Increased inflammatory signature in myeloid cells of non-small cell lung cancer patients with high clonal hematopoiesis burden

Hyungtai Sim,[1]* Hyun Jung Park,[2,3]* Geun-Ho Park, [4,5]*, Yeon Jeong Kim,[2] Woong-Yang Park,[2] Se-Hoon Lee, [4,5]** Murim Choi[1]**

*, ** These authors contributed equally

Affiliations:

[1]Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Republic of Korea

[2]Samsung Genome Institute, Samsung Medical Center, Seoul, Republic of Korea

[3]Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul National University, Seoul, Republic of Korea

[4]Division of Hematology-Oncology, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[5]Department of Health Sciences and Technology, Samsung Advanced Institute for Health Science & Technology (SAIHST), Sungkyunkwan University, Seoul, Republic of Korea

Corresponding authors:

Murim Choi, Department of Biomedical Sciences, Seoul National University College of Medicine,103 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea

Email: murimchoi@snu.ac.kr

Se-Hoon Lee, Division of Hematology-Oncology, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Republic of Korea
Email: sehoon.lee119@gmail.com

Running title: Increased inflammatory signature in NSCLC patients with CHIP

Keywords: Clonal Hematopoiesis, Immune checkpoint inhibitor, Inflammation, NSCLC

Abstract

Clonal hematopoiesis of indeterminate potential (CHIP) allows estimation of clonal dynamics and documentation of somatic mutations in the hematopoietic system. Recent studies utilizing large cohorts of the general population and patients have revealed significant associations of CHIP burden with age and disease status, including in cancer and chronic diseases. An increasing number of cancer patients are treated with immune checkpoint inhibitors (ICI), but the association of ICI response in non-small cell lung cancer (NSCLC) patients with CHIP burden remains to be determined. We collected blood samples from 100 metastatic NSCLC patients before and after ICI for high-depth sequencing of the CHIP panel and 63 samples for blood single-cell RNA sequencing(scRNA-seq). Whole exome sequencing (WES) was performed in an independent replication cohort of 180 patients. The impact of CHIP status on the immunotherapy response was not significant. However, metastatic lung cancer patients showed higher CHIP prevalence (44/100 for patients vs 5/42 for controls; $P = 0.01$). In addition, lung squamous cell carcinoma patients showed increased burden of larger clones compared to lung adenocarcinoma patients (8/43 for LUSC vs 2/50 for LUAD; $P = 0.04$). Furthermore, single cell RNA-seq analysis of the matched patients showed significant enrichment of inflammatory pathways mediated by NF-κB in myeloid clusters of the severe CHIP group. Our findings suggest minimal involvement of CHIP mutation and clonal dynamics during immunotherapy but a possible role of CHIP as an indicator of immunologic response in NSCLC patients.

Introduction

Recent studies have revealed that somatic mutations detected in hematopoietic malignancies are also commonly observed with clonal expansion in the individuals with no hematological conditions (1). Among such individuals, asymptomatic cases with a variant allele frequency (VAF) of ≥ 2% are defined as clonal hematopoiesis of indeterminate potential (CHIP) (2). In general, CHIP is strongly associated with aging; and mutagenic anticancer therapies and smoking can also induce clonal expansion in specific genes (1,3). Longitudinal monitoring of CHIP mutations offers important clues into the clonal dynamics of blood (4).

Hematopoietic stem and progenitor cells (HSPCs) carrying CHIP mutations serve as a major source of expanded clones, and even a few cells are sufficient to influence clonal homeostasis (5). Representative CHIP mutation-carrying genes, such as *DNMT3A*, *ASXL1*, and *TET2*, impact the myeloid population, including monocytes and granulocytes, and have been linked to hematologic malignancies such as myelodysplastic syndrome and acute myeloid leukemia through inflammation mediated by the NF-kB and NLRP3 inflammasome (6-8). The prevalence of CH in cancer patients is also significantly higher than that in healthy populations, and high CH burden has been implicated in a negative effect on overall survival in multiple cancer types (9,10). CHIP is also associated with diseases beyond hematologic malignancies, such as atherosclerosis, type 2 diabetes, and chronic obstructive pulmonary disease (11-13). Multiple studies, including both population and pan-cancer cohorts, have explored the correlation of CHIP with solid tumors (9,10,14,15). However, establishing a causal

relationship has proven challenging due to confounding factors such as history of smoking and anticancer therapy, both of which act as risk factors for CHIP (10). Non-small cell lung cancer (NSCLC) stands out globally as the solid tumor with the highest mortality burden (16). The difficulty in early detection of NSCLC significantly contributes to its mortality rate being high compared to its incidence rate. Recent studies examining CHIP within lung cancer cohorts and general population indicate a possible association of lung cancer risk with CH (15,17,18). As mentioned above, factors such as a patient's history of anticancer treatments and smoking serve as common risk factors for CH; however, the interplay between lung cancer and CHIP is not yet fully understood. Immune checkpoint inhibitors (ICI) targeting PD-(L)1 and CTLA-4 have revolutionized treatment of advanced NSCLC, both as monotherapies and in combination with chemotherapy (19). However, response rates to ICI treatments remain relatively low. The predictive ability of markers like PD-1 expression and tumor mutational burden (TMB) is insufficient, requiring the identification of additional markers. Recent studies suggest a marginal effect of ICI on the clonal dynamics of CHIP (20,21), but the current understanding of the role of CHIP in ICI-treated patients is not comprehensive. Also, only limited analysis has been performed concerning the impact of CHIP on cancer patients' blood. To understand the interplay between ICI and CHIP in NSCLC patients, we conducted CHIP-targeted panel sequencing and single-cell RNA sequencing (scRNA-seq) on blood samples collected before and 1-3 weeks after ICI administration. Our results indicate a marginal interaction between ICI and CHIP, but also validate that high CHIP burden manifests an inflammatory phenotype in metastatic NSCLC patients' myeloid cells.

Materials and Methods

Patients

Description of the discovery cohort is previously reported (22). Briefly, 100 patients (50 LUAD, 43 LUSC and 7 others) with metastatic NSCLC stage IV who were treated with anti-PD(L)1 (*e.g.*, Atezolizumab, Nivolumab, and Pembrolizumab) were recruited at Samsung Medical Center (SMC) under the permission of SMC Institutional Review Board (No. 2018-04-048, 2022-01-094). Whole blood sample was acquired before and one to three weeks after the treatment and used for genome sequencing and scRNA-seq. For replication, blood samples from additional 180 lung cancer patients (125 LUAD and 55 LUSC) were used for genome analysis (23).

