DYNAMICS OF CLONAL HEMATOPOIESIS IN A SUPER-CENTENARIAN

RUNNING TITLE: CLONAL HEMATOPOIESIS IN A SUPERCENTENARIAN

Erik B. van den Akker[1,2,3+#], Stavros Makrodimitris[3,5+], Marc Hulsman[3,4,5], Martijn H. Brugman[6], Tanja Nikolic[6], Ted Bradley[4], Quinten Waisfisz[4], Frank Baas[7], Marja E. Jakobs[4], Daphne de Jong[8], P. Eline Slagboom[2], Frank J.T. Staal[6], Marcel J.T. Reinders[1,3], and Henne Holstege[3,4,5#]

Affiliations

1. Leiden Computational Biology Center, Leiden University Medical Center, Leiden, The Netherlands
2. Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
3. Pattern Recognition & Bioinformatics, Delft University of Technology, Delft, The Netherlands
4. Department of Clinical Genetics, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands
5. Alzheimer Center Amsterdam, Department of Neurology, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands
6. Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands
7. Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
8. Department of Pathology, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, Amsterdam, The Netherlands

+Equally contributed

#Correspondence: Henne Holstege h.holstege@amsterdamumc.nl, Erik B. van den Akker e.b.van_den_akker@lumc.nl

Material & correspondence

Henne Holstege, PhD; Amsterdam Alzheimer Center & Clinical Genetics, Amsterdam UMC—location VUmc, de Boelelaan 1118, 1081 HZ Amsterdam, The Netherlands; Tel: +31 (0)20 444 08 16; E-mail: h.holstege@amsterdamumc.nl; website: www.100plus.nl

Erik B. van den Akker, PhD; Leiden Computational Biology Center, Leiden University Medical Center, Leiden, The Netherlands; Einthovenweg 20, 2333 ZC, Leiden, The Netherlands; Tel: +31 (0)71 526 85 57; Fax: +31 (0)71 526 82 80; E-mail: e.b.van_den_akker@lumc.nl; website: http://www.lcbc.nl

Key points

1. Longitudinal blood sampling from a woman aged 103-111 revealed a dynamic clonal hematopoiesis contributing to myeloid and lymphoid subsets
2. Despite the highly advanced age and extreme clonal hematopoiesis we observed a functional T-cell compartment

Abstract

Background: Age Related Clonal Hematopoiesis (ARCH) arises when a hematopoietic stem cell (HSC) acquires a somatic mutation that confers a competitive growth advantage. ARCH is associated with an increased risk of myeloid leukemia, type II diabetes and cardiovascular disease, yet is also frequently observed in elderly with no apparent disease. In fact, ARCH can become very extreme in the healthy oldest old, which raises questions about the impact of extreme ARCH on overall immuno-competence.

Methods: We sequenced a longitudinal blood sample and its sorted subsets from a healthy elderly individual at age 103, 110 and 111 years. In addition, we characterized the immune system by flow cytometry, T-cell receptor excision circle (TREC), and *in vitro* proliferation assays.

Results: Sequencing revealed that a single *DNMT3A*-mutated HSC produces over 75% of the peripheral blood at age 110. A longitudinal analysis spanning 9-years revealed an extensive subclonal architecture, including a recently evolved subclonal lineage that gained dominance within the timeframe of our sampling. The mutated HSC has a clear myeloid over lymphoid preference; yet contributes 22% of the $CD4^+$ T-cells. Despite clear signs of an aged immune system, we found that naive $CD4^+$ T-cells were capable of robust proliferation when challenged *in vitro.* Moreover, we observed a surprisingly high TREC content. Collectively, this indicates a functional thymic output at age 111.

Discussion: Our results comprise a unique account of an exceptionally long-lived individual with a hematopoietic system characterized by extreme ARCH, as well as a functional thymus supporting robust T cell responses.

Introduction

Age-related Clonal hematopoiesis (ARCH) is an inevitable physiological consequence of ageing[1], conferring a tenfold increased risk for the development of myeloid dysplasia or malignancies[2,3]. In addition to hematological disease, large prospective epidemiological studies have also established associations between ARCH and prevalent type 2 diabetes[4], prevalent chronic obstructive pulmonary disease[1], incident cardiovascular disease[5], and vascular or all-cause mortality[1-3,5]. While these associations suggest a potential role for ARCH not only in leukemia, but also in a broad spectrum of age-related low-grade inflammatory syndromes[6], its heterogeneity with respect to these clinical outcomes has also been noted[7,8]. For instance, ARCH has been postulated to constitute an early phase of myelodysplasia and leukemia[7], yet the absolute risk to progress to one of these clinical entities remains overtly small. In addition, while ARCH becomes increasingly prevalent with age[2,3,9], and can become very extreme in the exceptionally old[10] , its association with all-cause mortality seems to wane in these oldest old[11] . Collectively, these findings illustrate that we are far from understanding the molecular pathophysiological transitions that ARCH may initiate during ageing.

ARCH arises when an ageing hematopoietic stem cell (HSC) acquires a somatic mutation that confers a competitive growth advantage, leading to its gradual expansion[12]. ARCH-associated mutations typically target genes previously linked with myeloid aberrancies, most prominently in genes encoding epigenetic regulators such as *DNMT3A*, *TET2* and *ASXL1*[2,3,12]. In addition to ARCH caused by known candidate drivers (ARCH-CD), others have shown that a significant part of ARCH occurs in the absence of such known drivers (ARCH-UCD)[1,3]. These studies exploited the fact that HSCs accumulate somatic mutations over the course of a lifetime[13], thus effectively tagging each individual HSC with a unique genetic 'barcode'[1]. We previously applied this paradigm to trace the clonal origins of peripheral blood cells taken from a 115-year-old woman, which revealed that up to 65% of her peripheral blood was derived from a single expanded HSC[10]. While this study exemplified that extreme ARCH does not

preclude attaining an exceptional age, it also raises questions on the characteristics of this particular type of clonal hematopoiesis that sets it apart from other forms that eventually do lead to disease.

Studies on the ageing immune system and its relation to age-related disease have traditionally focused on the lymphoid branch, also referred to as 'immuno-senescence', which is predominantly characterized by the age-related loss of functionality of the T-cell compartment. Hallmarks of immuno-senescence involve the age-associated changes in circulating CD4+ and CD8+ T-cell populations, including the lack of naïve T-cells, increased numbers of memory T-cells, decreased complexity in the antigen-recognizing repertoire, and decreased cell proliferation upon encountering an antigen[14]. Nearly all of these changes, however, can be attributed to a reduced production of new T-cells in the thymus[15], which gradually ceases to function after puberty and is fully involuted and inactive at advanced age. Thymic proliferation can be assessed by quantifying the concentration of T-cell receptor excision circles (TRECs) in peripheral blood or in sorted immune subsets[16-18]. TREC content has been shown to decline sharply after middle age and becomes often undetectable at old age (>85 years)[19]. Despite the availability of methods to assess age-related changes in T-cell proliferation and age-related changes in hematopoietic clonality (ARCH), human studies on the ageing immune system typically do not include both these measures. Consequentially, it remains elusive to what extent immuno-senescence and ARCH may co-occur or even interact in the ageing human immune system.

We set out to address this question by performing a detailed study on peripheral blood and its sorted subsets drawn from an immuno-hematologically normal individual at age 103, 110 and 111 years. In accordance with our previous findings in a 115-year old individual, deep sequencing revealed extreme ARCH: over 75% of the peripheral blood cells shared their clonal origin. We subsequently employed this highly unique study sample to deeply characterize various aspects of extreme ARCH, including its subclonal architecture, its temporal stability, and its potential restriction to particular immune lineages. In parallel, we investigated various aspects of immuno-senescence using flow-

cytrometry, TREC assays and *in vitro* stimulation experiments. Collectively, this study represents a comprehensive exploration of the ageing immune system at the extreme end of human lifespan.

Methods

Full details of the methods used are given in the **Supplementary Methods**.

Study design and data collection

Subject of this study was W111, a Dutch female who lived for 111 years and 10 months whose blood showed no signs of cytopenia, dysplasia, or other clinical signs of haematological malignancies. W111 is enrolled as a participant of the 100-plus Study, a study aimed to investigate Dutch citizens who have attained an age of 100 years or older in good cognitive health[20]. Seven years prior to her participation in the 100-plus Study, W111 also enrolled in the Leiden Longevity Study, a study aimed to investigate the determinants of familial longevity and healthy ageing[21]. Upon learning that this individual was enrolled in both studies, we realized the unique opportunity to investigate blood samples collected across a time span of 8 years and 11 months. Blood samples were obtained at three time points, age 103 (timepoint 0), 110 (timepoint 1), and 111 (timepoint 2) respectively (**Figure 1**), and included peripheral blood (PB), its flow sorted subsets (granulocytes (G), monocytes (M), T-cells (T, T4, T8) and B-cells (B)). Additionally, a skin biopsy (S) was collected at age 110, and the subject agreed to post-mortem brain donation, allowing the investigation of brain tissue (cortex (C)). Throughout the remainder of the manuscript, particular biomaterials sampled at particular timepoints will be referred to by a combination of their tissue abbreviation, e.g. 'PB', and timepoint, e.g. '0', jointly creating: 'PB0', i.e. 'peripheral blood sampled at age 103 years'. Control peripheral blood samples were obtained from anonymized middle-aged female volunteers. The 100-plus Study and Leiden Longevity Study have been approved by their respective local Medical Ethical Committees and all participants gave informed consent. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Cataloguing mosaic somatic mutations in peripheral blood

To identify mutations that accumulated in the hematopoietic stem cell compartment, we conducted Whole Genome Sequencing (WGS) on peripheral blood and a skin biopsy taken at age 110 years and 3 months, using SOLiD sequencing technology (Life Technologies) at 80x mean depth (**Supplementary Methods 1**). To this end, Single Nucleotide Variations (SNVs) were called in both samples using GATK Haplotypecaller[22] and compared to identify mutations detectable in blood, but not in the skin biopsy. This comparison was performed using in-house software *Sommix*, that explicitly accounts for the potential contamination of the skin biopsy with blood cells (**Supplementary Methods 2**). Putative biallelic SNVs were prioritized by *Sommix* and assigned to tiers of increasing confidence (1-3). A total of 650 SNVs covering all three tiers were selected for validation by Ion Torrent amplicon sequencing at 6000x average read-depth (Proton Ampliseq sequencing, Thermo Fisher Scientific, **Table S1, Supplementary Methods 3**). Primer design was successful for 481 out of 650 SNVs, and for these SNVs we acquired the VAF in DNA samples derived from peripheral blood (PB0) and the *post-mortem* brain biopsy (C) (**Supplementary Methods 3**). Variants were called using the Torrent Variant Caller (TVC) version 5.0.0 (Life Technologies); mutations present in peripheral blood, but absent in brain tissue were considered to be confirmed somatic mutations (**Table S2**). In total we were able to confirm 307/481 somatic mutations. TruSight Myeloid panel (Illumina) was used to re-sequence 54 genes frequently mutated in myeloid neoplasia, using the PB1 blood sample on a MiSeq instrument (Illumina) at >500x (**Supplementary Methods 4**).

Impact prediction and mutational signatures

Identified somatic mutations were annotated with UCSC's Variant Annotation Integrator (https://genome.ucsc.edu/cgi-bin/hgVai, hg19, UCSC refgene definitions). Analyses of their tri-

nucleotide context and resemblance to previously published and curated mutational signatures (COSMIC, V2 - March 2015) was performed with the R package MutationalPatterns[23].

