nature cardiovascular research

Article

https://doi.org/10.1038/s44161-022-00206-6

TP53-mediated clonal hematopoiesis confers increased risk for incident atherosclerotic disease

Received: 30 March 2022

Accepted: 8 December 2022

Published online: 16 January 2023

Check for updates

Seyedeh M. Zekavat[1,2,3,35], Vanesa Viana-Huete[4,35], Nuria Matesanz[4,35], Saman Doroodgar Jorshery[5,6], María A. Zuriaga[4], Md Mesbah Uddin[3,7], Mark Trinder[3,7,8], Kaavya Paruchuri[3,7,9], Virginia Zorita[4], Alba Ferrer-Pérez[4], Marta Amorós-Pérez[4], Paolo Kunderfranco[10], Roberta Carriero[10], Carolina M. Greco[10,11], Alejandra Aroca-Crevillen[4], Andrés Hidalgo[4,12], Scott M. Damrauer[13,14], Christie M. Ballantyne[15], Abhishek Niroula[16,17,18], Christopher J. Gibson[17], James Pirruccello[19,20,21], Gabriel Griffin[3,22,23], Benjamin L. Ebert[17,24], Peter Libby[25], Valentín Fuster[4,26], Hongyu Zhao[27], Marzyeh Ghassemi[5,28,29], Pradeep Natarajan[7,9,30,36], Alexander G. Bick[31,36], José J. Fuster[4,32,36] ✉ & Derek Klarin[33,34,36] ✉

Somatic mutations in blood indicative of clonal hematopoiesis of indeterminate potential (CHIP) are associated with an increased risk of hematologic malignancy, coronary artery disease and all-cause mortality. Here we analyze the relation between CHIP status and incident peripheral artery disease (PAD) and atherosclerosis, using whole-exome sequencing and clinical data from the UK Biobank and the Mass General Brigham Biobank. CHIP associated with incident PAD and atherosclerotic disease across multiple beds, with increased risk among individuals with CHIP driven by mutation in DNA damage repair (DDR) genes, such as *TP53* and *PPM1D*. To model the effects of DDR-induced CHIP on atherosclerosis, we used a competitive bone marrow transplantation strategy and generated atherosclerosis-prone *Ldlr*$^{-/-}$ chimeric mice carrying 20% p53-deficient hematopoietic cells. The chimeric mice were analyzed 13 weeks after grafting and showed increased aortic plaque size and accumulation of macrophages within the plaque, driven by increased proliferation of p53-deficient plaque macrophages. In summary, our findings highlight the role of CHIP as a broad driver of atherosclerosis across the entire arterial system beyond the coronary arteries and provide genetic and experimental support for a direct causal contribution of TP53-mutant CHIP to atherosclerosis.

Peripheral artery disease (PAD) is a leading cause of cardiovascular morbidity and mortality worldwide[1], and age is among its strongest risk factors. PAD associates with an extremely high cardiovascular mortality and, if unmitigated, can progress to limb loss[2]. The age-related acquisition and expansion of leukemogenic mutations in hematopoietic stem cells (HSCs) has recently been associated with an increased risk of hematologic malignancy, coronary artery disease (CAD) and overall mortality[3,4]. This phenomenon, termed clonal hematopoiesis of

A full list of affiliations appears at the end of the paper. ✉e-mail: jjfuster@cnic.es; dklarin@stanford.edu

indeterminate potential (CHIP), is relatively common in asymptomatic older adults, affecting at least 10% of individuals older than 70 years of age[5]. CHIP mutations most frequently occur in epigenetic regulators *DNMT3A* and *TET2*, in DNA damage repair (DDR) genes *PPM1D* and *TP53* or in cell cycle and transcriptional regulator genes *JAK2* and *ASXL1* (ref. 6). CHIP associates with CAD, stroke and other cardiovascular conditions, but accumulating evidence suggests that the effects of CHIP depend, to a great extent, on the specific mutated gene and the specific cardiovascular condition[3,7–10]. Whether CHIP links with increased risk of atherosclerotic disease in the peripheral arterial bed is unknown.

In this study, we leveraged 50,122 whole-exome sequences from two genetic biobanks (UK Biobank (UKB) and Mass General Brigham Biobank (MGBB)) and tested whether CHIP was associated with increased risk of PAD and atherosclerosis across multiple arterial beds and, additionally, whether these associations varied by putative CHIP driver gene. Based on these results, we then performed functional analyses in *Ldlr*-null mice transplanted with 20% *Trp53*$^{-/-}$ bone marrow (BM) cells—a murine model of atherosclerosis and clonal hematopoiesis driven by *TP53* mutations.

Results

Study cohorts and risk of hematologic malignancy

After excluding individuals with a known history of hematologic malignancy at enrollment, we identified 37,657 unrelated individuals from the UKB and 12,465 individuals from the MGBB with whole-exome sequencing data available for downstream analysis. Using a previously validated somatic variant detection algorithm[11], we identified 2,194 (5.8%) and 657 (5.4%) CHIP carriers in the UKB and the MGBB, respectively (Supplementary Table 1). Demographic and clinical characteristics of these individuals, stratified by CHIP status, are depicted in Supplementary Table 2. CHIP carriers tended to be older, male, current or previous smokers and have a history of CAD, hypertension and hyperlipidemia (two-tailed chi-squared and Wilcoxon rank-sum $P < 0.05$).

We first replicated known CHIP associations[11] with white blood cell (beta 0.09 s.d.; 95% confidence interval (CI): 0.05–0.13; $P = 1.6 \times 10^{-5}$), monocyte (beta 0.05 s.d.; 95% CI: 0.01–0.09; $P = 0.009$), neutrophil (beta 0.10 s.d.; 95% CI: 0.06–0.14; $P = 2.1 \times 10^{-6}$) and platelet counts (beta 0.07 s.d.; 95% CI: 0.03–0.11; $P = 0.0005$) in the UKB, with larger CHIP clone size as measured by variant allele fraction (VAF) having stronger effects on blood counts (Extended Data Fig. 1). Consistent with the existing literature[6,11], CHIP also associated with incident hematologic malignancy (hazard ratio (HR) = 2.20; 95% CI: 1.70–2.85; $P = 1.8 \times 10^{-9}$), specifically for acute myeloid leukemia (HR = 8.08; 95% CI: 4.36–14.97; $P = 3.2 \times 10^{-11}$), myeloproliferative neoplasms (HR = 5.89; 95% CI: 3.69–9.89; $P = 9.7 \times 10^{-14}$) and polycythemia vera (HR = 12.37; 95% CI: 4.85–31.54; $P = 1.4 \times 10^{-7}$). This risk increased with larger VAF (Extended Data Fig. 2).

CHIP and incident PAD risk

We next tested the association of CHIP status with incident PAD. Using available electronic health record (EHR) data and a previously validated PAD definition[12], we identified 338 and 419 incident PAD cases in the UKB and the MGBB, respectively. CHIP associated with a 58% increased risk of incident PAD in the UKB ($HR_{UKB} = 1.58$, 95% CI: 1.11–2.25; $P = 0.01$; Fig. 1), results that were replicated in the MGBB (overall HR = 1.66, 95% CI: 1.31–2.11; $P = 2.4 \times 10^{-5}$). We then sought to evaluate whether those with larger CHIP clone sizes (that is, higher VAF) had greater risk for PAD, as larger CHIP clones associate more strongly with adverse clinical outcomes[7]. We observed a graded relationship between CHIP VAF and PAD, as those with a VAF > 10% had even greater risk for an incident PAD event (overall HR = 1.97, 95% CI: 1.44–2.71; $P = 2.3 \times 10^{-5}$; Fig. 1). Additional sensitivity analyses, including propensity score adjustment and a marginal structural Cox proportional hazards model estimated through stabilized inverse probability treatment weight (IPTW), revealed similar results in the UKB (Extended Data Fig. 3). Subsequent analyses showed no significant interaction between CHIP status and age, sex or smoking status on incident PAD risk. In sensitivity analyses accounting for a concomitant CAD diagnosis among those with incident PAD, the CHIP–PAD association remained significant (Supplementary Tables 3 and 4).

CHIP and incident atherosclerosis across multiple vascular beds

We next assessed whether CHIP was associated with eight other incident atherosclerotic diseases across multiple vascular beds. Using EHR-based disease definitions[13], we tested the association of CHIP with atherosclerotic disease across the mesenteric (acute and chronic), coronary and cerebral vascular beds as well as with aneurysmal disease (aortic and any other aneurysm). We observed significant associations for CAD (HR = 1.40, 95% CI: 1.20–1.63; $P = 1.9 \times 10^{-5}$), any aortic aneurysm (HR = 1.74; 95% CI: 1.21–2.51; $P = 0.0028$), other aneurysms (HR = 1.70; 95% CI: 1.23–2.34; $P = 0.0013$) and mesenteric ischemia (HR = 3.22; 95% CI: 2.01–5.17; $P = 1.3 \times 10^{-6}$) across both cohorts, with directionally consistent effect estimates observed for all the tested phenotypes (Fig. 2a). These associations were consistently stronger for large CHIP clones (Extended Data Fig. 4). We then created a composite, incident atherosclerosis outcome combining all nine atherosclerotic phenotypes ('pan-arterial atherosclerosis'; Supplementary Table 5). CHIP associated with this combined incident pan-arterial atherosclerosis endpoint (HR = 1.31, 95% CI: 1.14–1.49, $P = 9.7 \times 10^{-5}$), again with stronger effects conferred by large CHIP clones (HR = 1.45; 95% CI: 1.20–1.75; $P = 0.00013$) (Fig. 2b,c).

Gene-specific analyses of CHIP with incident atherosclerotic diseases

Next, we sought to understand whether the clonal hematopoiesis putative driver gene differentially affected the risk of acquiring atherosclerosis. Previous work focused primarily on the epigenetic regulators *DNMT3A* and *TET2* (refs. 4, 14), and whether DDR CHIP confers an increased risk of atherosclerosis is unknown. We stratified the CHIP–PAD, CHIP–CAD and CHIP pan-arterial atherosclerosis analyses by putative driver genes and specific mutations, focusing on *DNMT3A*, *TET2*, *ASXL1*, *JAK2*, the DDR genes *PPM1D* and *TP53* and mutations that specifically disrupt splicing factor genes (*LUC7L2*, *PRPF8*, *SF3B1*, *SRSF2*, *U2AF1* and *ZRSR2*)[15]. We observed an association of CHIP with PAD across three of the four common CHIP genes (*TET2*, *ASXL1* and *JAK2*), with significant heterogeneity of incident PAD effect sizes across the CHIP genes ($P_{heterogeneity} = 0.03$) (Fig. 3a). This heterogeneity persisted in sensitivity analysis after excluding *JAK2* carriers ($P_{heterogeneity} = 0.046$). A similar set of associations was observed for CHIP and CAD (Supplementary Table 6). These data also revealed the finding that DDR *TP53* and *PPM1D* CHIP associate with incident PAD (HR = 2.72; 95% CI: 1.20–1.75; $P = 0.00013$) and incident CAD (HR = 2.51; 95% CI: 1.52–4.13; $P = 0.00032$), with a numerically stronger effect on PAD conferred by *TP53* mutations (HR = 4.98; 95% CI: 1.23–20.09; $P = 0.024$) (Fig. 3a–c). Similar findings were observed for the incident pan-arterial atherosclerosis outcome when stratifying by putative driver gene (Extended Data Fig. 5). Further sensitivity analysis for DDR CHIP and incident PAD when excluding solid organ malignancy did not significantly change the associations ($P_{heterogeneity} > 0.05$).

Atherosclerosis development in p53$^{-/-}$ CHIP mice

Based on the finding of an association between somatic *TP53* mutations and atherosclerotic disease, we next characterized further the effects of *TP53*-mutant CHIP on atherosclerosis through experimental studies in mouse models. The most frequent somatic *TP53* variants in CHIP are mainly missense mutations that act in a dominant-negative manner and have an effect similar to that of full p53 deficiency in mouse studies[16]. Therefore, p53-deficient mice were used as a murine surrogate of human *TP53* mutations. To mimic CHIP and test whether the

https://doi.org/10.1038/s44161-022-00206-6

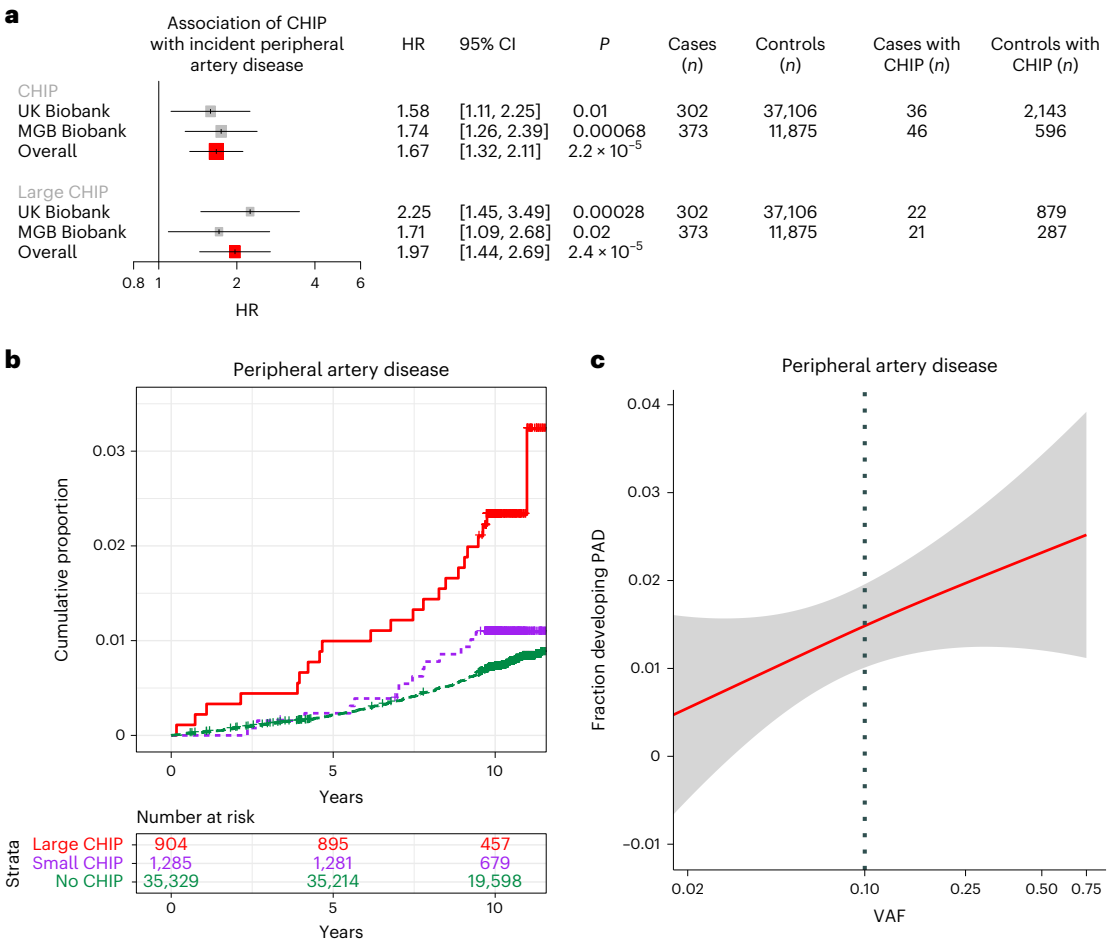


Fig. 1 | CHIP and incident PAD risk. a, Association of CHIP and large CHIP (VAF > 10%) carrier state with incident PAD events in the UKB (n = 37,657) and the MGBB (n = 12,465). Results were combined using an inverse variance-weighted fixed effects meta-analysis. Error bars are centered at the HR and show the 95% CI for estimates. **b**, Cumulative proportion of individuals developing PAD stratified by CHIP VAF clone size category in the UKB (n = 37,657). **c**, Fraction of individuals developing incident PAD by CHIP VAF in the UKB. The dotted vertical line at VAF = 0.10 represents the cutoff for the definition of a large CHIP clone. Error bands reflect the standard error of a generalized binomial additive model with integrated smoothness fit to the data.

expansion of p53-deficient hematopoietic cells contributes to atherosclerosis, a competitive bone marrow transplantation (BMT) strategy was used to generate atherosclerosis-prone *Ldlr*$^{-/-}$ chimeric mice carrying 20% *Trp53*$^{-/-}$ hematopoietic cells (20% KO-BMT mice) (Fig. 4a and Supplementary Fig. 1). This chimerism is consistent with the VAF of *TP53* mutations identified in our human study population (mean VAF = 0.13, that is ~26% mutant cells if monoallelic mutations; mean VAF = 0.23 in those defined as carrying large CHIP clones; Supplementary Table 7). Transplanted mice then consumed a high-fat/high-cholesterol (HF/HC) diet for 9 weeks to induce atherosclerosis development. To distinguish donor *Trp53*$^{-/-}$ and *Trp53*$^{+/+}$ cells in this experimental setting, *Trp53*$^{+/+}$ cells were obtained from mice carrying the CD45.1 variant of the CD45 cell surface marker, whereas *Trp53*$^{-/-}$ cells were obtained from mice carrying the CD45.2 variant of this protein. Control mice (20% WT-BMT) were transplanted with 20% CD45.2^{+}*Trp53*$^{+/+}$ cells and 80% CD45.1^{+}*Trp53*$^{+/+}$ cells. Flow cytometry analysis of CD45.2^{+} blood cells established that this BMT strategy led to a modest but significant expansion of donor *Trp53*$^{-/-}$ BM-derived cells compared to *Trp53*$^{+/+}$ cells in both BM hematopoietic stem/progenitor cells (LSK cells, defined as lineage^{-}Sca1^{+}c-Kit^{+}) and circulating white blood cells (Fig. 4b,c), consistent with previous studies[17–19]. Transplanted *Trp53*$^{-/-}$ BM cells expanded into all blood cell lineages to a similar extent, but this relative expansion did not affect absolute blood cell counts within the timeframe of the experiment (Extended Data Fig. 6a,b). Having validated this mouse model of p53-deficient CHIP based on a competitive BMT strategy, we next assessed whether this phenomenon affects the development of atherosclerosis or related metabolic abnormalities. The presence and expansion of *Trp53*$^{-/-}$ cells led to a significant ~40% increase in plaque size in the aortic root of male *Ldlr*$^{-/-}$ mice (Fig. 4d), without affecting body weight, spleen weight or serum cholesterol levels (Extended Data Fig. 6c–e). Similar results were obtained in female mice (Extended Data Fig. 6f–k). In contrast to previous findings on *DNMT3A/TET2* CHIP[4,14,20–22], the expansion of p53-deficient cells did not affect the expression of the pro-inflammatory cytokines IL-6 and IL-1β or that of the NLRP3 inflammasome in the atherosclerotic aortic wall (Extended Data Fig. 7a).

