status: CH negative (VAF < 1%), CH with VAF of 1-2%, CH with VAF 2-10% and CH with VAF > 10%.
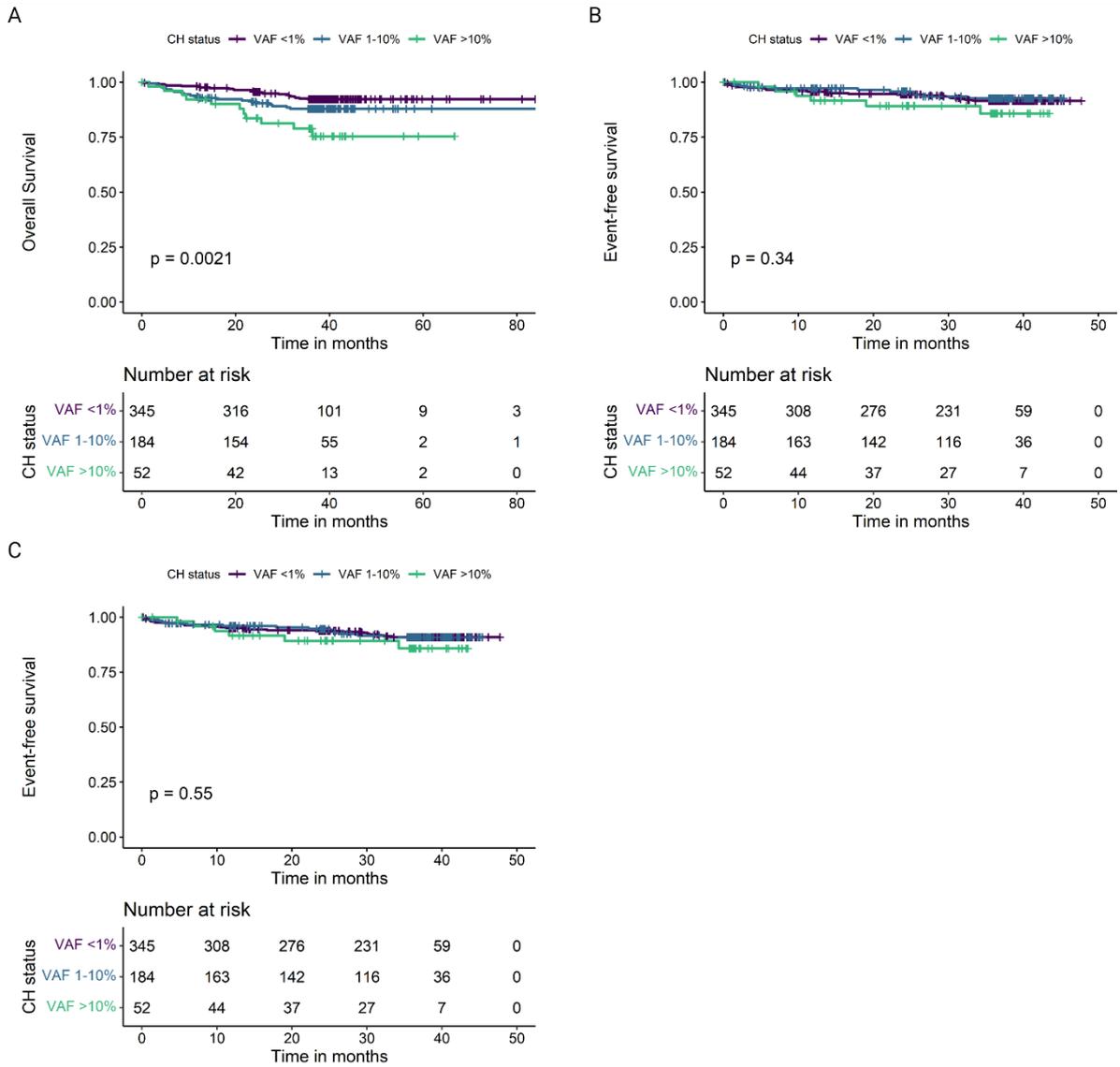
Figure S11

