nature medicine

Article

https://doi.org/10.1038/s41591-024-03213-1

Unidirectional association of clonal hematopoiesis with atherosclerosis development

Received: 19 May 2023

Accepted: 25 July 2024

Published online: 30 August 2024

Check for updates

A list of authors and their affiliations appears at the end of the paper

Clonal hematopoiesis, a condition in which acquired somatic mutations in hematopoietic stem cells lead to the outgrowth of a mutant hematopoietic clone, is associated with a higher risk of hematological cancer and a growing list of nonhematological disorders, most notably atherosclerosis and associated cardiovascular disease. However, whether accelerated atherosclerosis is a cause or a consequence of clonal hematopoiesis remains a matter of debate. Some studies support a direct contribution of certain clonal hematopoiesis-related mutations to atherosclerosis via exacerbation of inflammatory responses, whereas others suggest that clonal hematopoiesis is a symptom rather than a cause of atherosclerosis, as atherosclerosis or related traits may accelerate the expansion of mutant hematopoietic clones. Here we combine high-sensitivity DNA sequencing in blood and noninvasive vascular imaging to investigate the interplay between clonal hematopoiesis and atherosclerosis in a longitudinal cohort of healthy middle-aged individuals. We found that the presence of a clonal hematopoiesis-related mutation confers an increased risk of developing de novo femoral atherosclerosis over a 6-year period, whereas neither the presence nor the extent of atherosclerosis affects mutant cell expansion during this timeframe. These findings indicate that clonal hematopoiesis unidirectionally promotes atherosclerosis, which should help translate the growing understanding of this condition into strategies for the prevention of atherosclerotic cardiovascular disease in individuals exhibiting clonal hematopoiesis.

Somatic mutations that accumulate during the human lifespan are emerging as potential contributors to several disorders beyond their known role in cancer[1]. The most prominent example is that of clonal hematopoiesis (CH), a condition in which de novo somatic mutations in hematopoietic stem cells (HSCs) provide a selective advantage that leads to the clonal outgrowth of the mutant cell[2,3]. CH-associated mutations affect a limited set of genes—most frequently the epigenetic regulatory genes *DNMT3A* and *TET2*—that are also commonly mutated in hematological neoplasia. Accordingly, CH is associated with a higher risk of developing hematological cancer[4–8]. In addition, CH is also emerging as a shared risk factor for a growing list of nonhematological disorders[4,6,7,9–23]. Atherosclerotic cardiovascular disease (CVD) is the nonhematological condition most strongly linked to CH[2–4,9,13,19,24]. However, whether accelerated atherosclerosis is a cause or a consequence of CH remains a matter of debate. While sequencing studies in humans and experimental studies in mice support a direct causal contribution of some CH-related mutations to atherosclerosis by exacerbating inflammatory responses[4,9,13,19,25,26], mathematical models based on human and mouse data suggest that the association between CH and CVD may instead reflect reverse causality[27]. Specifically, it has been posited that

e-mail: vfuster@cnic.es; jjfuster@cnic.es

CH could simply be a direct consequence of increased HSC proliferation in individuals with atherosclerosis, resulting in accelerated expansion of existing mutant hematopoietic clones[27–29]. This reverse causality hypothesis could also apply to other CH-related associations, thereby explaining why CH is associated with a very diverse range of conditions[4,6,7,9–18,30–32]. The uncertain nature of the relationship between CH and atherosclerosis has particularly important clinical implications, as the specific atherogenic effects of CH mutations on leukocytes are being discussed as potential targets for the development of personalized strategies to reduce or prevent CVD in individuals with CH[3]. Such approaches would not be effective if the association between CH and atherosclerosis is merely due to the effects of atherosclerosis on the dynamics of CH.

Resolving the controversy surrounding the directionality of the CH–atherosclerosis association requires longitudinal analyses that combine deep genotyping of CH-related genes with deep phenotyping of study participants. Such longitudinal analyses were not feasible in earlier studies, which mainly relied on cross-sectional whole-exome sequencing (WES) datasets at a single time point[4,6,7,9,13,16]. To address this issue, here we take advantage of the Progression of Early Subclinical Atherosclerosis (PESA) study[33], a longitudinal cohort of healthy middle-aged individuals with serially collected blood DNA samples and deep cardiovascular and metabolic phenotyping, including extensive imaging-based assessment of subclinical atherosclerosis burden in multiple vascular territories and at multiple time points. Leveraging this resource, we used a high-sensitivity sequencing approach to investigate CH and its dynamics in a longitudinal manner, aiming to elucidate the interplay between CH and atherosclerosis.

Results

CH in healthy middle-aged individuals

Previous analyses of WES or whole-genome sequencing (WGS) datasets suggested that CH is relatively uncommon in middle-aged individuals, with frequencies ranging approximately from 2% to 3% in individuals aged between 40 and 55 years, compared with >10% in individuals older than 65 (refs. 4,6–8,34). However, these previous observations were limited by the low sensitivity of somatic mutation calling based on WES or WGS data, which hampers the detection of small mutant clones (for example those present with variant allele fraction (VAF) <5%, that is, 10% of mutant blood cells, assuming monoallelic mutations)[34]. Therefore, to elucidate the prevalence and characteristics of CH in this segment of the population, we performed high-sensitivity targeted sequencing of 54 CH-related genes (median depth 3,712×; Extended Data Fig. 1a) in blood samples collected at enrollment from 3,692 participants in PESA (40–55 years old at the time of enrollment, median 45; see Methods for details). We predefined a minimum VAF threshold of 0.2% to identify CH mutations, as our sequencing depth allowed us to detect CH variants at this VAF with a sensitivity greater than 90% (Methods and Extended Data Fig. 1b). Using this threshold and a previously described filtering and curation strategy[35,36] (Methods and Supplementary Table 1), we identified a total of 1,172 CH-related somatic mutations in 46 genes, with 1,079 variants (92%) occurring in 33 genes linked to myeloid CH (typically referred to as CH of indeterminate potential or CHIP) and 93 variants in 13 genes linked to lymphoid CH[8,37] (Extended Data Fig. 1a and Supplementary Table 2). Approximately 25% of individuals carried at least one CH-related mutation detectable with our sequencing approach. Mutations in *DNMT3A* were most frequent (657 mutations, driving CH in 14.8% of individuals) followed by mutations in *TET2* (153 mutations, driving CH in 3.9% of individuals), consistent with previous studies (Fig. 1a and Extended Data Fig. 1c)[4,6,8,34,38]. The basic features of CH mutations in the study population aligned well with those reported in previous WES or WGS analyses[4,6,34,38] (Extended Data Fig. 1d and Supplementary Table 2), with more than one-half of the single nucleotide variants corresponding to a C > T substitution, a mutational signature characteristic of aging and CH (Extended Data Fig. 1e).

The prevalence of CH mutations in this middle-aged population increased with advancing age (Fig. 1b). After adjustment for sex, each additional year of age was independently associated with a 9% higher relative risk of carrying detectable CH mutations (odds ratio (OR) 1.09, 95% confidence interval (CI) 1.07–1.11, $P < 0.001$), and the prevalence of CH was 33.5% in the top quartile of age, which included individuals between ~50 and ~55 years of age (Fig. 1b). Most mutation carriers (76%) had only one detectable mutation, although the frequency of carrying more than one mutation increased with age (Fig. 1c). Age was markedly associated with *DNMT3A*-mutant CH (OR 1.09, 95% CI 1.07–1.12, $P < 0.001$), *TET2*-mutant CH (OR 1.10, 95% CI 1.05–1.14, $P < 0.001$) and CH driven by mutations in other genes (OR 1.11, 95% CI 1.08–1.14, $P < 0.001$), although the increase in CH prevalence with aging differed between driver genes, with *DNMT3A*-mutant CH rising in prevalence earlier in life than CH driven by mutations in other genes (Fig. 1d).

As expected, the mutant hematopoietic clones identified in our study were generally of much smaller size than those identified in previous WES or WGS analyses[4,6,34,38] (median VAF 0.92%, range 0.2–39.9%; Fig. 1e, Extended Data Fig. 1f and Supplementary Table 2). In total, 78.8% of the detected mutations had a VAF smaller than 2%, the minimum clone size threshold used in most previous CH studies[4–8,10–17], and only ~4% of them had a VAF greater than 10%, which identified clones with significant impact on atherosclerotic CVD in previous large studies[4,7,9,13,16,39]. CH mutations over the typical 2% VAF threshold were detected in 6.0% of the study population and in ~9% of those between ~50 and ~55 years of age (Fig. 1f). Computational downsampling of sequencing depth suggests that this elevated prevalence of CH compared with previous WGS/WES studies[4,6–8,34] reflects the better sensitivity provided by our targeted sequencing approach (Extended Data Fig. 1g,h).

Consistent with recent reports[6,40], we found that *DNMT3A*-mutant CH was associated with female sex after adjusting for age, with women exhibiting a 64% higher risk of carrying detectable *DNMT3A* mutations than men (Fig. 1g). In total, 17.8% of women in the study population carried *DNMT3A* mutations, whereas 13.2% of men carried such mutations, and this higher frequency of *DNMT3A* mutations in women was evident across all quartiles of age (Fig. 1h). No significant sex-related differences were observed in the prevalence of CH driven by mutations in genes other than *DNMT3A* (Fig. 1g,h).

Clinical characteristics of CH mutation carriers

Having determined with high resolution the landscape of CH-related mutations in our study population, we next aimed to examine the pathophysiological effects of CH in these middle-aged individuals. Table 1 presents the demographic and clinical characteristics of the study population, stratified by CH status. The characteristics of individuals with clones with VAF ≥2% or gene-specific forms of CH are listed in Extended Data Table 1 and Supplementary Table 3, respectively. Carrying CH-related mutations was associated with higher absolute blood cell counts in all hematopoietic lineages in multivariate regression analysis adjusted for age and sex (Extended Data Table 2). However, this effect was mild overall, with 3.9% and 7.2% relative increases in median counts of circulating leukocytes in those carrying CH-related mutations with VAF <2% and ≥2%, respectively (Extended Data Table 3). No other associations with major hematological parameters were observed. Overall, these findings corroborate the known relationship between CH and blood cell counts[6,7,13,34]. CH mutation carriers also tended to exhibit higher blood pressure and higher levels of glycated hemoglobin HbA1C (Table 1), consistent with some previous reports[6,13,41], but these associations were not statistically significant after adjustment for age and sex (Supplementary Table 4).

No cross-sectional association between CH and atherosclerosis

We next investigated the cross-sectional relationship between CH and subclinical atherosclerosis burden assessed at enrollment through noninvasive vascular imaging across multiple vascular beds, including

https://doi.org/10.1038/s41591-024-03213-1

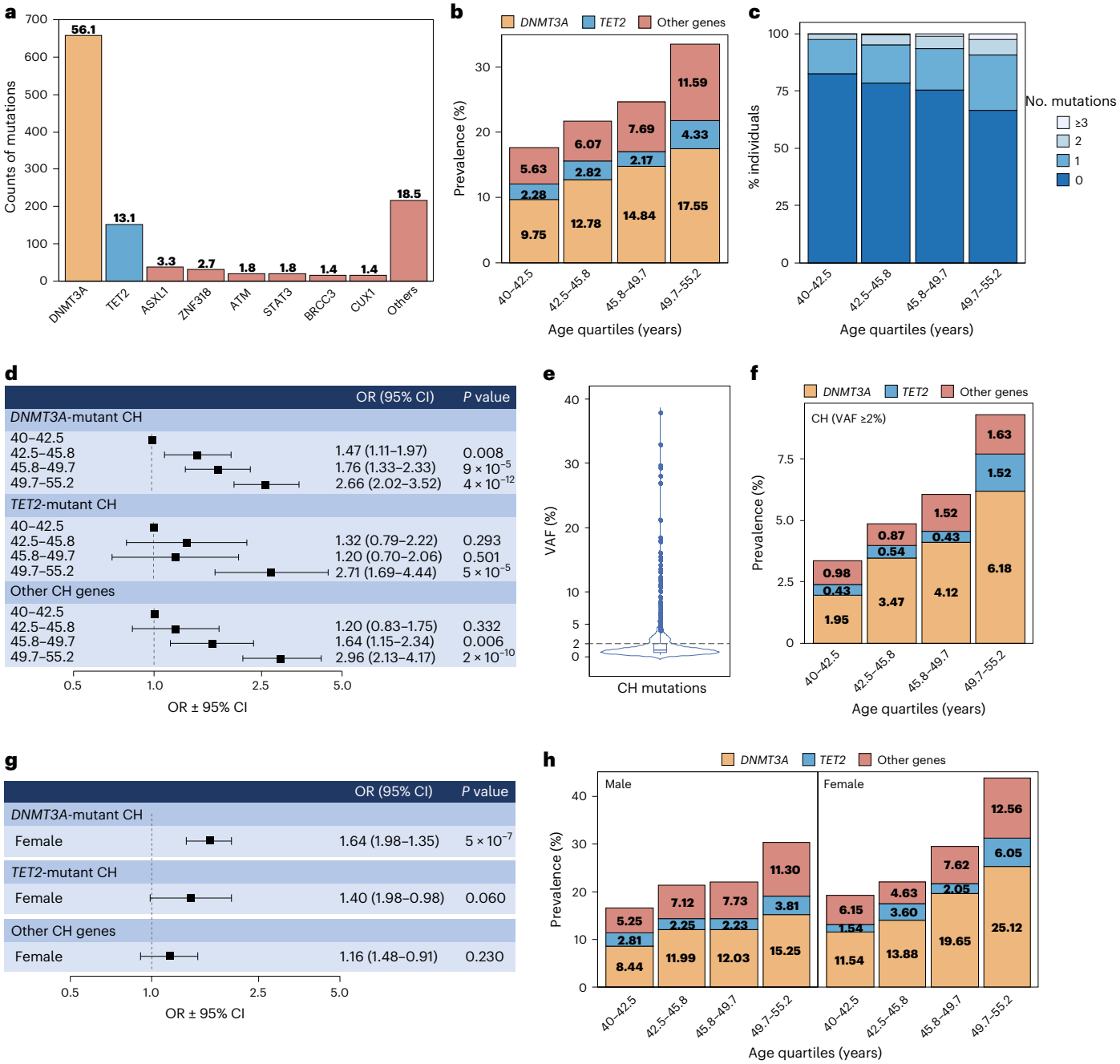


Fig. 1 | Prevalence and characteristics of CH in middle-aged individuals. We performed deep targeted sequencing to identify somatic mutations in a custom panel of 54 CH-related genes in 3,692 individuals from the PESA cohort. **a**, The number of CH driver mutations identified per gene. The values above the bars indicate the percentage of mutations affecting each specific gene. **b**, The CH prevalence across quartiles of age. **c**, The number of mutations per individual across quartiles of age. **d**, The association between advancing age (stratified as quartiles) and CH (analyzed separately as driven by mutations in *DNMT3A*, *TET2* or other genes) based on multivariate logistic regression analyses adjusted for sex. The bars indicate 95% confidence intervals centered in the mean value (square). **e**, The distribution of mutant clone size in the study population, assessed as VAF. The dashed line shows the 2% VAF threshold most typically used to identify CH. The box shows the 25th (Q1), 50th (median) and 75th (Q3) percentiles of the data. The whiskers represent Q1 − 1.5 × IQR at the minimum and Q3 + 1.5 × IQR at the maximum. **f**, The prevalence of CH with VAF ≥2% across quartiles of age. **g**, The association between gene-specific CH and female sex, based on multivariate logistic regression analyses adjusted for age. The bars indicate 95% confidence intervals centered in the mean value (square). **h**, The CH prevalence across quartiles of age stratified by sex. In **b**, **f** and **h**, CH status in individuals carrying more than one mutation was defined on the basis of the mutation with the highest VAF.