Expansion of p53-deficient macrophages in the murine atherosclerotic aorta

Increased atherogenesis in mice carrying *Trp53*$^{-/-}$ cells was paralleled by a substantial increase in plaque macrophage content, as assessed by immunohistological staining of Mac2 (Fig. 5a), suggesting a contribution of increased arterial macrophage burden to accelerated atherosclerosis in conditions of p53-deficient CHIP. No significant changes were observed in other plaque components, such as collagen content, vascular smooth muscle cell content, necrotic core extension or lipid content. (Fig. 5a and Extended Data Fig. 7b). Flow cytometry analysis of matched samples from blood and digested atherosclerotic aortas from

https://doi.org/10.1038/s44161-022-00206-6

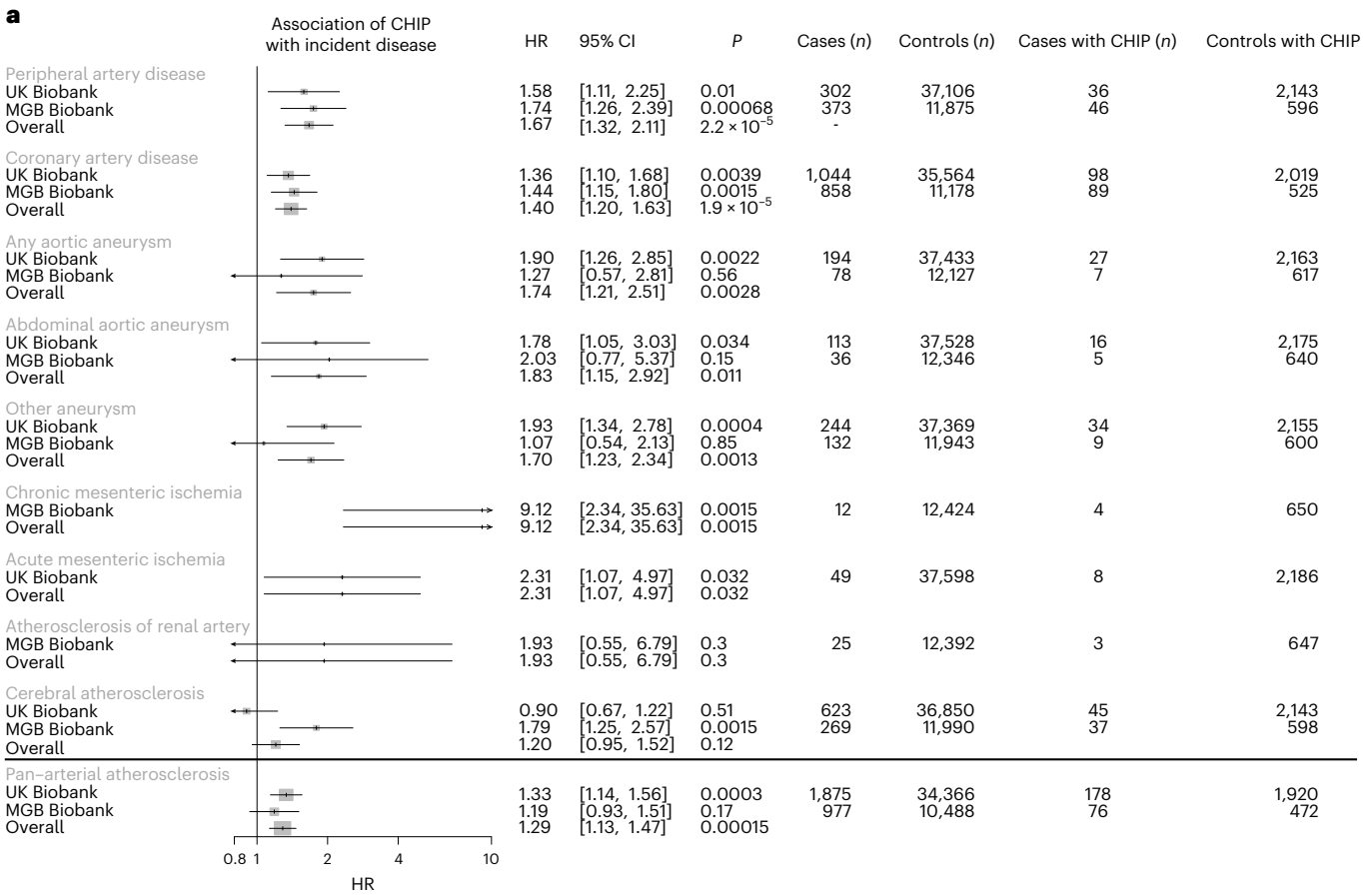


Fig. 2 | CHIP and incident pan-arterial atherosclerosis risk. a, Association of CHIP with nine total incident atherosclerotic diseases separately and combined in a 'pan-arterial atherosclerosis' phenotype in the UKB and the MGBB and meta-analyzed across both studies ('overall'). Error bars are centered at the HR and show the 95% CI for estimates. **b**, Cumulative risk of incident atherosclerosis across the composite 'pan-arterial atherosclerosis' phenotype stratified by no CHIP, small CHIP (VAF < 10%) and large CHIP (VAF ≥ 10%) carrier state in the UKB (n = 37,657). **c**, Association of CHIP VAF with fraction of individuals developing pan-arterial atherosclerosis in the UKB (n = 37,657). Error bands reflect the standard error of a generalized binomial additive model with integrated smoothness fit to the data.

20% KO-BMT mice revealed substantially higher chimerism in aortic macrophages (~76% *Trp53*$^{-/-}$) than in blood classical monocytes (~48%), the main source of plaque macrophages (Fig. 5b). Immunofluorescent staining of aortic root sections confirmed that plaque macrophages are predominantly *Trp53*$^{-/-}$ in 20% KO-BMT mice (Fig. 5c). These findings suggest that p53-deficient macrophages have a selective advantage to expand within the atherosclerotic plaque.

Increased proliferation of murine p53-deficient macrophages

Atherosclerotic plaques in mice carrying p53-deficient hematopoietic cells exhibited a significant increase in the proportion of cells positive for Ki-67, a marker of proliferating cells (Extended Data Fig. 8a). Flow cytometry analysis indicated that this is, at least in part, driven by increased proliferation of p53-deficient macrophages, as there was a two-fold increase in the frequency of Ki-67^{+} cells within

https://doi.org/10.1038/s44161-022-00206-6

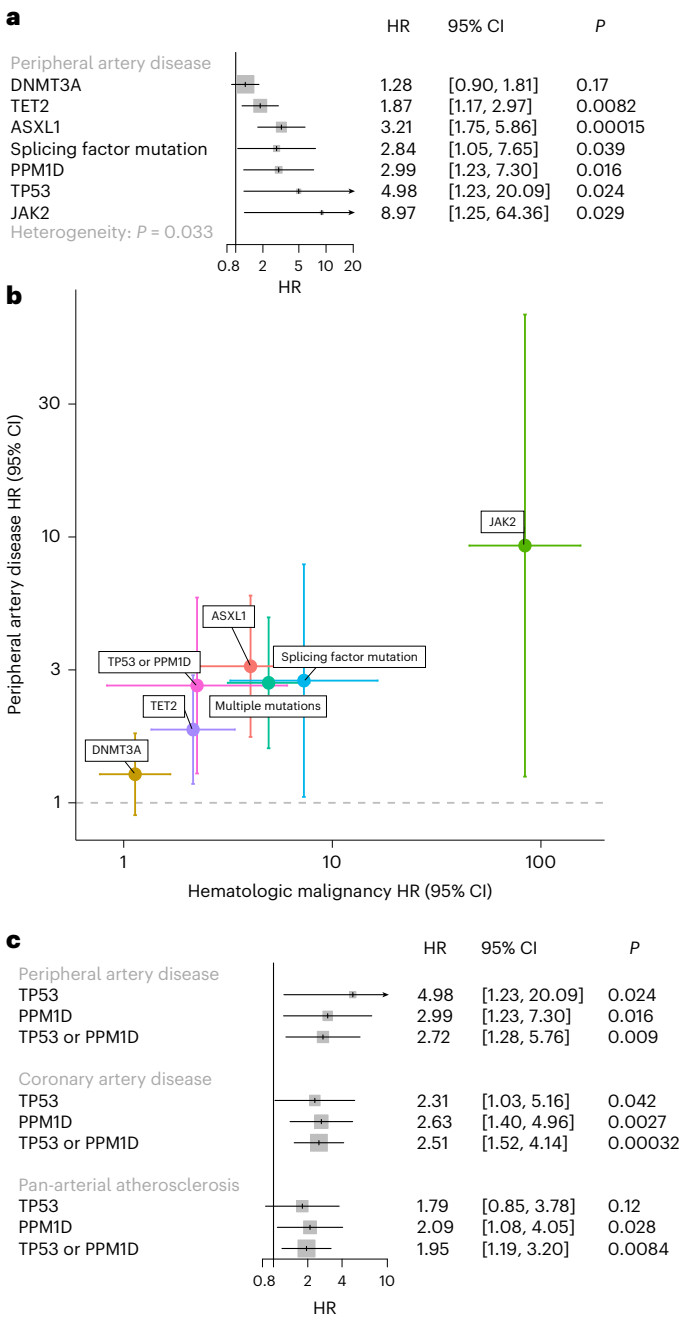


Fig. 3 | Gene-specific association of CHIP with incident PAD. a, CHIP–PAD association analyses stratified by putative CHIP driver gene. Results after meta-analysis across the UKB and the MGBB (n = 50,122) are shown. Error bars are centered at the HR and show the 95% CI for estimates. **b**, Gene-specific comparison of HR and 95% CI for hematologic malignancy (x axis) and PAD (y axis) in the UKB (n = 37,657). Error bars are centered at the HR and show the 95% CI for estimates. **c**, Association of DDR CHIP (*PPM1D or TP53*) with incident PAD, CAD and pan-vascular atherosclerosis. Results across UKB (n = 37,657) and MGBB (n = 12,465) were combined using an inverse variance-weighted fixed effects meta-analysis. Error bars are centered at the HR and show the 95% CI for estimates.

the CD45.2$^+$*Trp53*$^{-/-}$ aortic macrophage population compared to CD45.2$^+$*Trp53*$^{+/+}$ macrophages (Fig. 6a). Consistent with this in vivo observation, cultured *Trp53*$^{-/-}$ macrophages exhibited accelerated mitotic cell cycle progression, with a >2-fold increase in the percentage of DNA-replicating (S-phase) cells upon stimulation with macrophage colony stimulating factor (MCSF), a major determinant of plaque macrophage proliferation[23] (Fig. 6b). Furthermore, *Trp53*$^{-/-}$ macrophages showed an increased percentage of S-phase cells when proliferating asynchronously and higher rates of 5-bromo-2′-deoxyuridine (BrdU) incorporation when stimulated with MCSF (Extended Data Fig. 8b,c). p53 was expressed in quiescent macrophages at the transcript and protein level and further induced after MCSF stimulation (Fig. 6c,d), suggesting a central role of p53 in the normal regulation of macrophage cell cycle progression. Consistent with this possibility, transcriptomic profiling by RNA sequencing (RNA-seq) revealed a widespread alteration of gene expression in MCSF-stimulated *Trp53*$^{-/-}$ macrophages compared to WT controls (Extended Data Fig. 9). Functional annotation of differentially expressed genes showed that categories related to quantity of myeloid cells, cell cycle progression and immune cell proliferation were among the most upregulated in *Trp53*$^{-/-}$ macrophages (Extended Data Fig. 9). qPCR analysis at different timepoints of MCSF stimulation further confirmed that p53-deficient-macrophages exhibit major changes in the expression of pivotal regulators of cell cycle entry and progression, such as *Cdkn1a*/p21^{Cip1} and cyclin B1 (Fig. 6e). RNA-seq of cultured p53-deficient macrophages also revealed a significantly downregulated expression of several genes related to cell death and apoptosis (Extended Data Fig. 9). However, apoptosis rates were similar in p53-deficient and WT macrophages within atherosclerotic plaques (Supplementary Fig. 2), suggesting that differences in apoptosis are not a relevant contributor to increased macrophage burden and atherosclerotic plaque size in mice exhibiting p53-deficient CHIP. Accordingly, we found that p53 deficiency does not affect oxysterol-induced apoptosis in cultured macrophages (Supplementary Fig. 2). Similarly, neither the uptake of modified low-density lipoprotein (LDL), the uptake of apoptotic cells or the expression of central mediators of these cellular processes were affected by p53 deficiency (Supplementary Fig. 3). p53-deficient neutrophils isolated from 20% KO-BMT mice exhibited an expression of phenotypic markers and a production of reactive oxygen species (ROS) similar to those of WT neutrophils (Supplementary Fig. 4). Neutrophil extracellular traps were not detected in murine atherosclerotic plaques in the experimental conditions used in this study (Supplementary Fig. 5).

Discussion

This study combined exome sequencing data across two biobanks to detect somatic mutations in over 50,000 individuals and observed that the presence of CHIP was significantly associated with an increased risk of developing PAD and atherosclerosis across multiple arterial beds. Increased risk was differentially observed across CHIP driver genes with evidence of a graded relationship with CHIP VAF, with large CHIP clones conferring greater risk of disease. The scale of this study allowed a high-power analysis of gene-specific associations for a variety of CHIP genes, which provided evidence of an association between *TP53*-mutant CHIP and the risk of incident atherosclerotic disease, including both CAD and PAD. Lastly, experimental studies in mice demonstrated that p53-deficient CHIP results in increased size of aortic atherosclerotic plaques, accompanied by expansion of plaque macrophages, supporting a direct contribution of p53-mutant hematopoietic cells to accelerated atherogenesis.

These findings permit several conclusions. First, in humans, CHIP appears to promote atherosclerosis across the entire arterial system, but its effects vary by putative CHIP driver gene. Previous work linked CHIP with increased risk of CAD and early-onset myocardial infarction (MI)[3,4]. We build on these findings by demonstrating that CHIP is also associated with PAD and a composite pan-arterial atherosclerosis outcome reflective of an increased burden of atherosclerosis throughout the human arterial system. Our results demonstrate significant evidence of heterogeneity in CHIP effects based on putative driver gene, which is particularly evident when evaluating the effects of *DNMT3A* mutations. Jaiswal et al.[4] demonstrated an increased risk of

https://doi.org/10.1038/s44161-022-00206-6

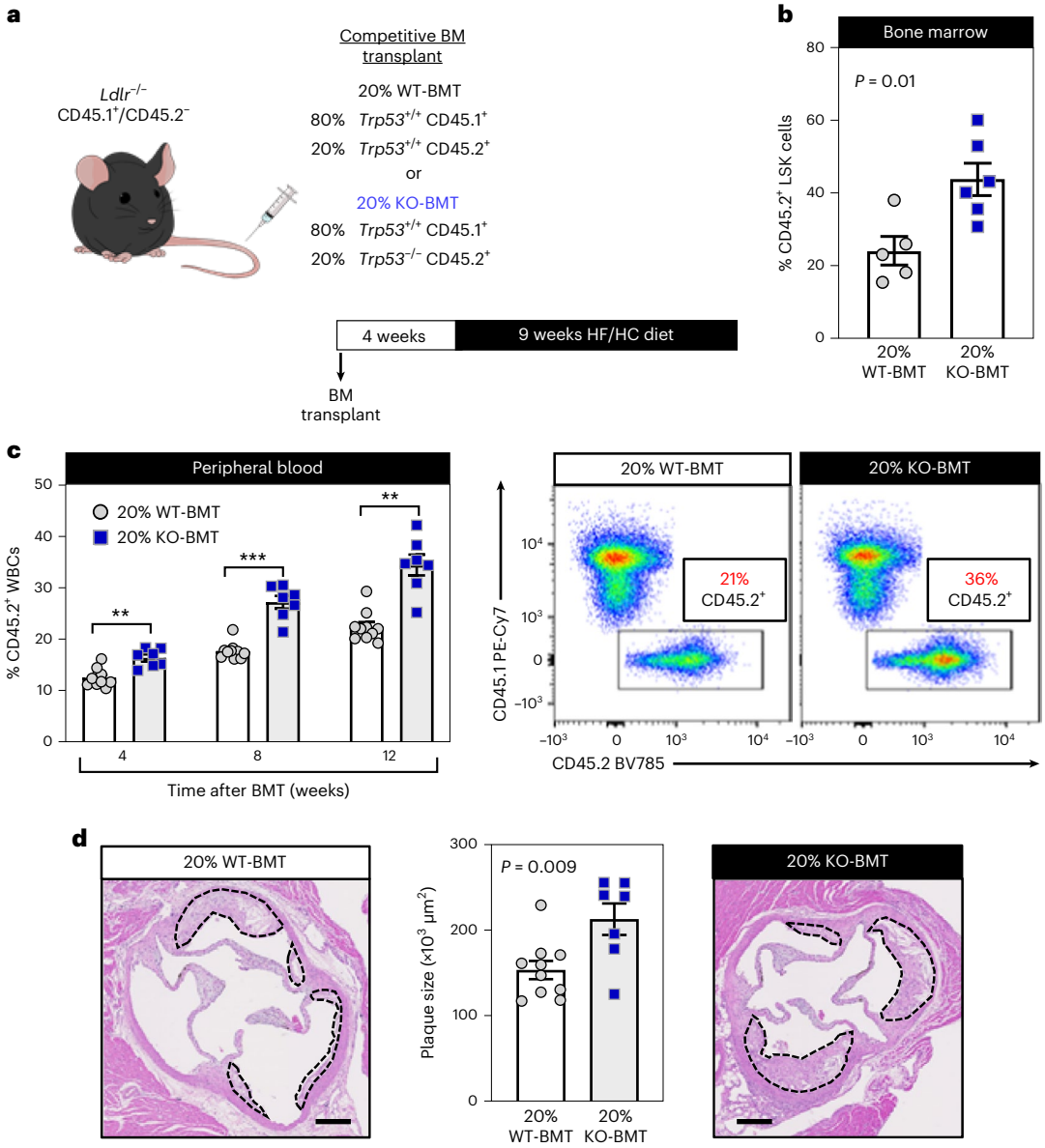


Fig. 4 | Accelerated atherosclerosis in a murine model of TP53 mutation-driven CHIP. a, Summary of the competitive BMT approach and the timeline of these studies. 20% KO-BM male mice and 20% WT-BMT controls were fed an HF/HC diet for 9 weeks, starting 4 weeks after BMT. **b**, Percentage of CD45.2$^+$ cells in the BM LSK population (lineage$^-$Sca1$^+$cKit$^+$ cells) after 9 weeks on an HF/HC diet (13 weeks after BMT), evaluated by flow cytometry (mean ± s.e.m.; $n = 5$ 20% WT-BMT mice and $n = 6$ 20% KO-BMT mice); a two-tailed unpaired t-test was used for statistical analysis. **c**, Percentage of CD45.2$^+$ cell in the white blood cell (WBC) population, evaluated by flow cytometry (mean ± s.e.m.; $n = 10$ 20% WT-BMT mice and $n = 7$ 20% KO-BMT mice); a two-way ANOVA with Sidak's multiple comparison test was used for statistical analysis (**$P < 0.01$ and ***$P < 0.001$). Representative CD45.1/CD45.2 dot plots are shown. **d**, Aortic root plaque size (mean ± s.e.m.; $n = 10$ 20% WT-BMT mice and $n = 7$ 20% KO-BMT mice); a two-tailed unpaired t-test was used for statistical analysis. Representative images of H&E-stained sections are shown; atherosclerotic plaques are delineated by dashed lines. Scale bars, 100 µm.