Targeted sequencing and whole exome sequencing (WES)

We designed a targeted sequencing panel comprising 167 selected genes associated with putative drivers of CHIP at Twist Bioscience (South San Francisco, CA). For WES, we utilized the SureSelect V5 panel from Agilent (Santa Clara, CA). CHIP-targeted sequencing and WES were both conducted at Samsung Genome Institute, respectively employing 50 ng and 200 ng of genomic DNA extracted from whole blood. The targeted sequencing panel was captured and sequenced using NextSeq 2000, with paired-end sequencing performed at a read length of 150 base pairs (bps).

Variant filtering and calling of CHIP variants

For somatic variant calling, the Sarek pipeline (version 3.2.1) implemented in Nextflow was used (24). In brief, the pipeline employed GATK's standard data preprocessing procedure and Mutect2 for somatic variant calling based on a BED file (25). All passed

calls were collected and annotated using Snpeff (version 4.3 and db version 105) and VEP (version 108) (26,27). To filter CHIP variants, a stepwise approach based on the following criteria was applied. First, variants should meet the following conditions: a) variant allele frequency (VAF) 2-35%, b) variant depth >500x for the discovery set and >40x for the WES replication set, c) 4 or more (2 or more for the WES replication set) variant-covering read pairs in each forward and reverse direction, d) global allele frequency from gnomAD <1e-5, or <1e-3 if a variant is in the COSMIC database (28)) allele count <5% in all processed samples, and f) no homopolymer signature found in 6 bp upstream or downstream of the candidate variant. Next, we called CHIP variants based on a) protein sequence alterations, including exonic splicing variants, and b) manual assessment of variant calls using IGV (29). Finally, we curated data from previous studies and the COSMIC database, assigning variants reported more than 10 times or matched within a predefined list as CHIP with putative drivers (called as clonal hematopoiesis of indeterminate potential with putative drivers, or CHIP-PD) (30).

Power estimation

The power of the Poisson binomial test was calculated as the probability that a variant would be detected exclusively in one of two arbitrary groups. In brief, the most prevalent CHIP variant, *DNMT3A* p.Arg882, constitutes approximately 10% of the *DNMT3A* CHIP cases. When this variant is totally absent in the one group, a minimum of 16 *DNMT3A* CHIP patients in the other group would be required to achieve a power of 0.8. Considering the current prevalence of *DNMT3A* CHIP in our cohort (8/50 CHIP samples), this suggests that a minimum of 100 CHIP samples per group is required for determining the prevalence of any specific variant.

scRNA-seq pre-processing and analysis

The scRNA-seq dataset was generated using the 10x Genomics 5' single-cell kit and was partially derived from a previous study (22). Reads were aligned to the hg38 reference using 10X CellRanger (ver. 7.0). Contaminant RNA barcodes were removed with cellbender and sample demultiplexing was performed by demuxlet (31,32). After filtering the cell expression matrix, we used Seurat (v 4.3.0) for downstream analysis (33). At the sample level, we filtered out cells with mitochondrial RNA content >15%, UMI counts <200 or >6,000, and calculated library-size corrected counts using SCTransform (34). Then, we computed PCA based on the normalized SCTransform values and integrated all samples using Harmony (35). Using the corrected PCA, we performed unsupervised clustering via Louvain algorithms and UMAP. For each cluster, we identified marker genes using the Wilcoxon rank-sum test with presto and validated annotations using Azimuth (33).

Gene set enrichment analysis (GSEA)

GSEA was employed to compare the expression patterns according to CHIP status based on their concordance to the hallmark pathways in the Molecular Signatures Database (MSigDB) (36). For GSEA of scRNA-seq data, we used fgsea (37). Briefly, the AUC rank of DEGs obtained with and without CHIP for each cluster was used as input. Significant gene set enrichment was determined only for adjusted $P < 0.05$.

scRNA-seq weighted correlation network analysis (WGCNA)

We utilized WGCNA to construct a gene co-expression network from the gene expression profiles of myeloid cells utilizing the R package hdWGCNA (38). In brief, the co-expression network was constructed using a soft power of 12 for SCT-transformed

cell counts. Within the constructed co-expression network, we explored the type and specificity of genes in each module, excluding unassigned genes. The functionality of the black module was validated through GSEA and by examining its overlap with DEGs identified between CHIP groups (38).

Gene regulatory network (GRN) inference

To further understand the altered gene modulation in major cell clusters, including B cells, NK cells, CD4 T cells, CD8 T cells, and monocytes, we constructed a GRN using pySCENIC (39). We created meta-cells for each sample using SEACells and employed these meta-cells for GRN calculation to address the sparse nature of scRNA expression and enhance resolution (40,41). The full transcriptome matrix and the GRN derived from meta-cells were then used to analyze regulon signatures for each cell. Finally, we mapped the regulon AUC values of each cell onto Seurat's UMAP and visualized the results using Seurat.

Cell to cell communication analysis

To understand the interactions between myeloid and other cell types, we utilized the Python package CellphoneDB (42), which curates human cell-cell interactions. The counts matrix, processed with SCTransform, was divided according to CHIP status and used as input for CellphoneDB. Significant interactions were then counted to determine the number of cell-cell interactions, and myeloid-related interactions were compared between samples of different CHIP status.

Statistical analysis

We present raw data if applicable. For scRNA-seq data, scaled and log-normalized raw counts from cells are displayed in violin plots. To determine statistical significance

between high-burden, CHIP, and negative groups, the Wilcoxon test or Kruskal-Wallis test was used for continuous variables and the Poisson binomial test or logistic regression for categorical variables. Statistical metrics from GESA were calculated using the fgsea package in R. All statistical analyses were conducted in R 4.2.1.

Data availability statement

Raw data for this study were generated at Samsung Genome Institute. Derived data supporting the findings of this study are available from the corresponding author upon request.

Results

CHIP profiles in metastatic NSCLC patients treated with ICI

To investigate the interplay between ICI treatment and CHIP in NSCLC, we collected blood samples before and after ICI treatment from 100 metastatic NSCLC patients (Fig. 1, Supplementary Table S1). These samples were utilized for a CHIP targeted sequencing panel (median depth ~600x) and scRNA-seq (**Fig. 1A** and Methods). At the baseline, we identified 67 CHIP mutations in 100 patients, 26 of them in known putative driver genes of hematopoietic cancer (CHIP-PD). After treatment, 68 CHIP and 32 CHIP-PD mutations were found in 91 patients who passed quality control assessment (**Fig. 1A**, Supplementary Fig. S1-S3 and Supplementary Table S3).