three-dimensional vascular ultrasound imaging (3DVUS) of carotid and femoral arteries, and computed tomography (CT)-based determination of coronary artery calcium scores (CACS). Although PESA participants who carried CH mutations exhibited greater subclinical atherosclerosis burden (Extended Data Fig. 2a–d), no significant association was observed after adjusting for age and sex, or age, sex and conventional cardiovascular risk factors (Extended Data Table 4). Similar results were obtained in gene-specific analyses (Supplementary Table 5).

CH is associated with de novo femoral atherosclerosis

We next took advantage of the longitudinal nature of the PESA study and evaluated whether individuals who carried CH mutations at enrollment

https://doi.org/10.1038/s41591-024-03213-1

Table 1 | Baseline characteristics of the study population stratified by the presence of CH mutations

	Total population (*n*=3,692)	No CH (*n*=2,792)	CH, any VAF (*n*=900)	*P* value
Age, years	45.0 [42.0–49.0]	45.0 [42.0–49.0]	47.0 [43.0–51.0]	<0.001
Men, %	2,357 (63.8)	1,812 (64.9)	545 (60.6)	0.020
Dyslipidemia, %	1,520 (41.2)	1,145 (41.0)	375 (41.7)	0.757
Total cholesterol, mg dl^{-1}	200.5±33.3	200.2±33.3	201.4±33.2	0.351
Low plasma HDL-C, %	1,194 (32.3)	904 (32.4)	290 (32.2)	0.963
LDL-C ≥160 mg dl^{-1}, %	627 (17.0)	476 (17.0)	151 (16.8)	0.891
HDL-C, mg dl^{-1}	49.1±12.2	49.0±12.1	49.7±12.8	0.156
LDL-C, mg dl^{-1}	132.4±29.6	132.3±29.7	132.5±29.2	0.861
Plasma triglycerides ≥150 mg dl^{-1}, %	406 (11.0)	303 (10.9)	103 (11.4)	0.665
Triglycerides, mg dl^{-1}	79.0 [60.0–113.0]	79.0 [59.0–112.0]	80.5 [61.0–114.0]	0.302
Serum lipoprotein (a), mg dl^{-1}	17.7 [6.8–42.9]	17.6 [6.7–43.0]	17.8 [6.9–42.6]	0.830
Hypertension, %	419 (11.3)	297 (10.6)	122 (13.6)	0.019
SBP, mmHg	116.2±12.5	116.0±12.2	116.9±13.2	0.073
DBP, mmHg	72.4±9.4	72.2±9.3	73.2±9.6	0.005
Diabetes mellitus, %	68 (1.8)	49 (1.8)	19 (2.1)	0.583
Fasting glucose ≥100 mg dl^{-1}, %	482 (13.1)	349 (12.5)	133 (14.8)	0.088
Fasting glucose, mg dl^{-1}	89.0 [84.0–95.0]	89.0 [84.0–95.0]	89.0 [84.0–95.0]	0.427
Insulin, μU ml^{-1}	5.1 [3.6–7.4]	5.1 [3.6–7.4]	5.1 [3.7–7.6]	0.171
HOMA-IR >2.5, %	366 (9.9)	266 (9.5)	100 (11.1)	0.187
HOMA-IR	1.1 [0.8–1.7]	1.1 [0.7–1.7]	1.1 [0.8–1.8]	0.178
Hemoglobin A1c, %	5.4 [5.2–5.6]	5.4 [5.2–5.6]	5.4 [5.2–5.7]	0.004
Current smoker, %	744 (20.2)	542 (19.4)	202 (22.4)	0.055
Past smoker, %	1,190 (45.2)	878 (43.7)	312 (49.9)	0.008
Ever smoker, %	2,248 (60.9)	1,661 (59.5)	587 (65.2)	0.003
Obesity (BMI ≥30 kg m^{-2}), %	506 (13.7)	375 (13.4)	131 (14.6)	0.418
BMI, kg m^{-2}	25.9±3.8	25.9±3.8	26.0±3.9	0.421
Central obesity, %	754 (20.4)	553 (19.8)	201 (22.3)	0.112
Waist circumference, cm	89.2±12.0	89.2±11.9	89.4±12.4	0.560
Metabolic syndrome, %	338 (9.2)	243 (8.7)	95 (10.6)	0.108
hsCRP, mg dl^{-1}	0.1 [0.0–0.2]	0.1 [0.0–0.2]	0.1 [0.0–0.2]	0.170
GOT, U l^{-1}	18.0 [16.0–22.0]	18.0 [15.0–22.0]	18.0 [16.0–22.0]	0.864
GPT, U l^{-1}	19.0 [14.0–27.0]	20.0 [14.0–27.0]	19.0 [14.0–27.0]	0.264
GGT, U l^{-1}	20.6 [14.0–31.0]	20.0 [14.0–31.0]	21.0 [14.0–30.0]	0.655

Data are expressed as *n*, %; mean±s.d.; or median [IQR], when appropriate. *P* values were calculated through independent two-tailed *t*-tests for continuous variables exhibiting a normal distribution, two-tailed Wilcoxon rank tests for continuous variables with a nonnormal distribution and χ^2 tests for categorical variables. HOMA-IR, Homeostatic Model Assessment for Insulin Resistance; GOT, glutamic-oxaloacetic transaminase; GPT, glutamic-pyruvic transaminase; GGT, gamma glutamyl transferase.

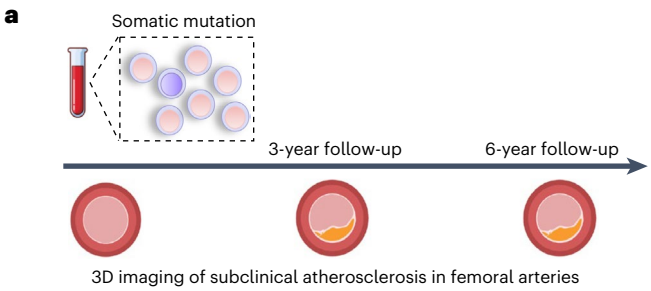
exhibited a higher risk of developing de novo atherosclerosis over subsequent years (Fig. 2a). To do this, we leveraged the availability of 3DVUS imaging data for the carotid and femoral arteries, collected at enrollment and the 3-year and 6-year follow-up imaging examinations in PESA (Fig. 2b). This imaging modality allows for the direct detection of atherosclerotic plaques, including the small plaques that are characteristic of early atherosclerosis and can be expected to develop in this timespan[42,43]. CH mutation carriers exhibited markedly higher rates of de novo femoral atherosclerosis development at the 3-year follow-up, which correlated with mutant clone size (Fig. 2c). Similar results were observed when restricting this analysis to mutations in CHIP genes or specific frequently mutated genes (Fig. 2c), or when evaluating de novo atherosclerosis development using imaging data from the 6-year follow-up (Fig. 2d).

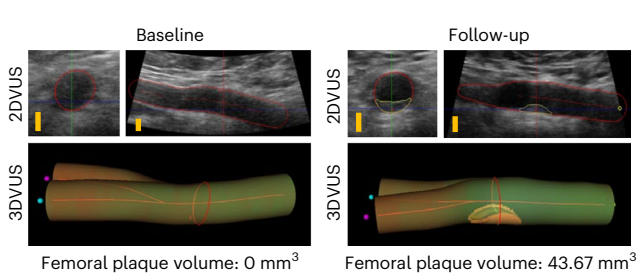
To account for the possible effects of the cumulative exposure to risk factors or the systemic burden of atherosclerosis[27–29], we next evaluated the effects of CH in statistical models adjusted for the area under the curve (AUC) of conventional CV risk factors, CACS (assessed by CT) and global atherosclerotic plaque volume (assessed by 3DVUS), in addition to age and sex (see Methods for details). Using this approach, CH was independently associated with a 1.5-fold higher risk of developing de novo femoral atherosclerosis at the 3-year follow-up among participants who lacked detectable atherosclerosis in this vascular region at baseline (Fig. 2e). The magnitude of this effect was dependent on mutant clone size, with smaller clones (VAF 0.2–2%) conferring a 1.4-fold higher risk of developing de novo femoral atherosclerosis, which did not reach statistical significance ($P = 0.072$), and larger clones leading to a 2.1-fold higher risk, an effect comparable to that of traditional modifiable risk factors such as dyslipidemia or smoking (Extended Data Fig. 3a). These associations were independent of blood cell counts (Extended Data Fig. 3b). The effect of CH was primarily driven by mutations in genes related to myeloid CH or CHIP, which were associated with 1.6-fold greater risk of de novo femoral atherosclerosis (Fig. 2e and Extended Data Fig. 3c,d). In gene-specific analyses, CH driven by mutations in *DNMT3A*, *TET2* and other genes were all associated with higher rates of de novo femoral atherosclerosis (Fig. 2e). The effect of CH mutations remained consistent after correction for multiple hypothesis testing (Supplementary Table 6). Similar, albeit milder, results were also obtained when evaluating femoral atherosclerosis development at the 6-year follow-up in PESA (Fig. 2f). No association was observed between CH and atherosclerosis development in the carotid artery (Supplementary Fig. 1), a region where atherosclerosis seems to develop at later ages based on previous analyses in PESA and other cohorts with multiterritorial vascular imaging data[42,44,45].

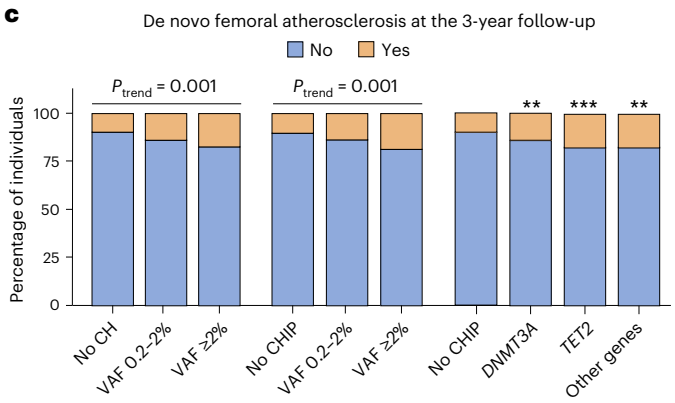
No effect of subclinical atherosclerosis on CH dynamics

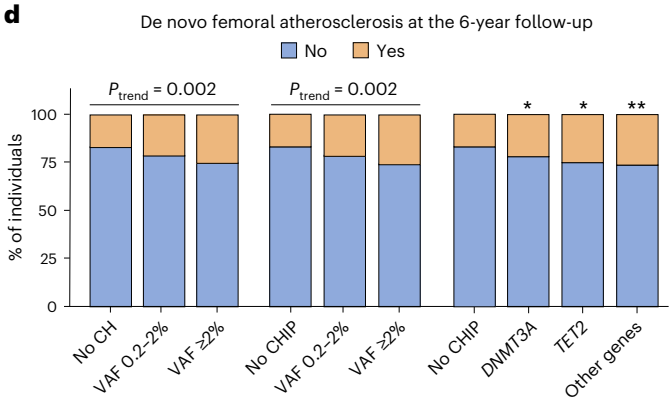
Finally, we determined the longitudinal dynamics of CH in healthy middle-aged participants in PESA to examine its potential modulation by the presence or extent of atherosclerosis. To do this, we sequenced blood samples collected at baseline and ~6 years after enrollment from 718 CH mutation carriers (Fig. 3a). We then used the serially determined VAF measurements as a surrogate for clone size to calculate the annual relative expansion rates (AER) of CH mutations in blood (see details in Methods). Among 602 CH-related mutations that were detected at both time points with comparable sequencing depth and quality (Supplementary Table 7), the median AER was 6.5% per year. CH-related mutations in *DNMT3A* (381 variants) and *TET2* (80 variants) expanded at median AER of 7.0% per year and 7.6% per year, respectively (Fig. 3b). Overall, mutations in epigenetic regulatory genes showed numerically slower expansion rates than genes encoding splicing regulators or involved in the DNA damage response (Supplementary Fig. 2), consistent with previous findings[46–48]. As a control, we also quantified the expansion of 271 nondeleterious somatic variants in CH-related genes identified at both time points in 223 individuals not carrying CH driver mutations. As expected, these neutral mutations generally

https://doi.org/10.1038/s41591-024-03213-1





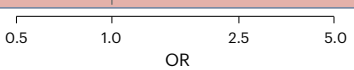




e

De novo femoral atherosclerosis development at the 3-year follow-up

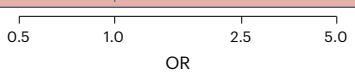
	Yes/No	OR (95% CI)	*P* value
No CH (Ref.)	178/1,615		
CH	81/473	1.53 (1.12–2.07)	0.007
CH (VAF 0.2–2%)	57/360	1.37 (0.97–1.93)	0.072
CH (VAF ≥2%)	24/113	2.06 (1.22–3.37)	0.005
No CHIP (Ref.)	183/1,651		
CHIP mutations	76/437	1.56 (1.13–2.13)	0.006
CHIP mutations (VAF 0.2–2%)	52/332	1.37 (0.95–1.95)	0.081
CHIP mutations (VAF ≥2%)	24/105	2.16 (1.27–3.55)	0.003
No CHIP (Ref.)	183/1,651		
DNMT3A	50/297	1.52 (1.04–2.18)	0.025
TET2	31/222	1.85 (0.98–3.29)	0.045
Other mutations	19/75	2.05 (1.24–3.28)	0.004



f

De novo femoral atherosclerosis development at the 6-year follow-up

	Yes/No	OR (95% CI)	*P* value
No CH (Ref.)	283/1,398		
CH	119/414	1.39 (1.06–1.81)	0.015
CH (VAF 0.2–2%)	85/317	1.29 (0.95–1.73)	0.093
CH (VAF ≥2%)	34/97	1.70 (1.07–2.65)	0.021
No CHIP (Ref.)	290/1,431		
CHIP mutations	112/381	1.45 (1.10–1.90)	0.008
CHIP mutations (VAF 0.2–2%)	79/291	1.35 (0.99–1.83)	0.053
CHIP mutations (VAF ≥2%)	33/90	1.72 (1.07–2.70)	0.021
No CHIP (Ref.)	290/1,431		
DNMT3A	73/259	1.39 (1.00–1.91)	0.044
TET2	47/196	1.73 (1.01–2.87)	0.037
Other mutations	26/63	1.73 (1.11–2.63)	0.013



did not show clear evidence of expansion within the timeframe of the study (median AER: 1.4% per year) (Fig. 3b).