Identification of candidate driver mutations

WGS data was also analysed for the presence of candidate driver mutations in 16 genes previously reported to be recurrently mutated in the blood of apparently healthy elderly individuals: *DNMT3A*, *TET2*, *ASXL1*, *TP53*, *JAK2*, *SF3B1*, *GNB1*, *CBL*, *SRSF2*, *GNAS*, *BRCC3*, *CREBBP*, *NRAS*, *RAD21*, *U2AF1*, *PPM1D*[2,3,9,11]. For this purpose, SNPs and short indels called by GATK Haplotypecaller were consecutively filtered according to the following five criteria. *i*) Positioned within or near, max 5 nucleotides, of the coding sequence of a transcript annotated to the aforementioned genes (hg19, UCSC refgene definitions). *ii*) A minimal read depth of the mutant allele of 3 and 6 for respectively SNVs and indels. *iii*) A protein altering predicted impact (UCSC Variant Annotation Integrator, NCBI Refseq curated subset). *iv*) Adhering to the gene specific mutational profiles, as previously compiled[5] (**Table S3**). *v*) Called in blood, but not in brain. A *DNMT3A* splice mutation was identified and confirmed by resequencing the peripheral sample blood drawn at 110 years and 3 months using the Illumina TruSight Myeloid panel and MiSeq instrument at 961x and visualized in Integrative Genome Viewer (IGV, Broad Institute, version 2.4.9).

Inference of subclonal architecture and subclonal lineage dynamics

Amplicon sequencing of the 307 confirmed somatic mutations in peripheral blood collected at ages at age 102 years and 11 months, 110 years and 3 months, and 111 years and 10 months, was used to infer the subclonal architecture and its longitudinal changes. For this purpose, SciClone version 1.1.0[24] was run at default settings to cluster the panel of confirmed somatic mutations on Variant Allele Frequencies (VAFs) measured in peripheral blood at these three time points (**Table S2**). Each identified cluster marks

a unique subclonal expansion events. Cluster annotations and VAFs were subsequently analysed with SCHISM version 1.1.2[25], at default settings, to infer a phylogenetic tree describing the most probable order in which these subclonal expansion events must have occurred. Each node of the obtained tree represents a subclonal lineage that arose by acquisition of one or multiple clonal events, as represented by the structure of the phylogenetic tree. The percentage of cells carrying a particular subclonal expansion event was computed using the median VAF across all mutations assigned to its respective cluster. To obtain percentages of cells belonging to particular subclonal lineages, i.e. cells carrying particular combinations of events, each subclonal expansion was corrected for its immediate child event in the tree, i.e. by substracting the median VAF of the child from the median VAF of the parent node.

Tracing the sub-clonal contribution to hematopoietic cell types

DNA isolates from sorted leukocyte subsets G1, M1,T1, B1 (derived from PB1) and G2, T8.2, T4.2, B2 (derived from PB2) were analyzed for carrying clonal events using amplicon sequencing of the 307 confirmed somatic mutations. Median VAFs over somatic mutations marking clonal events were computed for each sorted subset and corrected for its immediate child event in the previously inferred clonal evolution tree (see above). Apparent differences in the clonal contribution were further analyzed for each somatic mutation separately: the ratio of the VAF in a specific sorted cell type over the VAF within its corresponding peripheral blood sample was modelled as a function of the VAF in the peripheral blood sample, using linear models in R.

Immuno-phenotyping by flow cytometry

PBMCs were incubated with a titrated cocktail of antibodies for 30 minutes on ice in the dark. Cells were subsequently washed with FACS buffer (PBS supplemented with 2% FCS and Na-azide) prior to flow cytometry analysis using FACS Fortessa-X20 (BD Biosciences). The employed antibody mix consisted of:

FITC-anti-CD57 (clone HCD57); PE-anti-CD28 (clone CD28.2); APC-Cy7-anti-CD4 (clone RPA-T4); PE-Cy7-anti-CD197 (CCR7, clone 3D12); Pacific Blue-anti-CD27 (clone M-T271), BV650-anti-CD8 (clone SK1), all were from BD Pharmingen (USA). Alexa Fluor 700-anti-CD45RA (clone HI100); BV605-anti-CD31 (clone WM59) were from Biolegend (San Diego, USA). Phenotypic data were analyzed using FlowJo software (TreeStar).

TREC assays

T-cell receptor excision circle (TREC) assays are based on the rearrangement of the T-cell receptor occurring early during T-cell development[26]. This rearrangement results in the formation of a coding joint (CJ) which remains stably present in the genomic DNA, and a signal joint (SJ) on the corresponding excision circle. As with every cell division, SJ is diluted, while CJ is stably maintained in the genomic DNA, it is possible to derive the number of divisions the assayed cells have undergone from the difference between CJ and SJ[27]. TRECs analyses were essentially performed as previously described[28]. In short, DNA was extracted from peripheral blood or sorted T-cell subsets, after which two parallel real-time PCRs are performed that quantify the signal joint (SJ) or the coding joint (Cj) relative to an internal control of Albumin.

T-cell stimulation using a thymidine incorporation assay

Antigen-dependent T-cell proliferation capacity was determined using a mixed leukocyte reaction (MLR). PBMC's from test samples were cultured with irradiated HLA-mismatched PBMCs (3000 rad). To assess the general capacity of T-cells to proliferate, PBMCs were stimulated with Interleukin-2 (IL-2, 25 IU/ml). Cells were incubated in triplo in the 96-well round bottom plates in IMDM medium conditioned with Glutamine, Pen/Strep and 10% inactivated Human Serum (HS, Sanquin, Amsterdam) for 5 days. Radioactive thymidine (2 µCi/ml) was added in the last 18 hrs of the culture to enable incorporation in

the DNA of proliferating cells and the radioactivity in DNA recovered from the cells was measured using a scintillation beta-counter. Obtained data (counts per minute, CPM) were used to calculate stimulation index by dividing the CPM values of the test samples by the CPM of test PBMC cultured in medium alone.

Detection of clonal BCR / TCR gene recombinations

The presence of clonal BCR / TCR recombinations in peripheral blood may point to suspect lymphoproliferations, or a recent or sustained antigenic stimulation. The diversity in both T and B-cell receptor repertoire was investigated using IdentiClone™ PCR assays followed by quantification using an ABI Capillary Electrophoresis Instrument.

Results

Sequencing peripheral blood reveals extreme age-related clonal hematopoiesis

We identified 650 putative single nucleotide somatic variations (SNVs) that had accumulated in the hematopoietic stem cell compartment when W111 was 110 years old (**Figure 2A, Table S1**). Concurrent with previous work on the non-leukemic accumulation of somatic mutations in peripheral blood[10] or in cultures of single expanded hematopoietic stem cells[13,29,30] the vast majority of these mutations were non-coding (**Figure 2B**), did not exhibit any positional preferences in the genome (**Figure S1**), and were dominated by C>T and T>C transitions (**Figure 2D**). Moreover, the tri-nucleotide sequence context of the identified mutations (**Figure 2E, Figure S2**) could largely be explained by previously established 'clock-like' mutational signatures[30,31] (**Figure S3**). Importantly, while some mutational signatures are strongly linked to oncogenic processes underlying the aetiology of particular cancer types[32,33], mutations attributed to the 'clock-like' signatures 1 and 5 have been reported to correspond with the age of the sampled subject, and thus seem to represent relatively benign ageing processes[31].

Next, we screened the 650 identified SNVs for potential candidate driver mutations using the mutation definitions compiled by Jaiswal *et al.*[5] (**Methods,** definitions in **Table S3**) for 16 genes frequently mutated in the blood of healthy elderly individuals[11]. This analysis revealed a splice-donor site mutation in intron 11 of DNA (cytosine-5)-methyltransferase 3A (*DNMT3A,* NM_022552.4, chr2:25,469,028 C>T, c.1429+1 G>A). Targeted resequencing (TruSight Myeloid Sequencing Panel, Illumina, MiSEQ sequencing, Illumina, **Methods and Supplementary Methods 4**) confirmed this variant at 961x with an estimated variant allele frequency (VAF) of 0.378 (**Figure 2C**), suggesting that 75.6% of the peripheral blood cells are derived from a single clone carrying a *DNMT3A* splice-donor mutation.

To conclude, the number, nature and dominance of the identified somatic mutations provide strong indications that a major proportion of the peripheral blood production is derived from a single

hematopoietic stem cell. Hence, W111's hematopoietic system can thus be regarded as a case of extreme age-related clonal hematopoiesis (ARCH).

Longitudinal analysis of somatic mutations during a 9-year period reveals an extensive and dynamic subclonal architecture

Hematopoietic stem cells each accumulate randomly positioned somatic mutations during ageing[13], that can be employed as unique genetic barcodes[1] to trace the clonal origin of cells[34] in peripheral blood or sorted immune subsets. To obtain a high-confidence subset of somatic genetic markers suited for such applications of clonal tracing, we performed extensive validation experiments for the 650 putative SNVs that we identified (**Figure 2F**). More specifically, custom targeted re-sequencing panels (Ion Torrent Proton Ampliseq, Thermo Fisher Scientific) were successfully designed for 474 out of 650 (72.9%) identified somatic mutations. Using cortex as a control tissue, we observed replication rates ranging from 20.1% to 97.8% across the three confidence tiers, overall confirming 307 (64.8%) somatic mutations.

Amplicon sequencing revealed that all 307 mutations identified at age 110 years were also detectable in peripheral blood samples taken at ages 103 and 111 (**Figure 3A**). While Variant Allele Frequencies (VAFs) generally increased between ages 103 and 110 years, VAFs remained equal or lowered between ages 110 and 111 years. Nevertheless, VAFs were highly inter-correlated between timepoints (Pearson's $r_{103\text{->}110}$ = 0.983 and $r_{110\text{->}111}$ = 0.988, **Figure S4**), and VAF density distributions looked highly similar across time points. Overall, these results indicate the long-term stability of the identified somatic mutations, and thereby the long-term stability of extreme ARCH.

Density distributions of VAFs feature multiple peaks at all three timepoints, which is suggestive of an underlying subclonal architecture[13]. In accordance, a SciClone analysis[24] (**Methods**) on VAF data from all three time points, assigned the 307 confirmed somatic mutations to five independent clonal

events (A-E, **Figure 3A**). Subsequent modelling with SCHISM[25] (**Methods**) indicated that these five clonal events most likely have occurred consecutively within a single clonal lineage, that terminates into two independent events D and E (**Figure 3B**). The latter suggests the existence of two sister lineages that were derived from a shared ancestral subclonal lineage carrying events A-C, after which they diversified by acquiring either events D or E. Collectively, these results indicate the presence of an extensive subclonal architecture in the hematopoietic stem cell compartment (**Figure 3D**), that reflects the subclonal evolution that most likely occurred prior to our study.

When analyzing the temporal changes in clonal dominance, i.e. the changes in VAFs between ages 103, 110 and 111 years, we need to take into account that the somatic variants are inherited along the subclonal architecture **(Figure 3A).** For example, all somatic mutations in clone A are present in its clonal descendants B-E, and all somatic mutations in B are present in C-E, but not A. Consequently, strong interdependencies are observed between VAFs of subclones. Adjusting for these interdependencies **(Methods),** we observed that changes in dominance of clonal events A-C are largely explained by changes in dominance of event E (**Figure 3C**), and notably not by event D. In fact, while clonal event D exhibits a near equal contribution of approximately 16.5% of the cells to peripheral blood at age 103, 110 and 111 years, clonal event E nearly tripled its clonal contribution from approximately 6.1% at age 103 to 17.9% of the peripheral blood cells at ages 110 and 111. In addition, we observe that clone B becomes less dominant, as its contribution to peripheral blood decreases from approximately 15.7% at age 103 to 10.7% at ages 110 and 111. Hence, our results point to ongoing subclonal dynamics and possibly to a subclonal competition between sister-lineages readily occurring in the blood of an hematopoietically normal elderly individual.

Overall, our results indicate that the peripheral blood production has been governed by a single mutated stem cell and its subclonal descendants for a period spanning at least 9 years. Furthermore, our

data suggests that marked changes in the subclonal contribution to peripheral blood may occur in the blood of otherwise hematopoietically normal elderly individuals.

The founding clone and its descendants are myeloid biased, yet also produce lymphoid progeny with a preference for CD4$^+$ T-cells

Next, we investigated to what extent the somatic mutations were present in the various major cell subsets of peripheral blood sampled at age 110. To this end, the previously identified panel of 307 somatic mutations was re-sequenced in FACS-sorted granulocytes (G), monocytes (M), T-cells (T) and B-cells (B) (**Figure 1**). A comparison of VAFs between different cell subsets (**Figure 4A**) indicated a significantly higher presence in the myeloid branch, i.e. 87.4% of the granulocytes ($G_{VAF,\ age=110}$ = 0.437) and 77.8% of the monocytes ($M_{VAF,\ age=110}$ = 0.389), as compared to total peripheral blood, i.e. 67.0% of the cells ($PB_{VAF,\ age=110}$ = 0.335). Interestingly, somatic mutations were generally also observed in DNA derived from cell subsets of the lymphoid lineage, albeit at considerably lower VAFs. Nevertheless, approximately 10.6% of the T-cells ($T_{VAF,\ age=110}$ = 0.053), and 7.4% of the B-cells ($B_{VAF,\ age=110}$ = 0.037) carried mutations, demonstrating both the multipotency and myeloid bias of the mutated clone.