CAD among individuals with *DNMT3A*-mutant CHIP, but we did not observe a significant effect on the risk of CAD or PAD, despite the fact that *DNMT3A* mutations were the most frequent mutations in our study population. As the field of CHIP continues to evolve, shifting the focus from CHIP as a broad entity to deciphering the specific effects of individual CHIP driver genes will be critical.

Second, DDR gene CHIP appears to confer an increased risk of atherosclerotic CVD. In previous work, Jaiswal et al.[4] demonstrated an increased risk of CAD among individuals with CHIP through *DNMT3A*, *TET2*, *ASXL1* and *JAK2* somatic driver mutations. However, it remained unknown whether individuals who exhibit CHIP driven by mutations in DDR genes have a heightened risk of developing atherosclerotic disease. This gap in knowledge is relevant, as DDR CHIP can be detected in a substantial number of cancer-free individuals and is particularly frequent in patients with cancer and survivors treated with cytotoxic therapies, a population at high CVD risk[24]. In this context, the current study demonstrates that CHIP related to DDR genes (*TP53* and *PPM1D*) confers an increased risk of developing atherosclerosis, both in coronary and peripheral arteries. These findings lend human genetic support to the concept that patients undergoing post-cytotoxic therapy may benefit from surveillance for atherosclerotic conditions in addition to therapy-related myeloid neoplasms. Future studies will be required to assess whether DDR CHIP contributes to the increased CVD risk in patients with cancer and survivors.

https://doi.org/10.1038/s44161-022-00206-6

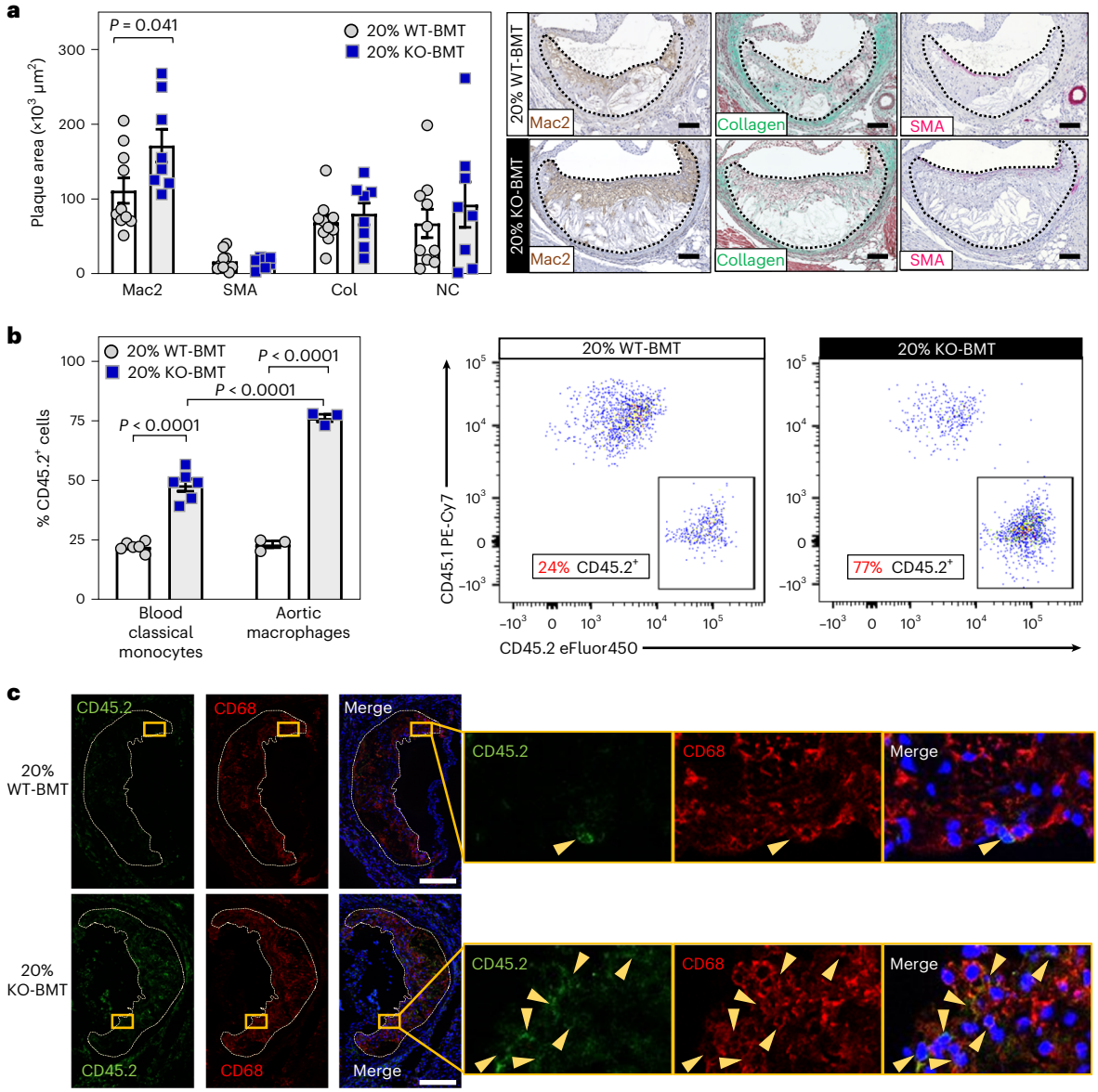


Fig. 5 | Expansion of p53-deficient macrophages within atherosclerotic plaques. a, Histological analysis of plaque composition in 20% KO-BMT female mice and 20% WT-BMT controls, quantified as absolute intimal content of macrophages (Mac2 antigen immunostaining), vascular smooth muscle cells (SMA immunostaining), collagen (Col) (Masson's trichrome staining) and necrotic core (NC) (collagen-free acellular regions); representative images are shown (mean ± s.e.m.; $n = 10$ 20% WT-BMT mice and $n = 8$ 20% KO-BMT mice). Two-tailed unpaired *t*-tests were used for statistical analysis. **b**, Percentage of $CD45.2^+$ cells within the aortic macrophage population ($CD3^-$, $Ly6G^-$, $CD11b^+$ and $F4/80^{High}$) and blood classical monocytes ($CD3^-$, $CD115^{High}$, $Ly6G^-$, $CD43^{Low}$ and $Ly6C^{High}$) of 20% KO-BMT mice and controls, evaluated by flow cytometry (mean ± s.e.m.; $n = 6$ independent mice per BM genotype in classical monocyte analyses and $n = 3$ pools of two aortic arches collected from independent mice in aortic macrophage analyses). A two-way ANOVA with Tukey's multiple comparison test was used for statistical analysis. Representative CD45.1/CD45.2 plots of aortic macrophages are shown. **c**, Immunofluorescent staining and confocal microscopy imaging of $CD45.2^+$ cells (green) and $CD68^+$ macrophages (red) in atherosclerotic plaques of 20% KO-BMT mice and 20% WT-BMT controls. DAPI-stained nuclei are shown in blue. Representative images are shown (sections from $n = 14$ 20% WT-BMT mice and $n = 12$ 20% KO-BMT mice from two independent experiments were examined). Yellow triangles indicate examples of CD45.2/CD68 double-positive cells. Atherosclerotic plaques are delineated by dashed lines. Scale bars, 50 µm.

Third, CHIP related to *TP53* mutations appears to drive atherosclerosis risk via expansion of p53-deficient macrophages within atherosclerotic plaques. Previous experimental studies assessed the role of p53 in atherogenesis using mice engineered to exhibit gain or loss of function of p53 in the whole body or in specific cell types[25–31]. Overall, these studies suggested that p53 exerts atheroprotective actions[27–29], although some null findings were also reported[25,30]. Here we demonstrate that carrying a fraction of p53-deficient blood cells is sufficient to accelerate atherosclerosis development. Although the contribution of additional mechanisms cannot be ruled out, this accelerated atherosclerosis seems mainly related to an increase in macrophage burden in the atherosclerotic plaque, as p53-deficient macrophage have a selective advantage to expand within the arterial wall. Mechanistically, this expansion of p53-deficient macrophages seems related mainly to increased proliferation, a major driver of macrophage burden in atherosclerotic plaques[32]. Although p53 expression is typically induced by DNA damage or other kinds of cellular stress, we found that it is expressed in resting conditions in cultured macrophages and further

https://doi.org/10.1038/s44161-022-00206-6

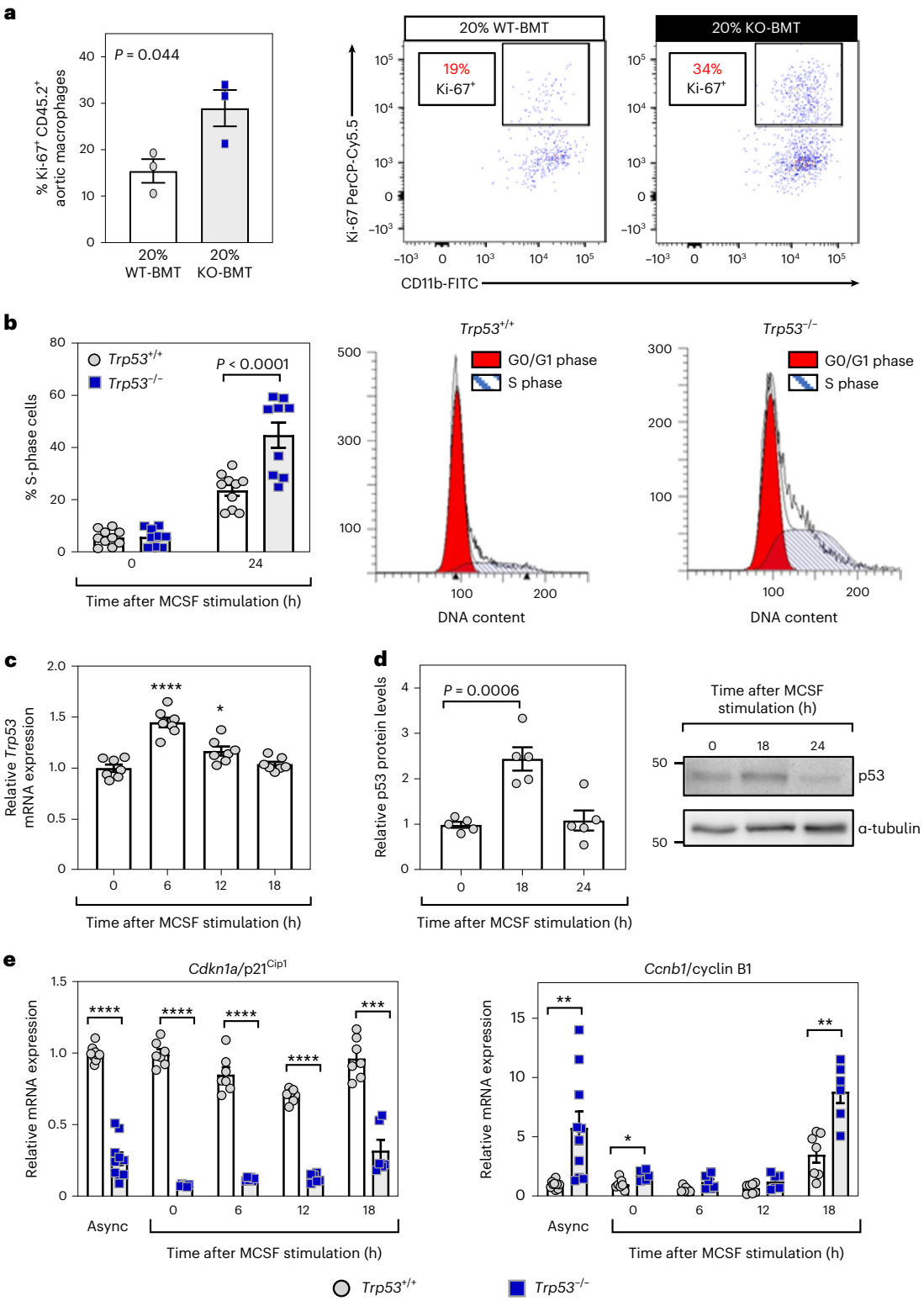


Fig. 6 | Increased proliferation of p53-deficient macrophages. a, Percentage of Ki-67$^+$ proliferating cells within the aortic CD45.2$^+$ macrophage population of 20% KO-BMT mice and controls, evaluated by flow cytometry (mean ± s.e.m.; $n = 3$ pools of two murine aortic arches per BM genotype); a two-tailed unpaired t-test was used for statistical analysis. Representative plots are shown. **b**, Percentage of S-phase cells in cultures of *Trp53*$^{-/-}$ and *Trp53*$^{+/+}$ murine BM-derived macrophages, evaluated by propidium iodide staining of cellular DNA content and flow cytometry. Treatment with MCSF was used to induce cell cycle entry and progression in quiescent G0-synchronized macrophages (mean ± s.e.m.; $n = 10$ *Trp53*$^{+/+}$ mice and $n = 9$ *Trp53*$^{-/-}$ mice). A two-way repeated-measures ANOVA with Sidak's multiple comparisons test was used for statistical analysis. **c**,**d**, qPCR (**c**, $n = 7$ mice) and western blot (**d**, $n = 5$ mice) analyses of *Trp53* expression in cultured primary macrophages after MCSF mitogenic stimulation. A representative blot is shown (mean ± s.e.m.; $^*P < 0.05$ and $^{****}P < 0.0001$ versus baseline); a one-way repeated-measures ANOVA with Tukey's multiple comparison test was used for statistical analysis. **e**, Expression of cell cycle regulators *Cdkn1a*/p21Cip1 and *Ccnb1*/cyclin B1 in cultured *Trp53*$^{-/-}$ and *Trp53*$^{+/+}$ macrophages proliferating asynchronously (Async, mean ± s.e.m.; $n = 10$ mice per genotype) or after MCSF stimulation (mean ± s.e.m.; $n = 7$ *Trp53*$^{+/+}$ mice and $n = 6$ *Trp53*$^{-/-}$ mice). A two-tailed unpaired t-test (with Welch's correction for *Ccnb1* analysis) and a two-way repeated-measures ANOVA with Sidak's multiple comparison test, respectively, were used for statistical analysis ($^{**}P < 0.01$, $^{***}P < 0.001$ and $^{****}P < 0.0001$).

induced by mitogenic stimulation, suggesting a physiological role in the regulation of macrophage cell cycle progression in the absence of stressors. Accordingly, p53 mutant macrophages exhibited accelerated cell cycle kinetics. Overall, these experimental findings provide support to the notion that *TP53* CHIP mutations contribute directly to accelerated atherosclerosis and highlight major differences in the mechanisms underlying accelerated atherosclerosis in CHIP driven by mutations in *TP53* or epigenetic regulators. In contrast to previous reports related to CHIP driven by mutations in the epigenetic regulatory genes *DNMT3A*, *TET2* and *ASXL1* (refs. 4, 14,20–22,33), we did not observe a significant effect of p53 deficiency on the expression of the pro-inflammatory cytokines IL-1β and IL-6. These results are consistent with previous human data showing that circulating levels of pro-inflammatory cytokines are significantly increased in carriers of CHIP driven by mutations in these epigenetic regulators but not in carriers of mutations in *TP53* or *PPM1D*[11]. These mechanistic differences between *TP53* CHIP and *DNMT3A/TET2/ASXL1* CHIP require consideration when designing preventive care strategies targeting the effects of CHIP on atherosclerosis. Thus, although the pathogenic effects of TET2 CHIP may be prevented by targeting IL-1β-driven inflammation[14], accelerated atherosclerosis associated with *TP53* mutations may be better tackled by other strategies. Additional experimental and clinical studies should evaluate these opportunities for personalized medicine in the context of CHIP.

Lastly, a recent publication has generated a novel hypothesis regarding the relationship between CHIP and atherosclerosis[34]. In their analysis, Heyde et al.[34] suggest that CHIP may be a symptom of the atherosclerosis trait complex (the interplay of chronic inflammation, hyperlipidemia and arterial plaque) rather than a causal risk factor, based on murine and human observations that atherosclerosis may accelerate HSC proliferation and mutant clone evolution. Although their study engenders an intriguing hypothesis, some of the observations in our current study contradict this interpretation. Heyde et al. state that the uniform HRs previously observed for the CHIP–CAD association provide supportive evidence for the notion that atherosclerotic disease precedes the development of clonal hematopoiesis. However, in our study, we identify significant heterogeneity of incident PAD and CAD effect sizes when stratifying by CHIP genes as well as a dose–response relationship between CHIP clone size (or VAF) and incidence of atherosclerosis, suggesting that CHIP genes vary in their effect on incident PAD risk and that expanded CHIP variants have a stronger influence on future atherosclerotic risk. Thus, given (1) the association of CHIP with incident atherosclerotic disease, (2) the dose–response relationship with CHIP clone size and heterogeneity of effect size by CHIP gene and (3) that mutations in *TP53*, *TET2* and *JAK2* (ref. 35) appear to drive atherosclerosis through disparate mechanisms, clonal hematopoiesis acting as a causal risk factor, at least when driven by certain mutations, appears to be a more parsimonious explanation to explain the association between CHIP and the risk of atherosclerotic disease

Several limitations merit mention. First, our atherosclerotic disease phenotypes are based on EHR data and may result in misclassification of case status. Such misclassification would likely reduce statistical power for discovery and, on average, bias results toward the null. Second, selection bias from differential loss of follow-up, volunteer bias and missingness in covariates may be present given the nature of the genetic biobanks used in this study. Third, although we maximized the number of participants in our analysis of CHIP and additional atherosclerotic diseases, it may still have been underpowered for certain phenotypes. Fourth, although *TP53* somatic mutations most commonly occur in patients undergoing post-cytotoxic chemotherapy, we did not specifically examine this subset of individuals given their small sample size, and future studies will be needed to validate these findings in this population. Lastly, our mouse model of CHIP was based on p53-deficient homozygous cells, whereas most cancer-free individuals with *TP53* mutant CHIP can be expected to carry monoallelic mutations. However, the most frequent *TP53* mutations linked to CHIP act in a dominant-negative manner and, even when monoallelic, lead to phenotypes similar to full biallelic *TP53* loss of function, as shown previously[16]. Thus, *Trp53*$^{-/-}$ cells provide an overall appropriate model of these CHIP-related mutations. Additional studies will be required to assess the effects of specific *TP53* mutations on atherosclerosis.