The AER of CH-related mutations was comparable across quartiles of age, consistent with previous findings (Fig. 3c)[48]. In contrast, a trend to faster expansion was observed in men (Fig. 3d) and among variants with lower baseline VAF (Fig. 3e). Importantly, the expansion of CH variants was not affected by the presence or the extent of subclinical atherosclerosis at baseline, assessed through various noninvasive imaging modalities (Fig. 3f–i and Table 2). Similarly, the cumulative burden of atherosclerosis over the timeframe of this longitudinal analysis, estimated as the AUC of several vascular imaging parameters, had no significant effect on the dynamics of CH (Table 2). In sensitivity analyses, we obtained comparable results when defining mutant cell expansion with two alternative mathematical approaches used in other studies of CH dynamics, namely the logarithm of the ratio of VAFs at both time points[49] and the exponential growth rate[50] (Table 2). Similarly, neither high-sensitivity C-reactive protein (hsCRP), a biomarker of systemic inflammation, nor several conventional cardiovascular risk factors, such as dyslipidemia, diabetes, smoking or obesity, were associated with significant changes in AER or other indicators of clonal

Fig. 2 | Longitudinal assessment of the effects of CH on de novo femoral atherosclerosis development. We investigated the association between CH at baseline and de novo development of femoral atherosclerosis ~3 years and ~6 years after enrollment among PESA participants who initially lacked detectable atherosclerosis in this region. 3DVUS imaging was used to determine the presence of femoral atherosclerosis. **a**, Summary of study design. **b**, Representative images from femoral atherosclerosis burden, assessed by 3DVUS, in an individual exhibiting de novo femoral atherosclerosis development (that is, absence of detectable atherosclerosis at the baseline evaluation (left) and plaque development at follow-up (right); scale bar, 5 mm). **c**,**d**, Incidence of de novo femoral atherosclerosis development at the 3-year follow-up (**c**, $n = 2,347$) or the 6-year follow-up (**d**, $n = 2,214$) in individuals free of CH (no CH) and in individuals exhibiting CH with VAF 0.2–2% or ≥2%. Incidence of de novo femoral atherosclerosis is also shown for myeloid CH (CHIP) or CH driven by specific CHIP genes. Statistical significance in the analyses of CH and CHIP was evaluated using univariate logistic regression models (*P* for trends are shown). In gene-specific analyses, statistical significance was examined through proportion tests relative to the non-CHIP carriers group (*$P < 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$). **e**,**f**, The association between CH or CHIP and de novo femoral atherosclerosis development at the 3-year follow-up (**e**, $n = 2,347$) and 6-year follow-up (**f**, $n = 2,214$) based on multivariate logistic regression analyses. Statistical models were adjusted for age, sex, smoking, lipid-lowering treatment and the AUC of SBP, fasting glucose, LDL-C, BMI, CACS and global atherosclerotic plaque volume assessed by 3DVUS imaging; the bars indicate 95% confidence intervals centered in the mean value (square).

expansion (Fig. 3j,k and Extended Data Table 5). Hypertension was associated with higher AER ($P = 0.048$; Fig. 3j), but this relationship was not statistically significant when evaluating other indicators of CH dynamics or the effect of baseline or cumulative systolic blood pressure (SBP) (Extended Data Table 5).

Discussion

CH has emerged in recent years as a potential shared driver of blood cancer and several nonhematological conditions, most notably atherosclerotic CVD[2,3]. However, the accumulating reports of positive associations between CH and a diverse range of conditions and outcomes[4,6,7,9–23,30,32] also lend support to the possibility that CH is simply a marker of a shared pathophysiological feature of these conditions, rather than a causal factor. The association between CH and atherosclerosis, in particular, has become a matter of active debate[27,29], hindering the translation of research findings into new strategies for the prevention of CVD in CH mutation carriers. In this context, our longitudinal assessment of the interplay between CH, atherosclerosis and related traits provides important new insights into the relationship between CH and atherosclerosis.

Most importantly, our findings shed light on the directionality of the association between CH and atherosclerosis. Previous research in both humans and mice has suggested that certain CH mutations directly contribute to the development of atherosclerotic CVD[4,9,13,25,26]. In contrast, recent mathematical models based on human and mouse data have proposed that the association between CH and CVD may instead reflect reverse causality, whereby CH would simply be a direct consequence of increased HSC proliferation in individuals with atherosclerosis, resulting in accelerated kinetics of mutant cell expansion[27,29]. This reverse causality hypothesis is also applicable to the association between CH and other conditions, such as diabetes[4,23,41], known to increase HSC proliferation[51]. The direction of these associations remains uncertain, partly due to the fact that most of them were identified through post hoc analyses of cross-sectional sequencing datasets[4,7,9–18,30–32], which provide only a static 'snapshot' of CH at a given time and, therefore, very limited insight into the dynamics of this condition and its association with disorders that develop subclinically over the course of years. In this context, our analysis of longitudinal sequencing and imaging data provides novel insights into the relationship between CH and atherosclerosis, demonstrating that: (1) carrying CH mutations at a given time confers a greater risk of developing femoral atherosclerosis over subsequent years, and (2) the presence or extent of subclinical atherosclerosis has no significant effect on the expansion of mutant hematopoietic clones. These findings argue strongly against the possibility that the association between CH and atherosclerosis is due to reverse causality, thus providing an essential prior step for the development of clinical trials that evaluate the efficacy of personalized preventive care strategies tailored to the effects of specific CH mutations. Our findings, however, do not rule out that certain cardiometabolic conditions may accelerate CH dynamics to some extent. Indeed, we found that hypertensive individuals exhibit a non-statistically significant trend to faster expansion of mutant cells, an observation that demands further exploration given the emerging evidence suggesting an impact of hypertension on the bone marrow niche and hematopoiesis[52–54]. In addition, we cannot rule out the possibility that clinically overt CVD (for example, acute ischemic heart disease) modifies the dynamics of CH. Although this possibility would not affect the primary conclusions of our study, it deserves further investigation as it may help understand the regulation and effects of CH in high-risk CVD patients. Overall, further studies are warranted to dissect the factors that determine the dynamics of CH and explore the underlying regulatory mechanisms.

Our study also provides a large high-sensitivity assessment of the prevalence of CH and the magnitude of its connection to atherosclerosis. Previous human genetic evidence linking CH to atherosclerotic CVD was mainly obtained by repurposing WES datasets[4,7,9,13,16,39], which were not intended initially to investigate somatic variants and have limited sensitivity to detect CH due to their modest sequencing depth. A previous report suggested that the typical depth of WES analyses only provides robust ability to identify mutant clones with VAF >5%, (that is, carriers of >10% mutant blood cells, if carrying monoallelic mutations)[34], which is consistent with our estimations based on downsampling of sequencing depth in our study population. Furthermore, slight differences in variant interpretation and filtering parameters can be a source of major variability when using WES data to detect CH, as illustrated by recent studies by independent groups, who reached substantially different conclusions about the prevalence of CH and the strength of its association with atherosclerotic CVD despite analyzing WES data

Fig. 3 | Longitudinal assessment of the dynamics of CH in middle-aged individuals. We monitored the expansion of 602 CH-related mutations through serial sequencing of blood samples from 494 individuals at baseline and ~6 years after enrollment. Then, we investigated the association between subclinical atherosclerosis burden or related traits at baseline and the AER of the mutant hematopoietic clones. **a**, Summary of study design. **b**, The AER of mutations in any CH gene ($n = 602$) or specifically in *DNMT3A* ($n = 381$), *TET2* ($n = 80$) or other genes ($n = 141$), compared with nondeleterious variants in CH-related genes ($n = 271$). Statistical significance was determined by two-sided Mann–Whitney–Wilcoxon tests (**$P \leq 0.01$, ****$P \leq 0.0001$). **c**, The AER of CH-related mutations stratified by age quartiles ($n = 602$). **d**, AER of CH-related mutations stratified by sex ($n = 602$). **e**, The correlation between AER of CH-related mutations and baseline variant allele frequency (VAF). The *P* value for two-sided Pearson correlation is indicated. **f**–**i**, The AER of CH-related mutations in individuals with no detectable atherosclerosis and across tertiles of plaque burden, based on CACS (**f**, $n = 600$), global plaque volume (**g**, $n = 578$), carotid plaque volume (**h**, $n = 597$) or femoral plaque volume (**i**, $n = 581$) measured by 3DVUS. **j**,**k**, The AER of CH-related mutations stratified by conventional modifiable risk factors (**j**, $n = 602$ for all variables, except for obesity, $n = 601$) and quartiles of hsCRP (**k**, $n = 600$). Statistical significance in **c**, **d** and **f**–**k** was determined by mixed-effects models adjusted for baseline VAF (**d**) or sex and baseline VAF (**c** and **f**–**k**). The boxes in **b**–**d** and **f**–**k** represent the 25th (Q1), 50th (median) and 75th (Q3) percentiles of the data. The whiskers represent Q1 − 1.5 × IQR at the minimum and Q3 + 1.5 × IQR at the maximum.

https://doi.org/10.1038/s41591-024-03213-1

from the same cohort[7,8,16]. Here, through high-sensitivity sequencing of almost 3,700 seemingly healthy individuals, we demonstrate that CH mutations are prevalent in middle-aged individuals, being detectable in ~25% of healthy individuals between 40 and 55 years of age, and in ~34% of those aged 50–55. Even when applying the commonly used 2% VAF threshold for CH detection, we found that 6% of the study population (age range 40–55 years old) and 9% of those aged 50–55 exhibit CH, indicating an approximately three times higher prevalence of CH in middle-aged individuals than previously reported[4,6–8,34]. Thus, our findings suggest that CH contributes to atherosclerosis development in a much greater proportion of the population than previously believed. Importantly, the vast majority of mutant clones in our study population were small, with 79% of the mutations exhibiting a VAF <2% and 96% of them exhibiting a VAF <10% at baseline. Yet, carrying

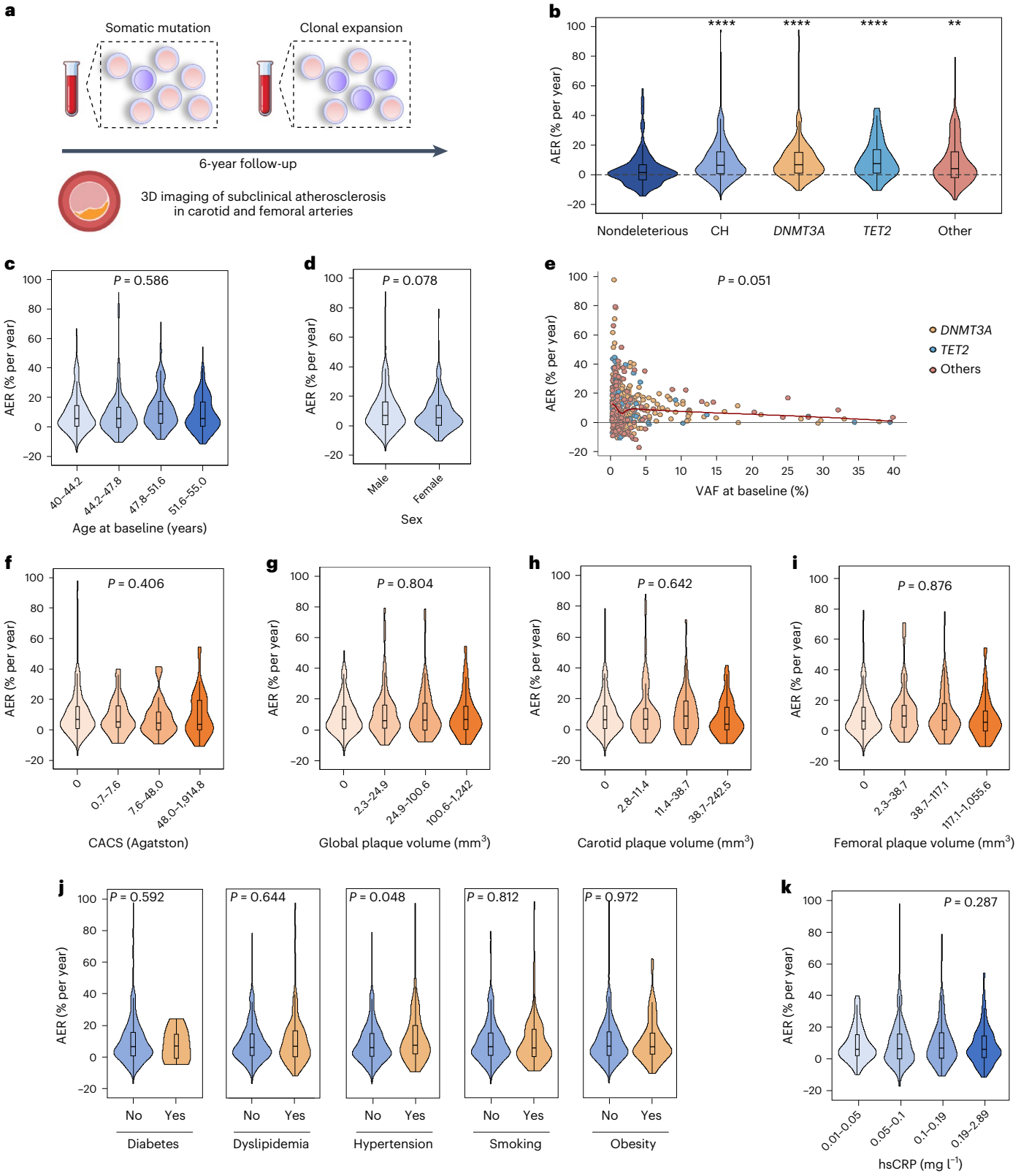


Table 2 | No effect of subclinical atherosclerosis on the expansion of mutant hematopoietic clones