Re-sequencing within the blood sample drawn at age 111 did not only confirm both the myeloid bias and the multipotency (**Figure 4A**), but also pointed to a striking bias present within the lymphoid branch. Median VAFs were significantly higher in CD4$^+$ T-cells (22.2% of the cells, $T4_{VAF,\ age=111}$ = 0.111), as compared to CD8 T-cells (6.4% of the cells, $T8_{VAF,\ age=111}$ = 0.032, p<0.001, Wilcoxon) and B-cells (6.0% of the cells, $B_{VAF,\ age=111}$ = 0.030, p<0.001, Wilcoxon). As *DNMT3A* mutations are not restricted to myeloid leukemias[35], but also have been reported in T-cell lymphomas[36] and T-cell leukemias[37], we verified the absence of T-cell malignancies by testing for clonal T-cell receptor gene recombinations[38]. This revealed a diverse T-cell receptor repertoire (**Figure S5**). Hence these results not only further substantiate the

multipotency of the founding clone and its descendants, it also points to biases within the lymphoid lineage, specifically for contributing to the $CD4^+$ T-cell pool.

Mapping the mutations measured in the subsets to the previously identified subclonal architecture **(Figure 3B)** revealed contributions of the founding clone (A) and its forthcoming descendants (B, C ,D and E) that generally seemed proportional to the contributions observed in peripheral blood samples of corresponding time points (**Figure 4B**). For T-cells, however, some consistent deviations were observed. For sorted T-cell subsets obtained at age 110 (**Figure 4B**, left), as well as at age 111 **(Figure 4B**, right**),** we observed a relatively low contribution of clone E relative to its corresponding peripheral blood sample. Interestingly, this disproportionality was more prominent at age 110 as compared to age 111, suggesting that subclone E increased it contribution to T-cells during this time span. When analyzing at the level of the individual somatic mutations, we observed that mutations assigned to smaller clones had a disproportionately lower VAF in the T-cell subsets as compared mutations assigned to larger clones (**Figure 4C-E**). No such trends were observed for B-cells or any other sorted subsets (**Figure S6**), suggesting that the observed bias in clonal blood production relates to T-cells specifically.

Overall, tracing the clonal contributions in various sorted immune subsets demonstrated the multipotency and temporal plasticity of the contributions of the founding clone and its subclonal descendants, and specifically pointed to their contribution to $CD4^+$ T-cells within the lymphoid immune lineage.

An immuno-competent naïve $CD4^+$ T-cell compartment

Ageing is associated with profound changes in immune subset composition, i.e. the thymus, responsible for de novo production of naive T cells, involutes[19,39], and aged HSCs produce many more myeloid cells than B lymphocytes. Accordingly, flow cytometry analyses of peripheral blood taken at age 110 show

increased fractions of senescent $CD4^+$ and $CD8^+$ T-cells, as apparent by the expression of CD57 (**Figure 5A**, **Figure S7**). Moreover, we observe high myeloid to B lymphocyte ratios, particularly due to lowered B-cell levels at age 111 years (**Figure 5B**). Hence, W111's peripheral blood shows clear signs of an aged immune system.

In addition to these hallmarks of an aged immune system, we also observed considerable fractions of naive $CD4^+$ and $CD8^+$ T-cells at ages 110 and 111 years. Compared to middle-aged controls, we observe only slightly reduced or comparable fractions of naive $CD4^+$ T-cells and naive $CD8^+$ T-cells respectively at age 111 years (**Figure 5B, Figures S8**). Moreover, we observe at age 110 years, that nearly 5% of the $CD4^+$ T-cells expressed CD45RA and CD31 (**Figure 5C, Figure S7**). Although this level is much lower as compared to the middle-aged controls, it nevertheless suggests that there are still recent thymic emigrants in W111's peripheral blood[40]. Indeed, assessment of the replication history of T-cells by T-cell receptor excision circles (TRECs) assays indicated that both $CD4^+$ and $CD8^+$ T-cells had undergone a number of divisions comparable to that of middle-aged healthy controls (**Methods**, **Figure 5D**), with TREC contents of 3-6%[19]. Hence, our results imply an active thymus at an exceptionally advanced age that is still capable of producing naive T-cells.

In parallel, we investigated the capacity of W111's T-cells to actually mount immune responses. Flow cytometry analyses of cells collected at age 110 indicated that both the $CD4^+$ and $CD8^+$ T-cell subsets contained considerable fractions of *in vivo* activated cells, evidenced by their high CD25 and CD69 expression (**Figure S8**). This suggests ongoing or recent T-cell directed immune responses. To further validate this observation, we performed two types of *in vitro* proliferation assays (**Methods**). In both the IL2/TCR-dependent and the allogeneic mixed-lymphocyte assay, $CD4^+$ T-cells taken from W111 outperformed those taken from middle-aged controls on a per cell basis (**Figure 5E**). Hence, although W111 has more senescent T-cells than middle-aged controls, those left in W111 are better in mounting immune responses on a per cell basis.

Collectively, these results indicate that at least a part of W111's T-cell compartment was not immune senescent and was still capable of mounting vigorous naive immune responses consistent with the presence of functional thymic activity in W111.

Discussion

We explored the properties of extreme age-related clonal hematopoiesis (ARCH) by studying peripheral blood of a immuno-hematologically normal elderly individual sampled at age 103, 110 and 111 years. In depth characterization of this highly unique study sample not only demonstrates the potential long-term stability of extreme ARCH, but also provides unique insights into its extensive subclonal architecture, its ongoing clonal evolution and its differential contribution to various immune subsets. Most notably, we observe that extreme ARCH can readily co-occur with a seemingly functional T-cell compartment, suggesting that clonal hematopoiesis and immune senescence confer different axis in the ageing immune system. Collectively, our findings suggest that extreme ARCH is neither a momentary flux of the hematopoietic system, nor a condition posing immediate and overt health risks to the exceptionally old.

ARCH is most frequently established by the identification of a candidate driver mutation (ARCH-CD), typically residing in *DNMT3A*, *TET2* or *ASXL1*[2,3,12]. However, ARCH can also be established by the presence of a disproportionate large number of somatic mutations in the absence of candidate driver mutations (ARCH-UD)[1,3]. The ARCH identified in our supercentenarian subject complied with both these criteria. Accumulated somatic mutations included a splice-donor mutation in intron 11 of *DNMT3A* (NM_022552.4, chr2:25,469,028 C>T, c.1429+1 G>A). This somatic mutation was previously reported in patients with hematopoietic or lymphoid malignancies (COSM5945645)[41-43], indicating that this event may indeed have driven the initiating clonal expansion. Apart from this mutation, no other clinical indications for hematologic malignancies could be established. A clinical examination of peripheral blood revealed neither cytopenias nor dysplastic morphologies. Moreover, sequencing-based diagnostic tests for suspect myeloid or lymphoid leukemias[38] **(Figure S5)**. Collectively, these observations fit the postulated inclusion criteria for Clonal Hematopoiesis of indeterminate Potential (CHiP)[7], also referred to as ARCH[44], to describe immuno-hematologically normal elderly individuals carrying a pre-leukemic mutation with unclear clinical implications.

The initiating clonal expansion could be attributed to a *DNMT3A* splice donor mutation. No additional candidate drivers were detected amongst the identified remaining somatic mutations that could explain an increased replicatory fitness required for any of the successive subclonal expansions to occur. While this result might be explained by undetected or incomplete knowledge on driver mutations[1], it could also point to non-genetic mechanisms, e.g. so-called 'epi-mutations'[45]. *DNMT3A* is a key epigenetic regulator orchestrating myelopoiesis[46] and responsible for maintenance of DNA methylation in HSCs[47]. Therefore, a mutation in DNMT3A may result in a gradual loss of epigenetic control within the mutated HSC possibly improving the replicatory fitness[48]. Yet another possibility is that the *DNMT3A* mutation may magnify the alternating temporal activity of HSCs either occurring under normal conditions[49], or under bone marrow stress, e.g. inflammation[50], or in response to environmental stimuli[51]. If the latter were true, the recovered clonal architecture would then represent the history of re-activation of otherwise quiescent HSCs, rather than the acquisition of (epi-)mutations that iteratively improved the replicatory fitness.

Clonal tracing across consecutive time points revealed intricate patterns of subclonal dynamics. By reconstructing clonal lineages and tracing them over time, we showed that a relatively recently evolved clonal lineage (E) quickly gained dominance in comparison to its competing sister lineage (D). While the mutations marking lineage E were on the border of detection-limit at the start of our study (median VAF = 3.0%), they rapidly grew in dominance in ensuing years to produce up to 18% of the peripheral blood (VAF = 9.0%) at age 110. This aggressive expansion parallels patterns of clonal evolution otherwise observed in overt myeloid disease, i.e. myelodysplastic syndromes[52,53], myeloproliferative neoplasms[54] or relapsed acute myeloid leukemia[55,56], in which subclonal lineages may completely be replaced. Whether these complete wipeouts occur within the healthy aged hematopoietic stem cell compartment remains to be determined. Nevertheless, our results indicate that the temporal plasticity and resulting subclonal architecture otherwise associated with dysplasia or

malignancies readily occurs in the hematopoietic stem cell compartment of immuno-hematologically normal elderly individuals, and could therefore be regarded as a process of normal ageing.

All subclonal events were observed in all sampled blood subsets, irrespective of myeloid or lymphoid descendance. This is in agreement with recent findings[57,58], independently indicating that *DNMT3A* mutations can be potent drivers of both the myeloid and lymphoid lineages, thereby giving rise to a so-called myeloid-lymphoid ARCH. Interestingly, while overall myeloid biased, we also observed a bias within the lymphoid lineage: for CD4+ T-cells over CD8+ T-cells and B-cells.-Similar to the myeloid-lymphoid bias, we speculate that the CD4+ over CD8+ T-cell bias might originate from a higher turnover rate of CD4+ versus CD8+ T-cells. We can visualize differences in turnover rates when assuming that the median VAF of a mutational event can serve as a proxy for its relative time of occurrence. We observe for all three T-cell subsets (**Figure 4C-E**) that older mutational events are present in disproportionally larger fractions of T-cells. This bias can be explained by the longevity of T-cells; mutations newly arising in T-cell producing progenitors would only become slowly apparent in the total pool of T-cells. Interestingly, this time-delayed contribution of *DNMT3A*-mutated HSCs to T-cells has previously been proposed to occur in AML patients[57,58]. Our data thus extents this time-delay to the pre-malignant setting ARCH. Moreover, we show that this time-delay may vary per T-cell subset; CD8+ T-cells had a stronger bias as CD4+ T-cells, which suggests that CD8+ T-cells may have a slower turnover rate than CD4+ T-cells, at least at this extremely advanced age. It thus seems that cases with *DNMT3A*-mutated clonal hematopoiesis, or its malignant successors, could potentially be stratified by a varying degree of T-cell involvement to indicate the rate of expansion of the mutated HSC.

Ageing of the T-cell compartment, i.e. immune senescence has been postulated as a major factor underlying a reduced life expectancy[59,60]. The thymus gradually ceases to function after puberty and is fully involuted and inactive at advanced age and the concentration T-cell receptor excision circles (TRECs) in peripheral blood, indicative of T-cell proliferation, decreases correspondingly, and is often

undetectable after 85 years[19]. Given this background, our observations in W111 are remarkable. TREC contents in W111's peripheral blood and sorted CD4 and CD8 T-cells at age 110 and 111 years were not only detectable, in fact, the observed TREC contents of 3-6% are comparable to that of middle-aged healthy controls. Together with the fact that W111 T-cells were still capable of mounting vigorous naive responses in *in vitro* proliferation assays, this suggests that W111 disposes over functional thymic activity.