Among the CHIP genes, *DNMT3A*, *PPM1D* and *TET2* were the most frequently mutated (**Fig. 1B**). *PPM1D* truncating mutations were also common, and have previously demonstrated association with patients' smoking habits and possible history with chemotherapy (3,9). Variants involving the *DNMT3A* p.Arg882 residue and *TP53*, which are frequent in CHIP, were not prevalent in our dataset, although our panel extensively covered these regions and these loci were manually inspected (data not shown).

Overall, we observed ICI treatment to effect minimal changes in CHIP burden (**Fig. 1B**). Also, VAF was highly correlated across all traced mutations and not dependent on clinical responses (**Fig. 1C** and **D**). However, we noted increased number of CHIP mutations with VAF >10% in patients with lung squamous cell carcinoma (LUSC). While the specific mutated genes differed between patients with different histology or responses, mutated gene burden did not significantly differ between groups.

Prevalence of high CHIP burden showed bias to LUSC patients

Next, we investigated the association of CHIP status with clinical parameters in the cohort (**Fig. 2**). CHIP prevalence was significantly higher in the lung cancer cohort compared to the control group (controls 5/42 vs. patients 44/100, age-adjusted logistic regression $P = 0.01$; **Fig. 2A**, Supplementary Table S1 and S2). However, the number of CHIP-positive patients did not differ before and after ICI treatment, nor in relation to treatment response (**Fig. 2B** and **C**). As previously reported, CHIP burden is heavily dependent on age (**Fig. 2D**). Smoking history may also impact CHIP prevalence, but we could not find a significant contribution of smoking, as almost all NSCLC patients in our study have a smoking history (CHIP prevalence of non-smoker 2/9 vs. smoker 42/91; Supplementary Fig. S4).

Subsequently, we examined clonal size in the cohort, as that has been reported as a risk factor of clinical outcome (43). Taking CHIP VAF as an indicator of clonal size, we observed that high CHIP burden was prevalent in LUSC, while multiple concurrent CHIP mutations were common in lung adenocarcinoma (LUAD) patients (**Fig. 2C** and Supplementary Fig. S5). To determine if this finding would reproduce in an independent cohort, we performed CHIP mutation calling using WES data of median ~80x depth from 180 additional peripheral blood mononuclear cell (PBMC) samples from an independent cohort of NSCLC patients (**Fig. 2E-G**, Supplementary Fig. S6, Supplementary Table S1, S3-S4 and Methods). The results showed similar patterns between the two cohorts, but the replication cohort alone was not powerful enough to achieve significance for difference in clonal sizes based on pathology (**Fig. 2E-G**).

High CHIP burden causes single-cell transcriptomic changes

Next, to test whether the presence of CHIP and ICI prognosis affects blood cell transcriptomes, we analyzed PBMC scRNA-seq data from 63 samples of the discovery cohort. After quality control processes, we retrieved 468,596 cells in a total of 26 clusters defined by systemic integration and clustering using the Louvain algorithm, then validated using Azimuth (**Fig. 3A**, Supplementary Fig. S7A-B, Supplementary Table S5, and Methods) (33). Cell composition analysis showed an increased proportion of NK cells and a decreased CD4 TEM population after ICI administration, but high variability between samples precluded systematic differences by CHIP status (Supplementary Fig. S7C). Subsequently, DEG analysis using Wilcoxon's rank-sum test according to CHIP VAF bins in each cluster (see Methods) highlighted genes important in immune response and transcriptional activation (**Fig. 3C**, Supplementary Fig. S8). To further implicate these DEGs functionally, a gene set enrichment test was performed using the AUC statistics from DEGs and the hallmark pathway gene sets in msigDB (36). The results indicated activation of the TNF signaling pathway via *NFKB* in the high CH burden group for most cell lineages, including both classical dendritic cells (DC) and NK cells (**Fig. 3D**, **3E**, Supplementary Fig. S9, and Supplementary Table S6; see Methods). To understand possible interactions between ICI treatment response and CHIP burden, we grouped samples by CHIP burden and ICI response (Supplementary Fig. S10A-B). However, the interactions mostly depended on CH status only, and not on response to ICI.

Implication of high CHIP burden in inflammatory signatures of myeloid cells

We performed a gene network analysis to comprehensively analyze the increased expression of inflammatory signaling genes in patients with higher CHIP burden. First, we used hdWGCNA to construct co-expression networks for genes in myeloid cells, which yielded a black module that exhibited myeloid specificity and represented most of the signaling-related DEGs identified via GSEA (Supplementary Fig. S6 and S11 and Supplementary Table S7) (38). We determined gene ontology (GO) term enrichment in this module and found the inflammatory and transcriptional activation pathways observed in DEGs to be clearly identified (Supplementary Fig. S11C). Additionally, further GSEA on the myeloid-specific module showed its function to resemble that of the overall set of DEGs (Supplementary Fig. S11D).

In addition to WGCNA, GRN analysis using SCENIC (39) (Supplementary Fig. S12) indicated that the regulatory activity of transcription factors in the AP-1 and NF-kB pathways was particularly notable in the myeloid lineage of high-burden patients (Supplementary Fig. S12A-B). Both the NF-kB regulon and the ATF3 regulon, which represents the AP-1 pathway, were activated in the myeloid cluster of high-burden patients (Supplementary Fig. S12C-E).

Recent studies have shown that expansion of HSPC-derived monocytes is the cause of CHIP, and representative CHIP genes, including *DNMT3A* and *TET2*, exert immune-regulatory effects through the NLRP3 inflammasome (5,8,44). Our network analysis also points to this pathway (Supplementary Fig. S12). Notably, a VAF of 10% in heterozygote mutations corresponds to approximately 20% of cells; thus, the presence

of such mutations in HSPC-derived monocytes of high-burden patients would have resulted in an overall transcriptional profile change for blood monocytes.

In addition to conducting gene network analysis, we explored the effect of the inflammatory signature on cell-cell interaction using CellPhoneDB (Supplementary Fig. S13) (42). The number of significant cell-cell interactions did not differ significantly between high-burden CHIP samples and negative samples, although the number of cells in each group seemed likely to have an impact (Supplementary Fig. S13A-B). However, when the interactions in the TNF and IL-1B pathways were analyzed separately, monocytes and DCs in high-burden cases were activated as receptors for TNF and Lymphotoxin A (Supplementary Fig. S13C). Collectively, our results suggest a sensitized response of monocytes to inflammation in lung cancer patients with high CHIP burden.

Effect of CHIP on clinical outcome

Finally, to understand whether clonal hematopoiesis status can affect the clinical outcomes of lung cancer patients undergoing ICI treatment, we conducted a survival analysis (Supplementary Fig. S14). Our results showed the progression-free survival (PFS) of patients after ICI treatment to be reduced in the high-burden CHIP group (825 vs 404, $P = 0.056$), with no difference in ICI response rates between the negative and high-burden CHIP groups. Previous studies have already observed a negative effect on life expectancy in CHIP patients with solid tumors. Therefore, it is possible that these results will be validated by analyzing additional samples.