	AER			log (VAF ratio)/time			Exponential growth rate		
	Estimate	Standard error	*P* value	Estimate	Standard error	*P* value	Estimate	Standard error	*P* value
Baseline presence of atherosclerosis									
CACS >0	−1.907	1.474	0.196	−1.161	0.893	0.194	−1.266	0.94	0.179
Global plaque volume >0	0.718	1.183	0.544	0.034	0.734	0.963	0.075	0.772	0.923
Carotid plaque volume >0	0.526	1.321	0.691	−0.071	0.798	0.929	−0.016	0.841	0.985
Femoral plaque volume >0	1.012	1.287	0.432	0.155	0.803	0.847	0.205	0.843	0.808
Baseline extent of atherosclerosis									
CACS (Agatston)	0.003	0.005	0.524	0.003	0.003	0.376	0.003	0.003	0.392
Global plaque volume (mm^3)	−0.007	0.007	0.304	-0.005	0.004	0.2	−0.005	0.004	0.214
Carotid plaque volume (mm^3)	−0.019	0.025	0.46	−0.005	0.014	0.742	−0.006	0.015	0.695
Femoral plaque volume (mm^3)	−0.013	0.008	0.127	-0.009	0.005	0.061	−0.009	0.005	0.069
Cumulative atherosclerosis burden over 6 years									
CACS (AUC)	1.816	10.08	0.857	3.044	6.121	0.619	2.867	6.444	0.657
Global plaque volume (AUC)	−7.095	6.955	0.308	−4.934	4.305	0.252	−5.38	4.53	0.236
Carotid plaque volume (AUC)	−9.817	14.13	0.488	−4.569	8.526	0.592	−5.268	8.99	0.558
Femoral plaque volume (AUC)	−11.047	8.223	0.18	−9.13	5.111	0.075	−9.53	5.371	0.077

We monitored the expansion of 602 CH-related mutations through serial sequencing of blood samples from 494 individuals at baseline and ~6 years after enrollment. We then investigated the association between baseline or cumulative atherosclerosis burden and different metrics of clonal expansion rates through mixed-effects models. The presence of subclinical atherosclerosis at baseline was defined as a categorical variable on the basis of CACS or 3DVUS imaging. The extent of subclinical atherosclerosis at baseline in those with detectable atherosclerosis in specific regions was defined as a continuous variable on the basis of CACS or 3DVUS imaging. The cumulative burden of subclinical atherosclerosis over the timeframe of this longitudinal analysis was estimated as the AUC for CACS and atherosclerotic plaque volumes, assessed by CT and 3DVUS imaging, respectively, at three different time points over 6 years.

clones in this size range was still pathophysiologically relevant, as it was associated with a higher risk of de novo femoral atherosclerosis development, particularly among individuals with clones expanded to VAF ≥2%. These findings contribute to clarify the threshold of VAF that defines pathophysiologically relevant CH, as large WES-based studies previously concluded that clones with VAF <10% had minimal impact on subclinical atherosclerosis, atherosclerotic CVD or CVD-related mortality[6,7,9,16,39], in contrast to smaller studies using more sensitive targeted sequencing approaches[19].

Our results also provide important insight into gene-specific effects of CH. While the effect of somatic *TET2* mutations on atherosclerotic CVD is supported by several sequencing studies[9,13,16], the effect of *DNMT3A* mutations has remained controversial, with previous WES-based studies reporting a minimal, almost null effect of these mutations on atherosclerotic CVD[7,13,16]. In this context, while additional studies are warranted, our gene-specific analysis reinforces the association between somatic *DNMT3A* mutations and atherosclerotic disease, consistent with experiments in mouse models[55] and recent targeted sequencing studies[56], and in contrast to WES-based analyses[7,13,16]. As *DNMT3A* mutations are the most frequent mutations linked to CH, these findings underscore the potential clinical significance of this condition.

Several limitations must be considered when interpreting our findings. First, the study population consisted entirely of Caucasian individuals residing in Madrid, Spain, which may limit the generalizability of our results to other racial and ethnic groups. Second, our findings related to the effects of CH on de novo femoral atherosclerosis cannot be directly extrapolated to other vascular beds; in this context, future studies with other imaging modalities and in other age ranges are warranted. Third, CH in this middle-aged population was markedly dominated by *DNMT3A* mutations, which limits our statistical power to assess the effects of CH driven by mutations in other genes; thus, gene-specific analyses in this work should be considered exploratory. Finally, our study was focused on examining the effects of atherosclerosis on the expansion of existing mutant hematopoietic clones; therefore, we cannot draw any conclusions regarding the potential effects of atherosclerosis or related traits on de novo somatic mutagenesis. Testing this possibility, which to our knowledge has not been postulated so far, would require more sensitive sequencing techniques, such as duplex sequencing.

In summary, our study elucidates the directionality of the association between CH and atherosclerosis in humans. These results provide critical new insight into the effects and regulation of CH, which should help translate our growing knowledge of this condition into personalized strategies for managing the risk of atherosclerotic CVD and, potentially, that of other age-related disorders.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-024-03213-1.

References

1. Mustjoki, S. & Young, N. S. Somatic mutations in "benign" disease. *N. Engl. J. Med.* **384**, 2039–2052 (2021).
2. Jaiswal, S. & Ebert, B. L. Clonal hematopoiesis in human aging and disease. *Science* **366**, 6465 (2019).
3. Tall, A. R. & Fuster, J. J. Clonal hematopoiesis in cardiovascular disease and therapeutic implications. *Nat. Cardiovasc. Res.* **1**, 116–124 (2022).
4. Jaiswal, S. et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).
5. Genovese, G. et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).
6. Kar, S. P. et al. Genome-wide analyses of 200,453 individuals yield new insights into the causes and consequences of clonal hematopoiesis. *Nat. Genet.* **54**, 1155–1166 (2022).
7. Kessler, M. D. et al. Common and rare variant associations with clonal haematopoiesis phenotypes. *Nature* **612**, 301–309 (2022).

8. Vlasschaert, C. et al. A practical approach to curate clonal hematopoiesis of indeterminate potential in human genetic datasets. *Blood* **141**, 2214–2223 (2023).
9. Jaiswal, S. et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *N. Engl. J. Med.* **377**, 111–121 (2017).
10. Miller, P. G. et al. Association of clonal hematopoiesis with chronic obstructive pulmonary disease. *Blood* **139**, 357–368 (2022).
11. Agrawal, M. et al. TET2-mutant clonal hematopoiesis and risk of gout. *Blood* **140**, 1094–1103 (2022).
12. Kim, P. G. et al. Dnmt3a-mutated clonal hematopoiesis promotes osteoporosis. *J. Exp. Med.* **218**, e20211872 (2021).
13. Zekavat, S. M. et al. TP53-mediated clonal hematopoiesis confers increased risk for incident atherosclerotic disease. *Nat. Cardiovasc. Res.* **2**, 144–158 (2023).
14. Shi, C. et al. Clonal haematopoiesis of indeterminate potential: associations with heart failure incidence, clinical parameters and biomarkers. *Eur. J. Heart Fail.* **25**, 4–13 (2023).
15. Yu, B. et al. Supplemental association of clonal hematopoiesis with incident heart failure. *J. Am. Coll. Cardiol.* **78**, 42–52 (2021).
16. Vlasschaert, C., Heimlich, J. B., Rauh, M. J., Natarajan, P. & Bick, A. G. Interleukin-6 receptor polymorphism attenuates clonal hematopoiesis-mediated coronary artery disease risk among 451 180 individuals in the UK Biobank. *Circulation* **147**, 358–360 (2023).
17. Kestenbaum, B. et al. Clonal hematopoiesis of indeterminate potential and kidney function decline in the general population. *Am. J. Kidney Dis.* **81**, 329–335 (2023).
18. Wong, W. J. et al. Clonal haematopoiesis and risk of chronic liver disease. *Nature* **616**, 747–754 (2023).
19. Zhao, K. et al. Somatic and germline variants and coronary heart disease in a Chinese population. *JAMA Cardiol.* **9**, 233–242 (2024).
20. Vlasschaert, C. et al. Clonal hematopoiesis of indeterminate potential is associated with acute kidney injury. *Nat. Med.* **30**, 810–817 (2024).
21. Schuermans, A. et al. Clonal haematopoiesis of indeterminate potential predicts incident cardiac arrhythmias. *Eur. Heart J.* **45**, 791–805 (2024).
22. Ahn, H. J. et al. Clonal haematopoiesis of indeterminate potential and atrial fibrillation: an East Asian cohort study. *Eur. Heart J.* **45**, 778–790 (2024).
23. Tobias, D. K. et al. Clonal hematopoiesis of indeterminate potential (CHIP) and incident type 2 diabetes risk. *Diabetes Care* **46**, 1978–1985 (2023).
24. Libby, P. et al. Clonal hematopoiesis: crossroads of aging, cardiovascular disease, and cancer: JACC review topic of the week. *J. Am. Coll. Cardiol.* **74**, 567–577 (2019).
25. Fuster, J. J. et al. Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science* **355**, 842–847 (2017).
26. Fidler, T. P. et al. The AIM2 inflammasome exacerbates atherosclerosis in clonal haematopoiesis. *Nature* **592**, 296–301 (2021).
27. Heyde, A. et al. Increased stem cell proliferation in atherosclerosis accelerates clonal hematopoiesis. *Cell* **184**, 1348–1361 e1322 (2021).
28. McAlpine, C. S. et al. Sleep exerts lasting effects on hematopoietic stem cell function and diversity. *J. Exp. Med.* **219**, e20220081 (2022).
29. Sanchez-Cabo, F. & Fuster, J. J. Clonal haematopoiesis and atherosclerosis: a chicken or egg question? *Nat. Rev. Cardiol.* **18**, 463–464 (2021).
30. Honigberg, M. C. et al. Premature menopause, clonal hematopoiesis, and coronary artery disease in postmenopausal women. *Circulation* **143**, 410–423 (2021).
31. Bohme, M. et al. Impact of clonal hematopoiesis in patients with cardiogenic shock complicating acute myocardial infarction. *J. Am. Coll. Cardiol.* **80**, 1545–1556 (2022).
32. Dharan, N. J. et al. HIV is associated with an increased risk of age-related clonal hematopoiesis among older adults. *Nat. Med.* **27**, 1006–1011 (2021).
33. Ibanez, B. et al. Progression of Early Subclinical Atherosclerosis (PESA) study: JACC Focus Seminar 7/8. *J. Am. Coll. Cardiol.* **78**, 156–179 (2021).
34. Bick, A. G. et al. Inherited causes of clonal haematopoiesis in 97,691 whole genomes. *Nature* **586**, 763–768 (2020).
35. Diez-Diez, M. et al. Clonal hematopoiesis is not prevalent in Hutchinson–Gilford progeria syndrome. *Geroscience* **45**, 1231–1236 (2023).
36. Pascual-Figal, D. A. et al. Clonal hematopoiesis and risk of progression of heart failure with reduced left ventricular ejection fraction. *J. Am. Coll. Cardiol.* **77**, 1747–1759 (2021).
37. Niroula, A. et al. Distinction of lymphoid and myeloid clonal hematopoiesis. *Nat. Med.* **27**, 1921–1927 (2021).
38. Buscarlet, M. et al. DNMT3A and TET2 dominate clonal hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. *Blood* **130**, 753–762 (2017).
39. Bick, A. G. et al. Genetic interleukin 6 signaling deficiency attenuates cardiovascular risk in clonal hematopoiesis. *Circulation* **141**, 124–131 (2020).
40. Kamphuis, P. et al. Sex differences in the spectrum of clonal hematopoiesis. *Hemasphere* **7**, e832 (2023).
41. Kim, M. J. et al. Clonal hematopoiesis as a novel risk factor for type 2 diabetes mellitus in patients with hypercholesterolemia. *Front. Public Health* **11**, 1181879 (2023).
42. Lopez-Melgar, B. et al. Subclinical atherosclerosis burden by 3D ultrasound in mid-life: the PESA study. *J. Am. Coll. Cardiol.* **70**, 301–313 (2017).
43. Lopez-Melgar, B. et al. Accurate quantification of atherosclerotic plaque volume by 3D vascular ultrasound using the volumetric linear array method. *Atherosclerosis* **248**, 230–237 (2016).
44. Fernandez-Friera, L. et al. Prevalence, vascular distribution, and multiterritorial extent of subclinical atherosclerosis in a middle-aged cohort: the PESA (Progression of Early Subclinical Atherosclerosis) study. *Circulation* **131**, 2104–2113 (2015).
45. Laclaustra, M. et al. Femoral and carotid subclinical atherosclerosis association with risk factors and coronary calcium: the AWHS study. *J. Am. Coll. Cardiol.* **67**, 1263–1274 (2016).
46. Robertson, N. A. et al. Longitudinal dynamics of clonal hematopoiesis identifies gene-specific fitness effects. *Nat. Med.* **28**, 1439–1446 (2022).
47. Fabre, M. A. et al. The longitudinal dynamics and natural history of clonal haematopoiesis. *Nature* **606**, 335–342 (2022).
48. van Zeventer, I. A. et al. Evolutionary landscape of clonal hematopoiesis in 3,359 individuals from the general population. *Cancer Cell* **41**, 1017–1031 e1014 (2023).
49. Bolton, K. L. et al. Cancer therapy shapes the fitness landscape of clonal hematopoiesis. *Nat. Genet.* **52**, 1219–1226 (2020).
50. Mack, T. et al. Cost-effective and scalable clonal hematopoiesis assay provides insight into clonal dynamics. *J. Mol. Diagn.* **26**, 563–573 (2024).
51. Hoyer, F. F. et al. Bone marrow endothelial cells regulate myelopoiesis in diabetes mellitus. *Circulation* **142**, 244–258 (2020).
52. Devesa, A. et al. Bone marrow activation in response to metabolic syndrome and early atherosclerosis. *Eur. Heart J.* **43**, 1809–1828 (2022).
53. Al-Sharea, A. et al. Chronic sympathetic driven hypertension promotes atherosclerosis by enhancing hematopoiesis. *Haematologica* **104**, 456–467 (2019).
54. Rohde, D. et al. Bone marrow endothelial dysfunction promotes myeloid cell expansion in cardiovascular disease. *Nat. Cardiovasc. Res.* **1**, 28–44 (2022).