In conclusion, CHIP, including *TP53* CHIP, is associated with incident atherosclerosis disease in humans, and experimental mouse studies suggest a direct contribution of somatic *TP53* mutations to atherosclerosis development through an expansion of plaque macrophages. These observations expand knowledge of the pathophysiological relevance of CHIP in atherosclerosis and highlight the value of single-gene analyses that shed light onto the effects of mutations in specific CHIP driver genes.

Methods

Cohorts and exclusion criteria

The UKB is a population-based cohort of approximately 500,000 participants recruited from 2006 to 2010 with existing genomic and longitudinal phenotypic data and a median 10-year follow-up[36]. Baseline assessments were conducted at 22 assessment centres across the UK with sample collections including blood-derived DNA. At the time of analysis, of 49,960 individuals with whole-exome sequencing data available, we analyzed 37,657 participants consenting to genetic analyses after our exclusion criteria. Use of the data was approved by the Massachusetts General Hospital institutional review board (protocol 2013P001840) and facilitated through UKB application 7089.

The MGBB contains genotypic and clinical data from >105,000 patients who consented to broad-based research across seven regional hospitals and a median 3-year follow-up[37]. Baseline phenotypes were ascertained from the electronic medical record and surveys. We analyzed 12,465 individuals consenting to genetic analysis after our exclusion criteria. Use of the data was approved by the Massachusetts General Hospital institutional review board (protocol 2020P000904).

Across both cohorts, we excluded individuals with prevalent hematologic cancer, individuals without genotypic–phenotypic sex concordance and one of each pair of first-degree or second-degree relatives at random. For the UKB, samples were further restricted to individuals with Townsend deprivation index, a marker of socioeconomic status, available for analysis. Follow-up time was defined as time from enrollment to disease diagnosis for cases or to censorship or death for controls.

Whole-exome sequencing and CHIP calling

UKB whole-exome sequencing was generated from whole-blood-derived DNA at the Regeneron Sequencing Center[38]. MGBB whole-exome sequencing was generated using whole-blood-derived DNA using Illumina sequencing (mean coverage 55×). Somatic CHIP variants were detected with GATK MuTect2 (version 1) software with parameters as previously described[7,11]. In brief, aligned short-read exome sequences were analyzed using GATK MuTect2 software to detect putative somatic genetic variants, which were then filtered to exclude common sequencing artifacts (including minimum variant read counts of three reads, required evidence of a variant on both forward and reverse reads and exclusion of frameshift mutations within homo-polymer repeat regions). Common germline variants and sequencing artifacts were excluded as before. Samples were annotated with the presence of any CHIP if MuTect2 identified one or more of a pre-specified list of pathogenic somatic variants reported in the literature and/or the Catalog of Somatic Mutations in Cancer (COSMIC, http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/), after manual curation, as previously described[3,4]. Additionally, samples were annotated with the presence of large CHIP (VAF > 10%), as larger CHIP clones have previously been more strongly associated with adverse clinical outcomes[7].

Phenotype definitions

Across both UKB and MGBB, PAD was defined by grouping together ICD-10 and ICD-9 billing codes for aortic atherosclerosis (I70), peripheral vascular disease (I73.8 and I73.9) and operative procedures, including amputation of leg (X09.3–5), bypass of artery of leg (L21.6, L51.3, L51.6, L51.8 and L59.1–8), endarterectomy or angioplasty of leg artery (L52.1–2, L54.1,4,8, L60.1–2 and L63.1,5) and other transluminal operations on leg arteries or peripheral stent placement (L63.9 and L66.7), as previously described[12]. Additionally, in the UKB, self-reported peripheral vascular disease, leg claudication/intermittent claudication, arterial embolism, femoral-popliteal leg artery bypass, leg artery angioplasty with or without stent or amputation of leg were also included (data fields 20002 and 20004), as performed previously[12]. CAD and cerebral atherosclerosis phenotype definitions for each cohort are detailed in Supplementary Tables 8 and 9. Other atherosclerotic conditions were defined using the Phecode Map 1.2 (ref. 39) ICD-9 (https://phewascatalog.org/phecodes) and ICD-10 (https://phewascatalog.org/phecodes_icd10) phenotype groupings for 'abdominal aortic aneurysm', 'aortic aneurysm', 'other aneurysm', 'chronic vascular insufficiency of intestine'/'acute vascular insufficiency of intestine' (combined into a mesenteric ischemia phenotype) and 'atherosclerosis of renal artery'. The composite atherosclerosis phenotype was created by combining all analyzed atherosclerosis phenotypes (including PAD, CAD, cerebral atherosclerosis, abdominal aortic aneurysm, aortic aneurysm, other aneurysm, chronic vascular insufficiency of intestine, acute vascular insufficiency of intestine and atherosclerosis of renal artery) into one phenotype, whereby the first instance across all of these phenotypes was used to determine time of first diagnosed atherosclerotic disease for survival analysis. Other phenotypic covariates (never/prior/current smoking status, hypertension, hyperlipidemia, principal components of ancestry, etc.) were used as previously defined[40].

Association analysis

In the UKB and the MGBB, a traditional Cox proportional hazards model was used using the Survival package in R-3.5, adjusting for age, age^2, sex, smoking status, normalized Townsend deprivation index as a marker of socioeconomic status (only available in the UKB) and the first ten principal components of genetic ancestry. Demographic and clinical characteristics found to differ between individuals with and without CHIP were tested using chi-squared (categorical) and Wilcoxon rank-sum (continuous) tests with a two-tailed $P < 0.05$ determining significance. Sensitivity analysis including other covariates (normalized body mass index, prevalent hypertension, prevalent type 2 diabetes and prevalent hyperlipidemia) was not found to significantly change associations with PAD in the UKB (Extended Data Fig. 10). Additional sensitivity analyses were used in the UKB, including propensity score adjustment as well as a marginal structural Cox proportional hazards model estimated through stabilized IPTW[41] to estimate the total causal effect of CHIP on PAD. Further details of propensity score methods and stabilized IPTW analysis are described below. For our primary outcome, PAD, results were combined across the UKB and the MGBB using an inverse variance-weighted fixed effects meta-analysis, and a two-tailed association $P < 0.05$ determined statistical significance. In our secondary analysis of CHIP with ten additional incident atherosclerotic diseases, a two-tailed Bonferroni P value threshold of $P < 0.05/10 = 0.005$ was used to declare statistical significance.

Epidemiologic causal inference methods used to assess the CHIP–PAD association

The propensity score is the probability of treatment assignment (that is, CHIP or large CHIP) conditional on the minimally sufficient set of confounders. Using a propensity score as a covariate allows observational studies to mimic some of the characteristics of a randomized controlled trial in the sense that, conditional on the propensity score, the distribution of covariates will be similar between the CHIP and the non-CHIP carriers. In our scenario, the propensity score will help assess the 'dimensionality problem' in that it will reduce the number of covariates from 15 covariates in the minimally sufficient set to one. Here, the propensity score is calculated using multivariable logistic regression where the outcome was CHIP and covariates consisted of the 15 covariates in the minimally sufficient set (age and age^2, included to fully account for age given significant associations in multivariate models, sex, smoking status and the first ten principal components of genetic ancestry). The propensity score was then predicted for all samples in the UKB for CHIP and, separately, for large CHIP.

Stabilized IPTW uses weights based on the propensity score to create a pseudo-population in which the distribution of measured covariates is independent of CHIP status—that is, removing the arrow from the confounders to the exposure. The conditional exchangeability assumption made in IPTW is that if the set of covariates is sufficient to block all backdoor paths from CHIP to PAD, then all confounding is eliminated in the pseudo-population, and, thus, the association between CHIP and PAD in the synthetic population is an estimate of the causal effect. To estimate the IPTW stabilized weight (SW) for each individual in the UKB, we used the following equation for CHIP carriers and controls, where PS reflects the propensity score for that individual:

$$SW^{CHIP+} = \frac{\Pr(CHIP = 1)}{\Pr(CHIP = 1|covariates)} = \frac{\Pr(CHIP = 1)}{PS}$$

$$SW^{CHIP-} = \frac{1 - \Pr(CHIP = 1)}{\Pr(CHIP = 0|covariates)} = \frac{\Pr(CHIP = 0)}{PS}$$

Mice

C57Bl/6J *Trp53*$^{-/-}$ mice were obtained from the Jackson Laboratory. *Ldlr*$^{-/-}$ mice carrying the CD45.1 isoform of the CD45 hematopoietic antigen were generated by crossing *Ldlr*$^{-/-}$ mice from the Jackson Laboratory and B6.SJL-*PtprcaPepcb*/BoyCrl mice from Charles River Laboratories. Animal experiments followed protocols approved by the institutional ethics committee at the Centro Nacional de Investigaciones Cardiovasculares and conformed to EU Directive 86/609/EEC and Recommendation 2007/526/EC regarding the protection of animals used for experimental and other scientific purposes, enforced in Spanish law under Real Decreto 1201/2005. All mice were maintained on a 12-hour light/dark schedule in a specific pathogen-free animal facility in individually ventilated cages and were given food and water ad libitum. Ambient temperature in the animal facility was 20–24 °C, and relative humidity was 45–65%.

Competitive bone marrow transplantation and atherosclerosis induction in mice

Next, 8–12-week-old CD45.1+ *Ldlr*$^{-/-}$ recipients were transplanted with suspensions of BM cells containing 20% CD45.2$^+$ *Trp53*$^{-/-}$ cells and 80% CD45.1$^+$ *Trp53*$^{+/+}$ cells (20% KO-BMT mice) or 20% CD45.2$^+$ *Trp53*$^{+/+}$ cells and 80% CD45.1$^+$ *Trp53*$^{+/+}$ cells (20% WT-BMT mice), similarly to previous studies[14] (Supplementary Fig. 1). BM cells were isolated from femurs and tibias of 6–12-week-old donor mice after euthanasia. Recipient *Ldlr*$^{-/-}$ mice were exposed to two doses of 450 rad 3 hours apart. After the second irradiation, each recipient mouse was injected with 10^7 BM cells intravenously. Water was supplemented with antibiotics for 7 days before transplant and for 14 days after transplant. Mice that did not recover full pre-irradiation body weight 28 days after transplant were excluded from further analysis. Starting 4 weeks after BMT—an interval sufficient for BM reconstitution in our experimental conditions—mice were fed an HF/HC Western diet (Harlan Teklad, TD.88137, Adjusted Calories Diet; 42% from fat, 0.2% cholesterol) to promote hypercholesterolemia and the development of atherosclerosis. Mice were maintained on an HF/HC diet for 9 weeks unless otherwise stated.

Analysis of aortic atherosclerosis burden in mice

Mice were euthanized, and plasma was collected in EDTA-coated tubes for cholesterol level quantification by an enzymatic assay (Cholesterol E, Wako Diagnostics). Aortas were removed after in situ perfusion with PBS injected through the left ventricle of the heart. Tissue fixation was achieved by immersion in 4% paraformaldehyde (PFA) in PBS overnight at 4 °C. Aortic tissue was then embedded in paraffin to obtain histological sections from the aortic root as determined by the location of the aortic valve leaflets. Alternatively, unfixed aortic root specimens were embedded in Tissue-Tek OCT compound. An operator who was blinded to genotype quantified plaque size in aortic root sections by computer-assisted morphometric analysis of microscopy images. For each mouse, atherosclerotic plaque size in aortic root cross-sections was calculated as the average of five independent sections separated by ~16 µm. Atherosclerotic plaque composition was examined by immunohistochemical techniques. Vascular smooth muscle cells were identified with an alkaline phosphatase-conjugated mouse anti-smooth muscle α-actin (SMA) monoclonal antibody (clone 1A4, Sigma-Aldrich) and Vector Red Alkaline Phosphatase Substrate (Vector Laboratories). Macrophages were detected with a rabbit anti-Mac2 monoclonal antibody (Cedarlane Labs), a biotin-conjugated goat anti-rat secondary antibody, streptavidin-HRP and DAB substrate (all from Vector Laboratories), with hematoxylin counterstaining. Ki-67 was detected with a rabbit anti-Ki-67 antibody (clone SP6, Abcam), a biotin-conjugated horse anti-rabbit secondary antibody and streptavidin-HRP. Collagen and lipid content were determined by a modified Masson's trichrome staining and Oil Red O (ORO) staining, respectively. Microscopy images were analyzed using Fiji ImageJ software using the Color Deconvolution plugin. Plaque cell apoptosis was quantified by TUNEL analysis (In situ Cell Death Detection Kit, Fluorescein, Roche) according to manufacturer instructions. Before apoptotic cell staining, sections were blocked with 3% goat serum in PBS and co-stained with Alexa Fluor 647-conjugated anti-CD45.2 antibody (BioLegend), anti-CD68 primary antibody (Bio-Rad) and Alexa Fluor 568-conjugated goat anti-rat IgG secondary antibody (Thermo Fisher Scientific). Nuclei were counterstained with DAPI. Images were acquired on a Leica SP8 Lightning-Navigator confocal microscope with LAS X 3.1.1.15751 software, using a z-stack of 5-µm-thickness frames. Results are shown as the percentage of TUNEL-positive cells within the total plaque cell population, the plaque macrophage population or the $CD45.2^+$ macrophage population in the aortic root. A list of antibodies and additional information, including antibody dilutions, can be found in Supplementary Table 10.

Flow cytometry analyses of blood and tissue samples

Peripheral blood was obtained from the facial vein and collected into EDTA-coated tubes. BM cells were flushed out of two femurs and two tibias per mouse. Aortic arches were digested for 45 minutes at 37 °C in RPMI containing 10% FBS and 0.25 mg ml^{-1} of Liberase TM (Roche Life Science). Red blood cells were lysed in all samples by treatment with 1× Red Blood Cell Lysis Buffer (eBioscience, Thermo Fisher Scientific) for 5 minutes on ice. BM $lineage^-$ cells were defined as negative for CD11b, Gr-1, Ter119, B220, CD3e and CD127. BM LSK was defined as $lineage^-Sca1^+c\text{-}Kit^+$. Blood classical monocytes were identified as $CD45^+$, $CD115^{High}$, $CD43^{Low}$ and $Ly6C^{High}$; patrolling monocytes as $CD45^+$, $CD115^{High}$, $CD43^{High}$ and $Ly6C^{Low}$; neutrophils as $CD45^+$, $CD115^{Int}$ and $Ly6G^+$; T lymphocytes as $CD45^+$, $CD115^-$, $B220^-$ and $CD3^+$; B lymphocytes as $CD45^+$, $CD115^-$, $CD3^-$ and $B220^+$; and aortic macrophages as $CD45^+$, $CD3^-$, $Ly6G^-$, $CD11b^+$ and $F4/80^{High}$. Gating strategies are summarized in Supplementary Fig. 6. Samples were stained with combinations of biotinylated and/or fluorescently labeled antibodies in PBS with 1% FBS for 30 minutes on ice. The following fluorescent antibodies were used for staining and flow cytometry analysis (additional information, including antibody dilutions, can be found in Supplementary Table 10): eFluor450-conjugated anti-CD45.2, FITC-conjugated anti-CD45.2, Pe-Cy7-conjugated anti-CD45.1, FITC-conjugated anti-CD11b, eFluor450-conjugated anti-CD11b, PE-conjugated anti-CD115, PE-eFluor610-conjugated anti-CD3 and FITC-conjugated anti-CD4 (from eBioscience); APC-Cy7-conjugated anti-B220, PerCP-Cy5.5-conjugated anti-Ly6G, BV510-conjugated anti-Ly6G, BV711-conjugated anti-CD43, PE-Cy7-conjugated anti-c-Kit and Alexa Fluor 647-conjugated anti-Sca-1 (from BD Biosciences); PE-conjugated anti-F4/80 (from R&D Systems); and PerCP-Cy5.5-conjugated anti-CD45.1, BV510-conjugated anti-CD8a and PerCP-Cy5.5-conjugated anti-mouse Ki-67 (from BioLegend). Fixation/permeabilization for Ki-67 intracellular staining was achieved using commercially available kits following the manufacturer's instructions (Foxp3 Transcription Factor Staining Buffer Set, eBioscience, Thermo Fisher Scientific). Dead cells were excluded from analysis by DAPI staining in unfixed samples and by LIVE/DEAD Fixable Near-IR staining (Thermo Fisher Scientific) in fixed samples. BD LSRFortessa and BD FACSymphony cytometers were used for data acquisition with BD FACSDiva software (BD Biosciences). Data were analyzed with FlowJo software (Tree Star).