Our observation of clonally expanded CD4 T-cell subset as a potential hallmark of the supercentenarian immune system is in line with recent findings by Hashimoto *et al.*[61]. While Hashimoto *et al.* establish the clonal expansion of CD4 T-cells with the use of T-cell receptor analyses, they conclude that the clonal expansion was most likely triggered by a sustained antigenic stimulation. However, they did not take into account that ARCH is practically unavoidable at an advanced age[1], and as such may offer a competing explanation. Supportive evidence that ARCH-associated mutations may indeed contribute to the clonal expansion of functional lymphoid subsets was recently provided by Fraietta *et al.*[62] in a paper elegantly reporting on an incidental finding in a study on CAR T-cell immunotherapy. In this study, CD8+ T-cells were genetically redirected to target tumor cells of patients with B-cell malignancies, so-called CAR T-cells. While transplantation should elicit a polyclonal T-cell response in the patient, thus clearing malignant B-cells, one patient showed a monoclonal outgrowth of CAR T-cells. Sequencing revealed that the procedure to genetically redirect the T-cell had also disrupted a gene associated with ARCH: *TET2*. Interestingly, *TET2*-deficient CAR T-cells were immuno-competent and the patient achieved full remission. In accordance with this study, our results indicate that the CD4 thymic output was descended from recent thymic immigrants, and moreover, was capable of eliciting a fierce naive response. Collectively, this raises the possibility that the unexpected functional $CD4^+$ T-cell immunity could be attributed to lymphoid descendants of the mutated stem cells.

This study also has limitations. Foremost, this study has been conducted with the material derived from only a single healthy supercentenarian. Hence, it is imperative that more of these highly unique individuals are investigated in order to generalize our findings to a broader context. Moreover, the nature of our study precludes any inferences on the causality between the two main observations within this study, namely the co-occurrence between an extreme *DNMT3A*-associated ARCH and an unexpectantly functional CD4 T-cell immunity. Nevertheless, we believe that the nature of the presented findings warrants future research into the causes of a functional $CD4^+$ immunity that is quintessential for achieving an exceptional age in good health.

Concluding, our findings open the possibility that extensive clonal evolution does not preclude and may even contribute to the replenishment of essential immune subsets that might otherwise be subject to age-dependent depletion. Our findings not only warrant future research into the underlying mechanisms driving clonal dynamics during normative aging of the hematopoietic system, they also stress the importance of methods such as clonal tracing[1,3,10], independent of known driver mutations.

Acknowledgements

The 100-plus Study was supported by Stichting Alzheimer Nederland (WE09.2014-03), Stichting Diorapthe (VSM 14 04 14 02) and Stichting VUmc Fonds. The Leiden Longevity Study has received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement number 259679. This study was supported by the Netherlands Consortium for Healthy Ageing (grant 050-060-810), in the framework of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO); by BBMRI-NL, a Research Infrastructure financed by the Dutch government (NWO 184.021.007). EvdA is funded by a personal grant of the Dutch Research Council (NWO; VENI: 09150161810095)

Author contributions

Competing financial interests

None of the authors have competing interests to declare.

References

1. Zink F, *et al.* Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. Blood. 2017; blood-2017-02-769869; doi: https://doi.org/10.1182/blood-2017-02-769869
2. Jaiswal S, *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med. 2014 Dec 25;371(26):2488-98. doi: https://doi.org/10.1056/NEJMoa1408617
3. Genovese G, *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med. 2014 Dec 25;371(26):2477-87. doi: https://doi.org/10.1056/NEJMoa1409405
4. Bonnefond A, *et al.* Association between large detectable clonal mosaicism and type 2 diabetes with vascular complications. Nat Genet. 2013 Sep;45(9):1040-3. doi: https://doi.org/10.1038/ng.2700
5. Jaiswal S, *et al.* Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. N Engl J Med. 2017 Jul 13;377(2):111-121. doi: https://doi.org/10.1056/NEJMoa1701719
6. Bowman RL, *et al.* Clonal Hematopoiesis and Evolution to Hematopoietic Malignancies. Cell Stem Cell. 2018 Feb 1;22(2):157-170. doi: https://doi.org/10.1016/j.stem.2018.01.011
7. Steensma DP, *et al.* Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood. 2015 Jul 2;126(1):9-16. doi: https://doi.org/doi:10.1182/blood-2015-03-631747
8. Buscarlet M, *et al.* DNMT3A and TET2 dominate clonal hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. Blood. 2017 Aug 10;130(6):753-762. doi: https://doi.org/10.1182/blood-2017-04-777029
9. Xie M, *et al.* Age-related mutations associated with clonal hematopoietic expansion and malignancies. Nat Med. 2014 Dec;20(12):1472-8. doi: https://doi.org/10.1038/nm.3733
10. Holstege H, *et al.* Somatic mutations found in the healthy blood compartment of a 115-yr-old woman demonstrate oligoclonal hematopoiesis. Genome Res. 2014 May;24(5):733-42. doi: https://doi.org/10.1101/gr.162131.113
11. van den Akker EB, *et al.* Uncompromised 10-year survival of oldest old carrying somatic mutations in DNMT3A and TET2. Blood. 2016 Mar 17;127(11):1512-5. doi: https://doi.org/10.1182/blood-2015-12-685925
12. Busque L, *et al.* Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. Nat Genet. 2012 Nov;44(11):1179-81. doi: https://doi.org/10.1038/ng.2413
13. Welch JS *et al.* The origin and evolution of mutations in acute myeloid leukemia. Cell. 2012 Jul 20;150(20):264-78. doi: http://doi.org/10.1016/j.cell.2012.06.023
14. Vicente R, *et al.* Cellular senescence impact on immune cell fate and function. Aging Cell. 2016 Jun;15(3):400-6. doi: https://doi.org/10.1111/acel.12455
15. Ciofani M, *et al.* The thymus as an inductive site for T lymphopoiesis. Annu Rev Cell Dev Biol. 2007;23:463-93. doi: https://doi.org/10.1146/annurev.cellbio.23.090506.123547
16. Aquino VM, et al. Evaluation of Thymic Output by Measurement of T-cell-receptor Gene Rearrangement Excisional Circles (TREC) in Patients who have Received Fludarabine. Leuk Lymphoma. 2003 Feb;44(2):343-8. doi: https://doi.org/10.1080/1042819021000029696
17. Hazenberg MD, *et al.* Thymic output: a bad TREC record. Nat Immunol. 2003 Feb;4(2):97-9. doi: https://doi.org/10.1038/ni0203-97
18. Douek D. Thymic Output and HIV Infection: On the Right TREC. Immunity. 2004 Dec;21(6):744-5. doi: https://doi.org/10.1016/j.immuni.2004.11.005

19. Nasi M, *et al.* Thymic output and functionality of the IL-7/IL-7 receptor system in centenarians: implications for the neolymphogenesis at the limit of human life. Aging Cell. 2006 Apr;5(2):167-75. doi: https://doi.org/10.1111/j.1474-9726.2006.00204.x
20. Holstege H, *et al.* The 100-plus Study of cognitively healthy centenarians: rationale, design and cohort description. Eur J Epidemiol. 2018 Dec;33(12):1229-1249. doi: https://doi.org/10.1007/s10654-018-0451-3
21. Schoenmaker M, *et al.* Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. Eur J Hum Genet. 2006 Jan;14(1):79-84. doi: https://doi.org/10.1038/sj.ejhg.5201508
22. McKenna A, *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010 Sep;20(9):1297-303. doi: https://doi.org/10.1101/gr.107524.110
23. Blokzijl F, et al. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. Genome Med. 2018 Apr 25;10(1):33. doi:https://doi.org/ 10.1186/s13073-018-0539-0
24. Miller MA, *et al.* SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. PLoS Comput Biol. 2014 Aug 7;10(8):e1003665. doi: https://doi.org/10.1371/journal.pcbi.1003665
25. Niknafs N, *et al.* SubClonal Hierarchy Inference from Somatic Mutations: Automatic Reconstruction of Cancer Evolutionary Trees from Multi-region Next Generation Sequencing. PLoS Comput Biol. 2015 Oct 5;11(10):e1004416. doi: https://doi.org/10.1371/journal.pcbi.1004416
26. Dik WA, *et al.* New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling. J Exp Med. 2005 Jun 6;201(11):1715-23. doi: https://doi.org/10.1084/jem.20042524
27. van Zelm MC, *et al.* Replication history of B lymphocytes reveals homeostatic proliferation and extensive antigen-induced B cell expansion. J Exp Med. 2007 Mar 19;204(3):645-55. doi: https://doi.org/10.1084/jem.20060964
28. van der Weerd K, *et al.* Combined TCRG and TCRA TREC analysis reveals increased peripheral T-lymphocyte but constant intra-thymic proliferative history upon ageing. Mol Immunol. 2013 Mar;53(3):302-12. doi: https://doi.org/10.1016/j.molimm.2012.08.019
29. Lee-Six H, *et al.* Population dynamics of normal human blood inferred from somatic mutations. Nature 2018 Sep;561(7724):473-478. doi: https://doi.org/10.1038/s41586-018-0497-0
30. Osorio F G, *et al.* Somatic Mutations Reveal Lineage Relationships and Age-Related Mutagenesis in Human Hematopoiesis. Cell Rep. 2018 Nov 27; 25(9): 2308–2316.e4. doi: https://10.1016/j.celrep.2018.11.014
31. Alexandrov LB, *et al.* Clock-like mutational processes in human somatic cells. Nat Genet. 2015 Dec;47(12):1402-7. doi: 10.1038/ng.3441. doi: https://doi.org/10.1038/ng.3441
32. Nik-Zainal S, *et al.* Mutational processes molding the genomes of 21 breast cancers. Cell. 2012 May 25;149(5):979-93. doi: https://10.1016/j.cell.2012.04.024
33. Alexandrov LB, *et al.* Deciphering Signatures of Mutational Processes Operative in Human Cancer. Cell Rep. 2013 Jan 31;3(1):246-59. doi: https://doi.org/10.1016/j.celrep.2012.12.008
34. Alemany M, *et al.* Whole-organism clone tracing using single-cell sequencing. Nature. 2018 Apr 5;556(7699):108-112. doi: https://doi.org/10.1038/nature25969
35. Ley TJ *et al.* DNMT3A mutations in acute myeloid leukemia. N Engl J Med. 2010 Dec 16;363(25):2424-33. doi: https://10.1056/NEJMoa1005143
36. Couronné L, *et al.* TET2 and DNMT3A Mutations in Human T-Cell Lymphoma, N Engl J Med. 2012 Jan 5;366(1):95-6. doi: https://doi.org/10.1056/NEJMc1111708