55. Rauch, P. J. et al. Loss-of-function mutations in Dnmt3a and Tet2 lead to accelerated atherosclerosis and concordant macrophage phenotypes. *Nat. Cardiovasc. Res.* **2**, 805–818 (2023).
56. Wang, S. et al. Prevalence and prognostic significance of DNMT3A- and TET2- clonal haematopoiesis-driver mutations in patients presenting with ST-segment elevation myocardial infarction. *EBioMedicine* **78**, 103964 (2022).

Miriam Díez-Díez[1,10], Beatriz L. Ramos-Neble[1,10], Jorge de la Barrera[1], J. C. Silla-Castro[1], Ana Quintas[1], Enrique Vázquez[1], M. Ascensión Rey-Martín[1], Benedetta Izzi[1], Lucía Sánchez-García[1], Inés García-Lunar[1,2,3], Guiomar Mendieta[1,4,5], Virginia Mass[1], Nuria Gómez-López[1], Cristina Espadas[1], Gema González[1], Antonio J. Quesada[1], Ana García-Álvarez[1,2,4,5,6], Antonio Fernández-Ortiz[1,2,7], Enrique Lara-Pezzi[1,2], Ana Dopazo[1,2], Fátima Sánchez-Cabo[1,2], Borja Ibáñez[1,2,8], Vicente Andrés[1,2], Valentín Fuster[1,9] ✉ & José J. Fuster[1,2] ✉

[1]Centro Nacional de Investigaciones Cardiovasculares, Madrid, Spain. [2]CIBER de Enfermedades Cardiovasculares, Madrid, Spain. [3]Cardiology Department, University Hospital La Moraleja, Madrid, Spain. [4]Servicio de Cardiología, Institut Clínic Cardiovascular, Hospital Clínic de Barcelona, Barcelona, Spain. [5]Institut d'Investigacions Biomèdiques August Pi I Sunyer, Barcelona, Spain. [6]Universitat de Barcelona, Barcelona, Spain. [7]Hospital Clínico San Carlos, Universidad Complutense, IdISSC, Madrid, Spain. [8]Cardiology Department, IIS-Fundación Jiménez Díaz University Hospital, Madrid, Spain. [9]Icahn School of Medicine at Mount Sinai, New York, NY, USA. [10]These authors contributed equally: Miriam Díez-Díez, Beatriz L. Ramos-Neble. ✉e-mail: vfuster@cnic.es; jjfuster@cnic.es

Methods

Study population

The study population included 3,692 participants in the PESA study (NCT01410318) with available blood DNA samples, multimodal vascular imaging data at multiple time points and written informed consent for DNA sequencing analyses. PESA is an ongoing observational prospective cohort study that is examining several imaging, biological and behavioral parameters potentially related to the presence and progression of early subclinical atherosclerosis in 4,184 apparently healthy middle-aged individuals. PESA participants were between 40 and 55 years old at enrollment (between June 2010 and February 2014), and have undergone three study visits so far (baseline and ~3- and ~6-year follow-up). All study visits included a clinical interview, physical examination, collection of blood samples and assessment of subclinical atherosclerosis by noninvasive imaging techniques, among other measurements[33]. Exclusion criteria included known CVD, cancer or immunological disorders, morbid obesity, chronic kidney disease, the presence of any disease expected to decrease life expectancy, and any condition that could affect adherence to the study procedures. The study protocol was approved by the Ethics Committee of Instituto de Salud Carlos III (CEI PI_52/2019), and all participants provided written informed consent.

Assessment of burden and progression of subclinical atherosclerosis and related traits

Atherosclerosis burden in the study population was assessed by noninvasive vascular imaging tests, in accordance with standard protocols in PESA, including 3DVUS-based assessment of the femoral and carotid arteries and noncontrast CT to determine CACS[33,42–44]. Individual-level 3DVUS data were collected using VPQ software v13 (Philips Healthcare) and CM2020 software (Philips Research Analysis). Atherosclerosis by 3DVUS was analyzed separately in carotid and femoral arteries as well as in an integrated manner (that is, global atherosclerotic plaque volume). Presence of subclinical atherosclerosis at baseline was defined as CACS >0 or plaque volume by 3DVUS >0 mm^3 (categorical variables). The extent of subclinical atherosclerosis at baseline was estimated on the basis of CACS and plaque volumes, either as continuous variables or categorized as 0 and tertiles. Representative 3DVUS images across tertiles of atherosclerotic plaque volume can be found in Extended Data Fig. 4. In longitudinal analyses, de novo atherosclerosis development was defined as transitioning from absence of detectable atherosclerosis by 3DVUS at baseline to detectable atherosclerosis with plaque volume >0 mm^3 at follow-up. Limited sensitivity and statistical power precluded an accurate examination of the effects of CH on the progression of prevalent atherosclerotic plaques. Modifiable risk factors and related traits were defined as continuous or categorical variables and evaluated on the basis of analyses of fasting blood samples and questionnaires. Diabetes was defined as exhibiting plasma fasting glucose ≥126 mg dl^{-1} or being treated with insulin or hypoglycemic medication. Hypertension was defined as exhibiting SBP ≥140 mmHg or diastolic blood pressure (DBP) ≥90 mmHg, or being treated with antihypertensive medication. Dyslipidemia was defined as exhibiting total cholesterol ≥240 mg dl^{-1}, low-density lipoprotein cholesterol (LDL-C) ≥160 mg dl^{-1} or high-density lipoprotein cholesterol (HDL-C) <40 mg dl^{-1}, or being treated with lipid-lowering drugs. Obesity was defined as exhibiting a body mass index (BMI) ≥30 kg m^{-2}. Low plasma HDL-C was defined as exhibiting plasma HDL-C <40 mg dl^{-1} in males or <50 mg dl^{-1} in females. Hematological analyses included counts of leukocytes and main leukocyte subsets (neutrophils, lymphocytes, monocytes, eosinophils and basophils), red blood cell count, hemoglobin, hematocrit, red blood cell distribution width and platelet counts.

High-sensitivity targeted sequencing to identify CH-related somatic mutations

Sample preparation and high-sensitivity DNA sequencing, as well as somatic variant calling and annotation, were performed as previously described[35,36]. In brief, a custom gene panel was designed to identify somatic mutations in 54 well-established CH-related genes (Extended Data Fig. 1 and Supplementary Table 1), and unique dual indexes and unique molecular identifiers were used to detect polymerase chain reaction and sequencing errors and, therefore, to improve identification of small mutant clones. After confirming the quality of blood DNA, DNA libraries were generated (KAPA Hyper plus kit, Roche) and genes of interest were captured using xGen Hybridization Capture reagents (Integrated DNA Technologies). Then, DNA libraries were paired-end sequenced on a HiSeq 4000 or a NovaSeq 6000 sequencing system (Illumina). Raw sequencing reads were mapped to human genome build GRCh38, and somatic variants were identified by using Genome Analysis Toolkit (GATK) Mutect2. We removed common sequencing artifacts according to several Mutect2 filters (multiallelic somatic calls and variants with deficient mapping or base quality) as well as those flagged by the position filter. Germline mutations were excluded, and we only considered variants with 'high' or 'moderate' impact according to Variant Effect Predictor, reference read depth ≥300, alternative depth ≥4 and evidence of the variant on both forward and reverse strands (F1R2 and F2R1 ≥2). We also filtered out variants preceded or followed by homopolymer regions, indels with ≥3 consecutive repeated nucleotides within the insertion or deletion, and those SNVs located in a ≥5-nucleotide-long homopolymer, except if found at multiple sequencing time points in the same individual. Candidate CH driver mutations were then identified on the basis of prespecified inclusion criteria (Supplementary Table 1), the presence of ≥3 times in the Catalogue of Somatic Mutations in Cancer (COSMIC, https://cancer.sanger.ac.uk/cosmic) in hematopoietic samples from at least three different studies, in silico pathogenicity predictors and previous publications using a similar deep sequencing strategy. A specific approach was required to identify mutations in *U2AF1*[57] because of an erroneous duplication in the region of this gene in the GRCh38reference genome[58].

We set a minimum VAF threshold of 0.2% to identify CH mutations, as our sequencing depth allowed us to detect CH variants at this VAF with a sensitivity greater than 90% (Extended Data Fig. 1b). In brief, we used a binomial distribution to estimate the sensitivity for the detection of somatic variants with a certain VAF at different sequencing depths, requiring a minimum alternative allele count of four reads. CH-related variants over the 0.2% VAF threshold were significantly associated with age (Supplementary Table 8), an essential feature of true CH variants[8].

Assessment of expansion rates of mutant hematopoietic clones

Blood samples collected ~6 years after enrollment from 718 CH mutation carriers were sequenced as described above. Among the 931 CH-related variants with VAF ≥0.2% identified in these individuals at baseline, 261 variants were not detected at follow-up. These variants were excluded from further analysis, as we were unable to ascertain if the apparent disappearance of these mutant clones is due to biological or technical reasons. We also excluded 68 additional variants for which sequencing depths differed substantially between both time points (ratio of baseline depth to follow-up depth exceeding the 90% percentile of the distribution). After applying this filtering strategy, 602 CH-related mutations in 494 individuals were detected at both time points supported by reads in both the forward and reverse strand (Supplementary Table 7). The AER of each of these mutations was calculated as the relative change in VAF over the follow-up period in years $(\Delta VAF/VAF_{baseline})/time \times 100$. As a sensitivity analysis, we also calculated two additional indicators of the expansion rate of mutant hematopoietic clones based on previous literature: the logarithmic ratio of VAFs, that is $\log(VAF/VAF_{baseline})/time \times 100$ (ref. 49), and the exponential growth rate, that is, $VAF = VAF_{baseline} \times (clonal\ growth + 1)^{time}$; clonal growth = $((VAF/VAF_{baseline})^{(1/(time))} - 1) \times 100$ (ref. 50). Time refers to the follow-up period measured in years. These three metrics showed a

perfect positive correlation with each other, with Spearman's rank correlation coefficients equal to 1.00 (Supplementary Fig. 3).

As a control, blood samples from 223 individuals not carrying CH-related mutations at baseline were also sequenced ~6 years after enrollment to estimate the expansion of nondeleterious somatic variants in CH-related genes, as a surrogate of the expansion related to stochastic neutral drift. We considered nondeleterious somatic variants those with impact 'low' (unlikely to change protein behavior, for example, synonymous variants) or 'modifier' (typically noncoding variants). Likely germline variants and common sequencing artifacts were filtered out, and only variants supported by reads on both the forward and reverse strands at both time points were included in the analysis, consistent with our approach to investigate the expansion of CH driver variants.

Statistical analysis

All statistical analyses were performed using RStudio (v2022.07.2 + 576) with R version 4.2.2, considering a statistical significance level of 0.05. Adjustments for multiple hypothesis testing were not performed unless otherwise indicated. Variables were expressed as frequency (%), mean and standard deviation, or median and interquartile range (IQR), as appropriate. CH was investigated as a composite of mutations in any CH-related gene, unless otherwise stated. Specific analyses evaluated the effects of mutations in *DNMT3A*, *TET2* and other genes separately. Differences in baseline characteristics between CH mutation carriers and individuals with no detectable CH mutation were evaluated through independent two-tailed *t*-tests for continuous variables exhibiting a normal distribution, Wilcoxon rank tests for continuous variables with a nonnormal distribution and χ^2 tests for categorical variables. Associations between CH and traits of interest were evaluated using univariate or multivariate linear or logistic regression models, adjusted for age and sex or age, sex, LDL-C, SBP, BMI, fasting glucose levels, lipid-lowering treatment and smoking, as specified in figure legends. Some specific analyses were adjusted for age, sex and cardiovascular risk factors defined as categorical variables, as indicated in figure legends. Analyses of the association between CH and de novo development of atherosclerosis were adjusted by the AUCs of LDL-C, SBP, BMI and fasting glucose to consider the cumulative exposure to this cardiovascular risk factors during follow-up. These analyses were also adjusted for the AUC of CACS and global plaque volume to account for the possibility that the possible cumulative effects of systemic atherosclerosis on CH[27] confound its relationship with de novo atherosclerosis development in specific regions. In multivariate analyses, only individuals with information for all the covariates included in the model were considered. In analyses concerning specific types of CH (for example, driven by mutations in a given gene or over a certain VAF), individuals not carrying detectable CH mutations were used as the reference category. In individuals carrying more than one CH mutation, all mutations were considered separately in gene-specific analyses, unless otherwise stated. In longitudinal analyses of CH dynamics, differences in the AER of gene-specific mutations compared with the expansion of nondeleterious variants were evaluated using two-tailed Mann–Whitney–Wilcoxon tests. The association between variables of interest and mutant clone expansion was evaluated through mixed-effects models (R package lme4 v1.1.34). Based on an estimation of statistical power (using the G*power software v3.1.9.7 and the R package pwrss v0.3.1), we had >80% power to detect as statistically significant ($\alpha = 0.05$) any effect of the presence of subclinical atherosclerosis leading to an absolute increase in AER greater than 3.2%, an effect size substantially smaller than that expected to result from the increased HSC proliferation reported in human individuals with atherosclerosis[27]. To prevent potential biases, the analysis of imaging data, the sequencing of DNA samples, the bioinformatic processing of sequencing data, the curation of somatic variants to identify CH driver mutations and the statistical analysis of associations were conducted independently by different investigators who were blinded to other data.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

CH-related mutations identified through high-sensitivity genomic DNA sequencing are listed in Supplementary Table 2. Other data from PESA participants are not publicly available for privacy and legal reasons. Access to these data can be requested to the PESA Scientific Committee via the corresponding authors or by e-mailing pesa-h@cnic.es. Data access will require a research proposal and approval by the PESA Scientific Committee, which meets every other month to evaluate such requests. To gain access, data requestors will need to sign a data access agreement and a nondisclosure agreement. Human genome build GRCh38 (GCF_000001405.40) is available at https://gatk.broadinstitute.org/hc/en-us/articles/360035890811-Resource-bundle. The present article includes all other data generated or analyzed during this study.