Murine BM-derived macrophage culture and cell cycle analysis

BM-derived macrophages (BMDMs) were obtained by differentiating femoral BM cells for 7 days in RPMI medium supplemented with antibiotics, 10% FBS and 100 ng ml^{-1} of MCSF. Cell cycle dynamics were analyzed both in asynchronously proliferating macrophages and after synchronization in G0 phase by MCSF deprivation. MCSF concentration in cell culture medium was decreased to 5 ng ml^{-1} for 48 hours to induce progressive synchronization of macrophages in G0 phase (that is, quiescence), and cell cycle re-entry was induced by treatment with 100 ng ml^{-1} of MCSF. Macrophages were trypsinized and collected by centrifugation for 5 minutes at 300*g*. After fixation in 80% ethanol for at least 1 hour at −20 °C, cells were incubated for at least 30 minutes with 50 µg ml^{-1} of propidium iodide containing 0.25 mg ml^{-1} of RNAse A (both from Sigma-Aldrich). Labeled cells were analyzed in a BD FACSCanto flow cytometer with FACSDiva software (BD Biosciences), and DNA histograms were fitted into cell cycle distributions using ModFit 3.0 software (Verity Software House). Macrophage proliferation was also assessed by evaluating BrdU incorporation into BMDM after synchronization in G0 phase by MCSF deprivation and induction of cell cycle re-entry by treatment with 100 ng ml^{-1} of MCSF together with 30 µM BrdU (Sigma-Aldrich) for 16 hours. To detect BrdU incorporation, BMDMs were fixed in ethanol at −20 °C for 20 minutes and subjected to acid treatment (10 mM citrate buffer pH 6.0). After neutralization, samples were then blocked with 5% goat serum in PBS–0.1% Tween and incubated with anti-BrdU primary antibody (Thermo Fisher Scientific), followed by Alexa Fluor 488-conjugated goat anti-mouse IgG secondary antibody (Thermo Fisher Scientific). Additional information on these antibodies, including antibody dilutions, can be found in Supplementary Table 10. Nuclei were counterstained with DAPI. Images were captured on a Nikon Eclipse microscope with Nikon AR 4.30.02 (b925) 64-bit software and analyzed using Fiji ImageJ software and the StartDist plugin.

Gene expression analysis by qPCR, mRNA sequencing and western blot

Total RNA from cultured macrophages or aortic arch tissue was isolated using TRIzol reagent and RNeasy kits (Qiagen). RNA was reverse transcribed with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems), and qPCR was performed with SYBR Green PCR Master Mix (Applied Biosystems) in an AB7900 real-time PCR system. Results were analyzed with the ΔΔCt method. The average of 36B4 and β-actin was used as reference for normalization. Primer sequences are listed in Supplementary Table 11.

Libraries for mRNA sequencing were prepared using the SMART-Seq version 4 Ultra Low Input RNA Kit (Clontech-Takara). Full-length cDNAs were processed using a Nextera XT DNA Library Preparation Kit (Illumina). Quality control was performed using a High Sensitivity DNA

ScreenTape Assay with a 4200 TapeStation System (Agilent). Libraries were then multiplexed in an equimolar pool and sequenced using a NextSeq 550 Platform (Illumina). Single-end unstranded libraries (75 bp) were first subjected to a quality check FastQC version 0.11.8 and aligned to the reference genome assembly mm10 with the following settings (STAR --outFilterMultimapNmax 20 --alignSJoverhangMin 8 –alignSJDBoverhangMin 1 --outFilterMismatchNmax 999 --outFilterMismatchNoverReadLmax 0.04 --alignIntronMin 20 --alignIntronMax 1000000 --alignMatesGapMax 1000000 --outSAMattributes NH HI NM MD).

RNA-seq quality assessment was performed using the RSeQC version 3.0.1 tool. Differential expression analysis was assessed with the edgeR version 3.30.0 R Bioconductor package (PubMed ID: 19910308), and only genes with a cutoff expression >1 counts per million (CPM) in at least three samples were included in the analysis. The TMM method was applied to normalize gene counts, and the glmQLFTest function was used to perform differential analysis, with a minimum cutoff of absolute fold change ≥1.5 and false dicovery rate (FDR) ≤ 0.05. A complete list of differentially expressed genes and statistical information (*P* value and FDR) for each condition are available at the Gene Expression Omnibus (GSE184420). Functional enrichment analysis was performed using Ingenuity Pathway Analysis software (IPA version 01–13, Qiagen). GOplot representations were used to combine and integrate the expression data with the results of the functional analysis[42].

For western blot analysis of protein levels, protein extracts from cultured macrophages were obtained using ice-cold lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM NA2EDTA, 1 mM EGTA and 1% Triton) supplemented with protease and phosphatase inhibitors (Roche Applied Science). Equal amounts of protein lysates were resolved by SDS-PAGE, and the following antibodies were used for immunoblotting: anti-p53 (Cell Signaling Technology) and HRP-conjugated anti-α tubulin (Abcam). Additional information on these antibodies, including antibody dilutions, can be found in Supplementary Table 10. An ImageQuant LAS 4000 biomolecular imaging system with ImageQuant LAS4000 software (GE Healthcare) was used for image acquisition, and Fiji ImageJ software was used for band densitometric analysis. p53 protein levels were normalized to α-tubulin.

Analysis of modified lipoprotein uptake and efferocytosis in cultured macrophages

Thioglycolate-elicited peritoneal macrophages were cultured in RPMI medium supplemented with antibiotics and 10% FBS. To asses modified LDL uptake, peritoneal macrophages were treated with 5 μg ml^{-1} of fluorescent acetylated LDL (DiI acLDL, Thermo Fisher Scientific) for 1 hour or 3 hours. Macrophage were then trypsinized, collected and analyzed by flow cytometry using a BD FACSymphony flow cytometer with BD FACSDiva software (BD Biosciences). Data were analyzed with FlowJo software (Tree Star). To assess uptake of apoptotic cells (that is, efferocytosis), thymuses from C57Bl/6J mice were collected, homogenized, filtered through a 70-μm cell strainer and labeled fluorescently with 5 μM CellTrace Violet (Thermo Fisher Scientific). Thymocyte apoptosis was induced by exposure to UV light for 10 minutes. One or two hours later, apoptotic thymocytes were added to macrophage cultures at a 5:1 ratio to allow apoptotic cell engulfment by macrophages. After 1 hour, macrophages were washed extensively, collected by trypsinization and stained with FITC-conjugated anti-CD11b (Invitrogen) and PE-conjugated anti-F4/80 antibodies (Tonbo Bioscience). Additional information on these antibodies, including antibody dilutions, can be found in Supplementary Table 10. Samples were analyzed by flow cytometry using a BD FACSCanto flow cytometer with BD FACSDiva software (BD Biosciences). Data were analyzed with FlowJo software.

Analysis of neutrophil phenotype and function

Neutrophil phenotype analyses and ROS quantification were performed in blood samples extracted from experimental mice and treated with RBC hypotonic lysis buffer (0.15 M NH_4Cl, 0.01 M $KHCO_3$ and 0.01 M EDTA in water). For ROS quantification, cells were stimulated with 133 nM of phorbol 12-myristate 13-acetate (PMA) for 20 minutes or PBS as control and then stained with 5 mM dihydrorhodamine 123 (DHR123, Thermo Fisher Scientific) for 20 minutes and with BUV737-conjugated anti-Ly6G antibody (BD Biosciences), PE-conjugated anti-CD11b antibody (Tonbo Biosciences), PE-Cy7-conjugated anti-CD45.1 antibody (BioLegend) and APC-conjugated anti-CD45.2 antibody (BioLegend) for 15 minutes. For surface marker analysis, samples were incubated with BUV737-conjugated anti-Ly6G antibody (BD Biosciences), PE-conjugated anti-CD62L antibody (eBioscience), PerCP-Cy5.5-conjugated anti-CXCR2 antibody (BioLegend), BV510-conjugated anti-CD11b antibody (BioLegend), PECy7-conjugated anti-CD101 antibody (Thermo Fisher Scientific), Alexa Fluor 647-conjugated anti-CD54 antibody (BioLegend), APC-Cy7-conjugated anti-CD45 antibody (BioLegend) and FITC-conjugated anti-CD45.2 antibody (BioLegend) for 15 minutes. Additional information on these antibodies, including antibody dilutions, can be found in Supplementary Table 10. Dead cells were excluded from analysis by DAPI staining. Samples were analyzed in a BD FACSymphony cytometer with BD FACSDiva software (BD Biosciences) for data acquisition. Analysis was performed using FlowJo software.

NETosis was assessed in cryopreserved aortic root sections from hematopoietic chimeric mice, and cryopreserved lung sections from mice were subjected to transfusion-related acute lung injury as positive control[43]. After fixing in 4% PFA for 15 minutes at room temperature, samples were blocked with 5% BSA, 1% normal goat serum, 5% FBS and 0.1% Triton X-100 in PBS in a humid chamber for 1 hour, followed by an incubation with antibodies against citrullinated histone 3 (Abcam), CD31 (Thermo Fisher Scientific) and biotinylated-MPO (R&D Systems) overnight at 4 °C. Sections were then incubated with goat anti-rabbit Alexa Fluor 568 (Life Technologies), goat anti-hamster Alexa Fluor 647 (Jackson ImmunoResearch), Alexa Fluor 488-conjugated streptavidin (BioLegend) and DAPI (Life Technologies) and mounted with Mowiol 4–88 (Mw 31,000, Sigma-Aldrich). Additional information on antibodies, including antibody dilutions, can be found in Supplementary Table 10. Imaging of NETs was performed using a Leica SP5 multi-line inverted confocal microscope with ×63 magnification and LAS-AF 2.7.3. build 9723 software. Images were analyzed using Fiji ImageJ software. An event was considered positive for NETosis when it was triple stained (MPO, cit H3 and DAPI).

Statistical analysis of data in experimental studies

Data are shown as mean ± s.e.m. unless otherwise stated. Statistical significance of differences in experiments with two groups and only one variable was assessed by unpaired Student's *t*-tests (with Welch correction for unequal variance when appropriate) or Mann–Whitney tests. Differences in experiments with more than one independent variable were evaluated by two-way ANOVA with post hoc Sidak's or Tukey's multiple comparison tests. All statistical tests were performed using GraphPad Prism software (GraphPad Software).

Reporting summary

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

Data availability

UKB individual-level data are available by request via application (https://www.ukbiobank.ac.uk). Individual-level MGBB data are available from https://personalizedmedicine.partners.org/Biobank/Default.aspx, only to Partners HealthCare investigators with appropriate approval from the Partners institutional review board. RNA-seq data are available at the Gene Expression Omnibus (GSE184420).

The present article includes all other data generated or analyzed during this study.

Code availability

Our CHIP-calling Terra pipeline using Mutect2 (version 1) is available at https://app.terra.bio/#workspaces/terra-outreach/CHIP-Detection-Mutect2. R-scripts for the observational epidemiologic associations are available at https://github.com/mzekavat/CHIP_PAD.

References

1. Song, P. et al. Global, regional, and national prevalence and risk factors for peripheral artery disease in 2015: an updated systematic review and analysis. *Lancet Glob. Health* **7**, e1020–e1030 (2019).
2. Conte, M. S. et al. Global vascular guidelines on the management of chronic limb-threatening ischemia. *J. Vasc. Surg.* **69**, 3S–125S (2019).
3. Jaiswal, S. et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).
4. Jaiswal, S. et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *N. Engl. J. Med.* **377**, 111–121 (2017).
5. Xie, M. et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **20**, 1472–1478 (2014).
6. Genovese, G. et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).
7. Bick, A. G. et al. Genetic interleukin 6 signaling deficiency attenuates cardiovascular risk in clonal hematopoiesis. *Circulation* **141**, 124–131 (2020).
8. 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).
9. Yu, B. et al. Supplemental association of clonal hematopoiesis with incident heart failure. *J. Am. Coll. Cardiol.* **78**, 42–52 (2021).
10. Bhattacharya, R. et al. Clonal hematopoiesis is associated with higher risk of stroke. *Stroke* **53**, 788–797 (2022).
11. Bick, A. G. et al. Inherited causes of clonal haematopoiesis in 97,691 whole genomes. *Nature* **586**, 763–768 (2020).
12. Klarin, D. et al. Genome-wide association study of peripheral artery disease in the Million Veteran Program. *Nat. Med.* **25**, 1274–1279 (2019).
13. Denny, J. C. et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat. Biotechnol.* **31**, 1102–1110 (2013).
14. Fuster, J. J. et al. Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science* **355**, 842–847 (2017).
15. Visconte, V., M, O. N. & H, J. R. Mutations in splicing factor genes in myeloid malignancies: significance and impact on clinical features. *Cancers (Basel)* **11**, 1844 (2019).
16. Boettcher, S. et al. A dominant-negative effect drives selection of *TP53* missense mutations in myeloid malignancies. *Science* **365**, 599–604 (2019).
17. Bondar, T. & Medzhitov, R. p53-mediated hematopoietic stem and progenitor cell competition. *Cell Stem Cell* **6**, 309–322 (2010).
18. Liu, Y. et al. p53 regulates hematopoietic stem cell quiescence. *Cell Stem Cell* **4**, 37–48 (2009).
19. TeKippe, M., Harrison, D. E. & Chen, J. Expansion of hematopoietic stem cell phenotype and activity in *Trp53*-null mice. *Exp. Hematol.* **31**, 521–527 (2003).
20. Sano, S. et al. Tet2-mediated clonal hematopoiesis accelerates heart failure through a mechanism involving the IL-1β/NLRP3 inflammasome. *J. Am. Coll. Cardiol.* **71**, 875–886 (2018).
21. Fuster, J. J. et al. TET2-loss-of-function-driven clonal hematopoiesis exacerbates experimental insulin resistance in aging and obesity. *Cell Rep.* **33**, 108326 (2020).
22. Abplanalp, W. T. et al. Clonal hematopoiesis–driver DNMT3A mutations alter immune cells in heart failure. *Circ. Res.* **128**, 216–228 (2021).
23. Sinha, S. K. et al. Local M-CSF (macrophage colony-stimulating factor) expression regulates macrophage proliferation and apoptosis in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **41**, 220–233 (2021).
24. Fuster, J. J. Clonal hematopoiesis and cardiovascular disease in cancer patients and survivors. *Thromb. Res.* **213**, S107–S112 (2022).
25. Boesten, L. S. et al. Macrophage p53 controls macrophage death in atherosclerotic lesions of apolipoprotein E deficient mice. *Atherosclerosis* **207**, 399–404 (2009).
26. Fuster, J. J. et al. Control of cell proliferation in atherosclerosis: insights from animal models and human studies. *Cardiovasc. Res.* **86**, 254–264 (2010).
27. Guevara, N. V., Kim, H. S., Antonova, E. I. & Chan, L. The absence of p53 accelerates atherosclerosis by increasing cell proliferation in vivo. *Nat. Med.* **5**, 335–339 (1999).
28. Mercer, J., Figg, N., Stoneman, V., Braganza, D. & Bennett, M. R. Endogenous p53 protects vascular smooth muscle cells from apoptosis and reduces atherosclerosis in ApoE knockout mice. *Circ. Res.* **96**, 667–674 (2005).
29. Merched, A. J., Williams, E. & Chan, L. Macrophage-specific p53 expression plays a crucial role in atherosclerosis development and plaque remodeling. *Arterioscler. Thromb. Vasc. Biol.* **23**, 1608–1614 (2003).
30. Sanz-Gonzalez, S. M. et al. Increased p53 gene dosage reduces neointimal thickening induced by mechanical injury but has no effect on native atherosclerosis. *Cardiovasc. Res.* **75**, 803–812 (2007).
31. van Vlijmen, B. J. et al. Macrophage p53 deficiency leads to enhanced atherosclerosis in APOE*3-Leiden transgenic mice. *Circ. Res.* **88**, 780–786 (2001).
32. Robbins, C. S. et al. Local proliferation dominates lesional macrophage accumulation in atherosclerosis. *Nat. Med.* **19**, 1166–1172 (2013).
33. Min, K. D., Polizio, A. H., Kour, A., Thel, M. C. & Walsh, K. Experimental *ASXL1*-mediated clonal hematopoiesis promotes inflammation and accelerates heart failure. *J. Am. Heart Assoc.* **11**, e026154 (2022).
34. Heyde, A. et al. Increased stem cell proliferation in atherosclerosis accelerates clonal hematopoiesis. *Cell* **184**, 1348–1361 (2021).
35. Fidler, T. P. et al. The AIM2 inflammasome exacerbates atherosclerosis in clonal haematopoiesis. *Nature* **592**, 296–301 (2021).
36. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
37. Smoller, J. W. et al. An eMERGE clinical center at Partners Personalized Medicine. *J. Pers. Med.* **6**, 5 (2016).
38. Van Hout, C. V. et al. Exome sequencing and characterization of 49,960 individuals in the UK Biobank. *Nature* **586**, 749–756 (2020).
39. Wu, P. et al. Mapping ICD-10 and ICD-10-CM codes to phecodes: workflow development and initial evaluation. *JMIR Med. Inform.* **7**, e14325 (2019).
40. Zekavat, S. M. et al. Hematopoietic mosaic chromosomal alterations increase the risk for diverse types of infection. *Nat. Med.* **27**, 1012–1024 (2021).
41. Hernan, M. A., Brumback, B. & Robins, J. M. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* **11**, 561–570 (2000).

42. Walter, W., Sánchez-Cabo, F. & Ricote, M. GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics* **31**, 2912–2914 (2015).
43. Adrover, J. M. et al. Programmed 'disarming' of the neutrophil proteome reduces the magnitude of inflammation. *Nat. Immunol.* **21**, 135–144 (2020).

Acknowledgements

D.K. is supported by the Veterans Administration (award 1IK2BX005759-01). P.N. is supported by a Hassenfeld Scholar Award from Massachusetts General Hospital; grants from the National Institutes of Health (NIH) National Heart, Lung, and Blood Institute (NHLBI) (R01HL1427, R01HL148565 and R01HL148050); and a grant from Fondation Leducq (TNE-18CVD04). S.M.Z is supported by the NIH NHLBI (1F30HL149180-01) and the NIH Medical Scientist Training Program Training Grant (T32GM136651). J.P.P. is supported by a John S. LaDue Memorial Fellowship. K.P. is supported by NIH grant 5-T32HL007208-43. P.L. receives funding support from the NHLBI (1R01HL134892), the American Heart Association (18CSA34080399), the RRM Charitable Fund and the Simard Fund. J.J.F. is supported by grant RYC-2016-20026 from the Spanish 'Ministerio de Ciencia e Innovación' (MICIN/AEI /10.13039/501100011033' and 'ESF Investing in your future'); a 2019 Leonardo Grant for Researchers and Cultural Creators from the BBVA Foundation; the European Research Area Network on Cardiovascular Diseases CHEMICAL (grant AC19/00133, funded by Instituto de Salud Carlos III and co-funded by the European Union, ERDF, 'A way to make Europe'); and the Leducq Foundation (TNE-18CVD04). A.G.B. is supported by NIH grant DP5 OD029586, a Burroughs Wellcome Fund Career Award for Medical Scientists and a Pew-Stewart Scholar for Cancer Research award, supported by the Pew Charitable Trusts and the Alexander and Margaret Stewart Trust. C.G. is supported by the European Research Area Network on Cardiovascular Diseases CHEMICAL. A.H. is supported by the Leducq Foundation (TNE-18CVD04). C.M.B. is supported by NIH grant R01-HL148050. C.J.G. is supported by NIH/NCI K08CA263555. P.L. receives funding support from the NHLBI (1R01HL134892 and 1R01HL163099-01), the American Heart Association (18CSA34080399), the RRM Charitable Fund and the Simard Fund. M.A.-P. is supported by grant FPU18/02913 funded by MCIN/AEI /10.13039/501100011033 and by 'ESF Investing in your future'. The project leading to these results also received funding from 'la Caixa' Foundation (ID 100010434), including fellowship code LCF/BQ/DR19/11740022 (to A.A.-C.) and agreement HR17-00267. Confocal microscopy was conducted at the Microscopy & Dynamic Imaging Unit at CNIC, ICTS-ReDib, co-funded by MCIN/AEI /10.13039/501100011033 and FEDER 'Una manera de hacer Europa' (ICTS-2018-04-CNIC-16). The CNIC is supported by the Instituto de Salud Carlos III (ISCIII), the MICIN 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 R. Moro for assistance with figure preparation and F. Sánchez-Cabo and M. Gómez for assistance with bioinformatics analysis of RNA-seq data. S.M.D. is supported by the Veterans Administration (award IK2-CX001780). We also thank participants and staff of the UK Biobank and the Mass General Brigham Biobank. UK Biobank analyses were conducted using application 7089.