37. Simon C. *et al.* A key role for EZH2 and associated genes in mouse and human adult T-cell acute leukemia. Genes Dev. 2012 Apr 1;26(7):651-6. doi: https://doi.org/10.1101/gad.186411.111
38. Van Dongen JJ, *et al.* Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia. 2003 Dec;17(12):2257-317. doi: https://doi.org/10.1038/sj.leu.2403202
39. Mitchell WA, *et al.* Tracing thymic output in older individuals. Clin Exp Immunol. 2010 Sep;161(3):497-503. doi: https://doi.org/10.1111/j.1365-2249.2010.04209.x
40. Kohler S, *et al.* Life after the thymus: CD31+ and CD31- human naive CD4+ T-cell subsets. Blood. 2009 Jan 22;113(4):769-74. doi: https://doi.org/10.1182/blood-2008-02-139154
41. Shin SY, *et al.* Mutation profiling of 19 candidate genes in acute myeloid leukemia suggests significance of DNMT3A mutations. Oncotarget. 2016 Aug 23;7(34):54825-54837. doi: https://doi.org/10.18632/oncotarget.10240
42. Wong TN, *et al.* Cellular stressors contribute to the expansion of hematopoietic clones of varying leukemic potential. Nat Commun. 2018 Jan 31;9(1):455. doi: https://doi.org/10.1038/s41467-018-02858-0
43. Zehir A, *et al.* Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med. 2017 Jun;23(6):703-713. doi: https://doi.org/10.1038/nm.4333
44. Shlush LI, *et al.* Age-related clonal hematopoiesis. Blood. 2018 Feb 1;131(5):496-504. doi: https://doi.org/10.1182/blood-2017-07-746453
45. Jost E, et al. Epimutations mimic genomic mutations of DNMT3A in acute myeloid leukemia. Leukemia. 2014 Jun;28(6):1227-34. doi: https://doi.org/10.1038/leu.2013.362
46. Challen GA, *et al.* Dnmt3a is essential for hematopoietic stem cell differentiation. Nat Genet. 2011 Dec 4;44(1):23-31. doi: https://doi.org/10.1038/ng.1009
47. Zhang X, *et al.* DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells. Nat Genet. 2016 Sep;48(9):1014-23. doi: https://doi.org/10.1038/ng.3610
48. Abdel-Wahab O, *et al.* Mutations in epigenetic modifiers in the pathogenesis and therapy of acute myeloid leukemia. Blood 2013;121(18):3563-3572. doi: https://doi.org/10.1182/blood-2013-01-451781
49. Bystrykh LV, *et al.* Counting stem cells: methodological constraints. Nat Methods. 2012 May 30;9(6):567-74. doi: https://doi.org/10.1038/nmeth.2043
50. Takizawa H, *et al.* Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation. J Exp Med. 2011 Feb 14;208(2):273-84. doi: https://doi.org/10.1084/jem.20101643
51. King KY, *et al.* Environmental Influences on Clonal Hematopoiesis. Exp Hematol. 2019 Dec 29. pii: S0301-472X(19)31240-8 doi: https://doi.org/10.1016/j.exphem.2019.12.005
52. da Silva-Coelho P, *et al.* Clonal evolution in myelodysplastic syndromes. Nat Commun. 2017 Apr 21;8:15099. doi: https://doi.org/10.1038/ncomms15099
53. Makishima H, *et al.* Dynamics of clonal evolution in myelodysplastic syndromes. Nat Genet. 2017 Feb;49(2):204-212. doi: https://doi.org/10.1038/ng.3742
54. Lundberg P, *et al.* Clonal evolution and clinical correlates of somatic mutations in myeloproliferative neoplasms. Blood. 2014 Apr 3;123(14):2220-8. doi: https://doi.org/10.1182/blood-2013-11-537167
55. Ding L, *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature. 2012 Jan 11;481(7382):506-10. doi: https://doi.org/10.1038/nature10738
56. Christen F, *et al.* Genomic landscape and clonal evolution of acute myeloid leukemia with t(8;21): an international study on 331 patients. Blood. 2019 Mar 7;133(10):1140-1151. doi: https://doi.org/10.1182/blood-2018-05-852822
57. Arends CM, *et al.* Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis. Leukemia. 2018 Sep;32(9):1908-1919. doi: https://doi.org/10.1038/s41375-018-0047-7

58. Buscarlet M. *et al.* Lineage restriction analyses in CHIP indicate myeloid bias for TET2 and multipotent stem cell origin for DNMT3A. Blood. 2018 Jul 19;132(3):277-280. doi: https://doi.org/10.1182/blood-2018-01-829937.
59. Aspinall R, *et al.* Immune senescence. Curr Opin Immunol. 2010 Aug;22(4):497-9. doi: https://doi.org/10.1016/j.coi.2010.07.002
60. Mekker A, *et al.* Immune senescence: relative contributions of age and cytomegalovirus infection. PLoS Pathog. 2012;8(8):e1002850. doi: https://doi.org/10.1371/journal.ppat.1002850
61. Hashimoto K, *et al.* Single-cell transcriptomics reveals expansion of cytotoxic CD4 T-cells in supercentenarians. Proc Natl Acad Sci U S A. 2019 Nov 26;116(48):24242-24251. doi: https://doi.org/10.1073/pnas.1907883116
62. Fraietta JA, *et al.* Disruption of TET2 promotes the therapeutic efficacy of CD19-targeted T cells. Nature. 2018;558(7709):307–312. doi: https://doi.org/10.1038/s41586-018-0178-z
63. Shapiro MB, *et al.* RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res. 1987 Sep 11;15(17):7155-74. doi: https://doi.org/10.1093/nar/15.17.7155

Figures and legends

Fig 1 Study design: At age 110 and 2 months, peripheral blood (PB) was drawn from W111 for DNA isolation prior and after cell sorting using FACS (granulocytes (G), CD3+ T-cells (T), CD19+ B cells (B)). In addition, DNA was isolated from a skin biopsy (S) performed two weeks after the initial blood draw. During follow-up at age 111 years and 10 months, one and a half year after inclusion, and three weeks prior to her death, peripheral blood (PB) was drawn and FACS sorted. DNA was isolated from peripheral blood (PB), granulocytes (G), CD3+CD4+ T-cells (T4), CD3+CD8+ T-cells (T8) and CD19+ B-cells (B). In addition, W111 gave consent for a post-mortem brain donation and DNA was isolated from the occipital cortex (C).

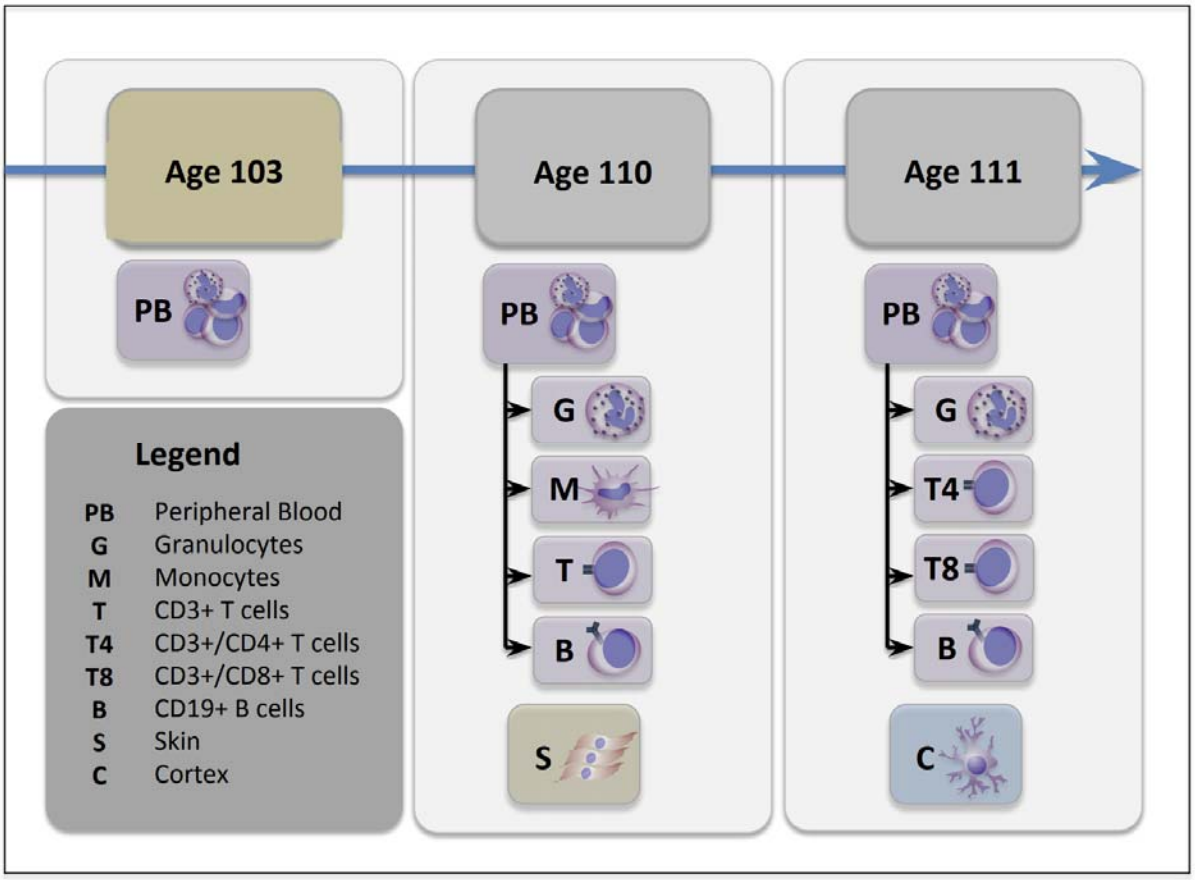


Figure 2: Cataloguing somatic mutations in peripheral blood drawn at age 110. [**A**] Whole genome sequencing of peripheral blood (PB) versus skin (S) and subsequent analysis with Sommix (**Methods**) identified 650 putative somatic mutations assigned to three confidence tiers (Tier I most confident). [**B**] Identified mutations mostly reside in non-coding genomic locations (UCSC's Variant Annotation Integrator, hg19, refgene definitions) [**C**] IGV plot of the validated splice-donor site mutation in *DNMT3A* (NM_022552.4, chr2:25,469,028 C>T, c.1429+1 G>A), altering the G nucleotide of the highly conserved GT intronic sequence[63]. Amplicon sequencing (TruSight Myeloid Sequencing Panel, Illumina) was performed at 961x and indicated a VAF of 37.8% for the mutant allele. [**D**] Identified mutations further stratified to observed nucleotide changes show a high frequency of C>T and T>C changes. Part of the C>T nucleotide changes coincide with a CpG site, potentially also affecting DNA methylation. [**E**] 3-Nucleotide sequence context of the identified mutations exhibits a high resemblance to the clock-like mutational signatures 1 and 5 (**Supplemental Figure S2**). [**F**] Validation of the identified somatic mutations using amplicon sequencing in peripheral blood (PB) versus Cortex (C), split per confidence Tier.

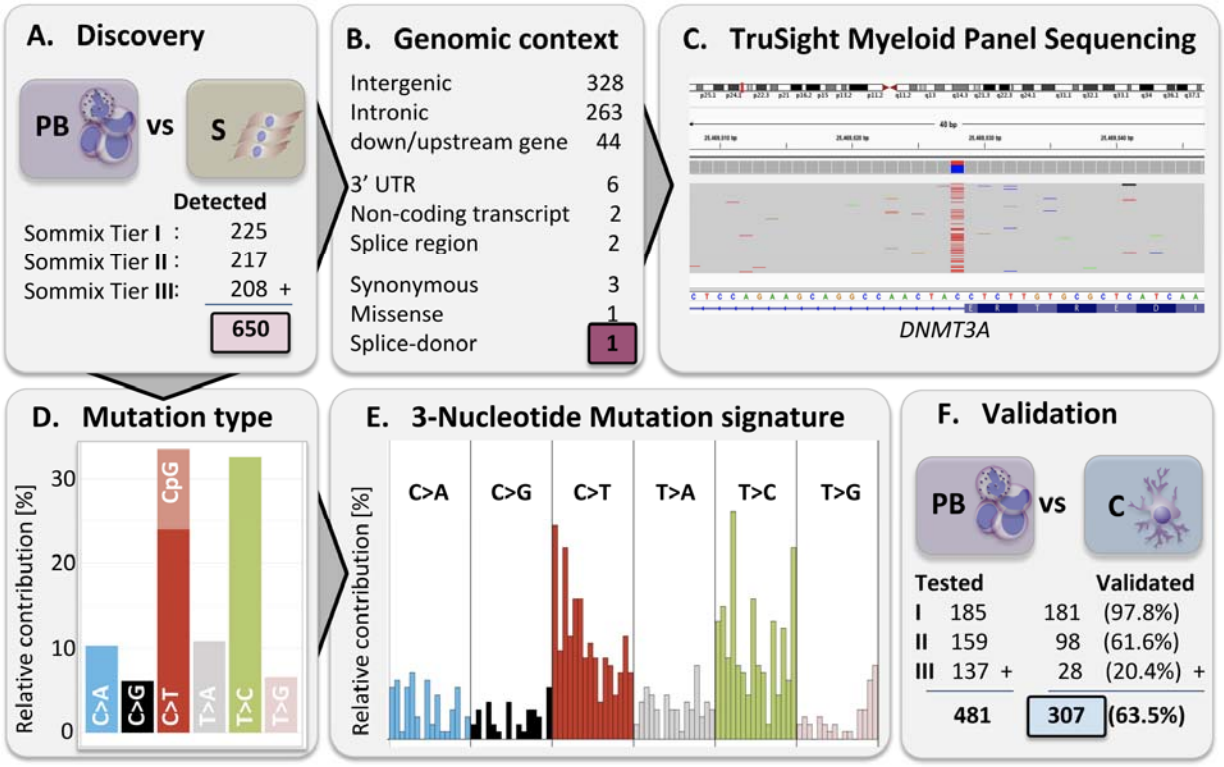


Figure 3: **Deep sequencing of longitudinal samples reveals the clonal architecture within the peripheral blood of elderly subject with age-related clonal hematopoiesis.** [**A**] Variant allele frequencies of the 307 confirmed somatic mutations at ages 103, 110 and 111. Lines connect the same mutations measured at different ages. Mutations were assigned to five independent clonal events (A-E) using SciClone[24]_and colored accordingly. [**B**] Subsequent analysis with SCHISM[25] revealed the most likely underlying clonal architecture. The number of somatic mutations supporting each clonal event are listed next to the clones. [**C**] Median VAFs after subtraction of the median VAF of the descendant clonal event indicates that clone E displays the most variation. [**D**] Reconstruction of subclonal evolution. Time frames A-E correspond to the periods in which passenger mutations (crosses) were accumulated until a clonal event driving expansion (thunder) was encountered. Widths of the time frames are roughly proportional to the number of mutations detected for each event. The y-axis reflects the relative contribution of a HSC to overall peripheral blood production. 'WINDOW' refers to our window of observation ranging from age 103 to 111, a 9-year period characterized by the expansion of clonal event E.