Figure S11: A) Kaplan-Meier analysis of A) overall survival, B) time to recurrent stroke or C) time to vascular event (stroke or myocardial infarction) in patients of the PROSCIS-B cohort stratified by CH status.
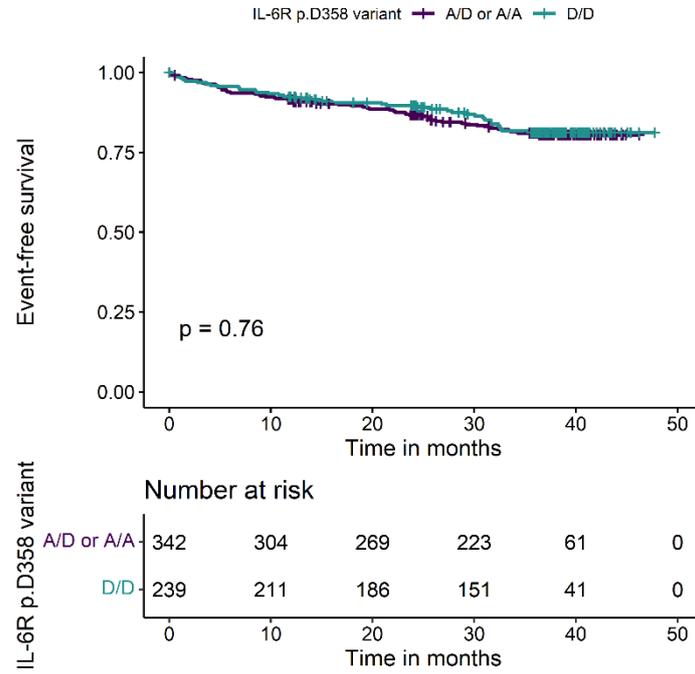
Figure S12

Figure S12: Kaplan-Meier analysis of the CEP stratified by IL-6R p.D358 variant.
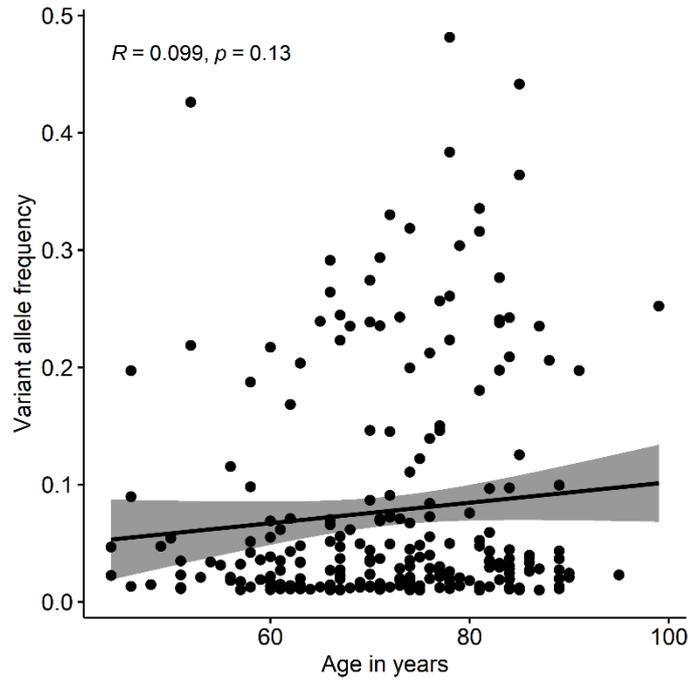
Figure S13

Figure S13: Scatterplot of variant allele frequency vs. patient age. R represents the Pearson correlation coefficient.

References

1.   Molder F, Jablonski KP, Letcher B, et al. Sustainable data analysis with Snakemake. *F1000Res.* 2021;10:33.
2.   Picard Tools. *Broad Institute GitHub repository, Accessed: 2020/08/20; version 2200 Available: http://broadinstitutegithubio/picard/*.
3.   Schneider VA, Graves-Lindsay T, Howe K, et al. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Res.* 2017;27(5):849-864.
4.   H L. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bioGN] Available: http://arxivorg/abs/13033997*. 2013.
5.   Fennell T HN. fgbio. *Available: https://githubcom/fulcrumgenomics/fgbio*.
6.   McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-1303.
7.   DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43(5):491-498.
8.   Lai Z, Markovets A, Ahdesmaki M, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res.* 2016;44(11):e108.
9.   Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164.
10.   Robertson NA, Latorre-Crespo E, Terradas-Terradas M, et al. Longitudinal dynamics of clonal hematopoiesis identifies gene-specific fitness effects. *Nat Med.* 2022;28(7):1439-1446.