Author contributions

S.M.Z., S.D.J., M.M.U., M.T., K.P., S.D., J.P., G.G. and A.N. designed, performed and analyzed population-level data from the UKB and the MGBB and interpreted data. S.M.Z. and D.K. interpreted data, created the figures and assisted in writing the respective parts of the paper. N.M. designed, performed and analyzed experiments; interpreted experimental data; created the figures; and assisted in writing the manuscript. V.V.-H. designed, performed and analyzed experiments; interpreted experimental data; and contributed to figure preparation. J.J.F., M.A.Z., M.A.-P. and A.A.-C. performed and analyzed experiments and interpreted experimental data. V.Z. and A.F.-P. provided technical support and performed experiments. P.K., R.C. and C.M.G. conducted RNA-seq studies and analyzed transcriptomics data. A.H., V.F., B.L.E., P.L., C.J.G., C.M.B., H.Z., M.G. and D.K. discussed the research strategy and the results and provided conceptual advice. P.N., A.G.B., J.J.F. and D.K. conceived the study, analyzed and interpreted data and created the figures.

Competing interests

D.K. is a scientific advisor for and accepts consulting fees from Bitterroot Bio. P.N. reports grant support from Amgen, Apple, AstraZeneca, Boston Scientific and Novartis; spousal employment and equity at Vertex; consulting income from Apple, AstraZeneca, Novartis, Genentech/Roche, Blackstone Life Sciences, Foresite Labs and TenSixteen Bio; and is a scientific advisory board member and shareholder of TenSixteen Bio and geneXwell, all unrelated to this work. P.L. is an unpaid consultant to, or involved in clinical trials for, Amgen, AstraZeneca, Baim Institute, Beren Therapeutics, Esperion Therapeutics, Genentech, Kancera, Kowa Pharmaceuticals, Medimmune, Merck, Novo Nordisk, Novartis, Pfizer and Sanofi-Regeneron. P.L. is a member of the scientific advisory board for Amgen, Caristo Diagnostics, Cartesian Therapeutics, CSL Behring, DalCor Pharmaceuticals, Dewpoint Therapeutics, Eulicid Bioimaging, Kancera, Kowa Pharmaceuticals, Olatec Therapeutics, Medimmune, Moderna, Novartis, PlaqueTec, TenSixteen Bio, Soley Therapeutics, and XBiotech. P.L.'s laboratory has received research funding in the last 2 years from Novartis. P.L. is on the board of directors of XBiotech. P.L. also has a financial interest in Xbiotech, a company developing therapeutic human antibodies; in TenSixteen Bio, a company targeting somatic mosaicism and CHIP to discover and develop novel therapeutics to treat age-related diseases; and in Soley Therapeutics, a biotechnology company that is combining artificial intelligence with molecular and cellular response detection for discovering and developing new drugs, currently focusing on cancer therapeutics. P.L.'s interests were reviewed and are managed by Brigham and Women's Hospital and Mass General Brigham in accordance with their conflict of interest policies. B.L.E. has received research funding from Celgene, Deerfield, Novartis and Calico and consulting fees from GRAIL. B.L.E. is a member of the scientific advisory board and a shareholder for Neomorph, TenSixteen Bio, Skyhawk Therapeutics and Exo Therapeutics. A.G.B. is a paid advisor and holds equity in TenSixteen Bio. S.M.D. has received research support via the University of Pennsylvania from RenalytixAI and Novo Nordisk. The other authors do not have any competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s44161-022-00206-6.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44161-022-00206-6.

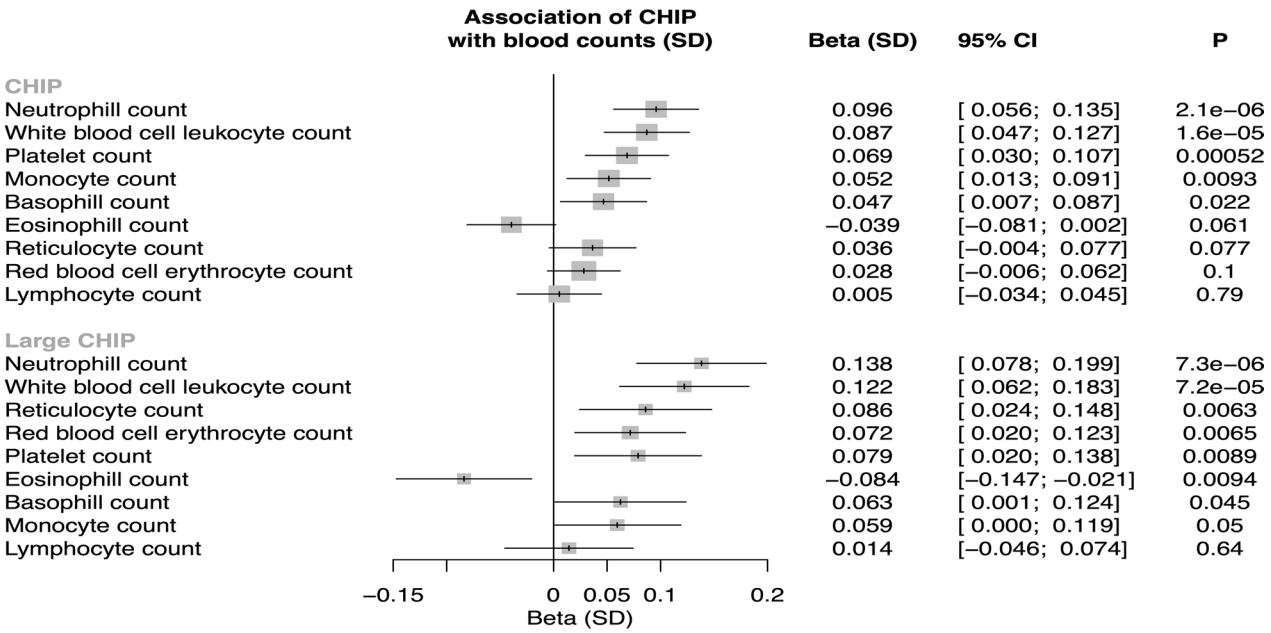
Correspondence and requests for materials should be addressed to José J. Fuster or Derek Klarin.

Peer review information *Nature Cardiovascular Research* thanks Alan Tall, Andrew Murphy, Gerard Pasterkamp and Steffen Massberg for their contribution to the peer review of this work.

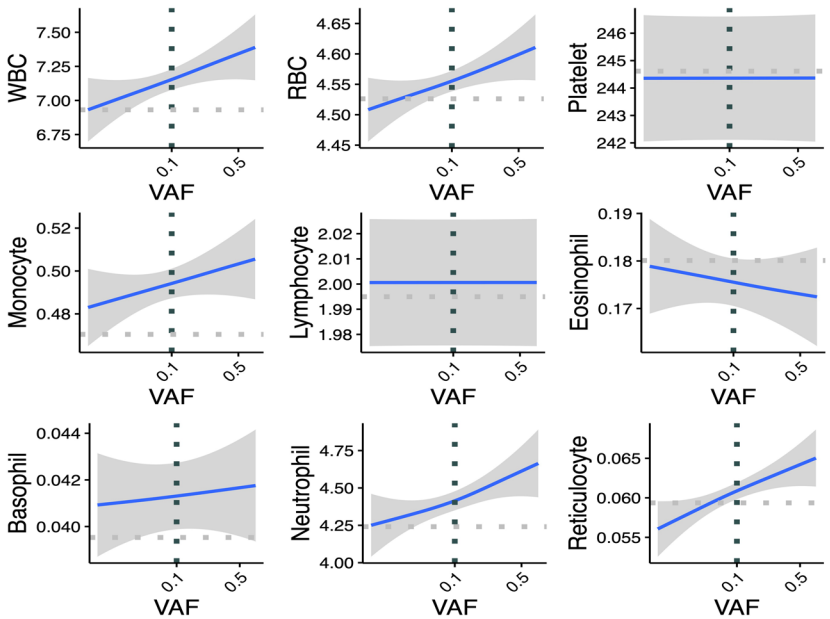
[1]Department of Ophthalmology, Massachusetts Eye and Ear Institute, Boston, MA, USA. [2]Computational Biology & Bioinformatics Program, Yale University, New Haven, CT, USA. [3]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [4]Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain. [5]Vector Institute for Artificial Intelligence, Toronto, ON, Canada. [6]Department of Electrical Engineering and Computer Science and Institute for Medical and Evaluative Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. [7]Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA. [8]Center for Heart Lung Innovation, University of British Columbia, Vancouver, BC, Canada. [9]Department of Medicine, Harvard Medical School, Boston, MA, USA. [10]IRCCS Humanitas Research Hospital, Milan, Italy. [11]Department of Biomedical Sciences, Humanitas University, Milan, Italy. [12]Vascular Biology and Therapeutics Program and Department of Immunobiology, Yale University School of Medicine, New Haven, CT, USA. [13]Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, USA. [14]Department of Surgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [15]Center for Cardiometabolic Disease Prevention, Baylor College of Medicine, Houston, TX, USA. [16]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [17]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. [18]Department of Laboratory Medicine, Lund University, Lund, Sweden. [19]Cardiology Division, Massachusetts General Hospital, Boston, MA, USA. [20]Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [21]Division of Cardiology, University of California, San Francisco, San Francisco, CA, USA. [22]Department of Pathology, Dana-Farber Cancer Institute, Boston, MA, USA. [23]Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA. [24]Howard Hughes Medical Institute, Boston, MA, USA. [25]Department of Medicine, Division of Cardiovascular Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. [26]Icahn School of Medicine at Mount Sinai, New York, NY, USA. [27]Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA. [28]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA. [29]Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA, USA. [30]Program in Medical and Population Genetics and the Cardiovascular Disease Initiative, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [31]Department of Medicine, Division of Genetic Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA. [32]CIBER en Enfermedades Cardiovasculares (CIBER-CV), Madrid, Spain. [33]VA Palo Alto Healthcare System, Palo Alto, CA, USA. [34]Department of Surgery, Stanford University School of Medicine, Palo Alto, CA, USA. [35]These authors contributed equally: Seyedeh M. Zekavat, Vanesa Viana-Huete, Nuria Matesanz. [36]These authors jointly supervised this work: Pradeep Natarajan, Alexander G. Bick, José J. Fuster, Derek Klarin. ✉e-mail: jjfuster@cnic.es; dklarin@stanford.edu

https://doi.org/10.1038/s44161-022-00206-6

a.



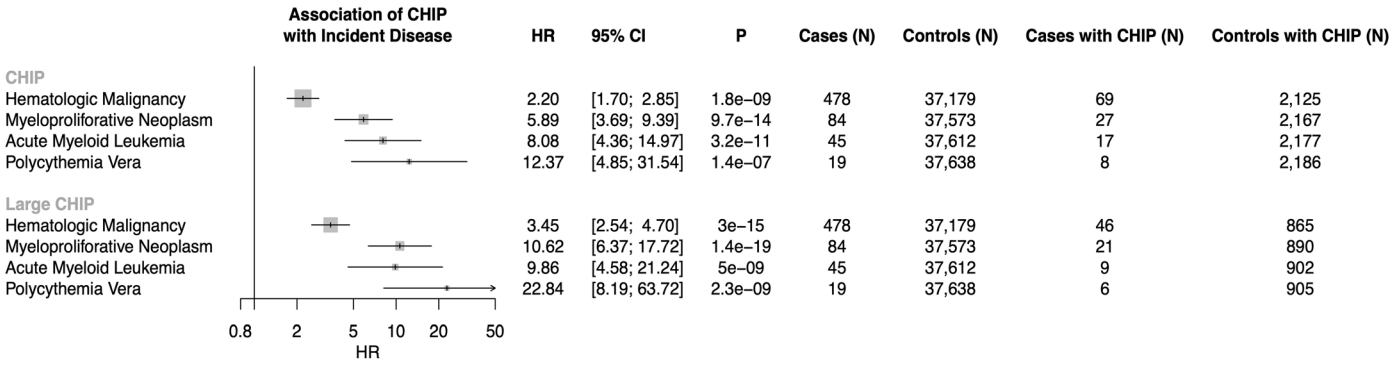
b.



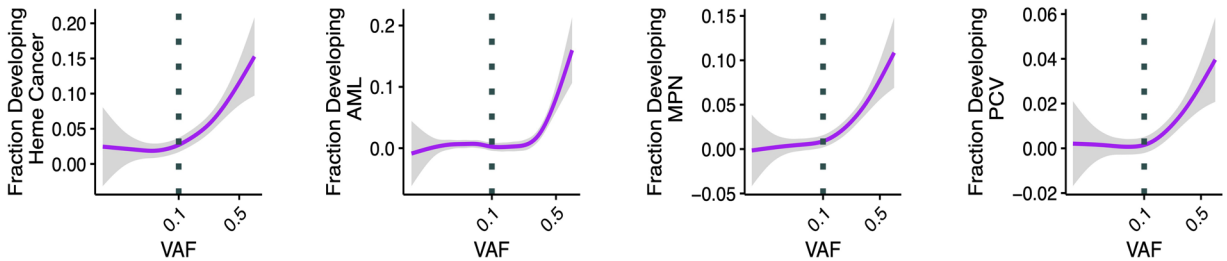
Extended Data Fig. 1 | Association of CHIP with blood counts among individuals without prevalent hematologic malignancy in the UK Biobank (n = 37,657). Blood counts were acquired at time of blood draw for whole exome sequencing. **a**) Association of CHIP and Large CHIP with normalized blood counts (SD). Associations are adjusted for age, age^2, sex, smoking status, and the first ten principal components of genetic ancestry. **b**) Association of CHIP variant allele frequency (VAF) with blood counts (in units of 10^9 cells/L). The gray horizontal dotted lines reflect average counts across non-CHIP carriers. The vertical black dotted line reflects the cutoff VAF for Large CHIP (VAF > 0.1). Error bands reflect the standard error of a generalized additive model with integrated smoothness fit to the data. CHIP = clonal hematopoiesis of indeterminate potential; VAF = variant allele fraction.

https://doi.org/10.1038/s44161-022-00206-6

a.



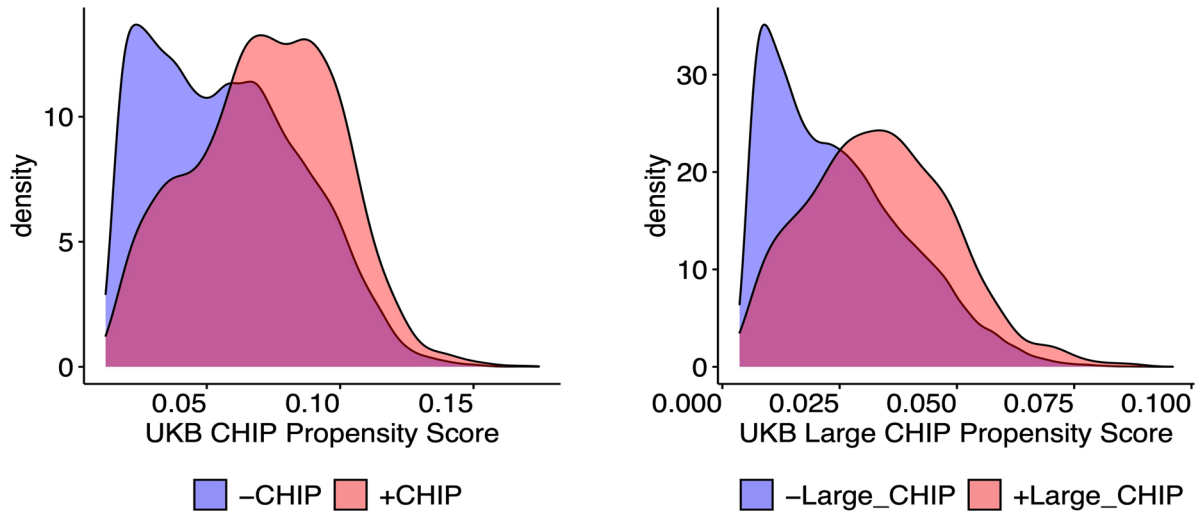
b.



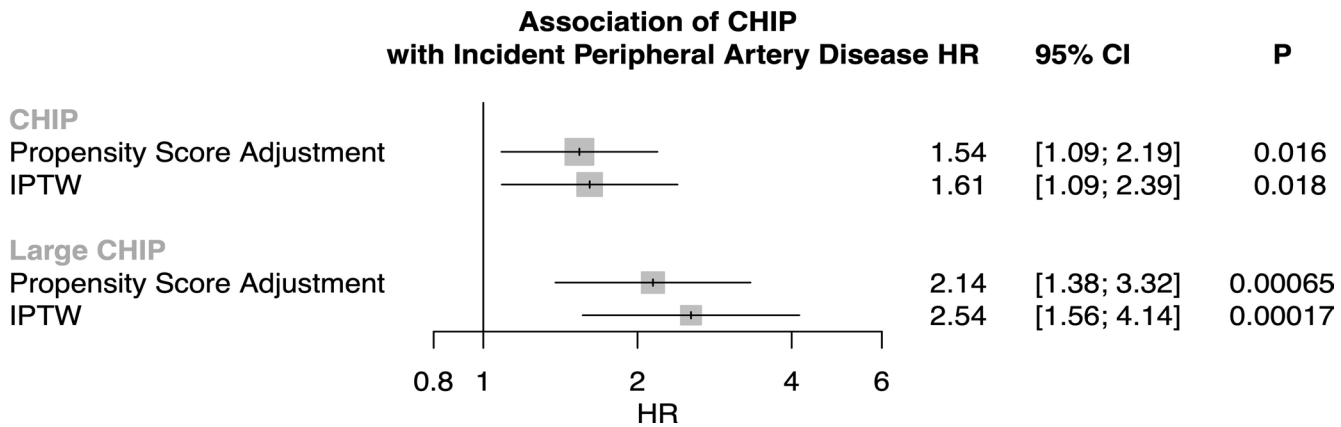
Extended Data Fig. 2 | Association of CHIP (a) and VAF (b) with incident hematologic malignancy among individuals without prevalent hematological malignancy in the UK Biobank (n = 37,657). Associations are adjusted for age, age^2, sex, smoking status, Townsend deprivation index, and the first ten principal components of genetic ancestry. (a) Error bars are centered at the HR and show the 95% CI for estimates. (b) Error bands reflect the standard error of a generalized binomial additive model with integrated smoothness fit to the data. CHIP = clonal hematopoiesis of indeterminate potential; VAF = variant allele fraction.

https://doi.org/10.1038/s44161-022-00206-6

a.

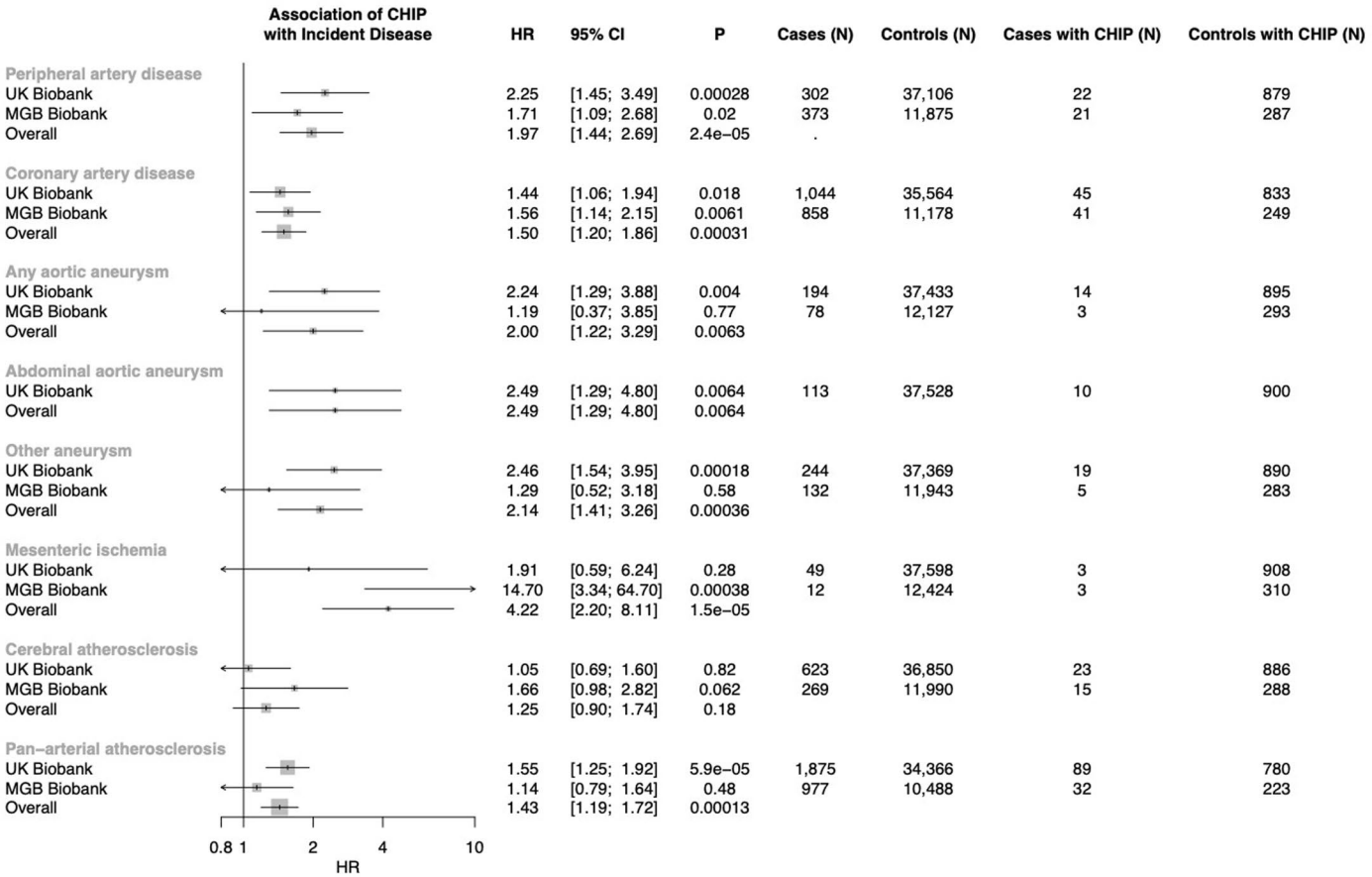


b.



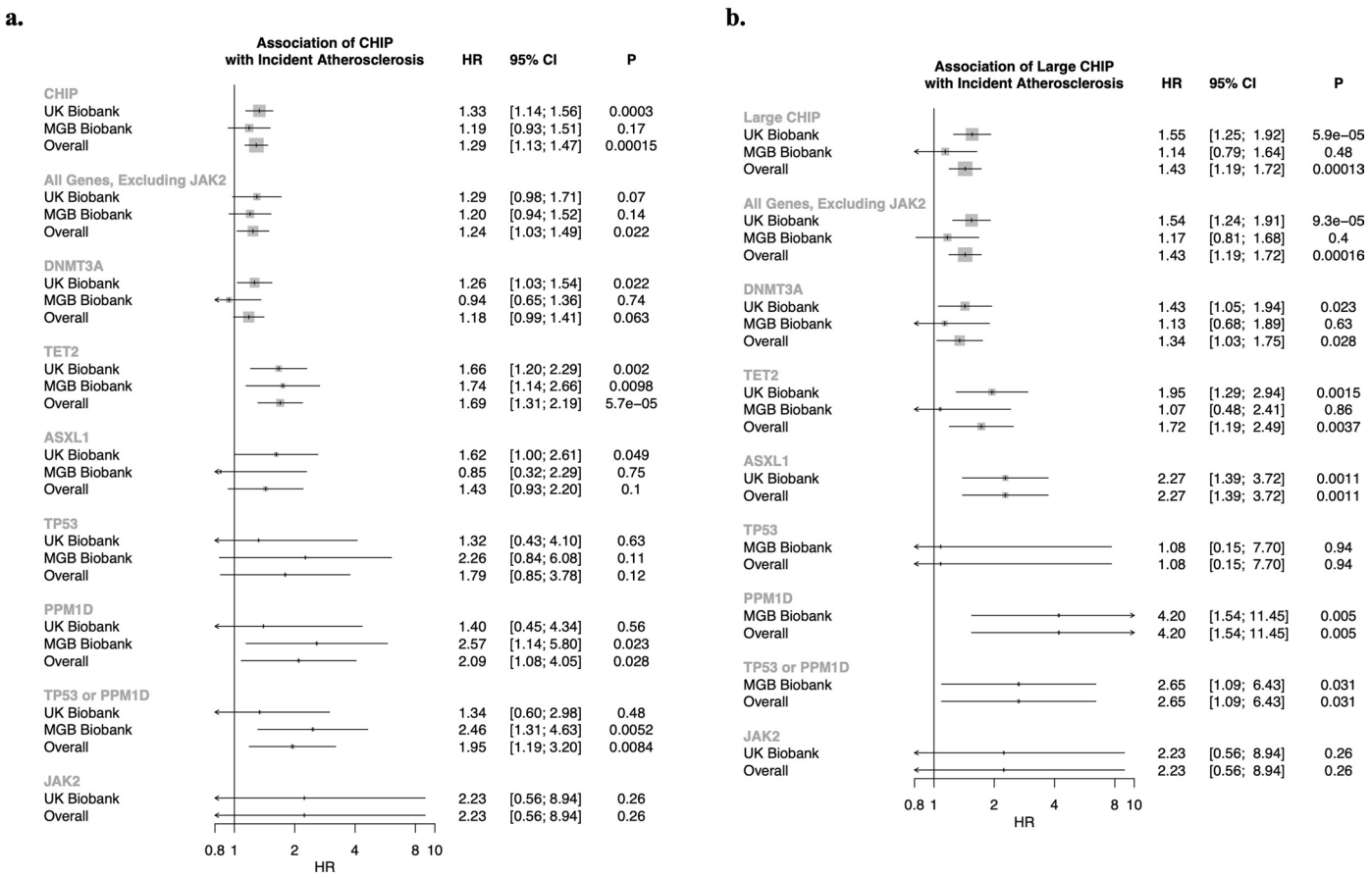
Extended Data Fig. 3 | Epidemiological causal inference analysis for CHIP on incident peripheral artery disease in the UK Biobank. a) Propensity scores by CHIP and Large CHIP status in the UKB (n = 37,657). **b)** Propensity score adjustment and stabilized inverse probability treatment weighting (IPTW) for the CHIP and Large CHIP association with incident PAD in the UKB. Error bars are centered at the HR and show the 95% CI for estimates. CHIP = clonal hematopoiesis of indeterminate potential; VAF = variant allele fraction; PAD = peripheral artery disease.

https://doi.org/10.1038/s44161-022-00206-6



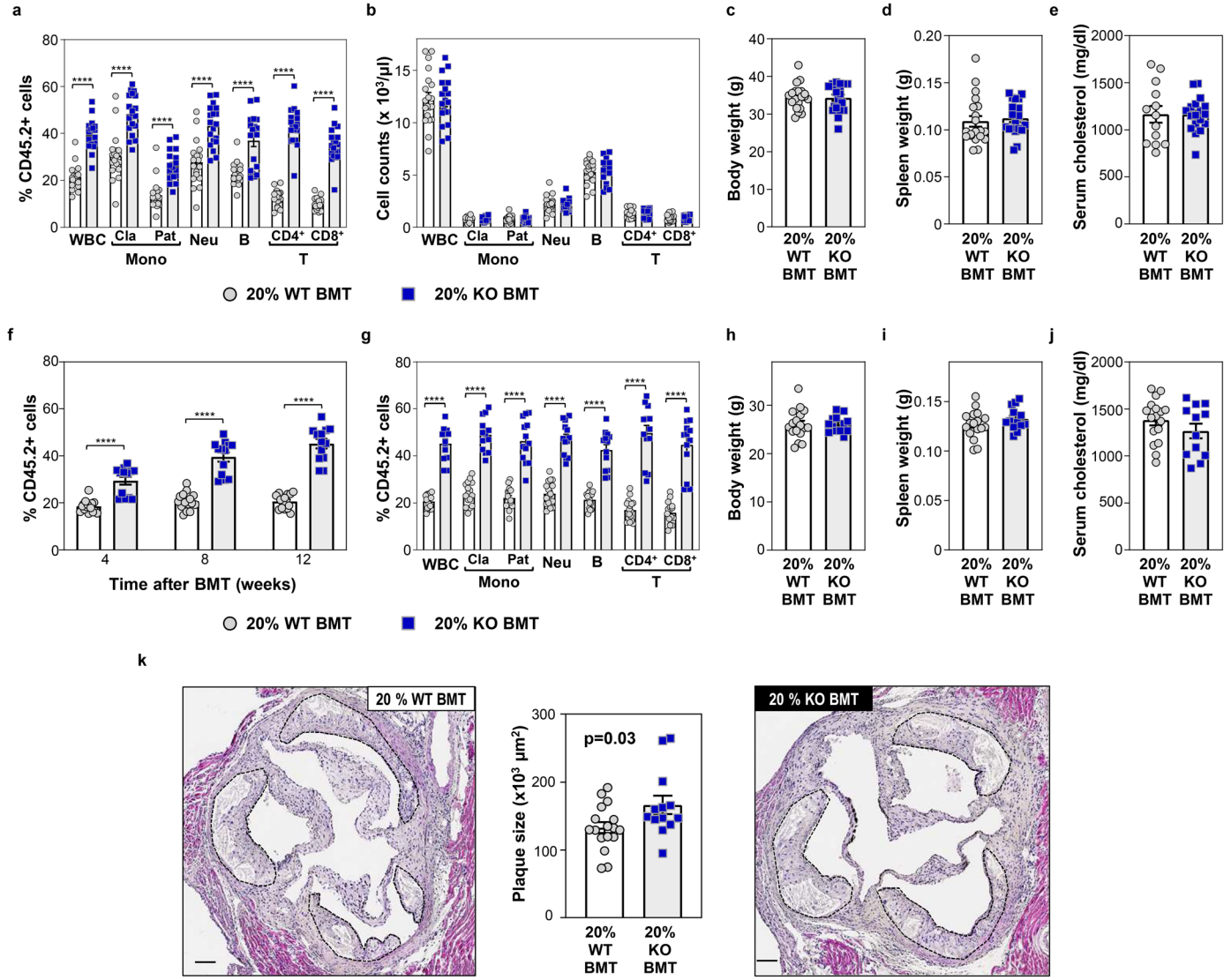
Extended Data Fig. 4 | Association of Large CHIP (VAF > 10%) with incident pan-arterial atherosclerosis, combined across peripheral artery disease, coronary artery disease, aneurysms, chronic and acute mesenteric ischemia, cerebral atherosclerosis, and renal artery stenosis. Error bars are centered at the HR and show the 95% CI for estimates. CHIP = clonal hematopoiesis of indeterminate potential; VAF = variant allele fraction.

https://doi.org/10.1038/s44161-022-00206-6



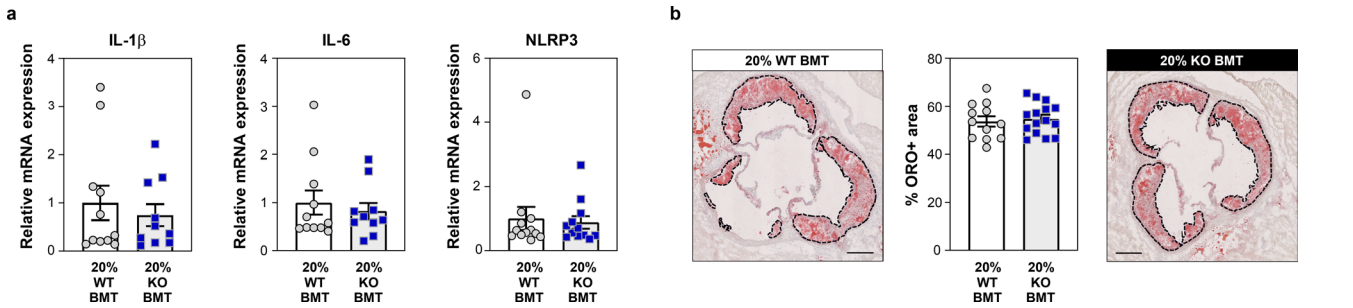
Extended Data Fig. 5 | Association of a) CHIP and b) Large CHIP genes with incident pan-arterial atherosclerosis, combined across peripheral artery disease, coronary artery disease, aneurysms, chronic and acute mesenteric ischemia, cerebral atherosclerosis, and renal artery stenosis. Error bars are centered at the HR and show the 95% CI for estimates. CHIP = clonal hematopoiesis of indeterminate potential; VAF = variant allele fraction.

https://doi.org/10.1038/s44161-022-00206-6



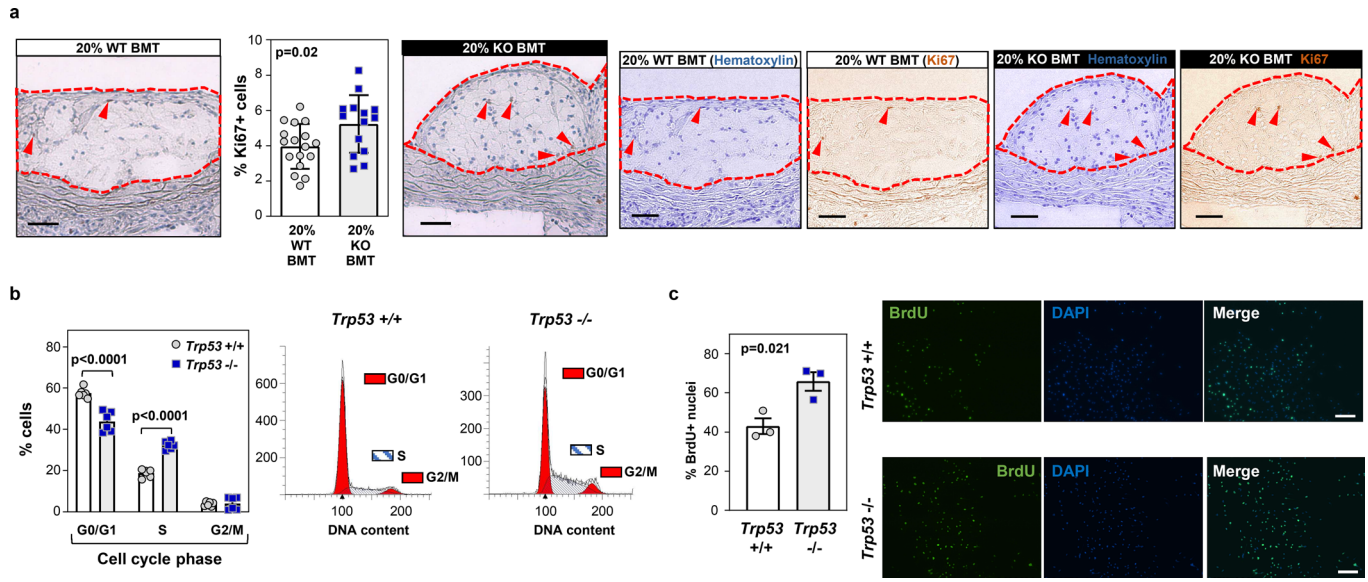
Extended Data Fig. 6 | Effects of experimental p53-deficient CHIP on atherosclerosis development in *Ldlr*-/- mice. a-e) 20% KO-BMT male mice (n = 19 mice) and 20% WT-BMT controls (n = 20 mice) were fed a high-fat/high-cholesterol (HF/HC) diet for 9 weeks, starting 4 weeks after BMT. **a)** Percentage of CD45.2+ cells in different white blood cell (WBC) lineages in peripheral blood, evaluated by flow cytometry. Two-tailed unpaired *t*-tests were used for statistical analysis (mean ± SEM, ****p < 0.0001). **b)** Absolute counts of main WBC sub-populations in peripheral blood, evaluated by flow cytometry (mean ± SEM). **c**) Body weight (mean ± SEM). **d**) Spleen weight (mean ± SEM). **e**) Total cholesterol level in serum, evaluated by enzymatic methods (mean ± SEM). **f-k**) 20% KO-BMT female mice and 20% WT-BMT controls were fed a high-fat/high-cholesterol (HF/HC) diet for 9 weeks, starting 4 weeks after BMT. **f**) Percentage of CD45.2+ cells in white blood cells at different timepoints, evaluated by flow cytometry. A two-way ANOVA with Sidak's multiple comparison test was used for statistical analysis (mean ± SEM,****p < 0.0001). **g)** Percentage of CD45.2+ cells in different WBC lineages in peripheral blood after 9 weeks on HF/HC diet (13 weeks post-BMT), evaluated by flow cytometry. Two-tailed unpaired *t*-tests were used for statistical analysis (mean ± SEM, n = 16 20% WT-BMT mice, n = 13 20% KO-BMT mice, ****p < 0.0001). **h)** Body weight (mean ± SEM, n = 17 20% WT-BMT mice, n = 13 20% KO-BMT mice). **i)** Spleen weight (mean ± SEM, n = 17 20% WT-BMT mice, n = 13 20% KO-BMT mice). **j)** Total cholesterol level in serum, evaluated by enzymatic methods (mean ± SEM, n = 17 20% WT-BMT mice, n = 12 20% KO-BMT mice) **k)** Aortic root plaque size. A two-tailed unpaired *t*-test was used for statistical analysis (mean ± SEM, n = 17 20% WT-BMT mice, n = 13 20% KO-BMT mice). Representative images of hematoxylin and eosin-stained sections are shown; atherosclerotic plaques are delineated by dashed lines. Scale bars, 100 µm.

https://doi.org/10.1038/s44161-022-00206-6



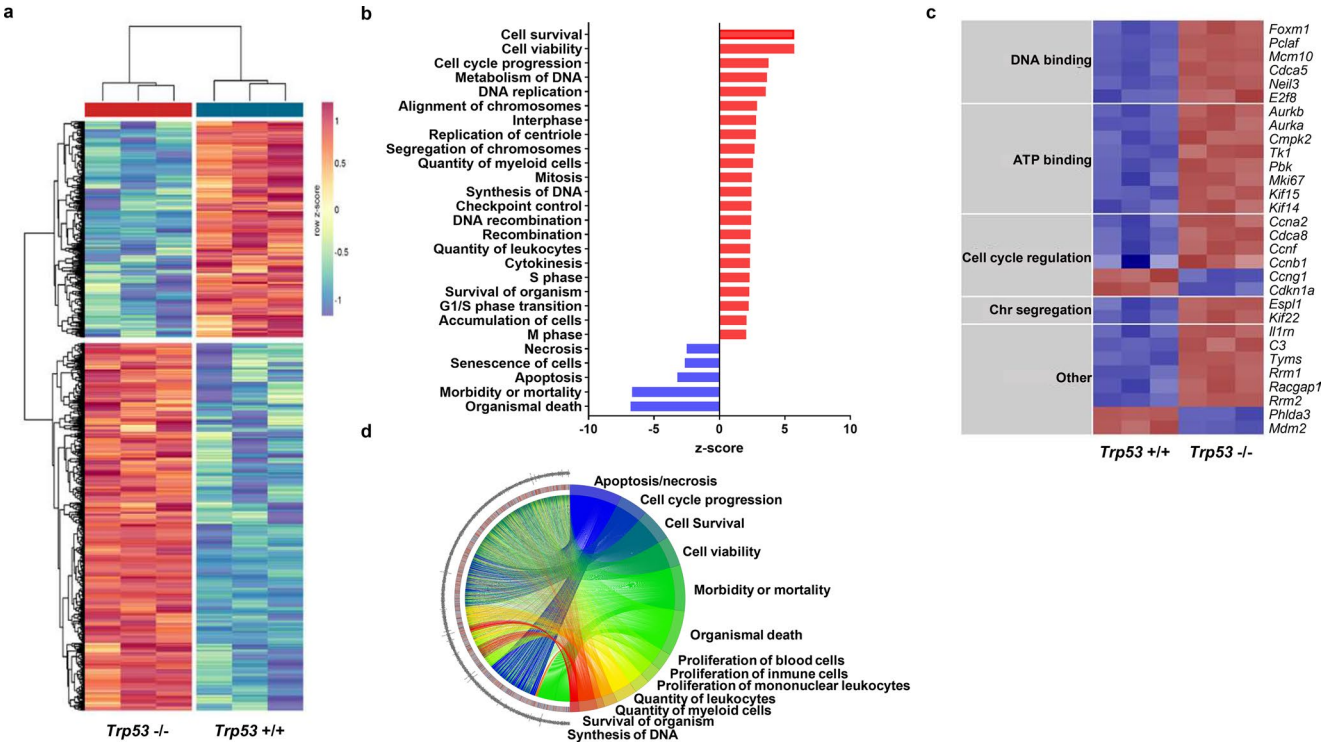
Extended Data Fig. 7 | No effect of p53-deficient CHIP on plaque lipid content or aortic expression of the proinflammatory cytokines IL-1b and IL-6 or the NLRP3 inflammasome. a) Aortic arch samples were obtained from HF/HC-fed 20% KO-BMT male mice and 20% WT-BMT controls, and gene expression was analyzed by qPCR (mean ± SEM, n = 11 20% WT-BMT mice, n = 10 20% KO-BMT mice). **b**) 20% KO-BMT male mice and 20% WT-BMT controls were fed a high-fat/high-cholesterol (HF/HC) diet for 9 weeks. Plaque lipid content was analyzed through Oil Red O (ORO) staining of cryostat sections (mean ± SEM, n = 12 20% WT-BMT mice, n = 15 20% KO-BMT mice). Representative images of ORO-stained sections are shown; atherosclerotic plaques are delineated by dashed lines. Scale bars, 200 µm.

https://doi.org/10.1038/s44161-022-00206-6



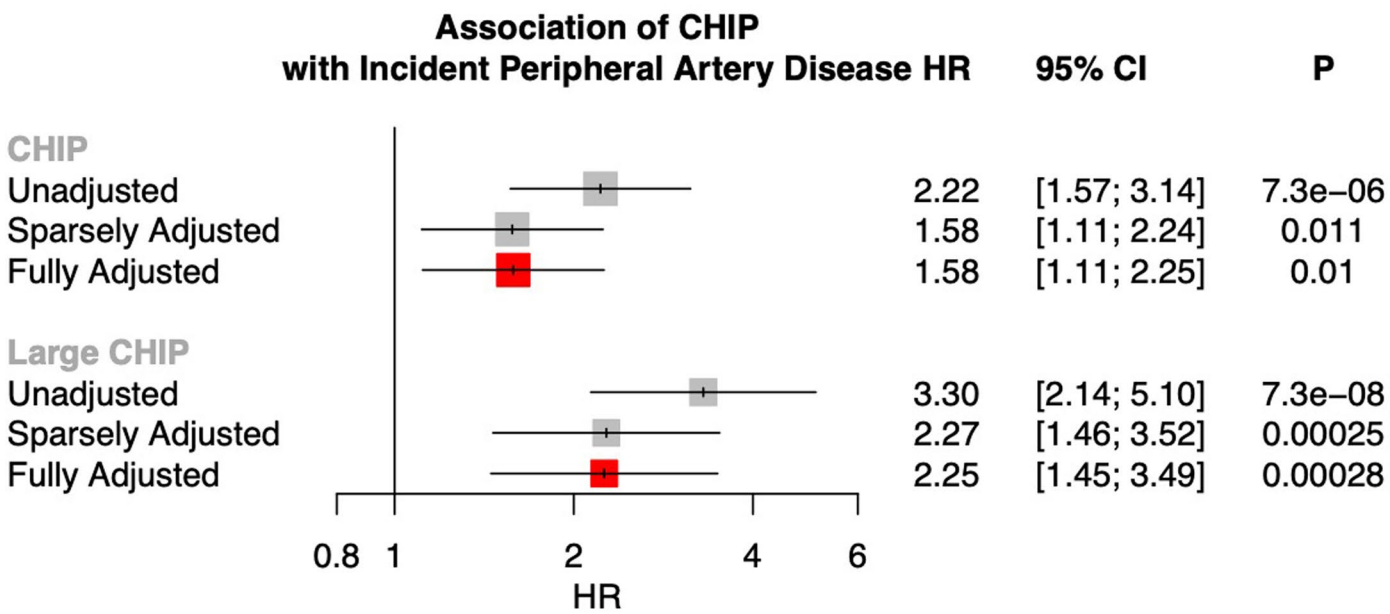
Extended Data Fig. 8 | Increased cell proliferation in conditions of p53-deficient CHIP. a) 20% KO-BMT female mice and 20% WT-BMT controls were fed a high-fat/high-cholesterol (HF/HC) diet for 9 weeks, starting 4 weeks after BMT. Plaque cell proliferation was estimated based on immunohistochemical staining of the Ki-67 proliferation marker (mean ± SEM, n = 17 20% WT-BMT mice, n = 14 20% KO-BMT mice). A two-tailed unpaired *t*-test was used for statistical analysis. Representative images of Ki-67-stained sections are shown; color deconvolution was applied to show separately the staining of hematoxylin (nuclei) and Ki-67. Atherosclerotic plaques are delineated by dashed lines. Scale bars, 50 µm. **b)** Cell cycle phase distribution of cultured *Trp53*-/- and + /+ bone marrow-derived macrophages proliferating asynchronously in the presence of 100 ng/ml MCSF, evaluated by propidium iodide staining of cellular DNA content and flow cytometry (mean ± SEM, n = 6 *Trp53* + /+ mice, n = 6 *Trp53*-/- mice). A two-way ANOVA with Sidak's multiple comparison test was used for statistical analysis. **c)** Analysis of the proliferation of cultured *Trp53*-/- and + /+ bone marrow-derived macrophages through immunostaining of BrdU incorporation into the DNA. Quiescent G0-synchronized macrophages were treated with MCSF to induce proliferation (mean ± SEM, n = 3 *Trp53* + /+ mice, n = 3 *Trp53*-/- mice). A two-tailed unpaired *t*-test was used for statistical analysis. A representative experiment is shown; three separate experiments were conducted. Scale bars, 100 µm.

https://doi.org/10.1038/s44161-022-00206-6



Extended Data Fig. 9 | Transcriptomic profiling of MCSF-stimulated p53-deficient macrophages. An 18h-treatment with MCSF was used to induce cell cycle entry and progression in quiescent G0-synchronized *Trp53*-/- (KO) and + /+ (WT) murine macrophages in culture (n = 3 per genotype). mRNA sequencing was used for transcriptomic profiling. **a)** Heatmap of differentially expressed genes with fold change (FC) ≥ 1.5. **b)** Functional categories enriched in the set of genes differentially expressed between *Trp53*-/- and + /+ macrophages, based on Ingenuity Pathway Analysis (cancer-related pathways are not shown). **c)** Heatmap of the most upregulated genes within selected functional categories that are enriched in genes differentially expressed in *Trp53*-/- macrophages with fold change ≥ 1.5. **d)** GoPlot of selected functional categories (right hand-side) and the logFC values of the most differentially expressed genes included in these categories (left hand-side).

https://doi.org/10.1038/s44161-022-00206-6



Extended Data Fig. 10 | Association of CHIP and Large CHIP (variant allele fraction > 10%) with PAD in the UKB (N = 37,657). under 1) unadjusted, 2) sparsely adjusted, and 3) fully adjusted models, where sparsely adjusted refers to the following covariates: age, age^2, sex, smoking status, Townsend deprivation index, and the first ten principal components of genetic ancestry, and the fully adjusted model additionally includes normalized BMI, prevalent hypertension, hyperlipidemia, and type 2 diabetes as covariates. Given the minimal difference between the sparsely adjusted and fully adjusted model, the sparsely adjusted model was moved forward for use in analysis. Error bars are centered at the HR and show the 95% CI for estimates.

nature portfolio

Corresponding author(s): Derek Klarin [dklarin@stanford.edu]

Last updated by author(s): Dec 2, 2022

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	Human data were previously collected by the Biobanks involved. Mouse flow cytometry data were collected with BD FACSDiva v.6.1.2 and v9.1 software (BD Biosciences). Histological and immunofluorescent images were acquired with LAS X Life Science Microscope Software v3.1.1, LAS-AF 2.7.3. build 9723 software or Nikon AR 4.30.02 (b925) software. Western blot images were acquired with Image Quant LAS4000 min v1.2 build 1.2.1.119 software. Somatic CHIP variants were detected with GATK MuTect2 (v1) software.
Data analysis	- UKB individual-level data are available for request by application (https://www.ukbiobank.ac.uk). Individual-level MGBB data are available from https://personalizedmedicine.partners.org/Biobank/Default.aspx, but restrictions apply to the availability of these data, which were used under IRB approval for the current study, and so are not publicly available. The present article includes all other data generated or analyzed during this study. Statistical analyses were performed in R v3.5. - Mouse histological images, immunofluorescent images and Western blot images were analyzed with Fiji ImageJ v1.53. - Mouse flow cytomety data was analyzed with FlowJo Software v10 (Tree Star). DNA histograms were fitted into cell cycle distributions using ModFit 3.0 software (Verity Software House). - Statistical analysis of mouse and cell culture experiments was done with GraphPad Prism 8

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

UKB individual-level data are available for request by application (https://www.ukbiobank.ac.uk). Individual-level MGBB data are available from https://personalizedmedicine.partners.org/Biobank/Default.aspx, but restrictions apply to the availability of these data, which were used under IRB approval for the current study, and so are not publicly available. Mouse macrophage RNA-seq data is publicly accesible at Gene Expression Omnibus (GEO: GSE184420). The present article includes all other data generated or analyzed during this study.

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	A total of 37,657 unrelated individuals from the UK Biobank (UKB) and 12,465 individuals from Mass General Brigham Biobank (MGBB) with whole exome sequencing data available for downstream analysis. In mouse experiments, sample sizes were estimated based on our previous work in the field (e.g. Fuster et al, Science 2017. doi: 10.1126/science.aag1381) and sample size in each experiment is indicated in the figure, figure legend and/or methods.
Data exclusions	A total of 50,122 unrelated individuals, passing genotype and CHIP quality control criteria, and without hematologic cancer at time of blood draw for genotyping were analyzed. In mouse studies, exclusion criteria were pre-established. Animals were excluded from the analysis if they did not develop hypercholesterolemia after bone marrow transplant and high cholesterol diet feeding, or if signs of sickness unrelated to atherosclerosis were observed, specifically solid tumors or dermatitis. In the histological assessment of mouse atherosclerosis, samples that exhibited evidence of artefactual tissue damage or abnormal orientation that could not be compensated by the analysis of multiple independent sections were excluded from analysis.
Replication	Replication of UKB analyses was performed in MGBB and were successful for the outcomes presented. In mouse studies, all the experiments were replicated at least two times and usually three times.
Randomization	Not directly applicable to human studies given study design; data was previously collected by the Biobanks.. Covariates across analyses are as specified. In mouse bone marrow transplantation studies, recipient mice were allocated randomly to p53-KO or WT bone marrow groups.
Blinding	Not directly applicable to human studies given study design; data was previously collected by the Biobanks. In mouse atherosclerosis studies, investigators were blinded to bone marrow genotype during data collection and analysis.

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
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used	A list of antibodies and additional information (e.g. supplier, catalog number, dilution) is provided in Supplementary Table 10.
Validation	All antibodies used in this study were previosuly validated for their indicated applications by the manufacturer.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	All mice were maintained on a 12-h light/dark schedule in a specific pathogen-free animal facility in individually ventilated cages and given food and water ad libitum. C57Bl/6J Trp53-/- mice were obtained from the Jackson Laboratory. Ldlr-/- mice carrying the CD45.1 isoform of the CD45 cell surface marker were generated by crossing LDLR-KO mice from the Jackson Laboratory and B6.SJL-PtprcaPepcb/BoyCrl mice from Charles River Laboratories. In bone marrow transplantation experiments, recipient and donor mice were 6-10 weeks old at the time of the transplant. Male and female mice were used, as described in figure legends and the main text.
Wild animals	No wild animals were used in the study.
Field-collected samples	No field-collected samples were used in the study
Ethics oversight	Animal experiments followed protocols approved by the Institutional Ethics Committee at the Centro Nacional de Investigaciones Cardiovasculares

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

Human research participants

Policy information about studies involving human research participants

Population characteristics	The UKB is a population-based cohort of approximately 500,000 participants recruited from 2006-2010. The MGBB contains genotypic and clinical data from >105,000 patients who consented to broad-based research. Age and demographic information is provided in supplementary table 2. Clinical outcomes data were defined by ICD and CPT procedure codes as outlined in the methods section.
Recruitment	The UKB is a population-based cohort of approximately 500,000 participants recruited from 2006-2010 with existing genomic and longitudinal phenotypic data and median 10-year follow-up. Baseline assessments were conducted at 22 assessment centres across the UK with sample collections including blood-derived DNA. The MGBB contains genotypic and clinical data from >105,000 patients who consented to broad-based research across 7 regional hospitals and median 3-year follow-up. Baseline phenotypes were ascertained from the electronic medical record and surveys
Ethics oversight	Use of the data was approved by the Massachusetts General Hospital Institutional Review Board (protocol 2020P000904) and facilitated through UK Biobank Application 7089.

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

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Peripheral blood was obtained from the facial vein and collected into EDTA-coated tubes. Bone marrow cells were flushed out of two femurs and two tibias per mouse. Aortic arches were digested for 45 minutes at 37°C in RPMI containing 10% FBS and 0.25 mg/ml Liberase TM (Roche Life Science).
Instrument	BD FACSCANTO, BD LSRFortessa or BD FACSymphony Cytometers (BD Bioscience)
Software	FlowJo v10 (Tree Star) ModFit LT v3 (Verity Software)
Cell population abundance	Representation of the abundance of the relevant cell populations identified through analytical flow cytometry can be visualized in the figures. Sorting flow cytometry was not used in the current study.
Gating strategy	Figures summarizing gating strategies are shown in Supplementary Figure 6.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.