Discussion

To address the influence of CH in lung cancer patients with differential prognosis following ICI, we sought to determine the following: (i) CH profiles from metastatic NSCLC patients, (ii) changes in PBMC gene expression by scRNA-seq, (iii) relationship of CH burden and transcriptome profile, and (iv) potential implication for the clinical outcomes. Here we report that CH prevalence was significantly higher in NSCLC patients and higher burden of mutations with VAF of greater than 10% was frequently observed in LUSC. Using scRNA-seq analysis, we found myeloid cells in NSCLC patients with high CH burden to exhibit an altered inflammatory milieu via increased NF-kB pathway expression. Our results provide possible explanations for heterogeneity in response to immunotherapy and suggest immune response to be altered due to CHIP in cancer patients.

Recent studies have employed mouse models and single-cell transcriptomics to understand the roles of specific mutations in CH and provided evidence of their implications for inflammatory response in the myeloid lineage (45,46). In particular, a study linking COVID-19 and CH showed an increased inflammatory response from monocytes in severe COVID-19 patients with CH, mediated via IFN-γ and TNF-α signaling signatures (47).Likewise, we observed that myeloid cells from NSCLC patients with CH tend to harbor increased inflammatory response, and this response is mediated by NF-kB pathway signaling (**Fig. 3**).

Our GSEA results specifically indicated the enrichment of TNF signaling and hypoxia pathways in most clusters of patients with severe CH (Supplementary Fig. S9). The

leading-edge genes from GSEA results showed core genes such as *FOS*, *DUSP1*, *JUN*, and *PPP1R15A,* which are known to play critical roles in the inflammatory phenotypes of immune cells, were shared between the TNF signaling and hypoxia pathways in all significant clusters (Supplementary Fig. S15). Furthermore, gene regulatory network analysis using SCENIC indicated a higher enrichment of inflammatory signatures in myeloid lineages (Supplementary Fig. S9), highlighting the pronounced inflammatory phenotype of CH clones in these cell lineages. To gain insight into the functional implications of myeloid-driven inflammatory gene pathways within CH in NSCLC patients, we conducted gene network analyses focusing on the myeloid lineage. The "black module" from the WGCNA analysis (Supplementary Fig. S9) and the NFKB1 and ATF3 GRN from SCENIC (Supplementary Fig. S11) revealed activation and adaptation of myeloid-specific inflammatory pathways (38,39). Additionally, cell-cell interaction analysis using cellphoneDB unveiled notable interactions between myeloid cells and other lineages mediated through the TNF pathway (Supplementary Fig. S12) (42).

As ICIs modulate pathways involved in immune response rather than cell proliferation, it is reasonable that ICIs do not exert a direct influence on CH development (21). Also, our results did not support significant difference of ICI outcome based on CH burden (**Fig. 2**). According to our power analysis, a minimum of 100 samples per group with CHIP is required to achieve a statistical power of 0.8 from pair-wise comparisons between each genes and variants (Methods for details). However, our findings do suggest that elevated inflammatory activity in the bloodstream may impact long-term patient survival pre- and post-ICI administration (Supplementary Fig. S14). Therefore, a future study with a larger sample size may yet clarify a causal relationship of CH with

ICI outcome. Notably, involvement of the IL-1 pathway in the aggressiveness and metastasis of lung cancer has been reported in experimental studies (48,49), and inhibition of IL-1β has proven effective in reducing lung cancer-related mortality, as shown in the CANTOS trial (50). Indeed, our analysis indicated a modest decrease in PFS in the ICI-responder group with high-CH patients (Supplementary Fig. S14, P = 0.056). Our study design is potentially limited by the relatively short observation period for defining ICI response Indeed, it is worth noting that more potent factors, such as smoking and anticancer therapies, may have a more substantial impact than immunotherapy (10). Consistent with this postulation, our analyses indicated minimal changes in in CH-related signatures before and after ICI treatment (**Fig. 3**, Supplementary S7, S9, and S10).

Also, the distinct characteristics of our cohort can be confounders for our results. Compared to control patients, our cohort was biased to slightly older ages, a higher prevalence of smoking habits, and with a higher proportion of males (mean age: 64.1 vs. 58.9; current smokers: 37/100 vs. 11/42; male/female: 91/9 vs. 18/24 Supplementary Figures S1 and S3). However, previous studies have reported similar prevalence rates of clonal hematopoiesis in NSCLC patients, aligned with our findings (9,51). Moreover, our most prevalent CH mutations, including *DNMT3A*, *TET2*, and *PPM1D*, were marginally affected by smoking, and this trend has been consistently observed in both healthy populations and NSCLC patients (10,51,52).

Here we have illustrated increased inflammatory activation in myeloid cells of NSCLC patients with high CH burden. Our scRNA-seq analysis has revealed that NF-kB signaling mediated by the IL-1β and the TNF superfamilies may enhance interactions of

myeloid cells with other cell lineages. These findings propose the potential utility of CH as a predictive marker for classifying NSCLC patients. Future investigations involving ICI-treated patients and their clinical outcomes may shed light on the impact of CH, offering a basis for the development of adjuvant anticancer strategies aimed at modulating CH.

Authors' contributions

H. Sim: Data curation, formal analysis, validation, investigation, writing-original draft. **H.-J. Park**: Conceptualization, data curation, formal analysis, methodology, investigation. **G.-H. Park**: Data curation, formal analysis, methodology, investigation. **Y. J. Kim**: Data curation, methodology, investigation. **W.-Y. Park**: Conceptualization, investigation, methodology. **S. H. Lee**: Conceptualization, investigation, writing-original draft. **M. Choi**: Conceptualization, investigation, writing-original draft.

Acknowledgements:

We thank George Vassiliou for critical comments. The authors also acknowledge the Korea Research Environment Open Network (KREONET) service and the usage of the Global Science Experimental Data Hub Center (GSDC) provided by Korea Institute of Science and Technology Information (KISTI). This work was supported in part by the grants from the Korean Research Foundation (NRF-2021R1A2C3014067, NRF-RS-2023-00207857 to M. Choi, and NRF-2020R1A2C3006535 to S.-H. Lee), the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) (HR20C0025 to S.-H. Lee), and Future Medicine 20*30 Project of the Samsung Medical Center (SMX1230041 to S.-H. Lee).

Conflict of interest:

Authors declare no conflict of interest.

Figures

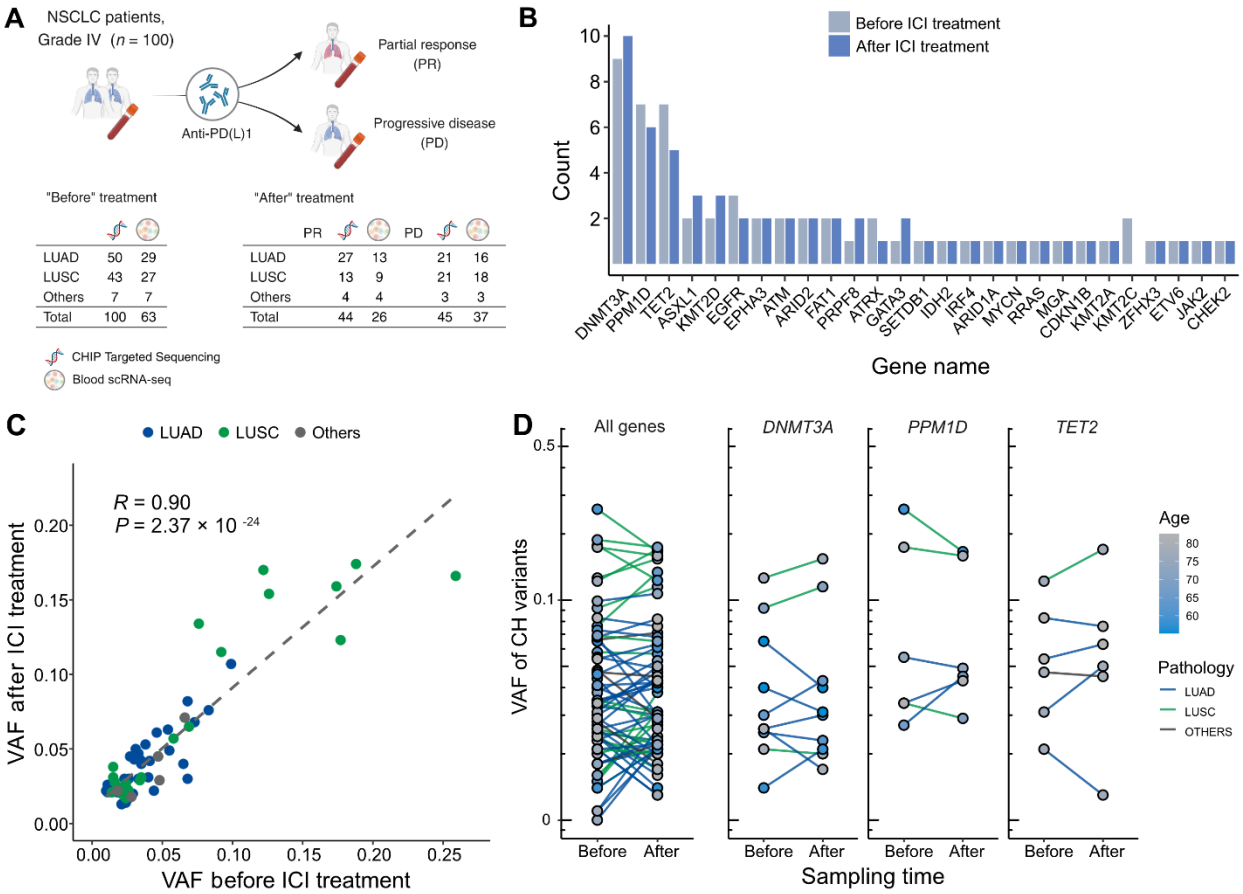

Figure 1. Profile of CH in NSCLC patients treated with ICI. **A,** Overall study design. The numbers indicate sample counts in each group. **B,** Frequency of CH variant detection in each gene in NSCLC samples before (light blue) and after (blue) ICI treatment. **C,** Comparison of variant allele frequency before and after ICI treatment, divided by pathology type (LUAD, LUSC, other). **D,** Effect of ICI treatment on the clonal landscape of all genes (left) and frequently detected CH genes.

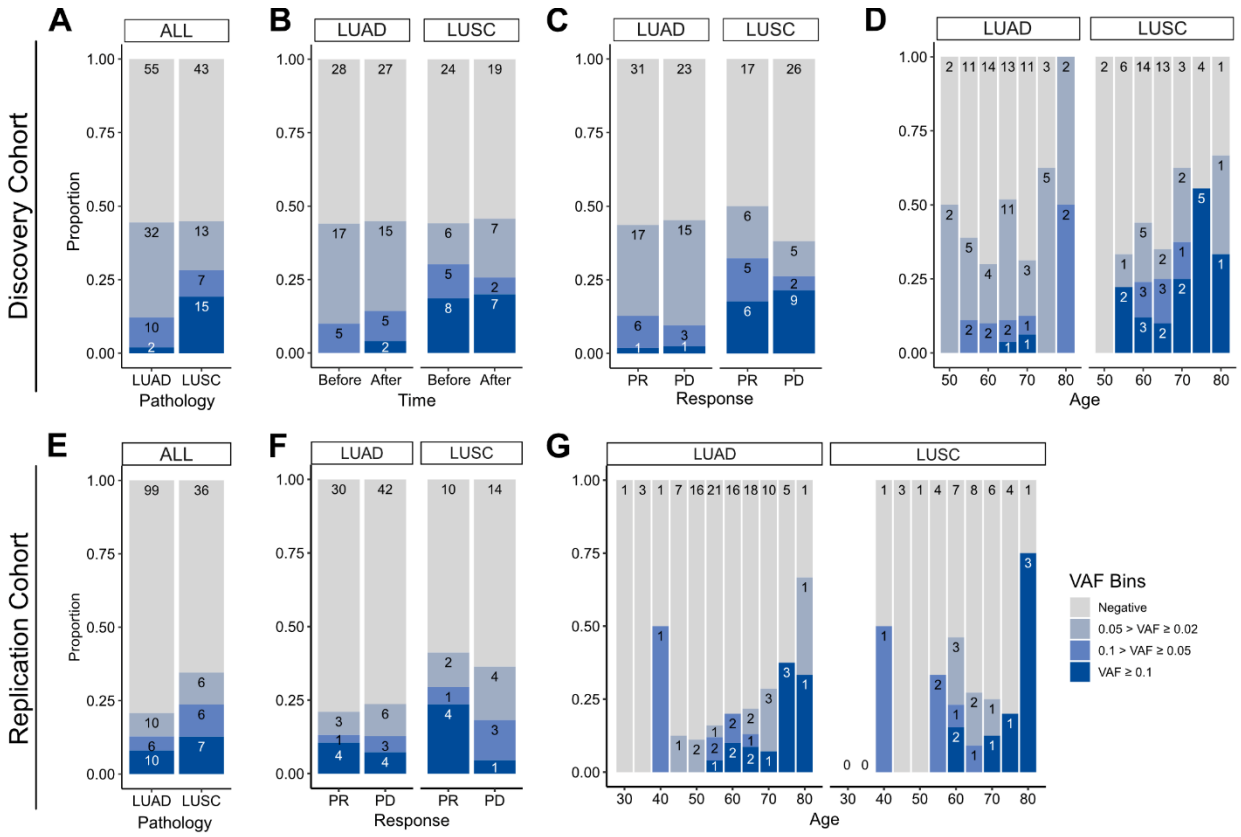

Figure 2. Prevalence of CH in relation to various parameters. **A-D,** discovery cohort (*n* = 100), **E-G,** replication cohort (*n* = 180; see Methods for description). **A**, **E,** CH prevalence by pathology (LUSC and LUAD). **B,** Effect of ICI treatment on CH. **C**, **F,** Effect of post-ICI prognosis to ICI on CH. **D**, **G,** Age distribution of the cohort, stratified by CHIP allele frequency and pathology status.

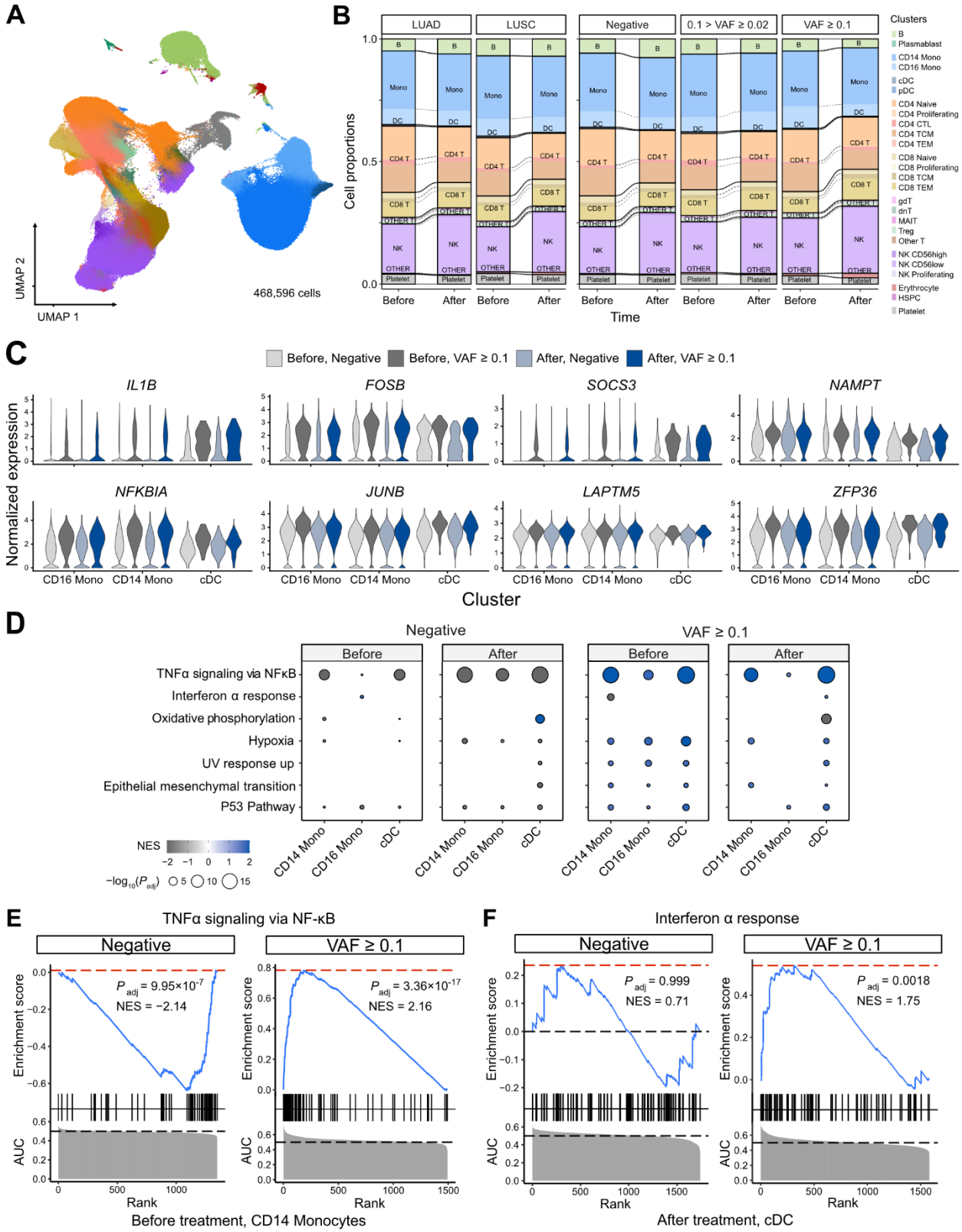
A
UMAP 2
UMAP 1
468,596 cells
B
LUAD
LUSC
Negative
0.1 > VAF ≥ 0.02
VAF ≥ 0.1
Cell proportions
Before
After
Time
Clusters
B
Plasmablast
CD14 Mono
CD16 Mono
cDC
pDC
CD4 Naive
CD4 Proliferating
CD4 CTL
CD4 TCM
CD4 TEM
CD8 Naive
CD8 Proliferating
CD8 TCM
CD8 TEM
gdT
dnT
MAIT
Treg
Other T
NK CD56high
NK CD56low
NK Proliferating
Erythrocyte
HSPC
Platelet
C
Before, Negative
Before, VAF ≥ 0.1
After, Negative
After, VAF ≥ 0.1
IL1B
FOSB
SOCS3
NAMPT
NFKBIA
JUNB
LAPTM5
ZFP36
Normalized expression
CD16 Mono
CD14 Mono
cDC
Cluster
D
Negative
VAF ≥ 0.1
Before
After
TNFα signaling via NFκB
Interferon α response
Oxidative phosphorylation
Hypoxia
UV response up
Epithelial mesenchymal transition
P53 Pathway
NES
−log10(Padj)
E
TNFα signaling via NF-κB
Negative
VAF ≥ 0.1
Padj = 9.95×10-7
NES = −2.14
Padj = 3.36×10-17
NES = 2.16
Enrichment score
AUC
Rank
Before treatment, CD14 Monocytes
F
Interferon α response
Negative
VAF ≥ 0.1
Padj = 0.999
NES = 0.71
Padj = 0.0018
NES = 1.75
After treatment, cDC

Figure 3. scRNA-seq analysis detected myeloid-specific inflammatory signatures in patients with high CH burden. **A,** UMAP plot of scRNA-seq data from the discovery cohort (*n* = 63). **B,** Effect of pathology (left) and VAF (right) on cell composition. **C,** Expression of selected genes in the NF-κB pathway from the scRNA-seq data. **D,** GSEA of DEGs from myeloid populations. Color represents normalized effect score (NES), and dot size represents adjusted *P*-values. **E,** GSEA plots of selected pathways from **C**.

References

1. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* 2014;**371**(26):2488-98 doi 10.1056/NEJMoa1408617.
2. Jaiswal S, Ebert BL. Clonal hematopoiesis in human aging and disease. *Science* 2019;**366**(6465):eaan4673 doi doi:10.1126/science.aan4673.
3. Hsu JI, Dayaram T, Tovy A, De Braekeleer E, Jeong M, Wang F, *et al.* PPM1D Mutations Drive Clonal Hematopoiesis in Response to Cytotoxic Chemotherapy. *Cell Stem Cell* 2018;**23**(5):700-13 e6 doi 10.1016/j.stem.2018.10.004.
4. Uddin MM, Zhou Y, Bick AG, Burugula BB, Jaiswal S, Desai P, *et al.* Longitudinal profiling of clonal hematopoiesis provides insight into clonal dynamics. *Immun Ageing* 2022;**19**(1):23 doi 10.1186/s12979-022-00278-9.
5. Welch JS, Ley TJ, Link DC, Miller CA, Larson DE, Koboldt DC, *et al.* The origin and evolution of mutations in acute myeloid leukemia. *Cell* 2012;**150**(2):264-78 doi 10.1016/j.cell.2012.06.023.
6. Muto T, Walker CS, Choi K, Hueneman K, Smith MA, Gul Z, *et al.* Adaptive response to inflammation contributes to sustained myelopoiesis and confers a competitive advantage in myelodysplastic syndrome HSCs. *Nat Immunol* 2020;**21**(5):535-45 doi 10.1038/s41590-020-0663-z.
7. Cai Z, Kotzin JJ, Ramdas B, Chen S, Nelanuthala S, Palam LR, *et al.* Inhibition of Inflammatory Signaling in Tet2 Mutant Preleukemic Cells Mitigates Stress-Induced Abnormalities and Clonal Hematopoiesis. *Cell Stem Cell* 2018;**23**(6):833-49 e5 doi 10.1016/j.stem.2018.10.013.
8. Abplanalp WT, Cremer S, John D, Hoffmann J, Schuhmacher B, Merten M, *et al.* Clonal Hematopoiesis-Driver DNMT3A Mutations Alter Immune Cells in Heart Failure. *Circ Res* 2021;**128**(2):216-28 doi 10.1161/CIRCRESAHA.120.317104.
9. Coombs CC, Zehir A, Devlin SM, Kishtagari A, Syed A, Jonsson P, *et al.* Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell* 2017;**21**(3):374-82 e4 doi 10.1016/j.stem.2017.07.010.
10. Bolton KL, Ptashkin RN, Gao T, Braunstein L, Devlin SM, Kelly D, *et al.* Cancer therapy shapes the fitness landscape of clonal hematopoiesis. *Nat Genet* 2020;**52**(11):1219-26 doi 10.1038/s41588-020-00710-0.
11. Miller PG, Qiao D, Rojas-Quintero J, Honigberg MC, Sperling AS, Gibson CJ, *et al.* Association of clonal hematopoiesis with chronic obstructive pulmonary disease. *Blood* 2022;**139**(3):357-68 doi 10.1182/blood.2021013531.
12. Tobias DK, Manning AK, Wessel J, Raghavan S, Westerman KE, Bick AG, *et al.* Clonal Hematopoiesis of Indeterminate Potential (CHIP) and Incident Type 2 Diabetes Risk. *Diabetes Care* 2023;**46**(11):1978-85 doi 10.2337/dc23-0805.
13. Jaiswal S, Natarajan P, Silver AJ, Gibson CJ, Bick AG, Shvartz E, *et al.* Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N Engl J Med* 2017;**377**(2):111-21 doi 10.1056/NEJMoa1701719.
14. Xie M, Lu C, Wang J, McLellan MD, Johnson KJ, Wendl MC, *et al.* Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med* 2014;**20**(12):1472-8 doi 10.1038/nm.3733.
15. Stacey SN, Zink F, Halldorsson GH, Stefansdottir L, Gudjonsson SA, Einarsson G, *et al.* Genetics and epidemiology of mutational barcode-defined clonal hematopoiesis. *Nat Genet* 2023 doi 10.1038/s41588-023-01555-z.

16. Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. *Nat Rev Clin Oncol* 2023;**20**(9):624-39 doi 10.1038/s41571-023-00798-3.
17. Hong W, Li A, Liu Y, Xiao X, Christiani DC, Hung RJ*, et al.* Clonal Hematopoiesis Mutations in Patients with Lung Cancer Are Associated with Lung Cancer Risk Factors. *Cancer Research* 2022;**82**(2):199-209 doi 10.1158/0008-5472.Can-21-1903.
18. Tian R, Wiley B, Liu J, Zong X, Truong B, Zhao S*, et al.* Clonal Hematopoiesis and Risk of Incident Lung Cancer. *J Clin Oncol* 2023;**41**(7):1423-33 doi 10.1200/JCO.22.00857.
19. Reck M, Rodriguez-Abreu D, Robinson AG, Hui R, Csoszi T, Fulop A*, et al.* Five-Year Outcomes With Pembrolizumab Versus Chemotherapy for Metastatic Non-Small-Cell Lung Cancer With PD-L1 Tumor Proportion Score >/= 50. *J Clin Oncol* 2021;**39**(21):2339-49 doi 10.1200/JCO.21.00174.
20. Miller PG, Sperling AS, Brea EJ, Leick MB, Fell GG, Jan M*, et al.* Clonal hematopoiesis in patients receiving chimeric antigen receptor T-cell therapy. *Blood Adv* 2021;**5**(15):2982-6 doi 10.1182/bloodadvances.2021004554.
21. Miller PG, Gibson CJ, Mehta A, Sperling AS, Frederick DT, Manos MP*, et al.* Fitness Landscape of Clonal Hematopoiesis Under Selective Pressure of Immune Checkpoint Blockade. *JCO Precis Oncol* 2020;**4** doi 10.1200/PO.20.00186.
22. Kim H, Park S, Han KY, Lee N, Kim H, Jung HA*, et al.* Clonal expansion of resident memory T cells in peripheral blood of patients with non-small cell lung cancer during immune checkpoint inhibitor treatment. *J Immunother Cancer* 2023;**11**(2) doi 10.1136/jitc-2022-005509.
23. Litchfield K, Reading JL, Puttick C, Thakkar K, Abbosh C, Bentham R*, et al.* Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell* 2021;**184**(3):596-614.e14 doi https://doi.org/10.1016/j.cell.2021.01.002.
24. Garcia M, Juhos S, Larsson M, Olason PI, Martin M, Eisfeldt J*, et al.* Sarek: A portable workflow for whole-genome sequencing analysis of germline and somatic variants. *F1000Res* 2020;**9**:63 doi 10.12688/f1000research.16665.2.
25. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A*, et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;**20**(9):1297-303 doi 10.1101/gr.107524.110.
26. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L*, et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* 2012;**6**(2):80-92 doi 10.4161/fly.19695.
27. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A*, et al.* The Ensembl Variant Effect Predictor. *Genome Biol* 2016;**17**(1):122 doi 10.1186/s13059-016-0974-4.
28. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N*, et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res* 2019;**47**(D1):D941-D7 doi 10.1093/nar/gky1015.
29. Robinson JT, Thorvaldsdottir H, Wenger AM, Zehir A, Mesirov JP. Variant Review with the Integrative Genomics Viewer. *Cancer Res* 2017;**77**(21):e31-e4 doi 10.1158/0008-5472.CAN-17-0337.
30. Niroula A, Sekar A, Murakami MA, Trinder M, Agrawal M, Wong WJ*, et al.* Distinction of lymphoid and myeloid clonal hematopoiesis. *Nat Med* 2021;**27**(11):1921-7 doi 10.1038/s41591-021-01521-4.
31. Fleming SJ, Chaffin MD, Arduini A, Akkad AD, Banks E, Marioni JC*, et al.* Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat Methods* 2023;**20**(9):1323-35 doi 10.1038/s41592-023-01943-7.

32. Kang HM, Subramaniam M, Targ S, Nguyen M, Maliskova L, McCarthy E, *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat Biotechnol* 2018;**36**(1):89-94 doi 10.1038/nbt.4042.
33. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A, *et al.* Integrated analysis of multimodal single-cell data. *Cell* 2021;**184**(13):3573-87 e29 doi 10.1016/j.cell.2021.04.048.
34. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* 2019;**20**(1):296 doi 10.1186/s13059-019-1874-1.
35. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 2019;**16**(12):1289-96 doi 10.1038/s41592-019-0619-0.
36. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;**1**(6):417-25 doi 10.1016/j.cels.2015.12.004.
37. Sergushichev AA. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv* 2016:060012 doi 10.1101/060012.
38. Morabito S, Reese F, Rahimzadeh N, Miyoshi E, Swarup V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep Methods* 2023;**3**(6):100498 doi 10.1016/j.crmeth.2023.100498.
39. Van de Sande B, Flerin C, Davie K, De Waegeneer M, Hulselmans G, Aibar S, *et al.* A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat Protoc* 2020;**15**(7):2247-76 doi 10.1038/s41596-020-0336-2.
40. Bravo Gonzalez-Blas C, De Winter S, Hulselmans G, Hecker N, Matetovici I, Christiaens V, *et al.* SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat Methods* 2023;**20**(9):1355-67 doi 10.1038/s41592-023-01938-4.
41. Persad S, Choo ZN, Dien C, Sohail N, Masilionis I, Chaligne R, *et al.* SEACells infers transcriptional and epigenomic cellular states from single-cell genomics data. *Nat Biotechnol* 2023 doi 10.1038/s41587-023-01716-9.
42. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* 2020;**15**(4):1484-506 doi 10.1038/s41596-020-0292-x.
43. Kessler MD, Damask A, O'Keeffe S, Banerjee N, Li D, Watanabe K, *et al.* Common and rare variant associations with clonal haematopoiesis phenotypes. *Nature* 2022;**612**(7939):301-9 doi 10.1038/s41586-022-05448-9.
44. Sano S, Oshima K, Wang Y, MacLauchlan S, Katanasaka Y, Sano M, *et al.* Tet2-Mediated Clonal Hematopoiesis Accelerates Heart Failure Through a Mechanism Involving the IL-1beta/NLRP3 Inflammasome. *J Am Coll Cardiol* 2018;**71**(8):875-86 doi 10.1016/j.jacc.2017.12.037.
45. Nam AS, Dusaj N, Izzo F, Murali R, Myers RM, Mouhieddine TH, *et al.* Single-cell multi-omics of human clonal hematopoiesis reveals that DNMT3A R882 mutations perturb early progenitor states through selective hypomethylation. *Nat Genet* 2022;**54**(10):1514-26 doi 10.1038/s41588-022-01179-9.
46. Fuster JJ, MacLauchlan S, Zuriaga MA, Polackal MN, Ostriker AC, Chakraborty R, *et al.* Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science* 2017;**355**(6327):842-7 doi 10.1126/science.aag1381.
47. Choi B, Kang CK, Park S, Lee D, Lee AJ, Ko Y, *et al.* Single-cell transcriptome analyses reveal distinct gene expression signatures of severe COVID-19 in the presence of clonal hematopoiesis. *Exp Mol Med* 2022;**54**(10):1756-65 doi 10.1038/s12276-022-00866-1.

48. Li R, Ong SL, Tran LM, Jing Z, Liu B, Park SJ*, et al.* Chronic IL-1beta-induced inflammation regulates epithelial-to-mesenchymal transition memory phenotypes via epigenetic modifications in non-small cell lung cancer. *Sci Rep* 2020;**10**(1):377 doi 10.1038/s41598-019-57285-y.
49. Zhang J, Veeramachaneni N. Targeting interleukin-1beta and inflammation in lung cancer. *Biomark Res* 2022;**10**(1):5 doi 10.1186/s40364-021-00341-5.
50. Ridker PM, MacFadyen JG, Thuren T, Everett BM, Libby P, Glynn RJ*, et al.* Effect of interleukin-1beta inhibition with canakinumab on incident lung cancer in patients with atherosclerosis: exploratory results from a randomised, double-blind, placebo-controlled trial. *Lancet* 2017;**390**(10105):1833-42 doi 10.1016/S0140-6736(17)32247-X.
51. Hong W, Li A, Liu Y, Xiao X, Christiani DC, Hung RJ*, et al.* Clonal Hematopoiesis Mutations in Patients with Lung Cancer Are Associated with Lung Cancer Risk Factors. *Cancer Res* 2022;**82**(2):199-209 doi 10.1158/0008-5472.CAN-21-1903.
52. Levin MG, Nakao T, Zekavat SM, Koyama S, Bick AG, Niroula A*, et al.* Genetics of smoking and risk of clonal hematopoiesis. *Sci Rep* 2022;**12**(1):7248 doi 10.1038/s41598-022-09604-z.