Code availability

The code is publicly available and can be found via GitHub at https://github.com/Unidad-Bioinformatica-CNIC/CHIP-candidate_mutations. The source code from the R-packages used in this study are freely available online (https://cran.r-project.org/).

References

57. Weinstock, J. S. et al. Aberrant activation of TCL1A promotes stem cell expansion in clonal haematopoiesis. *Nature* **616**, 755–763 (2023).
58. Miller, C. A. et al. Failure to detect mutations in U2AF1 due to changes in the GRCh38 reference sequence. *J. Mol. Diagn.* **24**, 219–223 (2022).

Acknowledgements

The PESA study is funded by the Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain, and Banco Santander, Madrid, Spain. The project leading to these results received funding from 'la Caixa' Foundation under the project codes LCF/PR/HR17/52150007 and LCF/PR/HR22/52420011. This research work was also supported by grant PLEC2021-008194, funded by MICIU/AEI/10.13039/501100011033 and by 'European Union NextGenerationEU/PRTR'; grant PID2021-126580OB-I00, funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU; and grant 202314-31, funded by Fundació 'La Marató TV3'. M.D.D. is supported by grant PRE2019-087463, funded by MICIU/AEI/10.13039/501100011033' and 'ESF Investing in your future'. B.L.R.-N. is supported by grant PRE2019-087829, funded by MICIU/AEI/10.13039/501100011033' and 'ESF Investing in your future'. B.Ib. is supported by the European Commission (grant nos. ERC-CoG 819775 and H2020-HEALTH 945118), the Spanish Ministry of Science and Innovation (PID2022-140176OB-I00) and the Red Madrileña de Nanomedicina en Imagen Molecular-Comunidad de Madrid (P2022/BMD-7403 RENIM-CM). B.Iz. is supported by the program Atracción de Talento of the Comunidad de Madrid (GN:2022-T1/BMD-23767). J.J.F. is supported by grant RYC-2016-20026, funded by MICIN/AEI/10.13039/501100011033' and 'ESF Investing in your future'. The CNIC is supported by the Instituto de Salud Carlos III (ISCIII), the Ministerio de Ciencia, Innovación y Universidades (MICIU) and the Pro CNIC Foundation and is a Severo Ochoa Center of Excellence (grant CEX2020-001041-S funded by MICIN/AEI/10.13039/501100011033). We thank J. Sanchez-González for his contribution to imaging analyses in PESA.

Author contributions

M.D.-D. and B.L.R.-N. analyzed somatic variant data, conducted association studies and contributed to writing the paper. A.Q., E.V., M.A.R-M., N.G-L., C.E. and A.D. conducted DNA sequencing analyses. J.B., J.C.S-C. and L.S-G. performed bioinformatics analyses of sequencing data. F.S.-C. supervised all bioinformatics and statistical analyses. B.Iz. provided methodological input. G.G. and A.J.Q. contributed to biological sample processing and revision. I.G.-L., G.M., V.M., A.G.-A., A.F.-O., E.L.-P., B.Ib., V.A. and V.F. contributed to the phenotyping of the PESA cohort and provided methodological input. V.F. is the principal investigator of the PESA study. J.J.F. conceived this study and its design, supervised all analyses and wrote the paper. All authors provided critical revision of the paper.

Competing interests

The authors declare no competing interests

Additional information

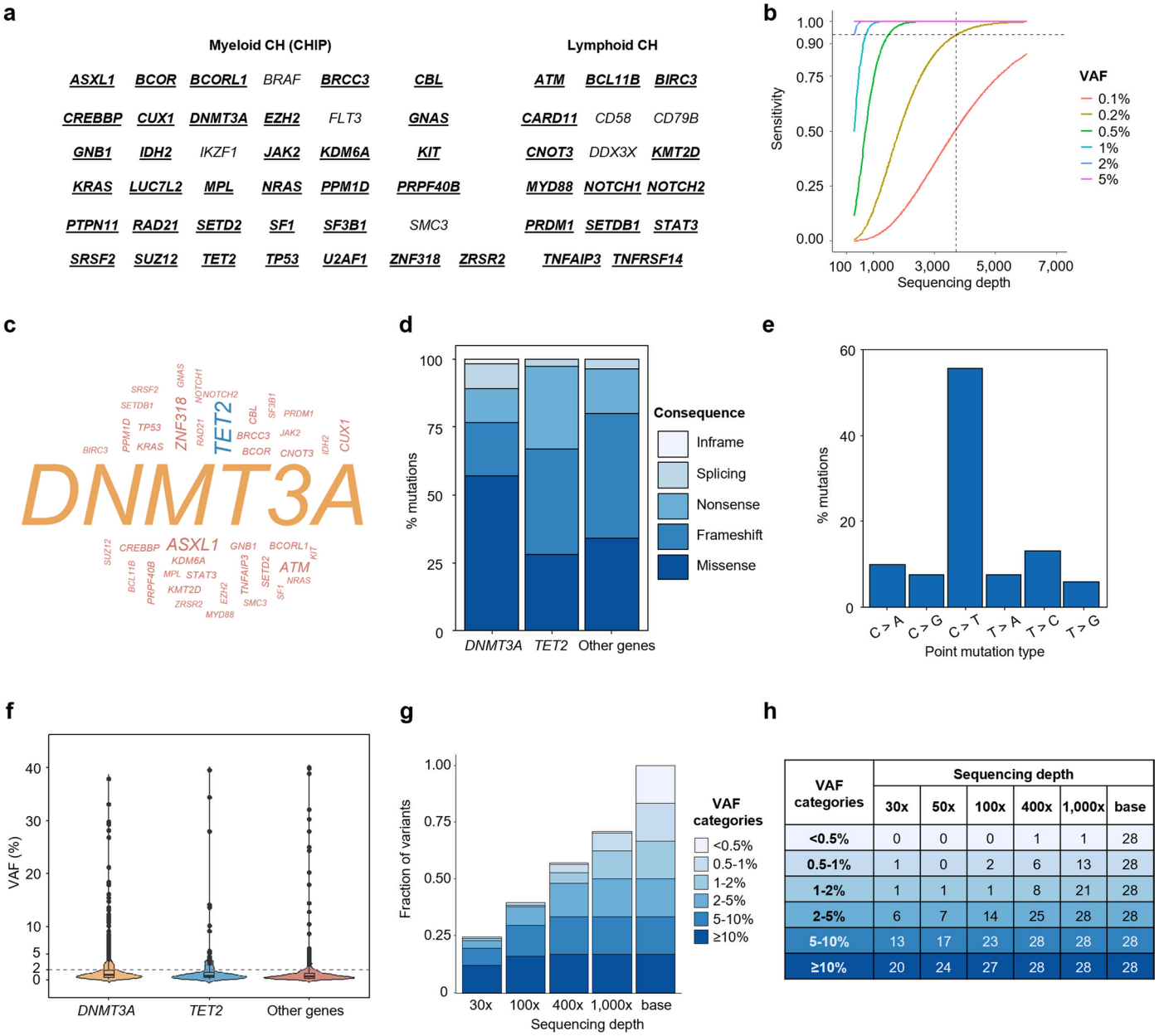
Extended data is available for this paper at https://doi.org/10.1038/s41591-024-03213-1.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-024-03213-1.

Correspondence and requests for materials should be addressed to Valentín Fuster or José J. Fuster.

Peer review information *Nature Medicine* thanks Alexander Bick, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Michael Basson, in collaboration with the *Nature Medicine* team.

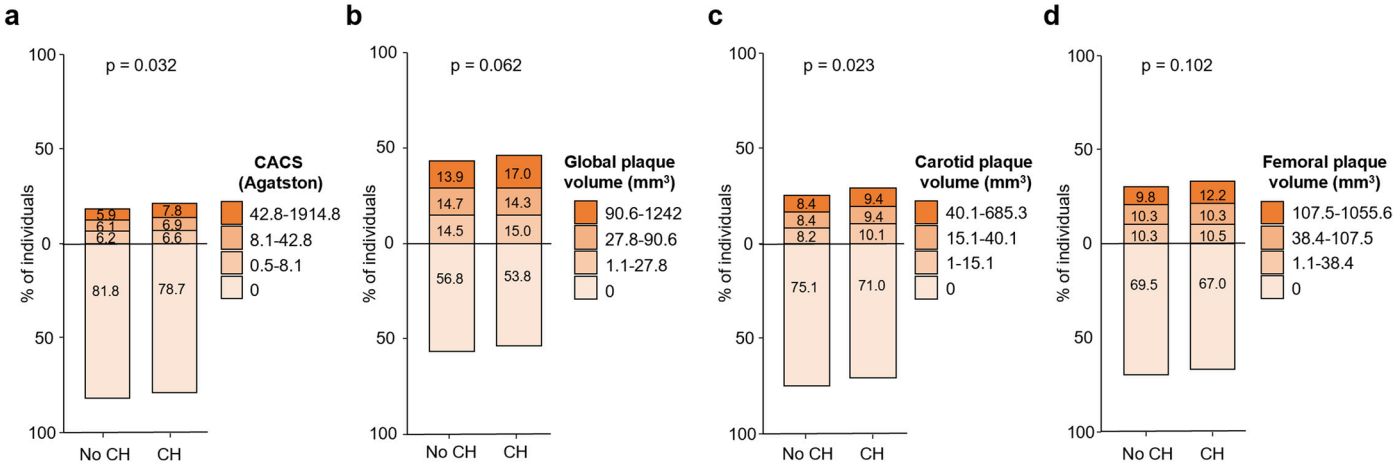
https://doi.org/10.1038/s41591-024-03213-1



VAF categories	Sequencing depth					
	30x	50x	100x	400x	1,000x	base
<0.5%	0	0	0	1	1	28
0.5-1%	1	0	2	6	13	28
1-2%	1	1	1	8	21	28
2-5%	6	7	14	25	28	28
5-10%	13	17	23	28	28	28
≥10%	20	24	27	28	28	28

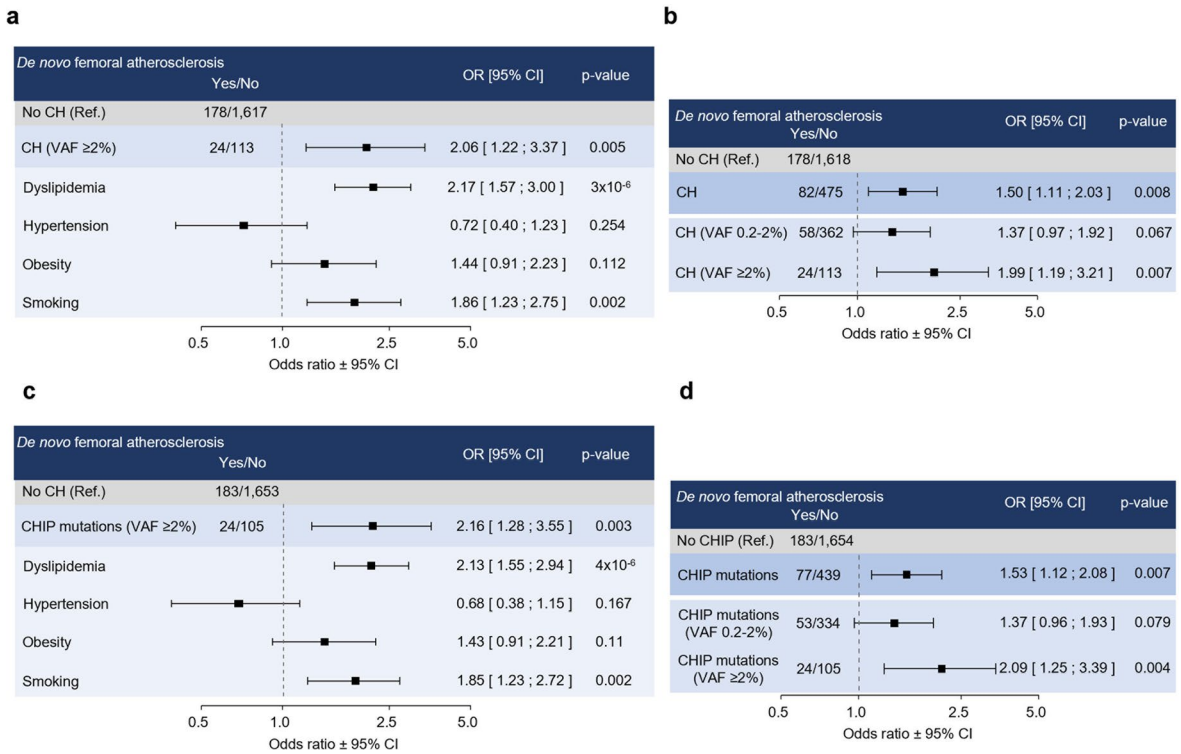
Extended Data Fig. 1 | Basic features of clonal hematopoiesis (CH)-related mutations and subclinical atherosclerosis burden in the study population. We performed deep sequencing to identify somatic mutations in a custom panel of 54 CH-related genes in 3,692 middle-aged individuals from the PESA cohort. **a**, CH-related genes included in the custom panel. Genes for which a CH mutation was found in at least one individual are underlined and in bold font. Genes typically considered as drivers of myeloid CH (frequently known as CHIP, clonal hematopoiesis of indeterminate potential) or lymphoid-CH are shown separately. **b**, Assay sensitivity for the detection of CH-related variants at different variant allele fractions (VAF). Dashed lines show the median sequencing depth of our study (3,712x; vertical line), and the sensitivity to detect CH variants at the minimum VAF threshold (0.2%) used to define CH (horizontal lines). **c**, Graphical representation of the CH-related genes found mutated in the study population; font size is proportional to the frequency of mutations. **d**, Proportions of different type of mutations in CH-related genes according to their functional consequence. **e**, Proportions of different types of single nucleotide substitutions. **f**, Distribution of mutant clone size, assessed as VAF, in *DNMT3A* (n = 657), *TET2* (n = 153) and other sequenced genes (n = 362). The dashed line shows the 2% VAF threshold typically used to identify CHIP. Boxes represent the 25th (Q1), 50th (median) and 75th (Q3) percentiles of the data. Whiskers represent Q1 - 1.5*IQR at the minimum and Q3 + 1.5*IQR at the maximum. **g**, Proportion of CH-related variants identified after downsampling of sequencing depth to 30x, 100x, 400x and 1,000x. "base" indicates the median sequencing depth for these variants using our original sequencing approach (3,625x); n = 168 CH variants, 28 for each VAF range. **h**, Number of CH-related variants identified after downsampling to specific sequencing depths.