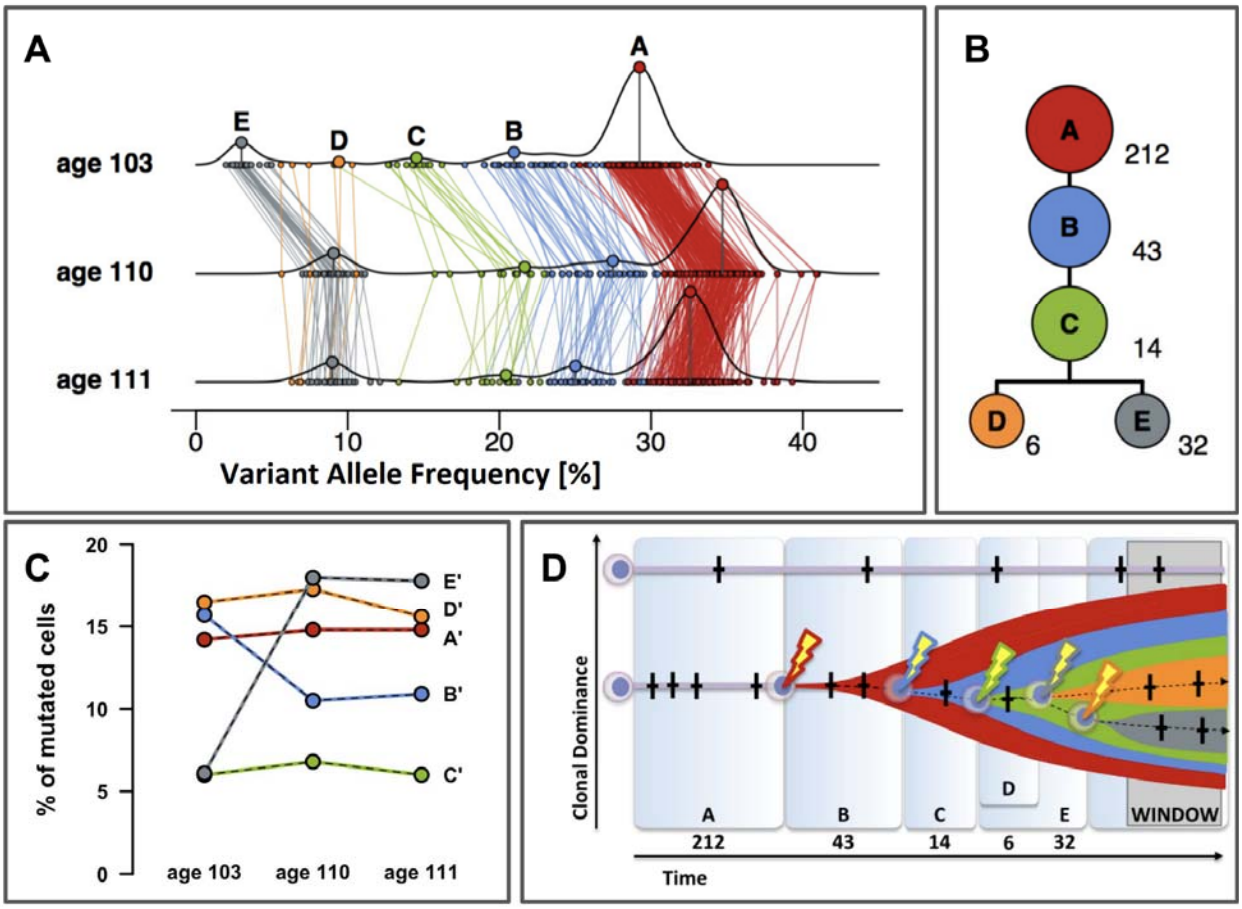


Figure 4: **Contributions of clonal events to the major subsets in peripheral blood.** [**A**] Violin plots of VAFs [%] in peripheral blood and its sorted subsets. M: Monocytes; G: Granulocytes; PB: Peripheral Blood; B: B-cells; T: T-cells; T4: CD4 T-cells; T8: CD8 T-cells. Numbers signify time points: 0: age 103; 1: age 110; 2: age 111. [**B**] Fractions of mutated cells per sorted cell subset derived from each clonal event. Barplots per subset add up to 100%. [**C-E**] For respectively T-cells, CD4 T-cells and CD8 T-cells, the ratio of VAFs for mutations within the sorted subset and the corresponding peripheral blood sample (y-axis) as a function of the VAFs in peripheral blood (x-axis). For all three T-cell subsets, clonal events with higher dominance have a disproportionately larger contribution to the production of the respective subsets.

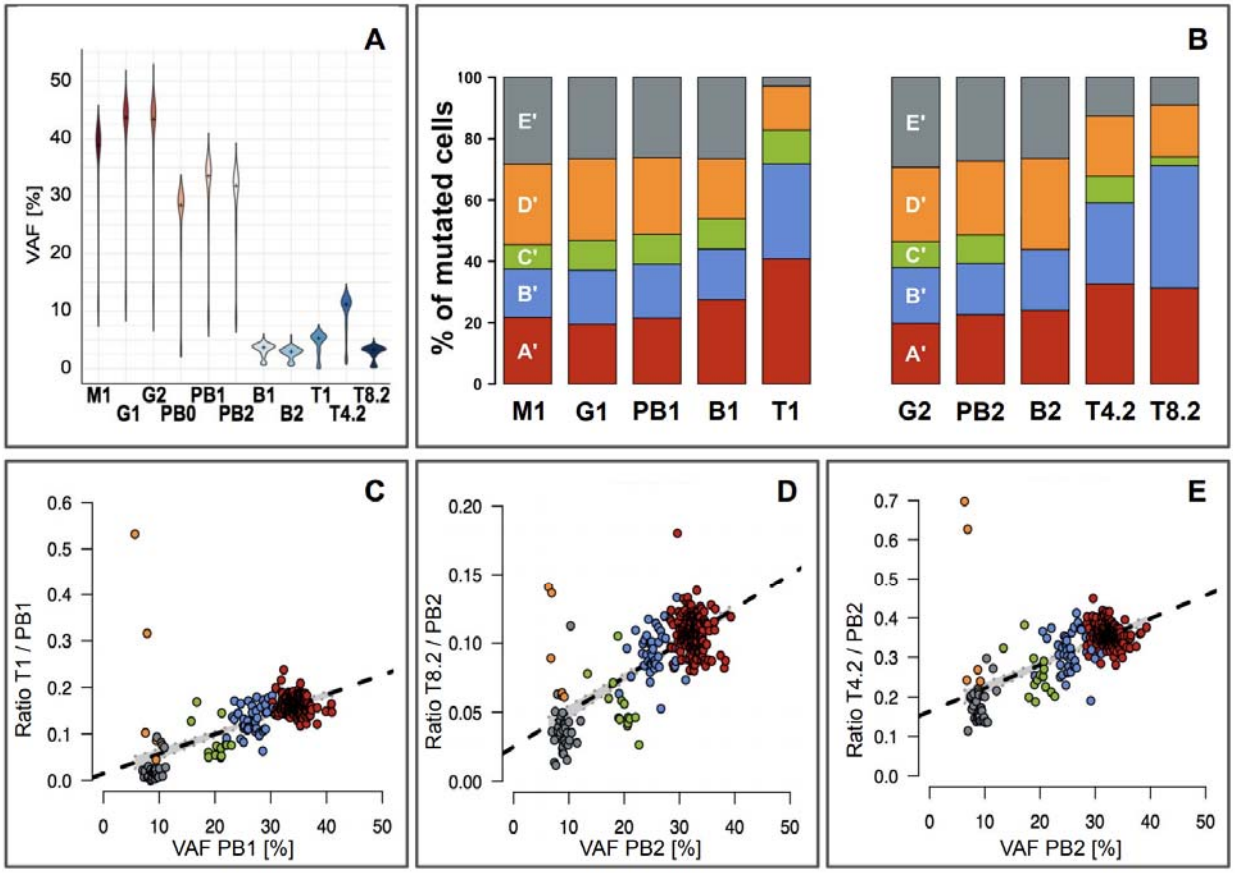
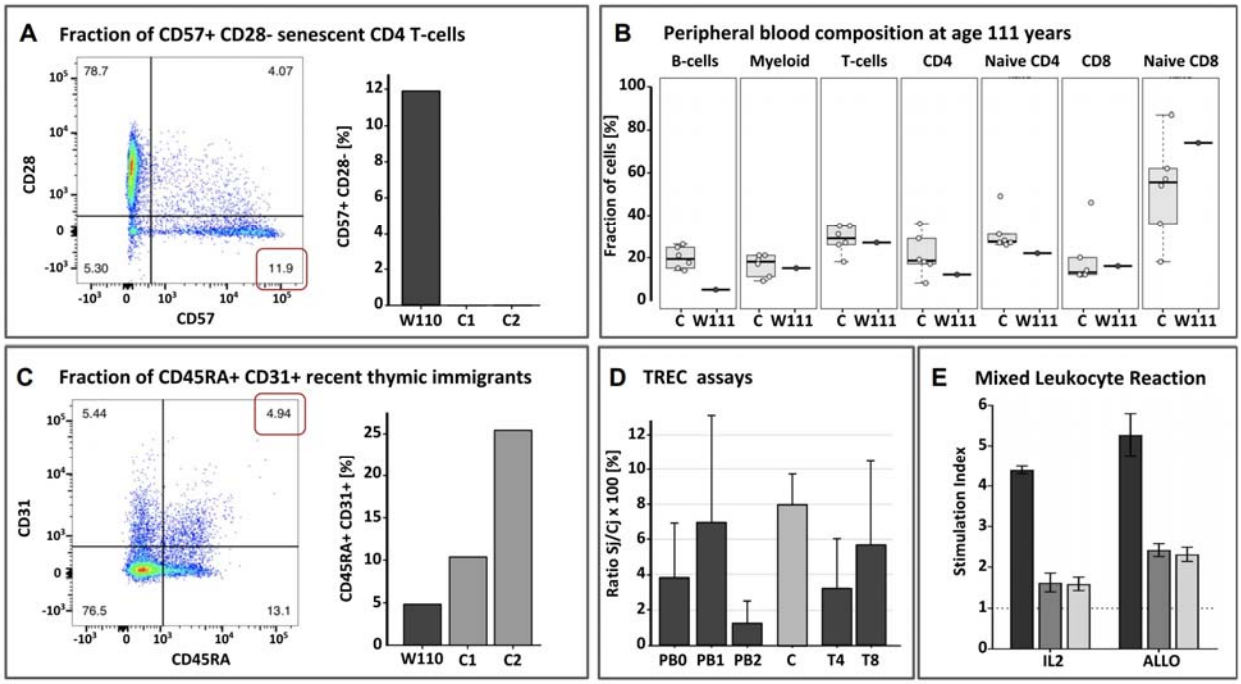
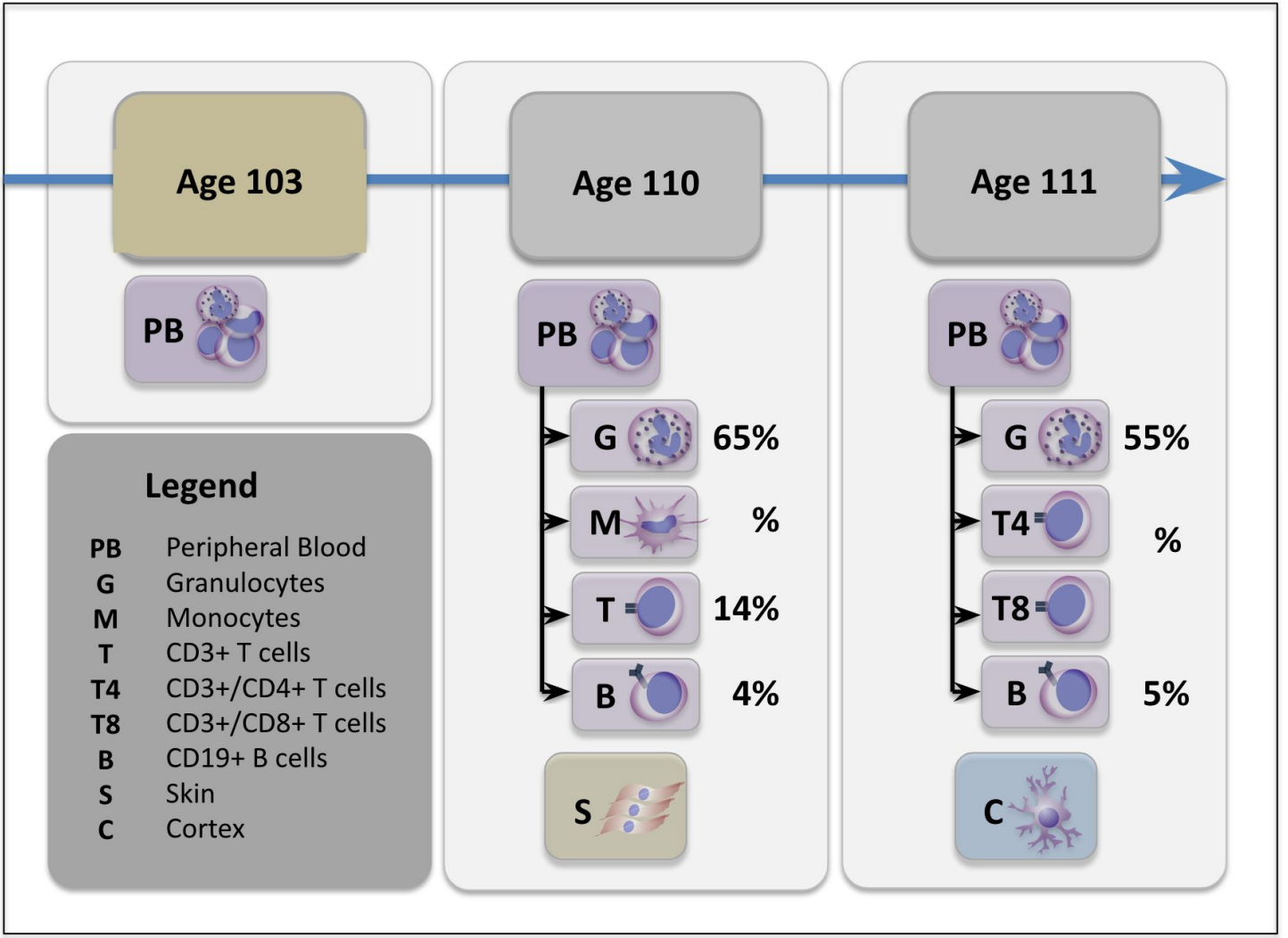
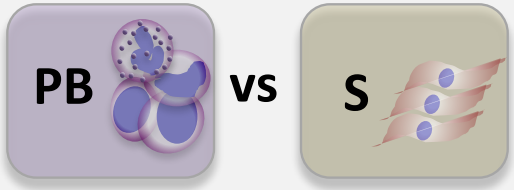


Figure 5: Immune characterization of W111. [**A**] Left: Sorting of CD57+ CD28- senescent CD4 T-cells at age 110 years; Right: W111 compared to middle-aged controls C1 and C2. [**B**] Proportions of sorted immune subsets at age 111 years (B-cells, Myeloid, T-cells, CD4+ T-cells, naive CD4+ T-cells, CD8+ T-cells, naive CD8+ T-cells) in peripheral blood of six middle-aged female controls (left) versus peripheral blood taken at age 110 years (right). [**C**] Left: Sorting of CD45RA+ CD31+ recent thymic immigrants at age 110 years (dark grey); Right: W111 compared to middle-aged controls C1 and C2. [**D**] Percentage of TREC positive cells in material derived from W111 (blue) or a middle-aged female control (grey). PB: Peripheral Blood cells; T: T-cells; T4: CD4 T-cells; T8: CD8 T-cells. Numbers signify time points: 0: age 103; 1: age 110; 2: age 111. [**E**] Stimulation Indices computed for IL2/TCR-dependent (IL2) and the allogenic (ALLO) assay of cultured CD4 T-cells (dark grey) versus two middle-aged females used as control (lighter shades of grey).




Age 103
PB
Age 110
PB
G
65%
M
%
T
14%
B
4%
S
Age 111
PB
G
55%
T4
%
T8
B
5%
C
Legend
PB Peripheral Blood
G Granulocytes
M Monocytes
T CD3+ T cells
T4 CD3+/CD4+ T cells
T8 CD3+/CD8+ T cells
B CD19+ B cells
S Skin
C Cortex

A. Discovery

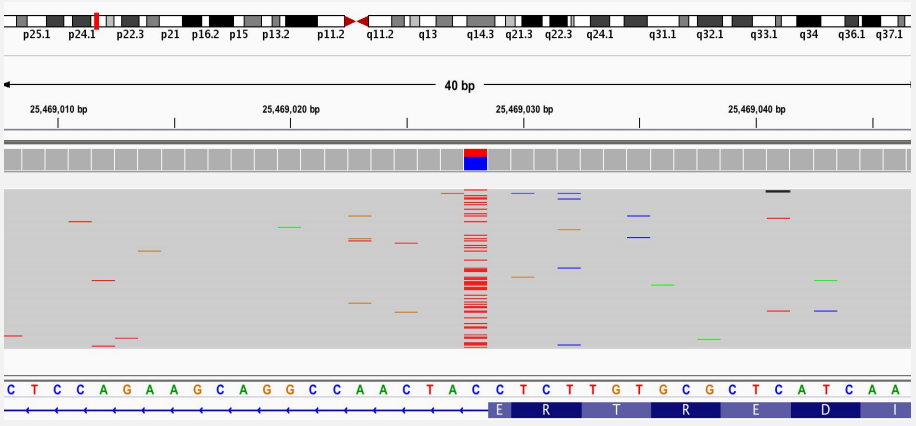


	Detected
Sommix Tier **I** :	225
Sommix Tier **II** :	217
Sommix Tier **III**:	208 +
	650

B. Genomic context

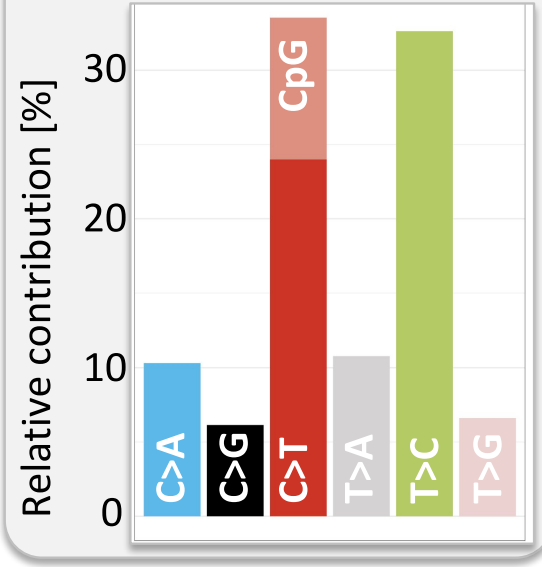
Intergenic	328
Intronic	263
down/upstream gene	44
3' UTR	6
Non-coding transcript	2
Splice region	2
Synonymous	3
Missense	1
Splice-donor	**1**

C. TruSight Myeloid Panel Sequencing

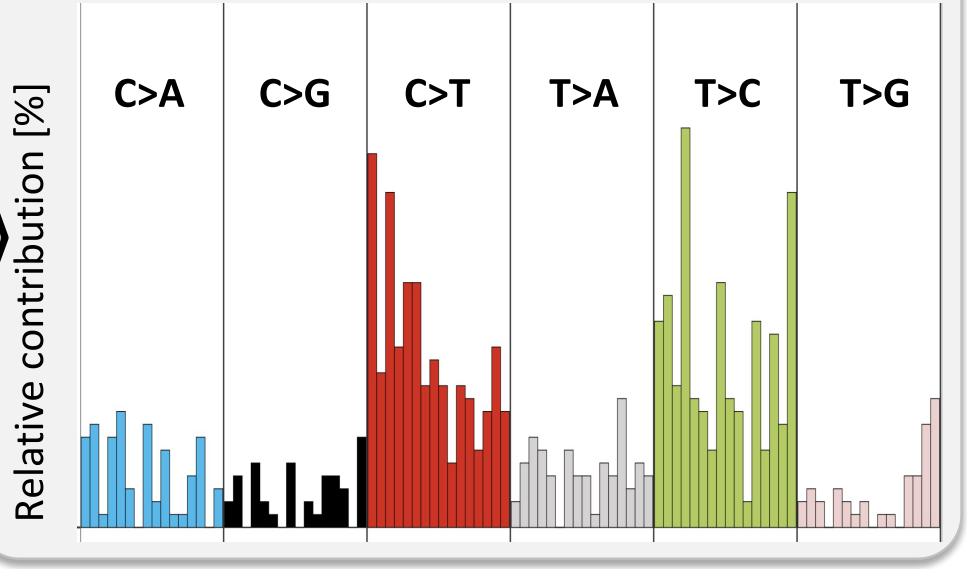


DNMT3A

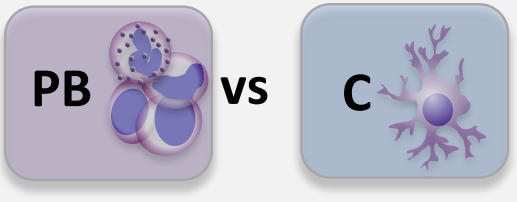
D. Mutation type



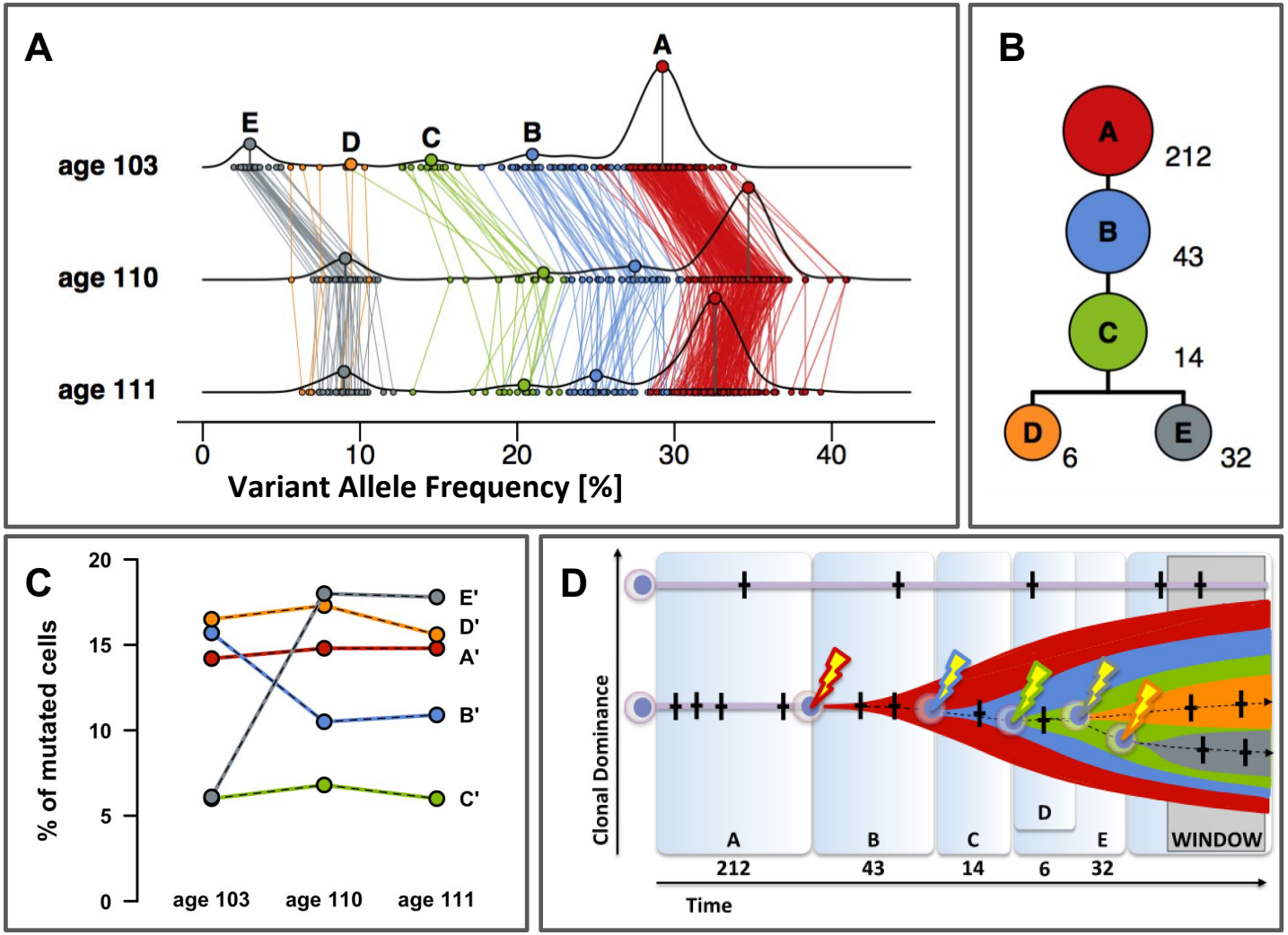
E. 3-Nucleotide Mutation signature

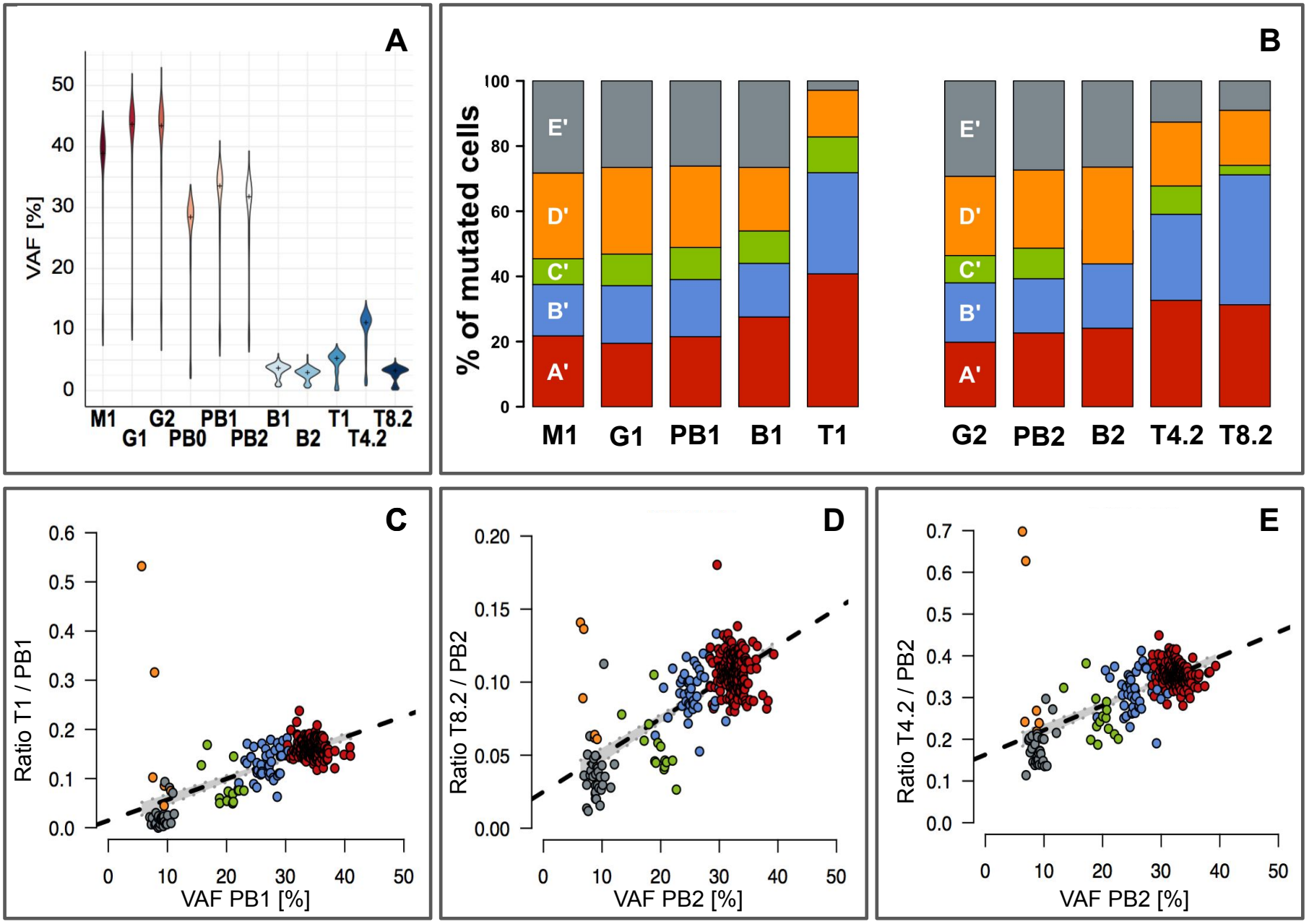


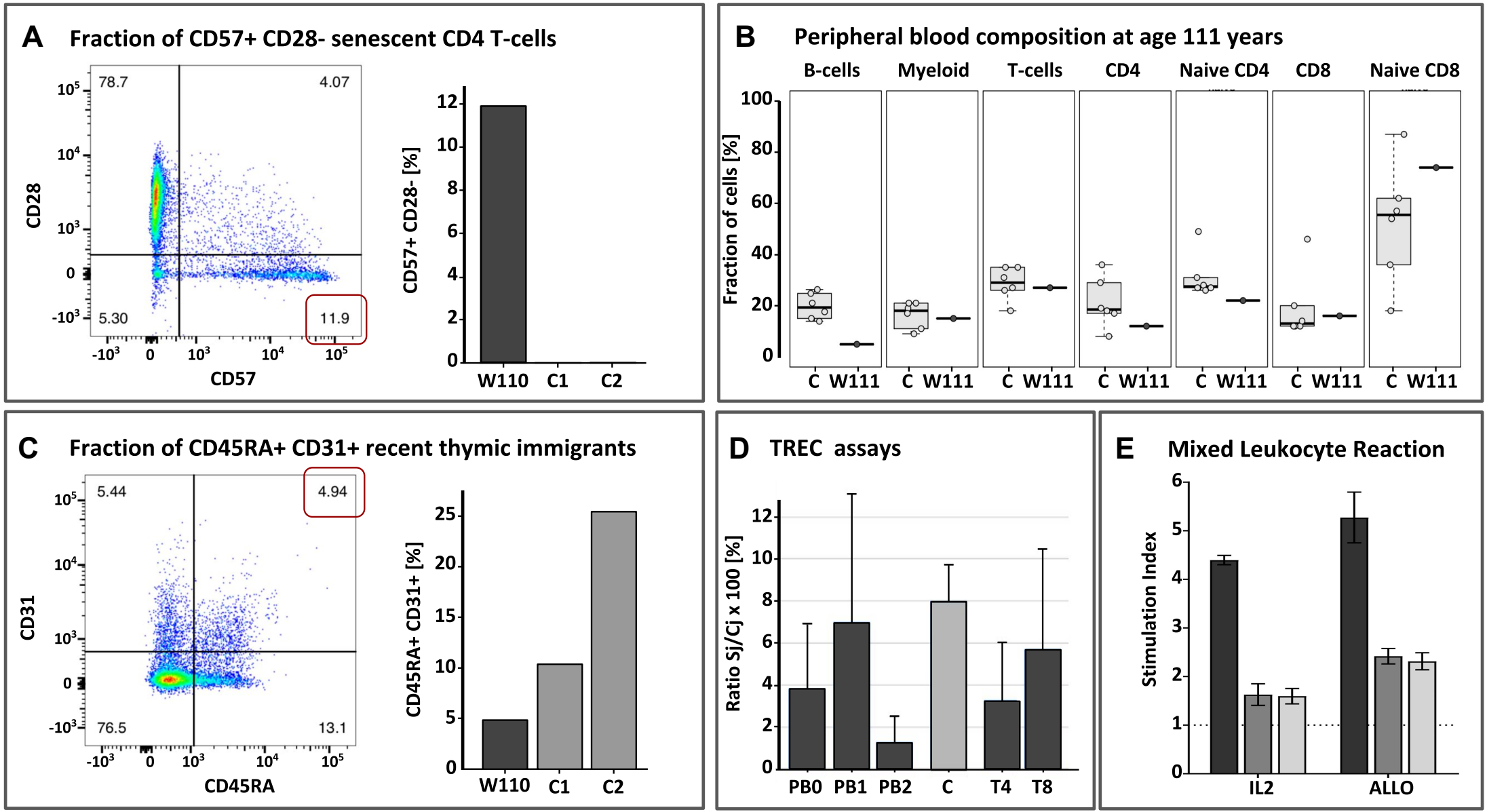
F. Validation



	Tested		Validated	
I	185		181	(97.8%)
II	159		98	(61.6%)
III	137	+	28	(20.4%) +
	481		**307**	**(63.5%)**


A
E
D
C
B
A
age 103
age 110
age 111
0
10
20
30
40
Variant Allele Frequency [%]
B
A
212
B
43
C
14
D
6
E
32
C
% of mutated cells
20
15
10
5
0
E'
D'
A'
B'
C'
age 103
age 110
age 111
D
Clonal Dominance
A
212
B
43
C
14
D
6
E
32
WINDOW
Time


A
VAF [%]
0
10
20
30
40
50
M1
G1
G2
PB0
PB1
PB2
B1
B2
T1
T4.2
T8.2
B
% of mutated cells
0
20
40
60
80
100
A'
B'
C'
D'
E'
M1
G1
PB1
B1
T1
G2
PB2
B2
T4.2
T8.2
C
Ratio T1 / PB1
0.0
0.1
0.2
0.3
0.4
0.5
0.6
VAF PB1 [%]
0
10
20
30
40
50
D
Ratio T8.2 / PB2
0.00
0.05
0.10
0.15
0.20
VAF PB2 [%]
0
10
20
30
40
50
E
Ratio T4.2 / PB2
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
VAF PB2 [%]
0
10
20
30
40
50


A
Fraction of CD57+ CD28- senescent CD4 T-cells
78.7
4.07
5.30
11.9
CD28
CD57
CD57+ CD28- [%]
W110
C1
C2
B
Peripheral blood composition at age 111 years
B-cells
Myeloid
T-cells
CD4
Naive CD4
CD8
Naive CD8
Fraction of cells [%]
C
W111
C
Fraction of CD45RA+ CD31+ recent thymic immigrants
5.44
4.94
76.5
13.1
CD31
CD45RA
CD45RA+ CD31+ [%]
W110
C1
C2
D
TREC assays
Ratio Sj/Cj x 100 [%]
PB0
PB1
PB2
C
T4
T8
E
Mixed Leukocyte Reaction
Stimulation Index
IL2
ALLO