https://doi.org/10.1038/s41591-024-03213-1



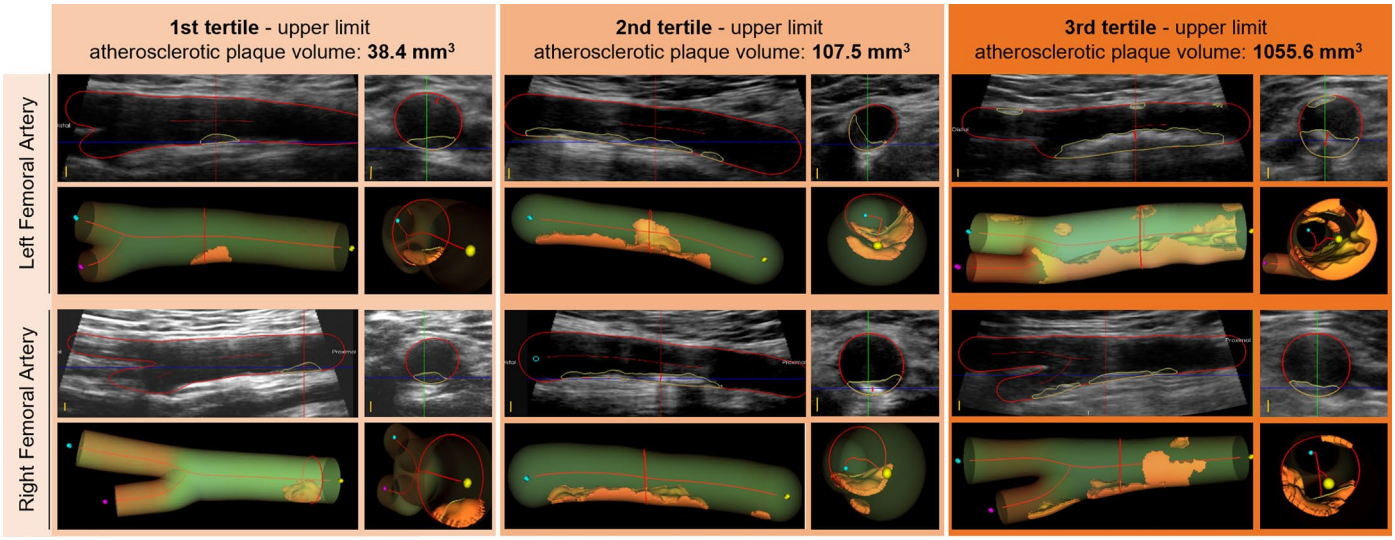
Extended Data Fig. 2 | Subclinical atherosclerosis burden at baseline stratified by CH status. We performed deep targeted sequencing to identify somatic mutations in a custom panel of 54 CH-related genes in 3,692 middle-aged individuals from the PESA cohort. CACS and 3DVUS imaging were used to determine subclinical atherosclerosis burden, assessed cross-sectionally at baseline. Graphs show the proportion of individuals with no detectable atherosclerosis and across tertiles of plaque burden, based on CAC score (**a**, n = 3,687), global plaque volume by 3DVUS (**b**, n = 3,543), carotid plaque volume by 3DVUS (**c**, n = 3,659) or femoral plaque volume by 3DVUS (**d**, n = 3,569). Statistical significance for differences in subclinical atherosclerosis burden, defined as 0 or tertiles, was evaluated using an ordinal logistic regression model.

https://doi.org/10.1038/s41591-024-03213-1



Extended Data Fig. 3 | Effects of clonal hematopoiesis (CH) on *de novo* atherosclerosis development in femoral arteries. We investigated the association between CH at baseline and *de novo* development of femoral atherosclerosis ~3-years after enrollment among individuals who initially lacked detectable atherosclerosis in that region, assessed by 3-dimensional-vascular ultrasound (3DVUS) imaging. **a**, Association between CH with VAF ≥ 2% and *de novo* atherosclerosis development in femoral arteries, based on multivariate logistic regression analyses adjusted for age, sex, conventional modifiable risk factors as categorical variables and the cumulative systemic burden of atherosclerosis across the timeframe of the study (AUC for CACS, assessed by CT imaging, and global atherosclerotic plaque volumes, assessed by 3DVUS imaging) (n = 1,932). **b**, Association between CH and *de novo* atherosclerosis development in femoral arteries, based on multivariate logistic regression analyses adjusted for age, sex, absolute counts of leukocytes, erythrocytes and platelets, and the cumulative systemic burden of atherosclerosis across the timeframe of the study (n = 2,353). **c**, Association between myeloid CH or CHIP with VAF ≥ 2% and *de novo* atherosclerosis development in femoral arteries, based on multivariate logistic regression analyses adjusted for age, sex, conventional modifiable risk factors and the cumulative systemic burden of atherosclerosis across the timeframe of the study (n = 1,965). **d**, Association between CHIP and *de novo* atherosclerosis development in femoral arteries, based on multivariate logistic regression analyses adjusted for age, sex, absolute counts of leukocytes, erythrocytes and platelets, and the cumulative systemic burden of atherosclerosis across the timeframe of the study (n = 2,353). Bars in each panel indicate 95% confidence intervals centered in the mean value (square).

https://doi.org/10.1038/s41591-024-03213-1



Extended Data Fig. 4 | Representative 3DVUS images from subclinical atherosclerosis burden in femoral arteries across tertiles of atherosclerotic plaque volume. Multiplanar views (upper rows, scale bar: 2 mm) and 3D reconstruction (lower rows) from the right and left femoral arteries. Images are representative of the upper limits of each tertile of atherosclerotic plaque volume. 3DVUS, 3-dimensional vascular ultrasound.

https://doi.org/10.1038/s41591-024-03213-1

Extended Data Table 1 | Baseline characteristics of the study population stratified by the presence of CH mutations with VAF 0.2-2% or ≥2%

	No CH (n=2,792)	CH, VAF 0.2-2% (n=682)	p-value	CH, VAF ≥2% (n=218)	p-value
Age, years	45.0 [42.0 - 49.0]	47.0 [43.0 - 50.0]	< 0.001	47.5 [44.0 - 52.0]	< 0.001
Men, %	1812 (64.9)	411 (60.3)	0.027	134 (61.5)	0.343
Dyslipidemia, %	1145 (41.0)	283 (41.5)	0.851	92 (42.2)	0.785
Total cholesterol, mg/dL	200.2 ± 33.3	201.5 ± 33.3	0.369	201.1 ± 32.9	0.696
Low plasma HDL-C, %	904 (32.4)	227 (33.3)	0.684	63 (28.9)	0.325
LDL-C ≥160 mg/dL, %	476 (17.0)	114 (16.7)	0.880	37 (17.0)	1.000
HDL-C, mg/dL	49.0 ± 12.1	49.5 ± 12.8	0.298	50.0 ± 12.5	0.229
LDL-C, mg/dL	132.3 ± 29.7	132.4 ± 29.4	0.967	133.0 ± 28.5	0.746
Plasma triglycerides ≥150 mg/dL, %	303 (10.9)	83 (12.2)	0.361	20 (9.2)	0.511
Triglycerides, mg/dL	79.0 [59.0 - 112.0]	80.0 [61.0 - 117.0]	0.198	81.0 [59.2 - 106.8]	0.905
Serum lipoprotein(a), mg/dL	17.6 [6.7 - 43.0]	17.3 [6.8 - 40.6]	0.594	20.4 [7.9 - 44.5]	0.135
Hypertension, %	297 (10.6)	93 (13.6)	0.031	29 (13.3)	0.269
SBP, mmHg	116.0 ± 12.2	116.6 ± 13.1	0.279	117.8 ± 13.4	0.056
DBP, mmHg	72.2 ± 9.3	73.1 ± 9.6	0.022	73.5 ± 9.5	0.051
Diabetes mellitus, %	49 (1.8)	13 (1.9)	0.916	6 (2.8)	0.426
Fasting glucose ≥ 100 mg/dL, %	349 (12.5)	102 (15.0)	0.100	31 (14.2)	0.528
Fasting glucose, mg/dL	89.0 [84.0 - 95.0]	89.0 [84.0 - 95.0]	0.559	90.0 [83.0 - 96.0]	0.498
Insulin, µU/mL	5.1 [3.6 - 7.4]	5.1 [3.7 - 7.6]	0.278	5.2 [3.7 - 7.6]	0.310
HOMA-IR > 2.5, %	266 (9.5)	74 (10.9)	0.332	26 (11.9)	0.301
HOMA-IR	1.1 [0.7 - 1.7]	1.1 [0.8 - 1.8]	0.297	1.1 [0.8 - 1.8]	0.295
Hemoglobin A1c, %	5.4 [5.2 - 5.6]	5.4 [5.2 - 5.7]	0.059	5.4 [5.2 - 5.8]	0.005
Current smoker, %	542 (19.4)	162 (23.8)	0.013	40 (18.3)	0.767
Past smoker, %	878 (43.7)	229 (49.2)	0.035	83 (51.9)	0.056
Ever smoker, %	1661 (59.5)	446 (65.4)	0.006	141 (64.7)	0.154
Obesity (BMI ≥30 kg/m^2), %	375 (13.4)	103 (15.1)	0.275	28 (12.8)	0.882
BMI, kg/m2	25.9 ± 3.8	26.0 ± 3.9	0.430	26.0 ± 3.6	0.761
Central obesity, %	553 (19.8)	149 (21.8)	0.256	52 (23.9)	0.178
Waist circumference, cm	89.2 ± 11.9	89.2 ± 12.5	0.892	90.1 ± 12.1	0.287
Metabolic syndrome, %	243 (8.7)	71 (10.4)	0.187	24 (11.0)	0.303
hs-CRP, mg/dL	0.1 [0.0 - 0.2]	0.1 [0.0 - 0.2]	0.183	0.1 [0.0 - 0.2]	0.578
GOT, U/L	18.0 [15.0 - 22.0]	18.0 [16.0 - 22.0]	0.851	19.0 [16.0 - 21.8]	0.977
GPT, U/L	20.0 [14.0 - 27.0]	19.0 [14.0 - 27.0]	0.600	18.0 [14.0 - 25.0]	0.129
GGT, U/L	20.0 [14.0 - 31.0]	21.0 [14.0 - 30.0]	0.631	21.0 [14.0 - 31.0]	0.925

Data is expressed as n (%), mean±SD or median [IQR] when appropriate. P-values were calculated through independent two-tailed t-tests for continuous variables exhibiting a normal distribution, two-tailed Wilcoxon rank tests for continuous variables with a non-normal distribution and χ2 tests for categorical variables. HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; SBP, systolic blood pressure; DBP, diastolic blood pressure; HOMA-IR, Homeostatic Model Assessment for Insulin Resistance; BMI, body mass index; hs-CRP: high sensitivity C-reactive protein test; GOT, glutamic-oxaloacetic transaminase; GPT, Glutamic-pyruvic transaminase; GGT, gamma glutamyl transferase.

https://doi.org/10.1038/s41591-024-03213-1

Extended Data Table 2 | Cross-sectional impact of CH mutations on blood cell counts

	CH, any VAF			CH, VAF 0.2-2%			CH, VAF ≥2%		
	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value
Leukocytes	0.23	0.06	< 0.001	0.25	0.07	< 0.001	0.16	0.11	0.139
Erythrocytes	0.04	0.01	0.004	0.03	0.01	0.018	0.05	0.02	0.025
Platelets	3.34	1.80	0.063	0.56	1.97	0.779	12.37	3.27	< 0.001
Monocytes	0.02	0.00	< 0.001	0.02	0.01	< 0.001	0.01	0.01	0.368
Neutrophils	0.15	0.05	0.001	0.17	0.05	0.002	0.11	0.08	0.178
Lymphocytes	0.06	0.02	0.003	0.06	0.02	0.003	0.03	0.03	0.317

CH status and blood cell counts were determined at baseline. Statistical associations were evaluated using multivariate regression models adjusted for age and sex. Results are shown for all CH mutations, mutations with variant allele fraction (VAF) between 0.2-2% and ≥2%.

https://doi.org/10.1038/s41591-024-03213-1

Extended Data Table 3 | Hematological parameters in the study population according to the presence of CH mutations with VAF 0.2-2% or ≥ 2%

	No CH (n=2,792)	CH, any VAF (n=900)	p-value	CH, VAF 0.2-2% (n=682)	p-value	CH, VAF ≥2% (n=218)	p-value
Leukocytes ($10^3/\mu L$)	5.5 [4.8 - 6.5]	5.8 [4.9 - 6.9]	< 0.001	5.7 [4.9 - 6.9]	< 0.001	5.9 [4.9 - 6.7]	0.008
Neutrophils ($10^3/\mu L$)	3.1 [2.6 - 3.9]	3.3 [2.7 - 4.1]	< 0.001	3.3 [2.7 - 4.1]	0.002	3.3 [2.7 - 4.1]	0.005
Lymphocytes ($10^3/\mu L$)	1.7 [1.4 - 2.1]	1.8 [1.5 - 2.1]	0.001	1.8 [1.5 - 2.2]	0.001	1.8 [1.5 - 2.0]	0.215
Monocytes ($10^3/\mu L$)	0.4 [0.3 - 0.5]	0.4 [0.3 - 0.5]	0.001	0.4 [0.3 - 0.5]	0.001	0.4 [0.3 - 0.5]	0.510
Eosinophils ($10^3/\mu L$)	0.1 [0.1 - 0.2]	0.1 [0.1 - 0.2]	0.627	0.1 [0.1 - 0.2]	0.593	0.1 [0.1 - 0.2]	0.937
Basophils ($10^3/\mu L$)	0.0 [0.0 - 0.1]	0.0 [0.0 - 0.1]	0.124	0.0 [0.0 - 0.1]	0.068	0.0 [0.0 - 0.1]	0.993
Neutrophils, %	57.5 [52.2 - 62.3]	57.5 [53.1 - 63.1]	0.117	57.1 [53.0 - 62.9]	0.412	58.5 [53.3 - 63.8]	0.049
Monocytes, %	7.2 [6.0 - 8.3]	7.1 [6.0 - 8.4]	0.656	7.2 [6.0 - 8.4]	0.964	7.0 [5.8 - 8.3]	0.276
Lymphocytes, %	31.4 [27.1 - 36.0]	31.4 [26.6 - 35.3]	0.156	31.6 [26.7 - 35.4]	0.416	30.6 [26.4 - 34.9]	0.100
Eosinophils, %	2.1 [1.3 - 3.3]	2.1 [1.3 - 3.3]	0.284	2.1 [1.3 - 3.3]	0.397	2.1 [1.2 - 3.1]	0.424
Basophils, %	0.9 [0.7 - 1.2]	0.9 [0.7 - 1.1]	0.421	0.9 [0.7 - 1.2]	0.912	0.9 [0.6 - 1.1]	0.110
Neutrophile/Lymphocyte	1.8 [1.4 - 2.3]	1.8 [1.5 - 2.4]	0.124	1.8 [1.5 - 2.3]	0.367	1.9 [1.5 - 2.4]	0.080
Erythrocytes ($10^6/\mu L$)	4.8 [4.4 - 5.0]	4.7 [4.4 - 5.0]	0.449	4.7 [4.4 - 5.0]	0.733	4.8 [4.4 - 5.1]	0.291
RDW, %	14.3 [13.7 - 14.9]	14.4 [13.8 - 14.9]	0.158	14.4 [13.8 - 14.9]	0.401	14.5 [13.8 - 14.9]	0.115
MCV (fL)	90.6 [88.4 - 92.9]	90.6 [88.3 - 92.8]	0.571	90.7 [88.4 - 92.9]	0.915	90.6 [88.2 - 92.4]	0.283
Hemoglobin (g/dL)	14.6 [13.6 - 15.4]	14.6 [13.6 - 15.4]	0.947	14.6 [13.6 - 15.4]	0.926	14.5 [13.6 - 15.4]	0.977
Hematocrit, %	43.1 [40.2 - 45.4]	43.0 [40.2 - 45.4]	0.886	43.0 [40.1 - 45.4]	0.871	42.9 [40.3 - 45.5]	0.987
Platelets ($10^3/\mu L$)	221.0 [194.0 - 253.0]	225.0 [195.0 - 257.0]	0.073	223.0 [194.0 - 253.8]	0.632	229.0 [200.0 - 267.8]	0.002

Data is expressed as median [IQR] when appropriate. P-values are derived from independent two-tailed t-tests relative to the non-CH group. VAF, variant allele fraction; RDW, red blood cell distribution width; MCV, mean corpuscular volume.

https://doi.org/10.1038/s41591-024-03213-1

Extended Data Table 4 | No association between CH and the presence or extent of subclinical atherosclerosis, assessed cross-sectionally at baseline

Model 1	CH, any VAF			CH, VAF 0.2-2%			CH, VAF ≥2%		
Baseline presence of atherosclerosis									
	OR	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
CACs > 0	1.07	0.87 ; 1.32	0.492	1.09	0.87 ; 1.37	0.437	1.04	0.71 ; 1.49	0.838
Global plaque volume >0	1.01	0.85 ; 1.19	0.923	1.09	0.91 ; 1.31	0.361	0.78	0.57 ; 1.07	0.128
Carotid plaque volume >0	1.1	0.92 ; 1.30	0.311	1.24	1.02 ; 1.50	0.029	0.73	0.51 ; 1.02	0.07
Femoral plaque volume >0	0.99	0.83 ; 1.19	0.948	1.03	0.84 ; 1.25	0.775	0.89	0.64 ; 1.24	0.509
Baseline extent of atherosclerosis									
	Estimate	**Std. Error**	**p-value**	**Estimate**	**Std. Error**	**p-value**	**Estimate**	**Std. Error**	**p-value**
CACS (Agatston)	12.47	16.25	0.443	22.18	18.08	0.22	-16.49	27.18	0.544
Global plaque volume (mm3)	0.26	7.67	0.972	-0.09	8.37	0.992	0.47	14.57	0.974
Carotid plaque volume (mm3)	-0.46	4.29	0.915	-1.26	4.64	0.786	2.88	8.8	0.743
Femoral plaque volume (mm3)	-1.38	8.7	0.874	-1.09	9.58	0.909	-3.00	15.91	0.851
Model 2	**CH, any VAF**			**CH, VAF 0.2-2%**			**CH, VAF ≥2%**		
Baseline presence of atherosclerosis									
	OR	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
CACs > 0	1.04	0.84 ; 1.29	0.7	1.04	0.82 ; 1.31	0.75	1.08	0.74 ; 1.55	0.695
Global plaque volume >0	0.99	0.83 ; 1.17	0.871	1.08	0.89 ; 1.31	0.443	0.74	0.54 ; 1.02	0.07
Carotid plaque volume >0	1.07	0.89 ; 1.28	0.463	1.23	1.01 ; 1.50	0.043	0.69	0.48 ; 0.97	0.036
Femoral plaque volume >0	0.96	0.79 ; 1.16	0.664	0.98	0.80 ; 1.21	0.866	0.9	0.64 ; 1.26	0.541
Baseline extent of atherosclerosis									
	Estimate	**Std. Error**	**p-value**	**Estimate**	**Std. Error**	**p-value**	**Estimate**	**Std. Error**	**p-value**
CACS (Agatston)	12.4	16.63	0.456	21.66	18.58	0.244	-17.98	27.54	0.514
Global plaque volume (mm3)	-4.21	7.67	0.583	-4.77	8.36	0.568	-3.84	14.51	0.791
Carotid plaque volume (mm3)	-1.35	4.33	0.756	-1.97	4.69	0.675	0.65	8.89	0.942
Femoral plaque volume (mm3)	-5.85	8.78	0.505	-5.64	9.68	0.561	-7.4	15.92	0.642

CH status and subclinical atherosclerosis burden were determined at baseline. Presence of subclinical atherosclerosis was defined as a categorical variable based on CACS or 3DVUS imaging. The extent of subclinical atherosclerosis in those with detectable atherosclerosis in specific regions was defined as a continuous variable based on CACS or 3DVUS imaging. Statistical associations were evaluated using multivariate regression models adjusted for age and sex (model 1), or age, sex and conventional cardiovascular risk factors as continuous variables (model 2).

https://doi.org/10.1038/s41591-024-03213-1

Extended Data Table 5 | Relationship between conventional cardiovascular risk factors and expansion rate of CH mutant clones

	Annual Expansion Rate			log (VAF ratio)/ time			Exponential growth rate		
	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value
CVRF (categorical)									
Hypertension	3.33	1.67	0.047	1.69	1.01	0.096	1.84	1.07	0.086
Diabetes	-2.58	4.81	0.591	-0.87	2.95	0.768	-1.04	3.10	0.738
Dyslipidemia	0.56	1.22	0.643	0.09	0.74	0.906	0.10	0.78	0.896
Obesity	0.06	1.72	0.971	0.08	1.05	0.936	0.07	1.11	0.952
Smoking	0.34	1.45	0.813	-0.08	0.88	0.93	-0.07	0.93	0.936
CVRF (continuous)									
Baseline SBP, mmHg	0.07	0.05	0.148	0.04	0.03	0.19	0.04	0.03	0.189
Baseline fasting glucose, mg/dL	0.05	0.06	0.368	0.04	0.04	0.205	0.04	0.04	0.245
Baseline LDL-C, mg/dL	0.01	0.02	0.701	0.00	0.01	0.858	0.00	0.01	0.892
Baseline BMI, kg/m2	0.12	0.17	0.493	0.09	0.10	0.374	0.09	0.11	0.413
CVRF (AUC)									
AUC SBP	10.29	5.49	0.062	5.54	3.31	0.095	5.87	3.49	0.093
AUC fasting glucose	11.75	10.40	0.259	10.77	6.29	0.088	10.51	6.63	0.114
AUC LDL-C	3.46	4.97	0.487	0.75	3.01	0.804	0.99	3.17	0.756
AUC BMI	2.42	4.15	0.559	2.47	2.50	0.325	2.34	2.64	0.376

We monitored the expansion of 602 CH-related mutations through serial sequencing of blood samples from 494 individuals at baseline and ~6 years after enrollment. We then investigated the association between traditional cardiovascular risk factors (CVRF) at baseline and the annual expansion rate (AER) of mutant hematopoietic clones through mixed-effects models, adjusted for sex and baseline variant allele fraction (VAF). CVRF were assessed as categorical and continuous variables, and as the area under the curve (AUC) at three different timepoints over 6 years. SBP, systolic blood pressure; LDL-C, low-density lipoprotein cholesterol; BMI, body mass index.

nature portfolio

Corresponding author(s): Valentin Fuster, José J. Fuster

Last updated by author(s): Jul 16, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Individual-level ultrasound imaging data were previously collected by investigators in the PESA study using VPQ software v13 (Philips Healthcare) and CM2020 software (Philips Research Analysis). No other software was used for data collection.
Data analysis	Statistical analyses were performed using RStudio (v2022.07.2+576) with R version 4.2.2. The code is publicly available and can be found at: https://github.com/Unidad-Bioinformatica-CNIC/CHIP-candidate_mutations. The source code from the R-packages used in this study are freely available online (https://cran.r-project.org/). The association between variables of interest and mutant clone expansion were evaluated through mixed-effects models (R package lme4 v1.1.34). Statistical power analyses were performed using the G*power software (v3.1.9.7) and the R package pwrss v0.3.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Clonal hematopoiesis-related mutations identified through high-sensitivity genomic DNA sequencing are listed in Supplementary Table 2. Other data from PESA participants are not publicly available for privacy and legal reasons. Access to these data can be requested to the PESA Scientific Committee via the corresponding authors or by e-mailing pesa-h@cnic.es. Data access will require a research proposal and approval by the PESA Scientific Committee, which meets every other month to evaluate such requests. To gain access, data requestors will need to sign a data access agreement and a non-disclosure agreement. Human genome build GRCh38 (GCF_000001405.40) is available at https://gatk.broadinstitute.org/hc/en-us/articles/360035890811-Resource-bundle. The present article includes all other data generated or analyzed during this study.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Sex was determined based on self-reporting, and validated based on sequencing data. 36% of the study population were women. Sex was considered in statistical analyses, as indicated in figure legends and main text.
Reporting on race, ethnicity, or other socially relevant groupings	All participants in the cohort are Caucasian individuals of middle age, residing in the same geographical region, and possessing similar demographic characteristics.
Population characteristics	Progression of Early Subclinical Atherosclerosis (PESA) is an observational, longitudinal and prospective cohort study of 4184 asymptomatic employees of the Santander Bank in Madrid who were 40 to 55 years old at the time of recruitment, between June 2010 and February 2014. Individuals with prior cardiovascular disease or any condition reducing life expectancy or affecting study adherence were not included. In the current study, the study population included 3,692 participants with available blood DNA samples, multimodal vascular imaging data at multiple timepoints and written informed consent for DNA sequencing analyses.
Recruitment	PESA study recruitment was based on volunteer participation, as described in https://doi.org/10.1016/j.ahj.2013.08.024
Ethics oversight	The study protocol was approved by the Instituto de Salud Carlos III Ethics Committee (CEI PI_52/2019), and all eligible participants provided written informed consent.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No a priori sample size calculation was performed. The analyses within this work included all participants in the PESA cohort with available blood DNA samples, multimodal vascular imaging data at multiple timepoints and written informed consent for DNA sequencing analyses.
Data exclusions	No exclusions
Replication	Not applicable to the current study, which is based on the integration of DNA sequencing data and multimodal non-invasive imaging data (including 3DVUS) collected at multiple timepoints in a population of apparently healthy individuals. No other comparable cohorts are currently available for replication.
Randomization	Not applicable to the analyses of the relationship between baseline CH and atherosclerosis burden. All PESA participants with available blood DNA samples, multimodal vascular imaging data at multiple time-points and written informed consent for DNA sequencing analyses were included in this part of the study. PESA participants included in the analysis of mutant clone expansion were selected randomly among CH mutation carriers at baseline.
Blinding	To prevent potential biases, the analysis of imaging data, the sequencing of DNA samples, the bioinformatic processing of sequencing data,

Blinding	the curation of somatic variants to identify CH driver mutations and the statistical analysis of associations were conducted independently by different investigators who were blind to other data.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging