Somatic mutations in lymphocytes in patients with immune-mediated bone marrow failure

Sofie Lundgren

Lääketieteen kandidaatti

Hematology Research Unit Helsinki, University of Helsinki and Helsinki University Hospital Comprehensive Cancer Center

Translational Immunology Research Program and Department of Clinical Chemistry and Hematology, University of Helsinki

Helsinki 15.5.2020

Tutkielma

sofie.lundgren@helsinki.fi

Ohjaajat: Prof. Satu Mustjoki ja LT Mikko Keränen

HELSINGIN YLIOPISTO

Lääketieteellinen tiedekunta

HELSINGIN YLIOPISTO HELSINGFORS UNIVERSITET

Tiedekunta/Osasto – Fakultet/Sektion – Faculty	Laitos – Institution – Department
Lääketieteellinen tiedekunta	Tutkimusohjelmayksikkö (RPU)

Tekijä – Författare – Author

Sofie Alexandra Lundgren

Työn nimi – Arbetets titel – Title

Somatic mutations in lymphocytes in patients with immune-mediated bone marrow failure

Oppiaine – Läroämne – Subject

Lääketiede

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä -Sidoantal - Number of pages
Tutkielma	15.5.2020	54

Tiivistelmä – Referat – Abstract

The role of somatic mutations in cancer development is undisputable, but our knowledge of the prevalence and function of somatic mutations in healthy cells is yet unresolved question. In immune-mediated bone-marrow failure (such as aplastic anemia and hypoplastic myelodysplastic syndromes) hematopoietic stem cells are destroyed by cytotoxic T cells, which have been shown to be oligoclonal and in some cases harbor *STAT3* mutations. We addressed clonality of non-malignant T cells by investigating somatic mutations with a custom-made gene panel of 2533 genes. We sequenced CD4+ and CD8+ cells of 29 patients with immune-mediated bone marrow failure and compared the results to 20 healthy controls as well as previously published data on patients with aplastic anemia. Somatic variants in lymphocytes were common both in patients and healthy controls, but especially enriched in aplastic anemia patients and CD8+ T cells. CD8+ T cells of aplastic anemia patients accumulated most mutations on JAK-STAT and MAPK pathways. The number of somatic mutations was associated with CD8+ T cell clonality, assessed by T cell receptor beta sequencing. To further understand the role of mutations in T cells, we performed single-cell RNA sequencing from 6 longitudinal samples of 2 aplastic anemia patients with *STAT3, KRAS, NFATC2* and *PTPN22* mutations in CD8+ T cell clones. Mutated clones showed cytotoxic phenotype, which was altered by successful immunosuppressive treatment. Our results suggest that somatic mutations in T cells are common, associate with clonality and may alter T cell phenotype. The role of somatic mutations in the initiation and persistence of autoreactive clones needs to be further investigated.

(258 words)

Tiivistelmä – Referat – Abstract

Somaattisten mutaatioiden rooli syövän kehityksessä on osoitettu kiistattomasti, mutta niiden yleisyys ja vaikutukset terveissä soluissa tunnetaan vielä puutteellisesti. Tutkimushypoteesimme mukaan somaattiset mutaatiot voivat edistää T-solukloonien kasvua ja selviytymistä kroonisessa inflammaatiossa, kuten autoimmuunisairauksissa. Tutkimme somaattisia mutaatioita immuunivälitteisissä luuydinkato-oireyhtymissä (aplastisessa anemiassa ja hypoplastisessa myelodysplastisessa oireyhtymässä), joissa sytotoksiset T-solut tuhoavat luuytimen kantasoluja. Aplastista anemiaa sairastavilla potilailla on aiemmin osoitettu esiintyvän *STAT3*-mutaatioita. Kartoitimme ei-pahalaatuisten T-solujen somaattisia mutaatioita 2533 geenin paneelilla. Sekvensoimme CD4+ ja CD8+ T-solut 29 potilaalta, jotka sairastivat immunologista luuydinkato-oireyhtymää, ja vertasimme tuloksia 20 terveeseen sekä aiemmin julkaistuun aineistoon aplastista anemiaa sairastavista potilaista. Somaattiset muutokset T-soluissa olivat yleisiä niin terveillä kuin potilaillakin, mutta ne rikastuivat aplastista anemiaa sairastavien potilaiden CD8+ T-soluihin. Näiden T-solujen mutaatiot kasautuivat voimakkaimmin JAK-STAT ja MAP kinaasi -signalointireiteille. Havaitsemiemme somaattisten mutaatioiden määrä oli yhteydessä T-solujen klonaalisuuteen, jota tutkimme sekvensoimalla T-solureseptoreita. Ymmärtääksemme somaattisten mutaatioiden roolia T-soluissa, analysoimme yksisolu-RNA-sekvensointimenetelmällä yhteensä 6 pitkittäisnäytettä kahdelta potilaalta, joiden CD8+ T-solukloneissa oli *STAT3, KRAS, NFATC2* ja *PTPN22* geenien mutaatioita. Mutatoituneet kloonit osoittautuivat fenotyypiltään sytotoksisiksi, ja muuttuivat immunosuppressiivisen hoidon myötä. Tutkimustulostemme mukaan somaattiset mutaatiot lymfosyyteissä ovat yleisiä, liittyvät klonaalisuuteen ja voivat muuttaa T-solujen fenotyyppiä. Lisätutkimusta tarvitaan selvittämään, ovatko somaattiset mutaatiot T-soluissa syy vai seuraus klonaalisesta kasvusta, ja mikä on näiden solujen rooli autoimmuniteetin kehittymisessä.
(197 sanaa)

Avainsanat – Nyckelord – Keywords

T-Cells; Autoimmunity; Bone Marrow Failure; Aplastic Anemia

Säilytyspaikka – Förvaringställe – Where deposited

Student saves electronic version to E-thesis.

Muita tietoja – Övriga uppgifter – Additional information

1 Introduction

Acquired aplastic (AA) anemia is an immune-mediated bone marrow failure syndrome (iBMF) in which cytotoxic T cells mediate the destruction of hematopoietic stem and progenitor cells (HSPC) (1,2). Interestingly, these target cells are not necessary healthy, but generally possess clonal features such as X-chromosome skewing (3), cytogenetic abnormalities (4,5), uniparental disomy of the 6p (6), somatic phosphatidylinositol glycan class A (PIGA) gene mutations leading to paroxysmal nocturnal hemoglobinuria (7) or other somatic mutations (8). It has been suggested that in AA, some clonal changes in HSPCs are driven by selective pressure from autoreactive T cells and clonal hematopoiesis (CH) serves as an immune escape mechanism (1,9). However, somatic mutations of CH may also drift to lymphocytes, which is a phenomenon called lympho-myeloid clonal hematopoiesis (LM-CH) (10,11). Somatic mutations can also arise *de novo* in mature lymphocytes.

The exact pathogenesis of AA is elusive, and the antigenic target on HSPC remains unknown. However, oligoclonal CD8+ T cells are detected in almost all AA patients and have shown to induce apoptosis of autologous HSPC (12,13). Furthermore, T cells in AA patients have been shown to be aberrantly active (14-16). We hypothesize that somatic events in lymphocytes could promote these alterations in AA.

Previously, we and others have detected functional somatic mutations in lymphocytes in patients with newly diagnosed rheumatoid arthritis (17), Felty syndrome immunodeficiency (18), multiple sclerosis (19) chronic graft-versus-host disease (20) and T-cell large granular lymphocytic (T-LGL) leukemia, where 40 % of T-LGL leukemia patients harbour activating *STAT3* mutations in CD8+T cells (21). Furthermore, in 11% of AA patients acquired *STAT3* mutations (22) were discovered, potentially explaining the overlapping characteristics of these related disorders.

We propose that mutations in mature lymphocytes may emerge during vigorous clonal expansion after recognition of their cognate antigen. This phenomenon could participate in the pathogenesis of autoimmune diseases. In the current study, we characterized the spectrum of somatic mutations in T cells in patients with AA and closely related hypoplastic myelodysplastic syndrome (hMDS). Furthermore, we evaluated their

origins and described their potential role at single-cell resolution at the diagnosis, relapse and the recovery of AA. We show that somatic mutations in lymphocytes are common in AA and hMDS patients, and even in healthy individuals. These findings extend our understanding of the role of somatic mutations in non-malignant diseases and provide basis for the further development of novel therapeutic approaches in lymphocyte-mediated autoimmune diseases.

2 Methods

2.1 Samples

Peripheral blood and bone marrow were collected from patients suffering from AA (n = 24) or hMDS (n = 5), see Table S1 for more detailed cohort description. All patients had given their written informed consent before sample collection. Healthy blood donor buffy coats were provided by the The national HLA laboratory, Finnish Red Cross Blood Service (n = 20). Mononuclear cells (MNC) from collected samples were separated with Ficoll centrifugation. CD8+ and CD4+ T-cell fractions were separated by immunomagnetic bead sorting (Miltenyi Biotec) or by flow cytometry (FacsAria). Patients were recruited from the Department of Hematology in the Helsinki University Hospital and from other participating university hospitals (Karonlinska Institute, Sweden and Cleveland Clinic, Cleveland, OH, USA).

2.2 Immunogene panel sequencing and data analysis

DNA from separated CD4+ and CD8+ T cells was isolated and fragmented for further analysis. We designed a custom gene panel based on 2533 candidate genes along pathways important in innate and adaptive immunity (18). A pool of custom probes captured the genomic region of interest from a fragmented genomic DNA sample. The target DNA fragments were sequenced using Illumina HiSeq 2500 with HiSeq hight output mode using v4 kits (Illumina) with 101-length paired-end reads. The sequencing was performed in Institute for Molecular Medicine Finland (FiMM). Sequencing coverage was computed using Samtools depth.

2.3 Other datasets

We received raw whole exome sequencing (WES) data from a previously published dataset of 37 AA patients from Yoshizato et al (8). In our analysis, we included CD3- and CD3+ MNC sample from each patient. To enable comparison to our in-house dataset, we preprocessed the sequencing data and called variants with same bioinformatics pipeline as our in-house data (see below).

We also utilized raw WES data from skin samples of 9 AA patients, sequenced earlier for clinical purposes with Nextera Rapid Exome Kit (Illumina, Cat#: FC-140-1083) as described previously (21). We preprocessed and called variants with same bioinformatics pipeline with other data and then used it as a part of Panel Of Normals (PON) for filtering variants detected with WES and immunogene panel sequencing (see below).

All cohorts used in somatic variant analysis are summarised in Table 1.

Table 1. Cohorts in somatic variant analysis

Dataset	Number of subjects	Purpose
Immunogene panel sequencing from iBMF patients	24 AA, 5 hMDS	To investigate somatic mutations in T cells
Immunogene panel sequencing from healthy individuals	20	PON for immunogene panel samples and comparison of adjusted mutation burden for AA and hMDS samples
WES skin from patients with AA	9	PON for WES samples and immunogene panel samples
CD3+ and CD3- MNC WES from patients with AA	37	To compare CH derived mutations with somatic mutations seen in T cells

2.4 Identification of somatic variants

Preprocessing of sequencing data was performed using Trimmomatic software for filtering low quality, adapter sequences, and read length less than 36bp in length, BWA-MEM for read alignment, SortSAM for alignment sorting and Picard toolkit (Broad Institute) for marking PCR duplicates. Somatic variant calling was done with Genome Analysis Toolkit (GATK), according to GATK somatic short variant best practice (version 3.5). First, variants from immunogene panel samples were identified with Mutect2 and filtered with default settings supplemented with requirement of read depth >= 10. For Finnish samples (sequenced with immunogene panel) we used both paired tumor-normal (with CD4+ cells as germline control sample for CD8+ cells and vice versa) and tumor-only (each sample separately without paired germline sample) variant calling strategy. We supplemented this list of variants with hotspot variants in genes recurrently mutated in AA(8) or T cell neoplasms (23). Hotspots were determined as variants detected in more than 30 COSMIC v90 (24) samples and constituting ≥1% of COSMIC v90 variants in each gene. We then performed genotyping of these variants from all samples (immunogene panel samples as well as CD3pos, CD3neg and skin WES samples) using GATK-4.1.3.0 Mutect2 program with --alleles parameter.

AA and hMDS immunogene panel variants were filtered against a PON consisting of variants detected in two or more immunogene panel healthy T cell samples or skin WES samples. CD3- and CD3+ WES samples were filtered against a PON consisting of variants detected in two or more skin WES samples. To enable differentiation of variants with a low variant allele frequency from technical or biological artefacts, datasets were filtered after variant calling for vector contamination as previously described (25). Variants were annotated and filtered using the Annovar tool15 against the RefGene database. At first, all variant calls were normalised using bcftools16. Variants other than those passing all MuTect2 filters with a TLOD ≥ 6.3 or a TLOD ≥ 5.0 were filtered. Variant data were then filtered for false-positives by removing variants in intronic and intergenic regions, with a total coverage ≤ 30, and not supported by at least one read in both directions as well as variants with variant quality value ≤ 40, variant allele frequency ≤ 2% or ≥30%, strand odd ratio for SNVs ≥ 3.00, strand odd ratio for indels ≥ 11.00, minor allele frequency ≥ 1% in the 1KG database, minor allele frequency ≥ 3% in the EPS database, minor allele frequency ≥ 0.1% in general, African, Finnish, Latino, East Asian, and Non-European ExAC, gnomAD exome, gnomAD

genome or gnomAD g3 databases and likelihood ratios score ≤ 2.00. Variants with a variant allele frequency 1-2% were accepted, if supported by five or more COSMIC17 samples.

For functional analyses, the previous variant call set was filtered further by removing synonymous mutations and non-frameshift variants. Variants that had been genotyped from both CD4+ and CD8+ T cells of same patient were interpreted to be derived from LM-CH, and the rest of the variants were categorized as CD8+ and CD4+ T cell variants.

2.5 Amplicon validation

Amplicon sequencing was performed to validate mutations found in the immunopanel sequencing (Supplementary table S3). Amplicons were amplified with 2-step PCR protocol. First PCR was done in a volume of 20 µL containing 10 ng of sample DNA, 10 µL of 2x Phusion High-Fidelity PCR Master Mix and 0.375 µM of each locus-specific primer. The reaction mix was brought to a final volume with water. The second PCR was done in a volume of 20 µL containing 1 µL of the amplified product from the first PCR, 10 µL of 2x Phusion High-Fidelity PCR Master Mix, 0.375 µM of index primer 1 and 0.375 µM of index primer 2. The reaction mix was brought to a final volume with water.

DNA Engine Tetrad 2 (Bio-Rad Laboratories) or G-Storm GS4 (Somerton) thermal cyclers were used to cycle the 1-step PCR samples according to the program: initial denaturation at 98 °C 30 s, 30 cycles at 98 °C for 10 s, at 67 °C for 30 s, and at 72 °C for 15 s, and the final extension at 72 °C for 10 min. Two-step amplifications were done with the same thermal cyclers. The second PCR (Index PCR) was done according to the program: initial denaturation at 98 °C 30 s, 8 cycles at 98°C for 10s, at 65°C for 30s, and at 72°C for 20s, and the final extension at 72 °C for 5 min. The amplified samples were pooled together and the pool was purified with Agencourt AMPure XP beads (Beckman Coulter, CA, USA) twice using 1x volume of beads compared to the sample pool volume. Agilent 2100 Bioanalyzer (Agilent Genomics, CA, USA) was used to quantify amplification performance and yield of the purified sample pools. Sample pools were sequenced with Illumina HiSeq System using Illumina HiSeq Reagent Kit

v4 100 cycles kit or Illumina MiSeq System using MiSeq 600 cycles kit (Illumina, San Diego, CA, USA).

Sequencing reads alignment was performed with Bowtie2, and GATK IndelRealigner was used for local realignment near indel. A variant was called if variant base frequency was 0.5% of all reads covering a given a position. All variants with the base quality frequency ratio (ratio of number of variant calls/numbers of all bases and quality sum of variant calls/ quality sum of all bases at the position) ≥ 0.9 were considered as true somatic variants.

2.6 Somatic mutation burden

As described above, healthy T cell samples were used as part of PON when filtering variants. Hence, patients' variant burden (how many variants were detected per Mb) was not directly comparable to that of healthy individuals. To overcome this, we calculated adjusted somatic mutation burden for each sample by counting only variants, which were detected in one individual ("single variants") in the whole immunogene panel cohort supplemented with 9 healthy skin WES samples. Adjusted somatic mutation burden was calculated separately for variants seen only in CD4 cells or CD8 cells in each patient. Total number of variants was divided by the number of bases sequenced with the minimum coverage of 30x in each sample.

2.7 Pathogenicity predictions

Pathogenicity of detected variants was predicted with 8 tools: SIFT (26), Polyphen-2 (HVar and HDiv) (27), likelihood ratio test (28), MutationTaster (29), MutationAssessor (30), FATHMM (31), VEST3 (32,33), CADD (34,35), and 3 conservation scores: SiPhy (36), phyloP (placental and vertebrate) (37). We then selected pathogenic variants in accordance with majority of these tools (if more than 50% of predictions categorized variant as "damaging", "pathogenic", "possibly pathogenic", "medium" or high", it was predicted to be pathogenic).

2.8 Search for oligoclonal T cell populations

Sorted CD8+ T cell fractions were analyzed with a multiplexed PCR assay that targets the variable CDR3 region of the rearranged TCRβ locus (38). Next-generation sequencing was performed and data was analyzed with the ImmunoSEQ analysis tools provided by Adaptive Biotechnologies. Clonality was calculated as normalized measure of Shannon's entropy according to the formula

$$Sample\ clonality = 1 - \frac{-\sum_{i=1}^{N} p_i \log_2(p_i)}{\log_2(N)}$$

where p_i is the proportional abundance of the rearrangements *I* and N is the total number of rearrangements.

2.9 TCR Vβ family based flow cytometry analysis and sorting

TCR Vβ families were analyzed from frozen MNC by flow cytometry-based antibody staining using IOTest® Beta Mark TCR Vβ Repertoire Kit recognizing 24 members of TCR β chain, which covers about 70% of the normal human TCR Vβ repertoire (cat. no: IM3497, Beckman Coulter). Briefly, MNC samples were stained after thawing with anti-CD3 (SK7, BD, cat. no 345767), anti-CD4 (SK3, BD, cat. no 345770), and anti-CD8 (SK-1, BD, cat. no 335822) and the panel of TCR Vβ antibodies. All antibodies were used according to the manufacturers' instructions. Stained cells were further analysed and sorted using FACSAria II (Becton Dickinson).

2.10 Signature analysis

All synonymous and non-synonymous variants were included in the analysis. Identification of mutational signatures was done using the deconstructSigs18 software with default parameters and using cancer profiles downloaded from the COSMIC web site on September 2017. Function mapSeqlevels from the package GenomeInfoDb was used to convert EnsEMBL chromosome nomenclature to UCSC nomenclature.

2.11 Single-cell gene expression and V(D)J transcript profiling

Frozen MNC from PB or BM were sorted with BD Influx Cell sorter and the gene and V(D)J transcript profiles were studied with 10x Genomics Chromium Single Cell V(D)J

and 5' Gene Expression platform. After sorting, the sample processing and sequencing was done at the Institute of Molecular Medicine Finland. Briefly, single-cells were partitioned using a Chromium Controller (10X Genomics) and scRNA-seq and TCRαβ-libraries were prepared using Chromium Single Cell 5' Library & Gel Bead Kit (10X Genomics), according to manufacturer's instructions (CG000086 Rev D). 12 000 cells from each sample, suspended in 0.04% BSA in PBS, were loaded on the Chromium Single Cell A Chip. During the run, single-cell barcoded cDNA is generated in nanodroplet partitions. The droplets were subsequently reversed and the remaining steps were performed in bulk. Full length cDNA was amplified using 14 cycles of PCR (Veriti, Applied Biosystems). TCR cDNA was further amplified in a hemi-nested PCR reaction using Chromium Single Cell Human T Cell V(D)J Enrichment Kit (10X Genomics). Finally, the total cDNA and the TCR-enriched cDNA was subjected to fragmentation, end repair and A-tailing, adaptor ligation, and sample index PCR (14 and 9 cycles, respectively). All libraries were sequenced using NovaSeq 6000 system (Illumina), S1 flowcell with the following read length configuration for gene expression libraries: Read1=26, i7=8, i5=0, Read2=91. Length configurations used for TCR-enriched libraries: Read1=150, i7=8, i5=0, Read2=150. The raw data was processed using Cell Ranger 3.0.1 pipelines. Pipeline called "cellranger mkfastq" was used to produce FASTQ (raw data) files, "cellranger count" to perform alignment, filtering and UMI counting for the 5' gene expression data and "cellranger vdj" to perform V(D)J sequence assembly and paired clonotype calling for the V(D)J data. Mkfastq was run using the Illumina bcl2fastq v2.2.0 and alignment was done against human genome GRCh38.

2.12 Single-cell RNA-sequencing data analysis

All cells were subject to quality control. Cells with high amount of mitochondrial transcripts (>15% of all UMI counts) or ribosomal transcripts (>50%), cells with less than 100 genes or over than 4500 genes expressed, cells expressing low or high (<25% or >60%) amount of house-keeping genes. or cells with low or high read depth (<500 or >30 000) were excluded from the analyses. We captured 7822 – 15 651 cells per sample.

To overcome batch-effect, we used a recently described probabilistic framework to overcome different nuisance factors of variation in an unsupervised manner with deep generative modelling as described elsewhere (39). Briefly, the transcriptome of each cell is encoded through a nonlinear transformation into a low-dimensional, batch corrected latent embedding. The latent embedding was then used for graph-based clustering implemented in Seurat (3.0.0) and UMAP-dimensionality reduction. Differential expression analyses were performed based on the t-test, as suggested by Robinson et al (40). Clusters were annotated using differentially expressed genes, comparison to bulk-RNAseq based on sorted immune subsets and canonical markers. Gene Set Enrichment Analysis (GSEA) (software.broadinstitute.org/gsea/index.jsp) between groups was based on ordered gene lists by fold-change. Overlap with GO and HALLMARK-categories was assessed and the False Discovery Rate (FDR) calculated while the number of permutations was 1000.

3 Results

3.1 CD8+ T cells from AA patients had the highest mutation burden

We performed immunogene panel deep sequencing of CD4+ and CD8+T cells from age-matched groups of 24 AA patients, 5 hMDS patients and 20 healthy controls (age median[range] respectively: 63[23-84]; 52[34-68]; 60.5[44-66], Fig S1). Mean sequencing depth was 295x [248x, 370x and 359x for AA, hMDS and healthy samples respectively] (Figure S2A). We divided the variants to 3 categories: somatic variants detected only in CD4 or CD8 sample and LM-CH variants (see Methods). We detected altogether 842 somatic variants in 583 genes in patients' and healthy controls' CD4+ and CD8+ T cells, with mean[median] of 9.1[8] variants per sample (1.60 [1.49] variants/Mb) and mean variant allele frequency of 4.5%. Density plot of variant allele frequencies is presented in Figure S3A.

Adjusted somatic mutation burden (see Methods) was highest in CD8+ T cells in patients with AA, then hMDS and healthy control CD8+ T cells. In every group, CD4+ cells had lower mutation burden that CD8+ T cells (mean[median] for each group: AA

CD8 1.52[1.59], AA CD4 0.98[0.87], hMDS CD8 1.43[1.34], hMDS CD4 1.06[0.85], healthy CD8 1.00[0.78], healthy CD4 0.71[0.65], Kruskal-Wallis p-value = 0.0046, Figure 1). Adjusted somatic mutation burden was not correlated with sequencing depth (p = ns, Figure S2B). We validated selected set of mutations with amplicon sequencing, and results are shown in Table S3. From variants detected with immunopanel sequencing, 15/16 were confirmed with amplicon sequencing from at least one sample, if several samples from same patient were tested.

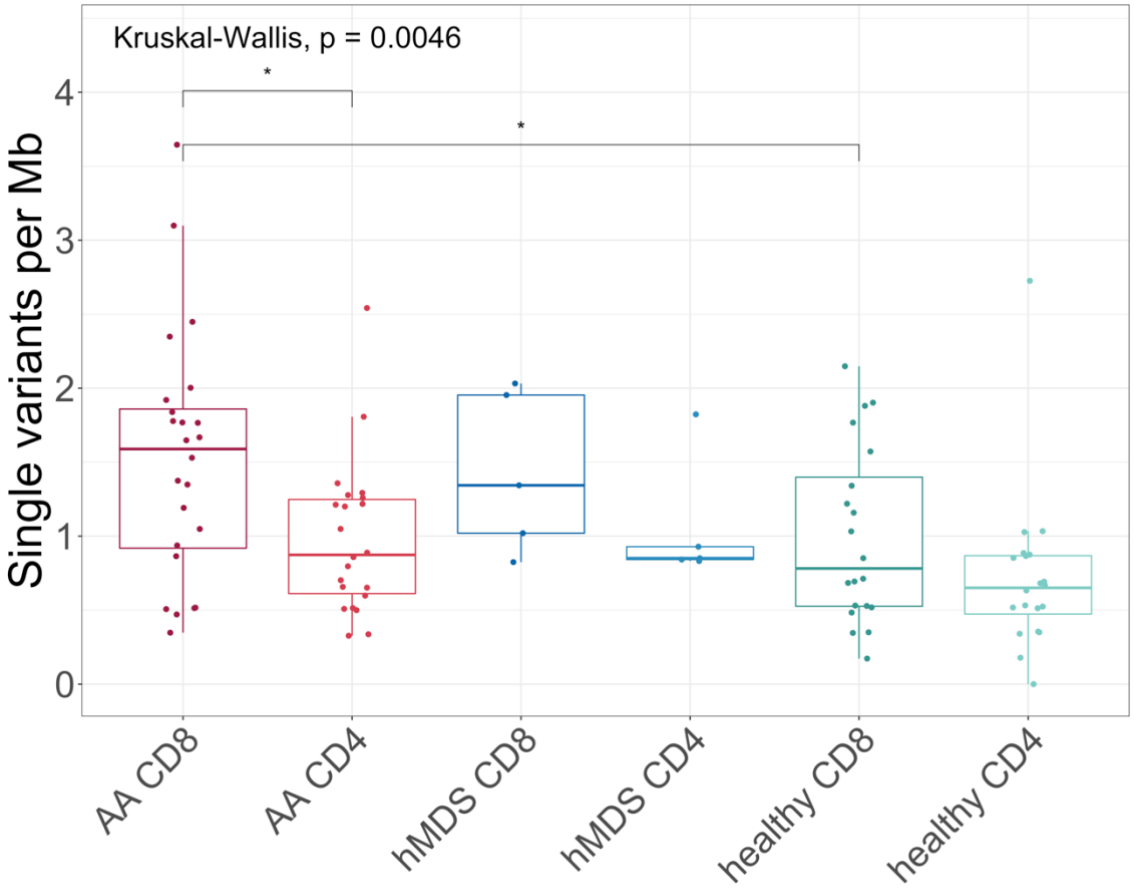


Figure 1. Adjusted somatic mutation burden. AA patient's cytotoxic T cells harbored most mutations per Mb. Only statistically significant differences are marked. Each dot is one patient.

3.2 Clonal T cell populations were associated with somatic mutation burden in CD8+ T cells

With TCRB sequencing, we defined T cell clonality index for each CD8+ T cell sample (see Methods) and compared the results with somatic mutations. Adjusted somatic mutation burden was correlated with CD8+ T cell clonality (Spearman ρ = 0.47, p = 0.005) but not with age (p = ns) (Figure 2A and B). When patient groups were analyzed individually, the correlation was seen in AA CD8+ T cells, although statistical significance was not achieved due to low number of samples (Figure 2A).

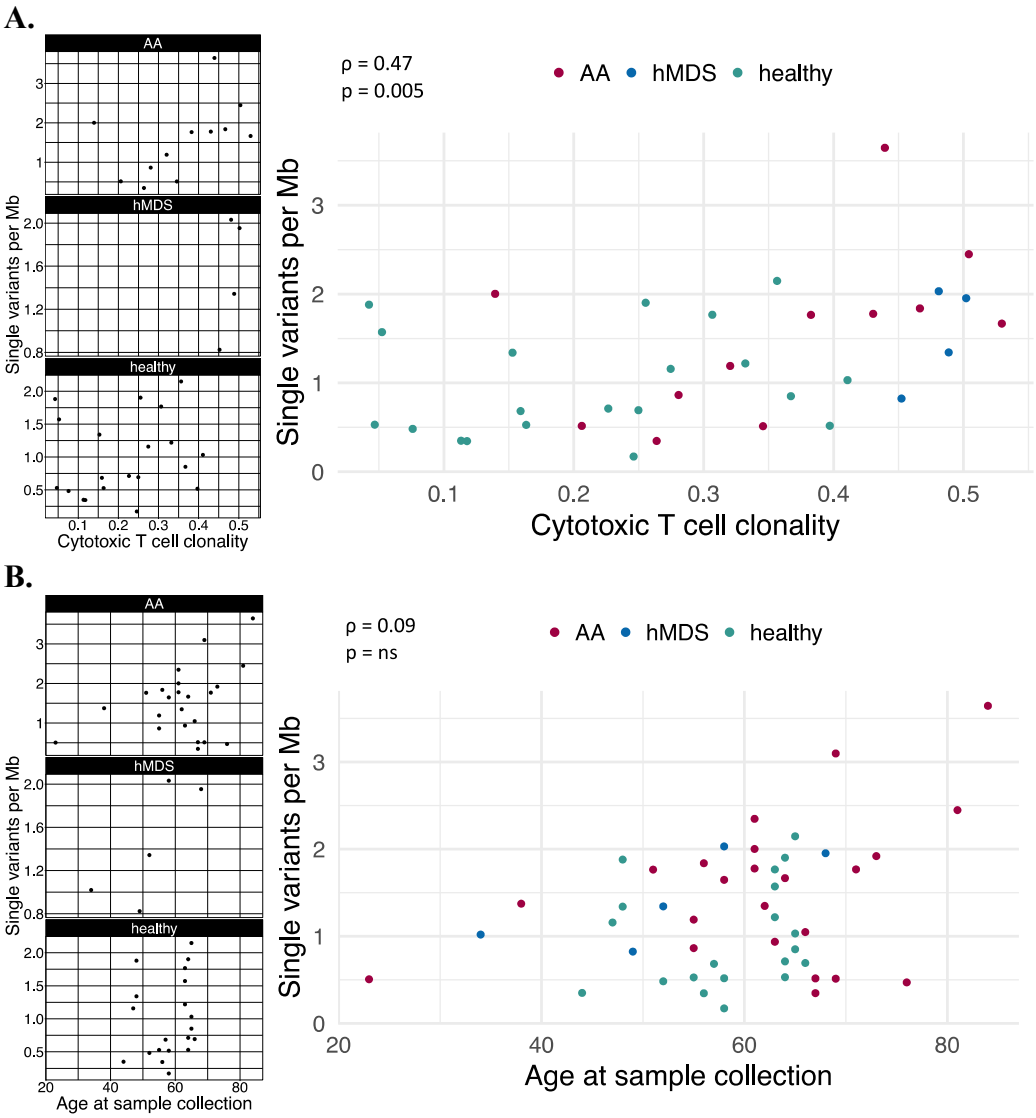


Figure 2A. Mutation burden in CD8+ T cells was associated with T cell clonality (Spearman test value ρ = 0.47, p = 0.005). Each group is shown separately on the left panel (Spearman test values for AA: ρ = 0.48, p = 0.11; hMDS: ρ = 0.4, p = 0.75; healthy: ρ = 0.21, p = 0.37). **B.** On the contrary, number of somatic mutations in CD8+ T cells did not have correlation with age at sample collection.

3.3 Somatic mutations of CD8+ T cells in AA accumulated into key immune regulatory pathways

The complete list of mutations is shown in the supplementary table S1. CD8+ T cells in AA harboured most mutations in the JAK-STAT signaling and MAP kinase signaling pathways (Figure 3 and S4A). Mutations in these pathways were also detected in AA CD4+ T cells and hMDS patients.

Mutations in JAK-STAT signaling pathway were detected in 18/24 (75 %) AA patients. In CD8+ T cells in AA patients, *JAK1* and *STAT3* were recurrently mutated (both 8.3 %, 2/24 AA patients' CD8+ T cells). *JAK1* mutation was also detected in 4/24 AA patients' CD4+ T cells, in hMDS-5 CD8+ T cells and in one healthy control (HC-4) CD8+ T cells, while *STAT3* mutations were not detected in AA patient's CD4+ T cells, hMDS nor in healthy controls. Of note, all *JAK1* mutations were located in homopolymer regions (Figure 3, homopolymer regions marked with white dots), which are challenging for reliable NGS analysis. Both *STAT3* mutations were p.Y640F, which is a mutation hotspot in *STAT3* exon21.

We predicted the pathogenicity of the mutations as described in the Methods. From JAK-STAT pathway, genes with pathogenic-predicted mutations included *CNTFR, CREBBP, EP300, IFNG, IL11RA, IL2, JAK3, PTPN11, PTPN2, SOS2* and *STAT3*. We discovered only one pathogenic-predicted JAK-STAT pathway mutation per sample in AA and hMDS CD4+ or CD8+ T cells.

Recurrently mutated genes detected on MAP kinase signaling pathway included *CACNA1A, ARRB2, CACNA1E, CACNG4, FLNB, NLK, PPP3CC* and *TGFBR2*.All detected variants on MAPK pathway genes are presented in Figure S4A.

A.

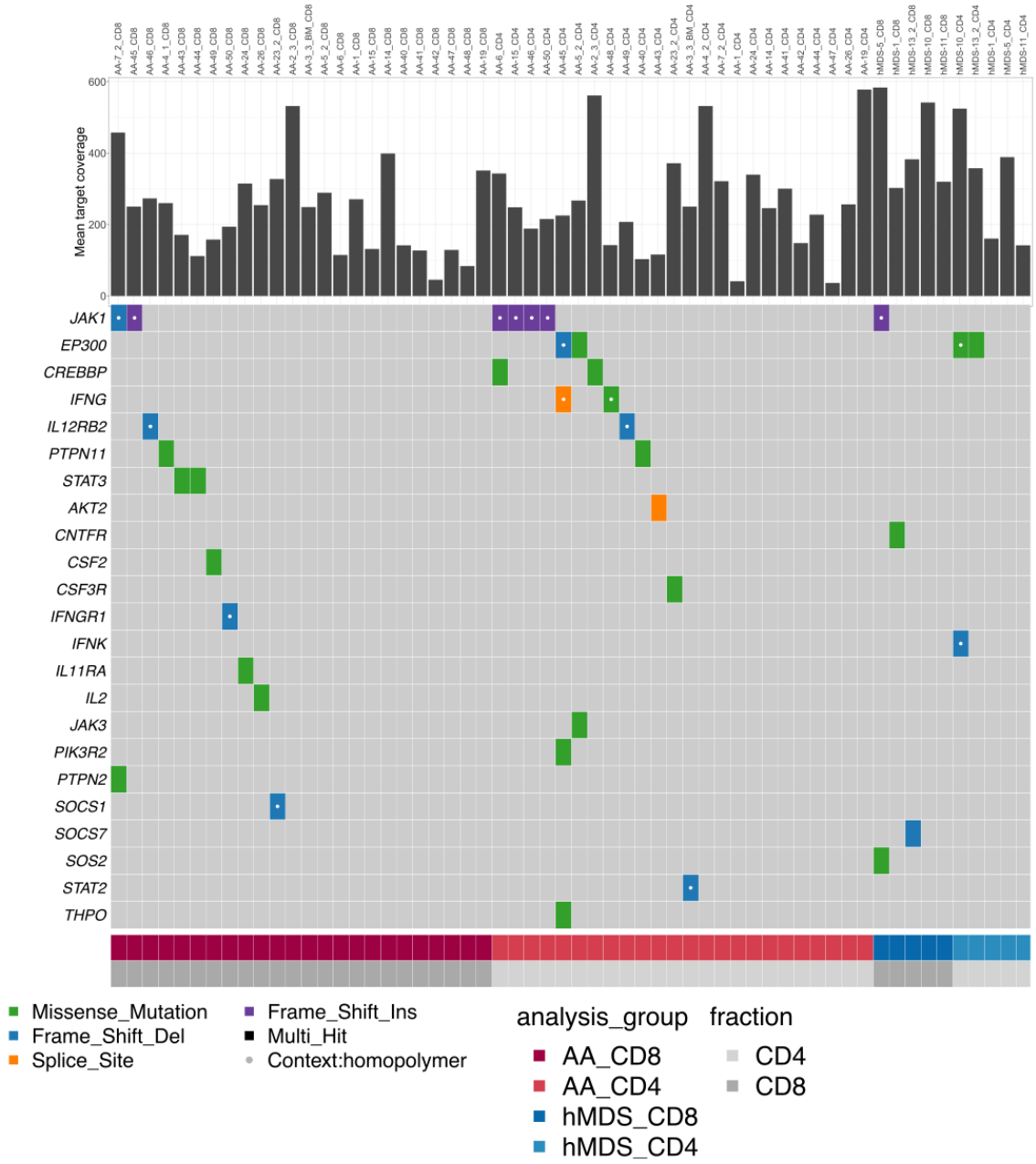


Figure 3 Somatic mutations in JAK-STAT pathway genes in AA and hMDS patients. Mean target coverage of samples is presented in the top panel. Homopolymer regions are marked with white dots.

To investigate possible mechanisms inducing somatic mutations in T cells, we performed mutational signature analysis of the variants. We included both synonymous and non-synonymous single-nucleotide variants (SNVs) in the analysis and pooled together the variants from CD4+ and CD8+ T cells in each group. Age-related signature 1 and defective mismatch-repair signature 15 accounted for a majority of mutations in

AA patients (Figure 4). Also signature 29, which has has been linked to guanine damage, was detected in AA T cells. In addition to these, we detected signature 8 (etiology unknown) and adenine damage-related signature 22 in hMDS patients. Signatures 1 and 15 were seen in healthy controls.

When looking at SNV types across different groups, most frequent alteration was C>T, which is a characteristic transversion type of replication as well as aging-associated mutagenesis Figure S3B). Most non-silent mutations were missense mutations (Figure S3C). There were no significant differences between groups (chi-squared test p-value = ns).

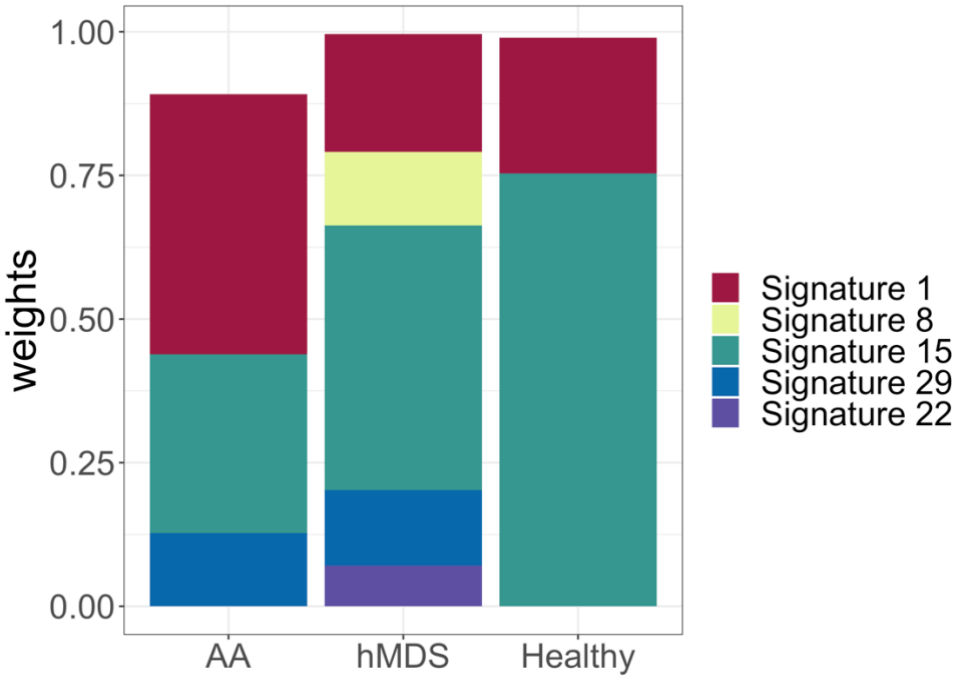


Figure 4. Mutational signatures from pooled CD4+ and CD8+ T cell variants.

3.4 Index patients

To understand the functional impact of the variants and clonal T cell expansions, we performed paired single-cell RNA and TCR αβ chain sequencing analysis on two index patients with mutations on MAPK and JAK-STAT signaling pathways. Patient AA-4 was a previously healthy female at the age of 52. She was diagnosed with very severe AA 17 months after being hospitalized for H1N1-influenza. She was first treated only with cyclosporine A (CsA) and corticosteroids (CS) with no response, and then two months later with equine anti-thymocyte globulin (ATG). She got a response to ATG with complete normalization of blood counts within four months. First sample was taken just before ATG treatment and follow-up samples at 5 and 37 months. At last sampling the patient had been without any immune-suppressive treatment with normal blood counts for 25 months. Patient AA-3 was a previously healthy female, at the age of 58. She was diagnosed with severe AA after two years of macrocytosis and mild thrombocytopenia. The initial treatment only CsA and CS with no response, and by 3 months after the diagnosis she was treated with equine ATG-based immunosuppresion. No response was observed by six months, and rabbit ATG-based immunosuppressive therapy was repeated. A year later hemoglobin was normal, and thrombocytes and neutrophils steadily recovering. However, the blood counts declined during the second year. The patient achieved second remission following the re-initiation of cyclosporine and prednisone and has since been on constant low-dose immunosuppression with moderate thrombocytopenia. The first sample was taken before the first ATG, and the second sample before the second ATG (refractory situation). The third sample was taken before the retreatment at the first relapse.

We found somatic *STAT3* mutation p.Y640F in CD8+ cells of AA-4, which was confirmed to be restricted in one CD8+ T cell clone (Vbeta 5.1) with flow cytometry sorting and amplicon sequencing. The variant was not detected in Vbeta 5.1 negative cells nor Vbeta 14 positive cells, corresponding to other expanded T cell TCRBV05-01 clone at diagnosis (Figure 5B). The variant allele frequency in sorted Vbeta 5.1 clone was 46% indicating heterozygosity, and the variant allele frequency of the mutation detected with immunogene panel sequencing decreased from 16% to 1% corresponding to shrinkage of expanded T cell clone TCRVB05-01 with TCRB sequencing after successful immunosuppression (Figure 5A).

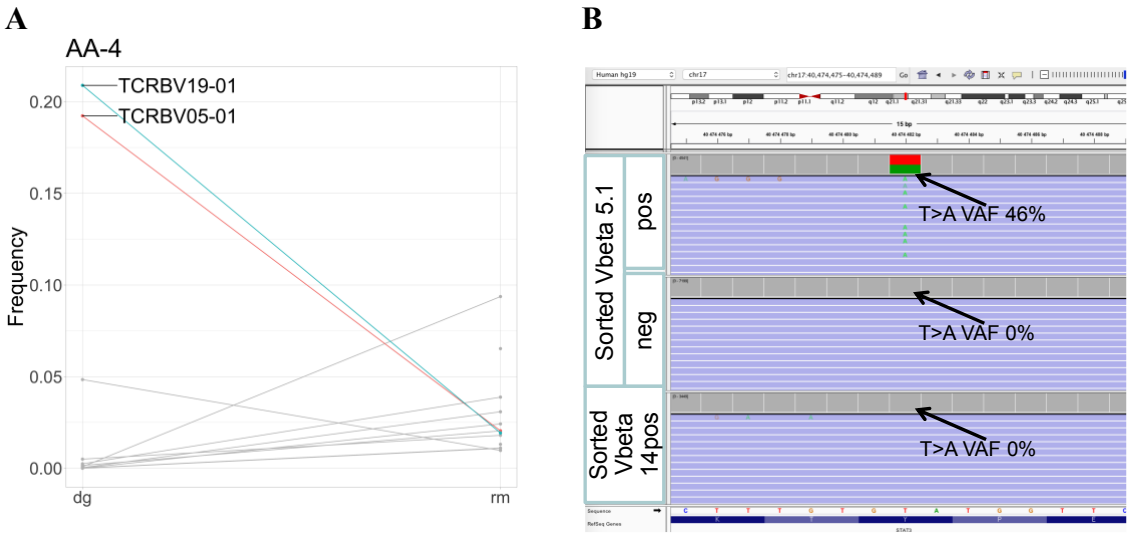
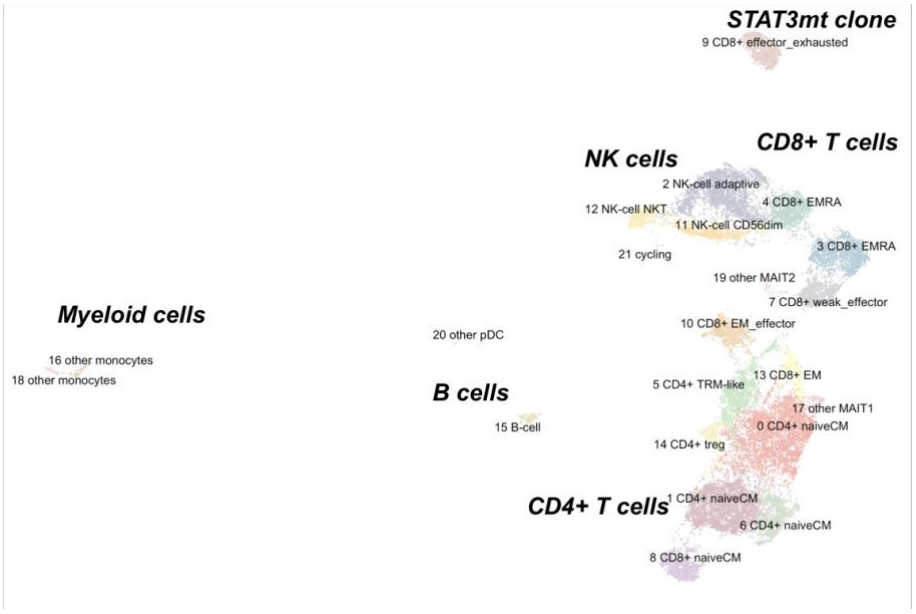


Figure 5 A. Index patient AA-4. Most expanded CD8+ T cell clones of AA-4 at diagnosis (dg) and remission (rm), based on TCRb sequencing. On the x axis are different time points and y axis shows the frequency of each TCRb clone from whole CD8+ TCRb repertoire. At diagnosis, there were two ~20% CD8+ T cell clones, which significantly diminished during immunosuppressive treatment. **B.** Somatic *STAT3* p.Y640F mutation was validated with amplicon sequencing of sorted Vbeta5.1pos fraction, but the mutation was not detected in Vbeta14pos fraction nor Vbeta5.1neg fraction. Picture of BAM files opened at variant site (chr17:42322464) in integrative Genomics Viewer.

In the single-cell analysis from this patient's CD45+ lymphocyte cells sampled from peripheral blood from three time points, we identified 21 different clusters including 7 CD8+ T cell clusters and 5 CD4+ T clusters. Of interest, almost all cells from Vbeta 5.1 clone clustered as its own cluster, named 9 CD8+ effector / exhausted cluster (Figure 6A). We compared the Vbeta 5.1 cluster against other CD8+ clusters, and the upregulated genes included cytotoxic genes (such as *PRF1, NKG7, DUSP1*), genes associated with T cell exhaustion (*LAG3, TIGIT*) and cytokines included in cell migration (*CCL3, CCL4, CCL3L1* and *CCL4L1*, Figure 6B).

A.



B.

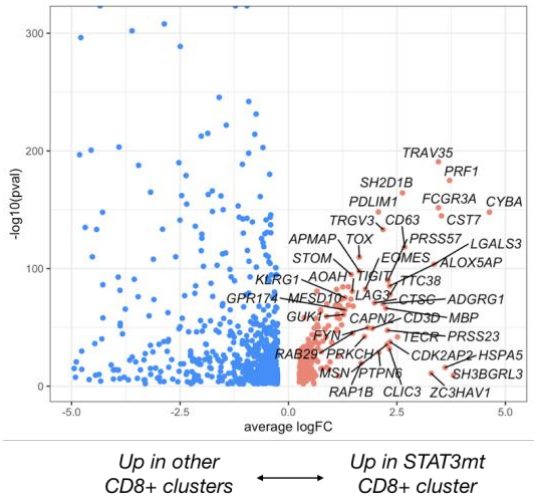


Figure 6. AA-4 scRNAseq analysis. A. Two-dimension UMAP-projection of clustered CD45+ lymphocyte cells pooled from three time points from peripheral blood. A total of 24 000 cells are annotated in 21 distinct clusters, seven of which that can be annotated as CD8+ T cells. **B**. Volcano plot showing differentially expressed genes between the Vbeta 5.1 associated cluster (to the right, significant shown in red) against other CD8+ clusters (to the left, significant shown in blue).

At baseline, the Vbeta 5.1 associated cluster expressed the highest level of *STAT3* and genes associated with 4kb of its binding site (Figure 7A and B). After treatment initiation, the Vbeta 5.1. cluster diminished significantly which is in line with the decrease of the *STAT3* VAF (Figure 7C). The therapy did also change the phenotype of the cluster, as it lost its expression of multiple cytotoxic genes, including *PRF1, GZMB* and *GNLY* (Figure 7D).

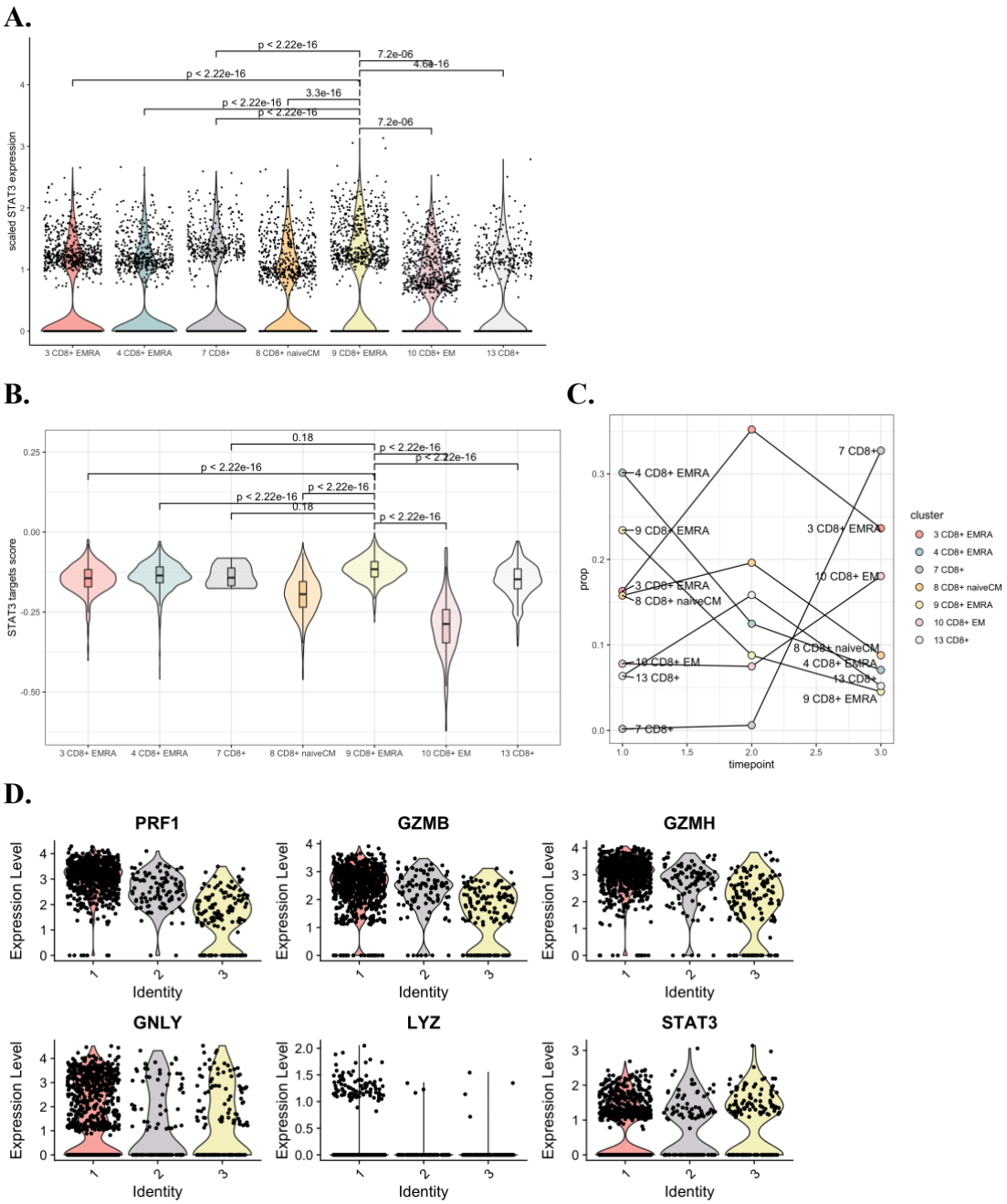


Figure 7A. Violin plot showing the scaled *STAT3* expression in CD8+ clusters at baseline **B.** Violin plot showing the score of genes having at least one occurrence of the transcription factor binding site within 4kb of *STAT3* binding site at baseline **C.** The relative abundances of CD8+ clusters at different timepoints. **D.** Violin plot showing the significantly downregulated cytotoxic genes (PRF1, GZMB, GZMH, GNLY, LYZ) and significantly downregulated *STAT3*.

Patient AA-3 had somatic mutations in several genes, including *KRAS, NFATC2, PTPN22, TNFAIP3* and *JAGN1*s. *KRAS* A146P, *JAGN1* T473A and *TNFAIP3* D212fs mutations were predicted to be pathogenic, and same *KRAS* mutation has been found in several hematopoietic neoplasms (genomic mutation ID in COSMIC v91: COSV55541748). As in AA-4, these mutations were confirmed to be restricted to single

CD8+ T cell clone (Vbeta 7.2, corresponding to TCRBV0403), as the variants were not detected in sorted monocytes, B cells nor Vbeta 7.2 negative T cells. Vbeta7.2 clone was increasing despite immunosuppressive treatment (Figure 8A). Variant allele frequencies in sorted Vbeta 7.2 clone at two different time points are shown in Figure 8B. Based on amplicon sequencing data, possible subclonal architecture of the mutations is presented in Figure 8C.

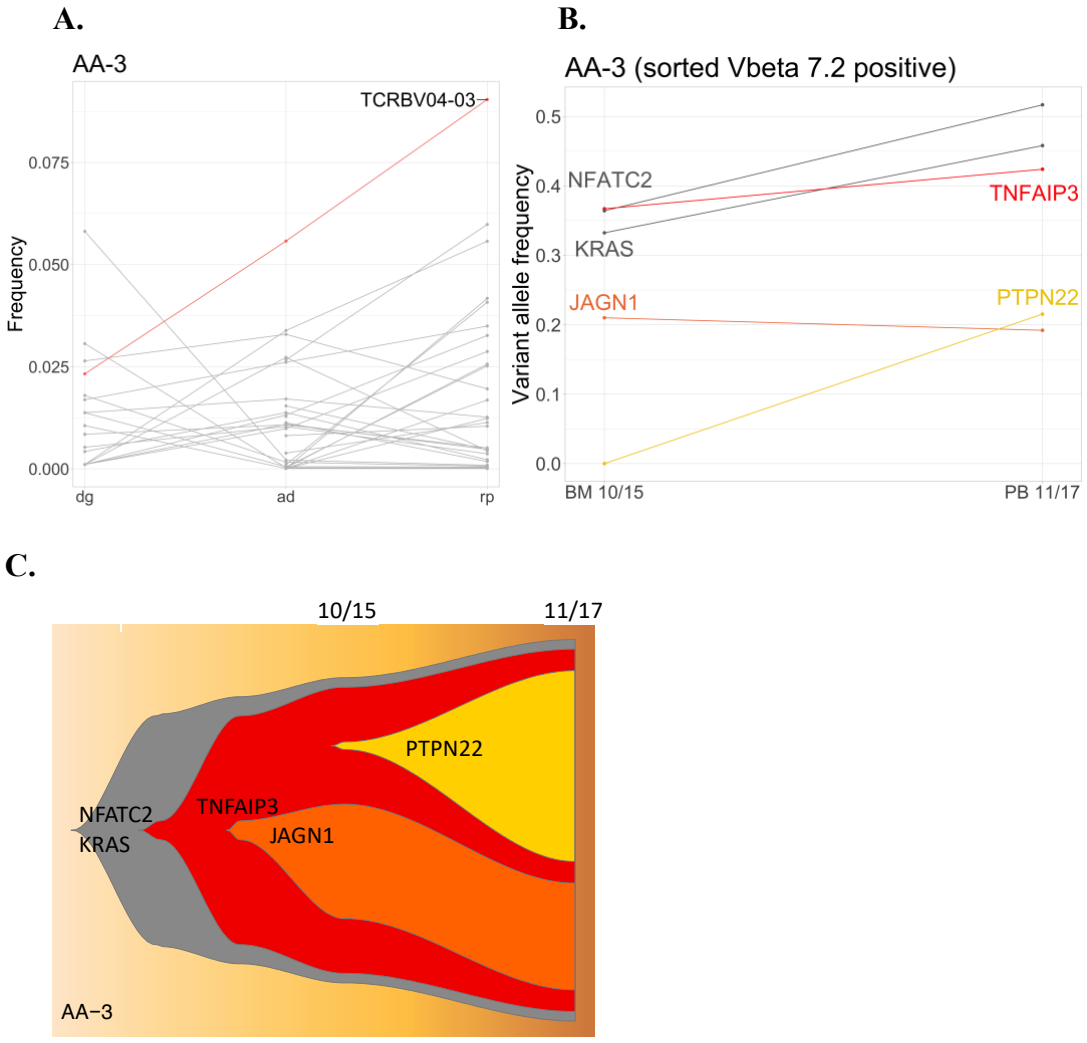
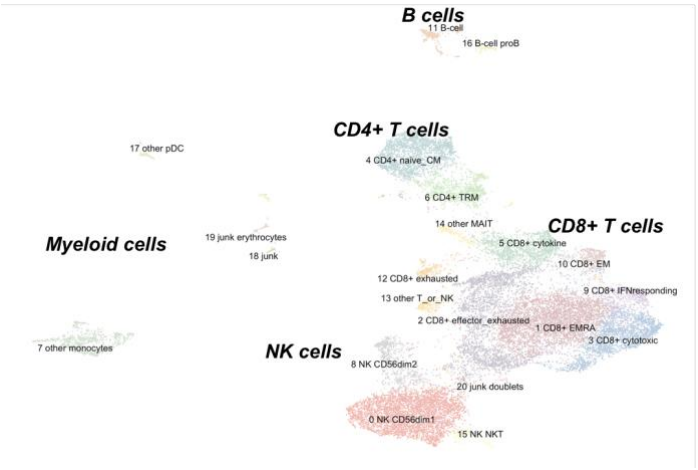


Figure 8A. In patient AA-3, TCRb sequencing revealed that TCRBV04-03 clone was expanding from diagnosis to relapse. Y axis shows frequency of top clones in CD8+ T cells. Dg = diagnosis, ad = active disease, rp = relapse. **B.** Amplicon sequencing confirmed all mutations in Vbeta7.2pos fraction of CD8+ T cells, corresponding to TCRBV04-03 clone. Y axis shows variant allele frequency of variants in sample 3 and 4 (shown at x axis). **C.** Possible subclonal architecture of patient AA-3, based on variant allele frequencies from amplicon sequencing.

In the single-cell analysis from this patient's CD45+ lymphocyte cells sampled from the bone-marrow at three time points, we identified 21 different clusters including 7 CD8+

T cell clusters but only 2 CD4+ T clusters (Figure 9A). We identified 350 cells with matching CDR3 nucleotide sequence of the Vbeta 7.2 clone. Most of the cells from the clone were of CD8+ terminally differentiated effector memory cells (TEMRA), supporting the cytotoxic role of the Vbeta 7.2 clone (Figure 9B).

A.



B.

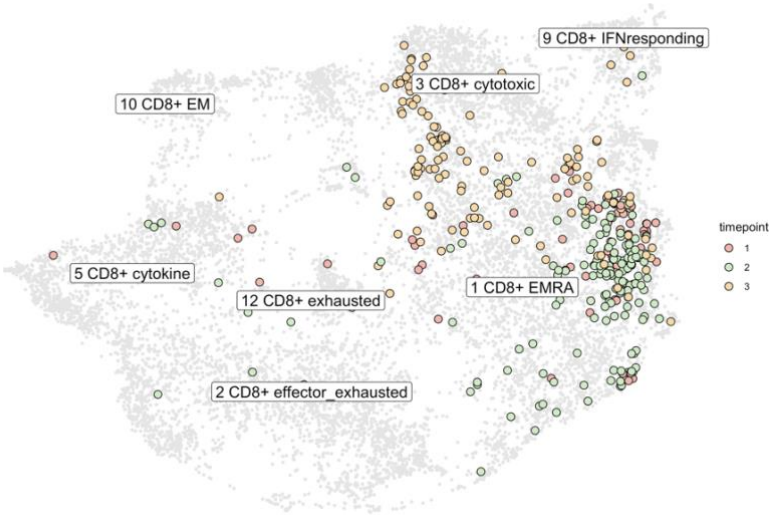


Figure 9 AA-3 scRNAseq analysis. A. Two-dimension UMAP-projection of clustered CD45+ lymphocyte cells pooled from three time points from bone marrow. A total of 22 407 cells are annotated in 21 distinct clusters, seven of which that can be annotated as CD8+ T cells but only two as CD4+ T cells. **B.** Focused two-dimension UMAP-projection of re-embedded CD8+ T cells. The encircled dots are the Vbeta 7.2 clone colored with different time points, where red is the diagnostic, green is after the treatment and yellow is from relapse. Most of the Vbeta 7.2. clonotype cells can be classified as 1 CD8+ TEMRA cells.

Interestingly, the Vbeta 7.2. clone increased after the CsA, CS and repeated ATG administration, and it was highest in the relapse sample, explaining 4.8% of the CD8+ repertoire (Figure 10A). Also, the phenotype of the Vbeta 7.2 clone was slightly altered at relapse (Figure 10B).

A. **B.**

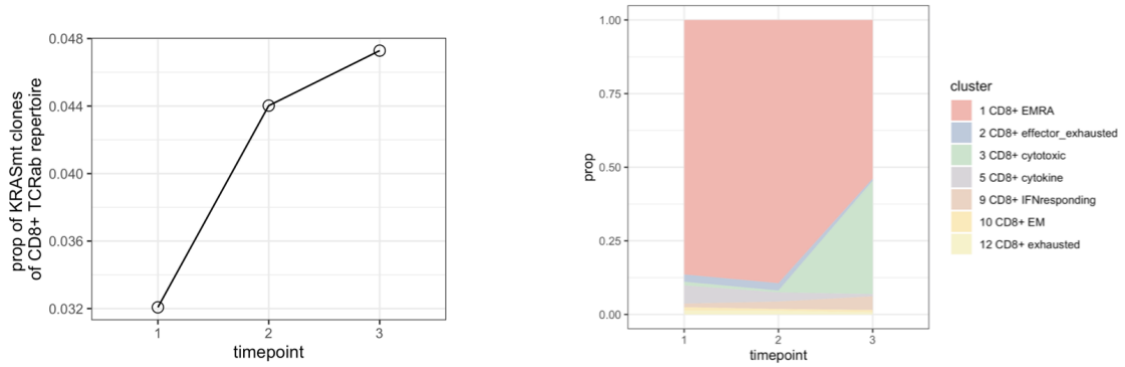


Figure 10A. The proportion of Vbeta 7.2 clonotype cells from the total CD8+ repertoire from the single-cell TCRab data. **B.** The relative abundances of Vbeta 7.2 clones phenotype at sample 1, 2 and 3.

To understand the effect of the somatic mutations on the phenotype of the Vbeta 7.2 clonotype, we performed differential expression (DE) analysis between the TEMRA cells, comparing the cells from the Vbeta 7.2 clonotype against the other cells in TEMRA cluster. In the relapse sample, the amount of differentially expressed genes between the Vbeta 7.2 clonotype in comparison to other TEMRA cells was the highest (56 genes up in Vbeta 7.2 clonotype and 77 genes up in other TEMRA cells) including several genes associated with immune respose like *STAT4, NFKB1, IFNGR1* upregulated in the Vbeta 7.2 clonotype (Figure S5A). In the gene set enrichment analysis (GSEA) the upregulated pathways included the type I interferon production, NFkappaB signaling and lymphocyte activation (Figure S5B).

3.5 Clonal hematopoiesis mutations transmit to T cells

We were interested in the origin of somatic mutations in T cells, and whether they stem from lympho-myeloid clonal hematopoiesis (LM-CH). We looked at mutations occurring concurrently both in CD4+ and CD8+ T cells and compared the results to previously published exome sequencing data. In our data, we found altogether 24 somatic variants in 14 AA and 2 hMDS patients occurring concurrently in both CD4+ and CD8+ T cells (Figure 11). With the exception of *NF1*, mutated genes have not been previously linked to clonal hematopoiesis. However, CD4+ and CD8+ T cells harbored many mutations in genes linked to clonal hematopoiesis (e.g. *DNMT3A, BCORL1, TERT* and *LAMB4*), without concurrent detection in both T cell fractions (Figure S4B).

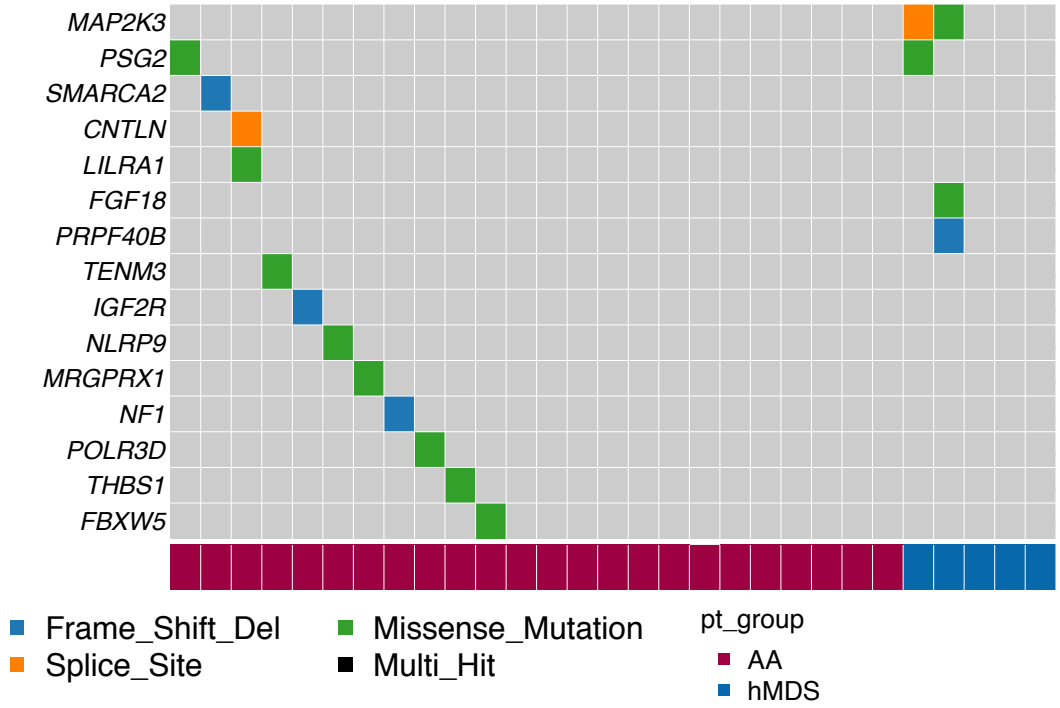


Figure 11. LM-CH variants in AA and hMDS patients

When taking a closer look at genes linked to clonal hematopoiesis in reanalyzed WES data from Yoshizato et al (8), we could see mutation transmission to T cells in several genes. With exome sequencing of CD3+ and CD3- MNCs, most frequently found mutations in CD3- fraction were *DNMT3A, BCOR/BCORL1* and *ASXL1* mutations (mutations in CD3- fraction were found in 7, 6 and 6 patients, respectively). In some cases, these mutations were seen in CD3+ cells of same patient, but also exclusively in CD3+ cells (*BCOR* and *BCORL1*, Figure S6). We did not detect *ASXL1* or *DNMT3A* mutations in CD3+ cells without concurrent emergence in CD3- cells, possibly due to their clonal advantage in myeloid cells. *JAK* mutations (n=2) were found in both CD3- and CD3+ fractions. One *STAT*3 mutation (c.G1012T, p.V338F) was detected in CD3- MNC with 3.3% VAF. This particular mutation was not found in COSMIC v90 (24).

4 Discussion

Accumulation of somatic mutations is inherently associated with normal cell division and aging. Mutations may lead to abnormal cell function and uncontrolled cell proliferation, which has been well established in cancer. Recently, it has also been shown that somatic mutations occur in various healthy human tissues (41-44) and play a role in the pathogenesis of other non-malignant diseases (45,46). Little is known about the mechanisms by which somatic mutations alter our immunity.

Somatic mutations in certain genes have been associated with increased risk of cardiovascular disease and this link has been explained by inflammatory mechanisms (47-49). It has been shown that CH may be related to enhanced adaptive immune response in bone marrow transplant recipients, as patients with CH have lower relapse rate but higher rates for graft-versus-host disease (50). Furthermore, Fraietta et al showed that a TET2-deficient CAR-T cell clone had reduced T cell exhaustion in response to stimulus and hence killed tumor cells more efficiently (51).

With gene panel sequencing of T cells, we characterized somatic variants in healthy individuals and iBMF syndrome patients. We found that somatic variants in immune-related genes are strikingly common in T cells, which emphasizes unresolved question of the role of somatic mutations in non-malignant cells of adaptive immune system. In the current paper, we are addressing the key questions of somatic mutations in T cells: in which subset of the cells they occur, at which state of immune cell development they emerge, which genes and pathways they hit and finally, do somatic mutations have functional impact on T cells.

We found that CD8+ T cells have higher mutation burden compared to CD4+ T cells. This observation is in line with previous findings (17-19) and can result from both technical and biological reasons. CD8+ T cells have been shown to form larger clonal expansions (52), especially following strong antigen stimulus like EBV and CMV infections (53-55). Rapid proliferation makes cells more susceptible to mutations. On the other hand, larger clones have higher variant allele frequencies, which make mutations more easily detectable with NGS. Notably, we showed association between clonality and mutation burden within AA patients' CD8+ T cells and all samples together, but not in healthy CD8+ T cells alone. This observation implies that better technical detectability could not alone explain higher mutation rate in clonal T cells.

In index patients as well as our recent papers using similar technology (17,18), somatic mutations were restricted to single T cell clone, indicating post-thymic emergence. However, here we also showed that mutations in genes related to CH were common in lymphocytes, and CH mutation transmission to lymphocytes was frequently seen in previously published data on CH. The vast majority of publications on CH are based on whole MNC sequencing, which limits comparison to our findings. Sequencing data from myeloid cells would be needed to authenticate origin of T cell mutations, and studies on divergence of mutations on lymphoid and myeloid cells would elucidate the mechanisms of clonal advantage in different immune cell lineages.

In T cells from patients with immune-mediated BMF, we found that somatic mutations accumulated to JAK-STAT and MAP kinase pathways. These pathways are highly interesting in T cell biology: JAK-STAT pathway plays a central in cytokine signaling, transmitting activation as well as proliferation signals in T cells. MAPK pathway is one of major pathways induced by TCR stimulation, and it also regulates apoptosis. However, to enable adequate coverage to detect low allele fraction mutations, we had pre-selected genes to our panel. This precludes us from drawing conclusions on whole somatic mutation spectrum of T cells. Analysis of all human genes with whole exome or whole genome sequencing in large cohorts would be required to authenticate the enrichment of T cell mutations to immune-related pathways.

We studied functional impact of T cell somatic mutations with single-cell sequencing and showed that *STAT3* p.Y640F mutated CD8+ T cells had higher expression of *STAT3* target genes. This is consistent with previous studies showing *STAT3* activation following p.Y640F mutation in LGL leukemia, anaplastic large cell lymphoma and inflammatory hepatocellular adenomas (21,56,57). *STAT3* has also been shown to mediate effector T cell resistance to suppression in patients with type 1 diabetes (58).

Finally, we compared T cell repertoire and mutated T cell clones between two patients with different clinical outcomes. Our single-cell resolution data confirmed the central role of CD8+ T cells and cytotoxic properties of mutated CD8+ T cell clones in 2 AA patients. We showed that immunosuppressive treatment may alter not only the clone size but also the cytotoxic function of mutated CD8+ T cells, and that these changes could be linked to patients' clinical outcomes. To our knowledge, this kind of longitudinal data on the effect of immunosuppression on T cell repertoire and

phenotype has not yet been published. More studies need to be done to understand the normal and abnormal immune cell responses following these treatments, and to conclude if somatic mutations can explain different clinical outcomes.

5 Author contributions

This thesis is based on manuscript in preparation, to which many people have contributed. Sofie Lundgren (S.L), Satu Mustjoki[1,2] (S.M.) and Mikko A.I. Keränen[1,2,3] (M.A.I.K.) designed the study and experiments. S.L. prepared biological samples for immunogene panel sequencing and sorted cells for single-cell RNA sequencing. S.L. and Hanna Rajala[1,2] sorted cells and extracted DNA for amplicon sequencing. Matti Kankainen[1,2,4] (M.K.) was responsible for somatic variant calling pipeline. Somatic variant filtering was done by M.K and S.L. S.L. analyzed somatic variant data with the help of M.A.I.K, M.K. and S.M. Jani Huuhtanen[1,2,5] (J.H.) performed the single-cell analysis with the help of S.L. Clinical data was provided by M.A.I.K, Gunilla Walldin[6] and Cassandra Hirsch[7]. S.L. wrote the thesis with the help of J.H, M.A.I.K. and S.M.

6 Acknowledgements

I cordially thank my supervisors Satu Mustjoki and Mikko Keränen for endless patience and support during this project. Huge thanks belong also to Matti Kankainen for optimizing pipeline for somatic variant calling and Jani Huuhtanen for beautiful single-cell analyses. The research was funded with research grants from the European Research Council (M-IMM project), Academy of Finland, Sigrid Juselius Foundation, Cancer Foundation and Intrumentarium.

[1] Hematology Research Unit Helsinki, University of Helsinki and Helsinki University Hospital Comprehensive Cancer Center, Helsinki, Finland
[2] Translational Immunology Research program and Department of Clinical Chemistry and Hematology
[3] Department of Hematology, Helsinki University Hospital Comprehensive Cancer Center, Helsinki, Finland;
[4] Medical and Clinical Genetics, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[5] Department of Computer Science, Aalto University School of Science, Finland
[6] Center for Hematology and Regenerative Medicine, Department of Medicine, Karolinska Institutet, Karolinska University Hospital, Stockholm
[7] Translational Hematology and Oncology Department, Taussig Cancer Center, Cleveland Clinic, Cleveland, OH, USA.

References

1. Young NS, Calado RT, Scheinberg P. Current concepts in the pathophysiology and treatment of aplastic anemia. Blood. American Society of Hematology; 2006 Oct 15;108(8):2509–19.

2. Luzzatto L, Risitano AM. Advances in understanding the pathogenesis of acquired aplastic anaemia. Br J Haematol. John Wiley & Sons, Ltd; 2018 Sep;182(6):758–76.

3. Mortazavi Y, Chopra R, Gordon-Smith EC, Rutherford TR. Clonal patterns of X-chromosome inactivation in female patients with aplastic anaemia studies using a novel reverse transcription polymerase chain reaction method. Eur J Haematol. John Wiley & Sons, Ltd; 2000 Jun;64(6):385–95.

4. Afable MG, Wlodarski M, Makishima H, Shaik M, Sekeres MA, Tiu RV, et al. SNP array-based karyotyping: differences and similarities between aplastic anemia and hypocellular myelodysplastic syndromes. Blood. 2011 Jun 23;117(25):6876–84.

5. Mikhailova N, Sessarego M, Fugazza G, Caimo A, De Filippi S, van Lint MT, et al. Cytogenetic abnormalities in patients with severe aplastic anemia. Haematologica. 1996 Sep;81(5):418–22.

6. Katagiri T, Sato-Otsubo A, Kashiwase K, Morishima S, Sato Y, Mori Y, et al. Frequent loss of HLA alleles associated with copy number-neutral 6pLOH in acquired aplastic anemia. Blood [Internet]. American Society of Hematology; 2011 Dec 15;118(25):6601–9. Available from: http://www.bloodjournal.org/content/118/25/6601.abstract

7. Tichelli A, Gratwohl A, Würsch A, Nissen C, Speck B. Late haematological complications in severe aplastic anaemia. Br J Haematol. John Wiley & Sons, Ltd; 1988 Jul;69(3):413–8.

8. Yoshizato T, Dumitriu B, Hosokawa K, Makishima H, Yoshida K, Townsley D, et al. Somatic Mutations and Clonal Hematopoiesis in Aplastic Anemia. N Engl J Med. 2015 Jul 2;373(1):35–47.

9. Ogawa S. Clonal hematopoiesis in acquired aplastic anemia. Blood. American Society of Hematology; 2016 Jul 21;128(3):337–47.

10. Shlush LI, Zandi S, Mitchell A, Chen WC, Brandwein JM, Gupta V, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature. 2014 Feb 20;506(7488):328–33.

11. Thol F, Klesse S, Köhler L, Gabdoulline R, Kloos A, Liebich A, et al. Acute myeloid leukemia derived from lympho-myeloid clonal hematopoiesis. Leukemia. Nature Publishing Group; 2017 Jun;31(6):1286–95.

12. Risitano AM, Maciejewski JP, Green S, Plasilova M, Zeng W, Young NS. In-vivo dominant immune responses in aplastic anaemia: molecular tracking of putatively pathogenetic T-cell clones by TCR beta-CDR3 sequencing. Lancet. 2004 Jul;364(9431):355–64.

13. Giudice V, Feng X, Lin Z, Hu W, Zhang F, Qiao W, et al. Deep sequencing and flow cytometric characterization of expanded effector memory CD8+CD57+ T cells frequently reveals T-cell receptor Vβ oligoclonality and CDR3 homology in acquired aplastic anemia. Haematologica. Haematologica; 2018 May 1;103(5):759–69.

14. Li B, Guo L, Zhang Y, Xiao Y, Wu M, Zhou L, et al. Molecular alterations in the TCR signaling pathway in patients with aplastic anemia. J Hematol Oncol. BioMed Central; 2016 Mar 31;9(1):32–9.

15. Li B, Liu S, Niu Y, Fang S, Wu X, Yu Z, et al. Altered expression of the TCR signaling related genes CD3 and FcεRIγ in patients with aplastic anemia. J Hematol Oncol. BioMed Central; 2012 Mar 8;5(1):6–7.

16. Sheng W, Liu C, Fu R, Wang H, Qu W, Ruan E, et al. Abnormalities of quantities and functions of linker for activations of T cells in severe aplastic anemia. Eur J Haematol. John Wiley & Sons, Ltd; 2014 Sep;93(3):214–23.

17. Savola P, Kelkka T, Rajala HL, Kuuliala A, Kuuliala K, Eldfors S, et al. Somatic mutations in clonally expanded cytotoxic T lymphocytes in patients with newly diagnosed rheumatoid arthritis. Nat Commun. Nature Publishing Group; 2017 Jun 21;8:15869.

18. Savola P, Brück O, Olson T, Kelkka T, Kauppi MJ, Kovanen PE, et al. Somatic STAT3 mutations in Felty syndrome: an implication for a common pathogenesis with large granular lymphocyte leukemia. Haematologica. Haematologica; 2018 Feb;103(2):304–12.

19. Valori M, Jansson L, Kiviharju A, Ellonen P, Rajala H, Awad SA, et al. A novel class of somatic mutations in blood detected preferentially in CD8+ cells. Clin Immunol. 2017 Feb;175:75–81.

20. Kim D, Park G, Huuhtanen J, Lundgren S, Khajuria RK, Hurtado AM, et al. Somatic mTOR mutation in clonally expanded T lymphocytes associated with chronic graft versus host disease. Nat Commun. Nature Publishing Group; 2020 May 7;11(1):2246–17.

21. Koskela HLM, Eldfors S, Ellonen P, van Adrichem AJ, Kuusanmäki H, Andersson EI, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. N Engl J Med. 2012 May 17;366(20):1905–13.

22. Jerez A, Clemente MJ, Makishima H, Rajala H, Gómez-Seguí I, Olson T, et al. STAT3 mutations indicate the presence of subclinical T-cell clones in a subset of aplastic anemia and myelodysplastic syndrome patients. Blood. American Society of Hematology; 2013 Oct 3;122(14):2453–9.

23. Greenplate A, Wang K, Tripathi RM, Palma N, Ali SM, Stephens PJ, et al. Genomic Profiling of T-Cell Neoplasms Reveals Frequent JAK1 and JAK3 Mutations With Clonal Evasion From Targeted Therapies. JCO Precis Oncol. 2018;2018.

24. 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 Jan 8;47(D1):D941–7.

25. Adnan Awad S, Kankainen M, Ojala T, Koskenvesa P, Eldfors S, Ghimire B, et al. Mutation accumulation in cancer genes relates to nonoptimal outcome in chronic myeloid leukemia. Blood Adv. 2020 Feb 11;4(3):546–59.

26. Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res. 2012 Jul;40(Web Server issue):W452–7.

27. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. John Wiley & Sons, Ltd; 2013 Jan;Chapter 7(1):Unit7.20–7.20.41.

28. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome Res. Cold Spring Harbor Lab; 2009 Sep;19(9):1553–61.

29. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. Nature Publishing Group; 2014 Apr;11(4):361–2.

30. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res. 2011 Sep 1;39(17):e118.

31. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GLA, Edwards KJ, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat. John Wiley & Sons, Ltd; 2013 Jan;34(1):57–65.

32. Carter H, Douville C, Stenson PD, Cooper DN, Karchin R. Identifying Mendelian disease genes with the variant effect scoring tool. BMC Genomics. BioMed Central; 2013;14 Suppl 3(3):S3–16.

33. Douville C, Masica DL, Stenson PD, Cooper DN, Gygax DM, Kim R, et al. Assessing the Pathogenicity of Insertion and Deletion Variants with the Variant Effect Scoring Tool (VEST-Indel). Hum Mutat. John Wiley & Sons, Ltd; 2016 Jan;37(1):28–35.

34. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. Nature Publishing Group; 2014 Mar;46(3):310–5.

35. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. 2019 Jan 8;47(D1):D886–94.

36. Garber M, Guttman M, Clamp M, Zody MC, Friedman N, Xie X. Identifying novel constrained elements by exploiting biased substitution patterns. Bioinformatics. 2009 Jun 15;25(12):i54–62.

37. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. Cold Spring Harbor Lab; 2010 Jan;20(1):110–21.

38. Robins HS, Campregher PV, Srivastava SK, Wacher A, Turtle CJ, Kahsai O, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. Blood. 2009 Nov 5;114(19):4099–107.

39. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. Nat Methods. Nature Publishing Group; 2018 Dec;15(12):1053–8.

40. Soneson C, Robinson MD. Bias, robustness and scalability in single-cell differential expression analysis. Nat Methods. Nature Publishing Group; 2018 Apr;15(4):255–61.

41. Martincorena I, Campbell PJ. Somatic mutation in cancer and normal cells. Science. American Association for the Advancement of Science; 2015 Sep 25;349(6255):1483–9.

42. Martincorena I, Fowler JC, Wabik A, Lawson ARJ, Abascal F, Hall MWJ, et al. Somatic mutant clones colonize the human esophagus with age. Science. American Association for the Advancement of Science; 2018 Nov 23;362(6417):911–7.

43. Blokzijl F, de Ligt J, Jager M, Sasselli V, Roerink S, Sasaki N, et al. Tissue-specific mutation accumulation in human adult stem cells during life. Nature. 2016 Oct 3;538(7624):260–4.

44. Yokoyama A, Kakiuchi N, Yoshizato T, Nannya Y, Suzuki H, Takeuchi Y, et al. Age-related remodelling of oesophageal epithelia by mutated cancer drivers. Nature. Nature Publishing Group; 2019 Jan 1;565(7739):312–7.

45. Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P, et al. Universal Patterns of Selection in Cancer and Somatic Tissues. Cell. 2017 Oct 18.

46. Nikolaev SI, Vetiska S, Bonilla X, Boudreau E, Jauhiainen S, Rezai Jahromi B, et al. Somatic Activating KRAS Mutations in Arteriovenous Malformations of the Brain. N Engl J Med. 2018 Jan 18;378(3):250–61.

47. 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 Dec 25;371(26):2488–98.

48. 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 Jul 13;377(2):111–21.

49. 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. American Association for the Advancement of Science; 2017 Feb 24;355(6327):842–7.

50. Frick M, Chan W, Arends CM, Hablesreiter R, Halik A, Heuser M, et al. Role of Donor Clonal Hematopoiesis in Allogeneic Hematopoietic Stem-Cell Transplantation. J Clin Oncol. 2018 Nov 7;:JCO2018792184.

51. Fraietta JA, Nobles CL, Sammons MA, Lundh S, Carty SA, Reich TJ, et al. Disruption of TET2 promotes the therapeutic efficacy of CD19-targeted T cells. Nature. Nature Publishing Group; 2018 Jun;558(7709):307–12.

52. Qi Q, Liu Y, Cheng Y, Glanville J, Zhang D, Lee J-Y, et al. Diversity and clonal selection in the human T-cell repertoire. Proc Natl Acad Sci USA. National Academy of Sciences; 2014 Sep 9;111(36):13139–44.

53. Callan MF, Steven N, Krausa P, Wilson JD, Moss PA, Gillespie GM, et al. Large clonal expansions of CD8+ T cells in acute infectious mononucleosis. Nat Med. Nature Publishing Group; 1996 Aug;2(8):906–11.

54. Maini MK, Gudgeon N, Wedderburn LR, Rickinson AB, Beverley PC. Clonal expansions in acute EBV infection are detectable in the CD8 and not the CD4 subset and persist with a variable CD45 phenotype. J Immunol. American Association of Immunologists; 2000 Nov 15;165(10):5729–37.

55. Khan N, Shariff N, Cobbold M, Bruton R, Ainsworth JA, Sinclair AJ, et al. Cytomegalovirus seropositivity drives the CD8 T cell repertoire toward greater clonality

in healthy elderly individuals. J Immunol. American Association of Immunologists; 2002 Aug 15;169(4):1984–92.

56. Crescenzo R, Abate F, Lasorsa E, Tabbo F, Gaudiano M, Chiesa N, et al. Convergent mutations and kinase fusions lead to oncogenic STAT3 activation in anaplastic large cell lymphoma. Cancer Cell. 2015 Apr 13;27(4):516–32.

57. Pilati C, Amessou M, Bihl MP, Balabaud C, Nhieu JTV, Paradis V, et al. Somatic mutations activating STAT3 in human inflammatory hepatocellular adenomas. J Exp Med. 2011 Jul 4;208(7):1359–66.

58. Ihantola E-L, Viisanen T, Gazali AM, Näntö-Salonen K, Juutilainen A, Moilanen L, et al. Effector T Cell Resistance to Suppression and STAT3 Signaling during the Development of Human Type 1 Diabetes. J Immunol. American Association of Immunologists; 2018 Aug 15;201(4):1144–53.

Supplementary figures and tables

Attachment 1: Figures S1-S6

Attachment 2: Table S1 Patient characteristics

Attachment 3: Table S2 Somatic variants in CD4+ or CD8+ T cells with immunogene panel sequencing

Attachment 4: Table S3 Amplicon validation summary

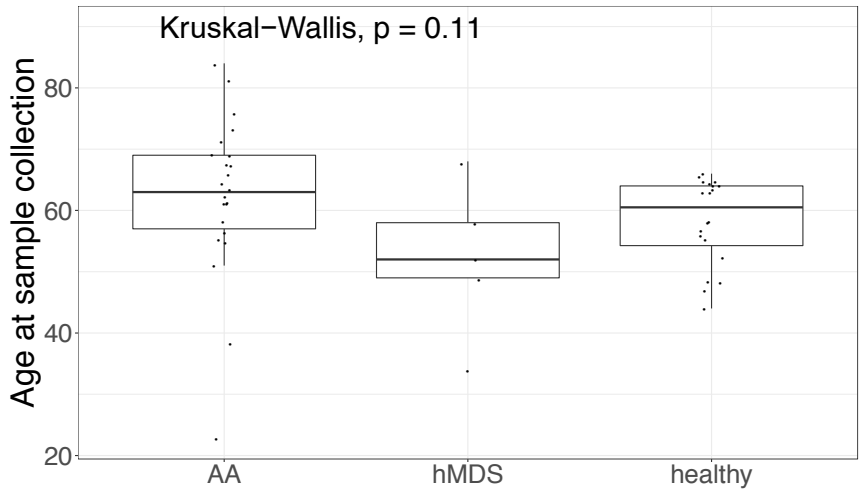


Figure S1. Age in cohorts analysed with immunogene panel sequencing

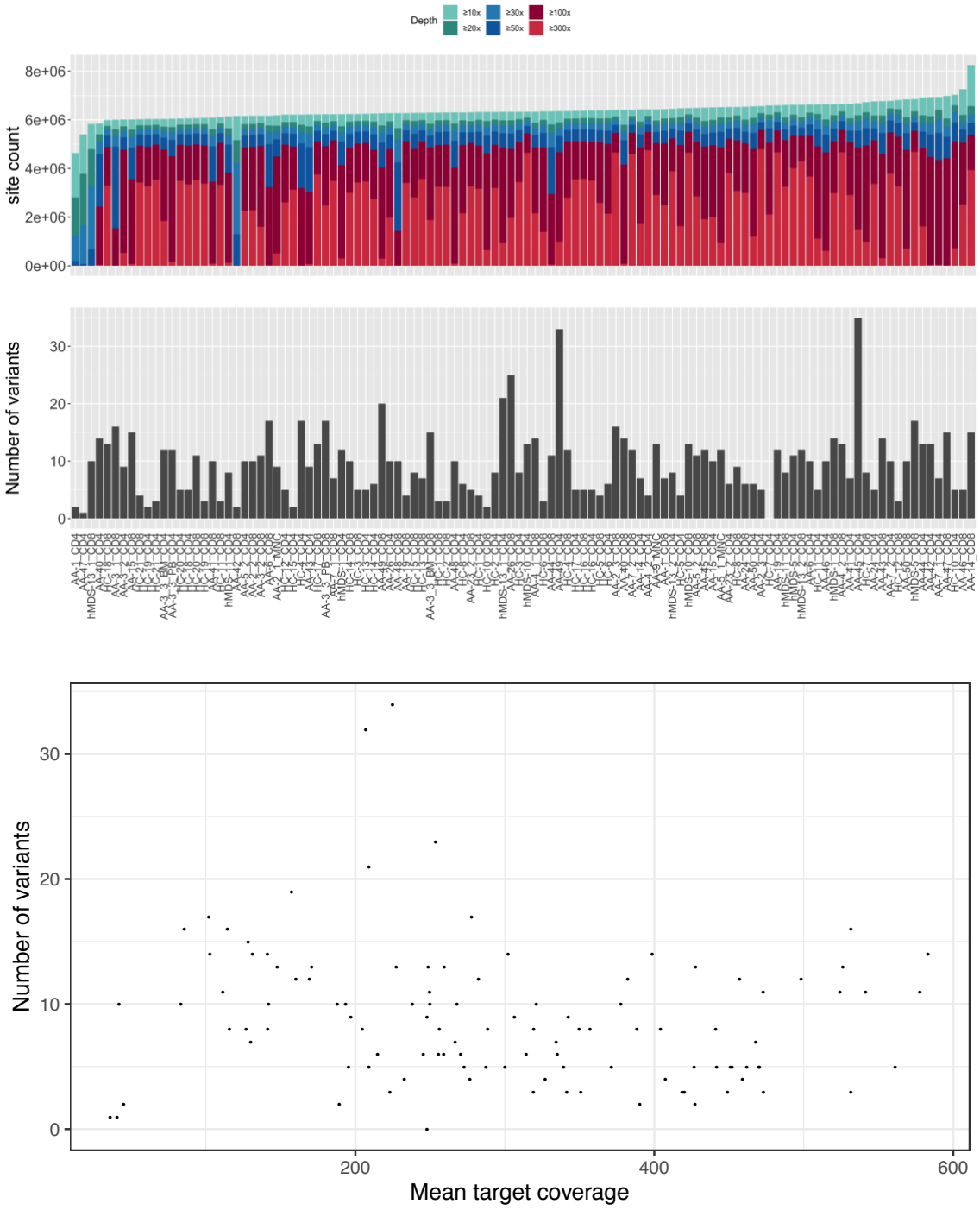


Figure S2 Coverages in immunogenepanel sequencing. A. In upper panel is shown the number of genomic sites covered >10/20/30/50/100/300x in all samples. In the panel below is number of variants in corresponding sample. **B.** Number of variants in AA and hMDS samples did not correlate with sequencing coverage.

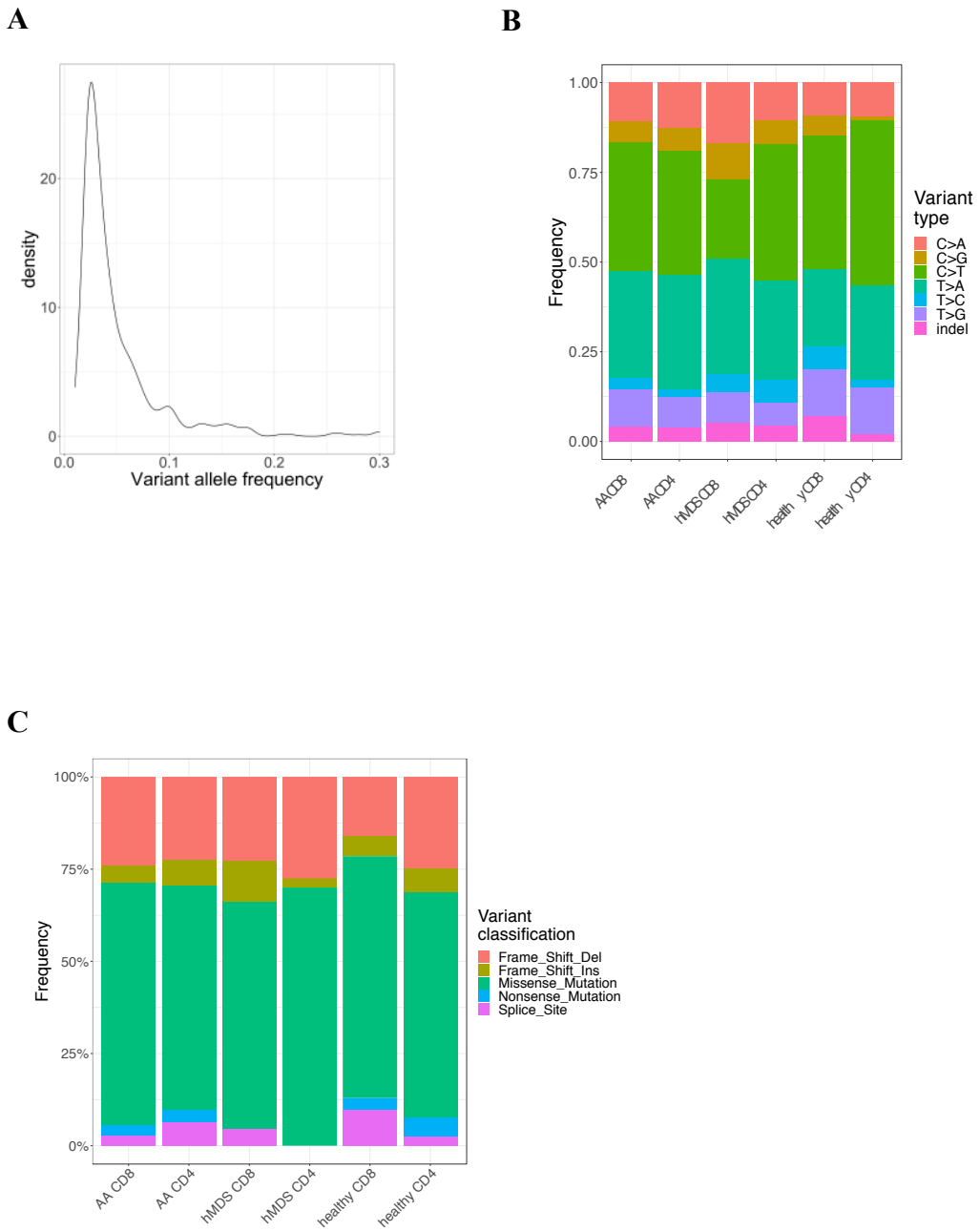


Figure S3 A. Density plot of variant allele frequencies of variants detected with immunogene panel sequencing. **B.** Single-nucleotide types of variants detected with immunogene panel sequencing and **C.** Functional variant types of non-silent variants detected with immunogene panel sequencing.

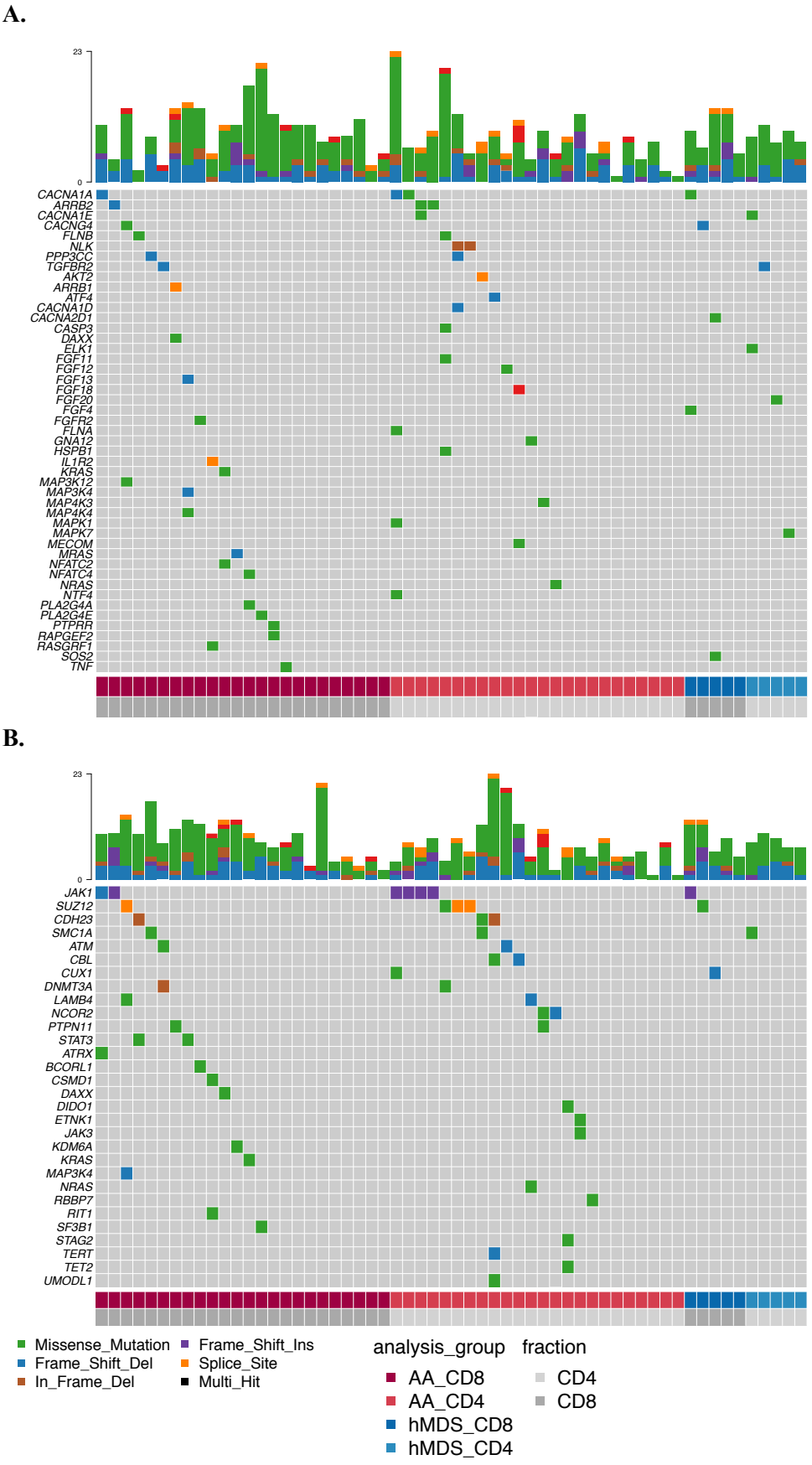
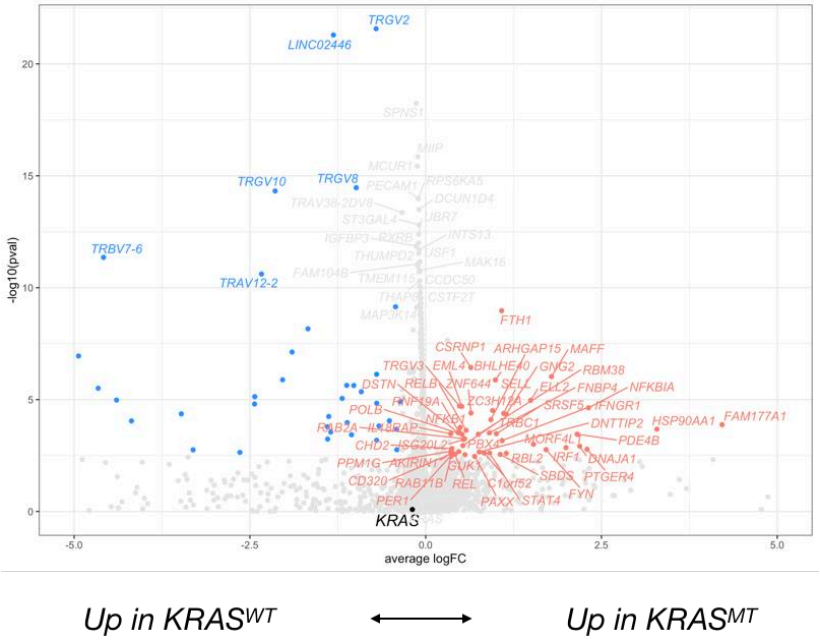


Figure S4 Somatic mutations in CD4+ and CD8+ T cells in iBMF patients. A. MAPK pathway genes. **B.** Genes related to CH.

A.



B.

Figure S5A. Volcano plot showing differentially expressed genes between clonotype of interest (to the right, significant shown in red) and cells from other clonotypes from the 1 CD8+ TEMRA cluster (to the left, significant shown in blue). **B.** Gene Set Enrichment Analysis (GSEA) results from the differential expression analysis. Shown here are GO-categories enriched (FDR qval < 0.05) to the Vbeta 7.2 clone.

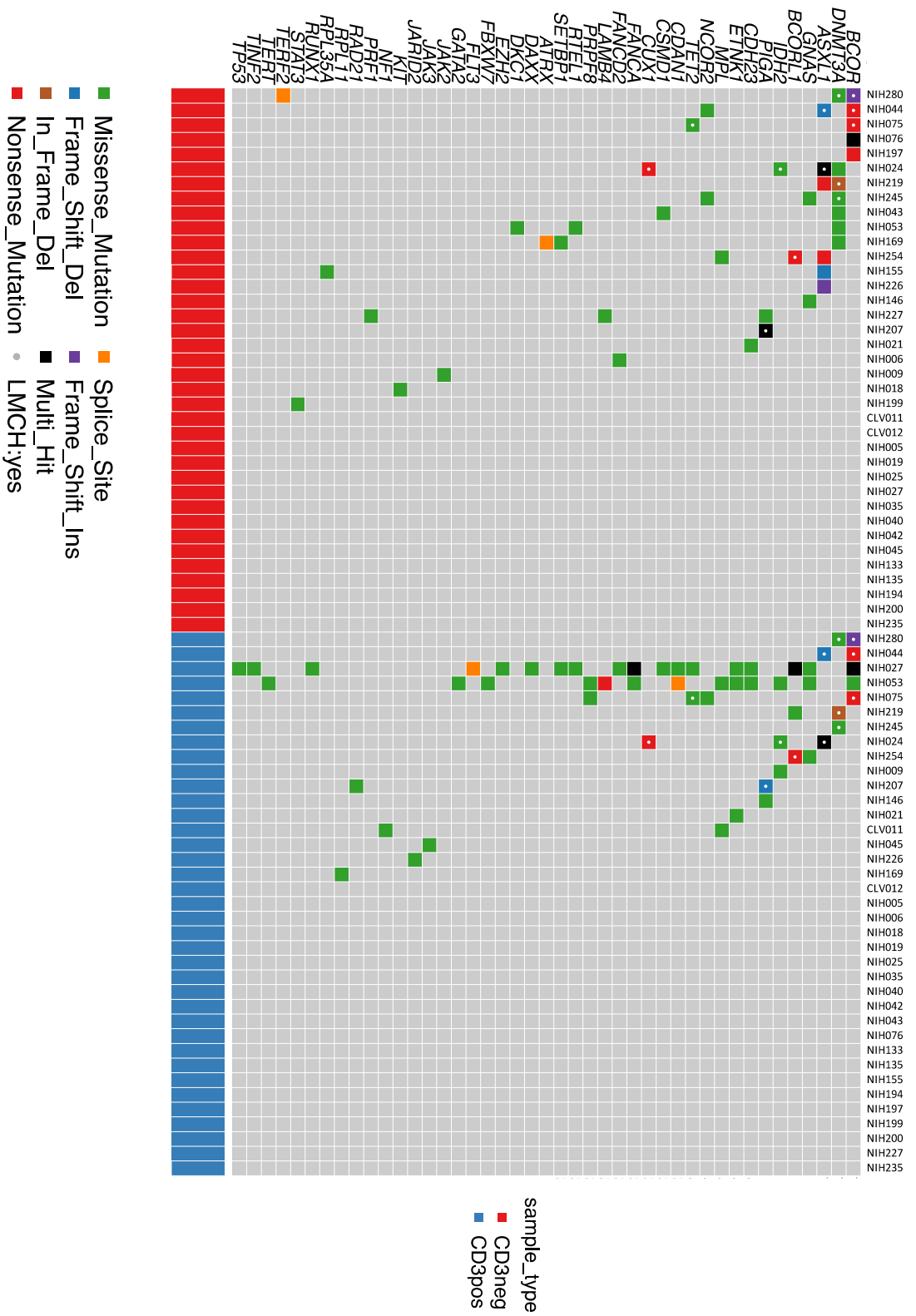


Figure S6. Somatic mutations in WES of CD3+ and CD3- MNC. Here we only present variants in genes previously linked to clonal hematopoiesis. Mutations detected concurrently in both CD3+ and CD3- cells are marked with white dots.

Table S1 Patient characteristics

	AA	hMDS	healthy
# of cases	24	5	20
Age at diagnosis (years old)	63 (17-83)	45 (27-63)	
Age at sample collection (years old)	63 (23-84)	52 (34-68)	60.5 (44-66)
Time from diagnosis to sample collection (months)	28.2 (0.119)		
Gender (Male/Female)	9 / 15	1 / 4	10 / 10
Timing of sample collection			
At diagnosis / active disease	7		
At remission	8		
Recovering	5		
At relapse	3		
Not available		5	
Treatment status at sample collection			
Pre-treatment	3		
Post-treatment	19	3	
Not available	2	2	
Response to first ATG			
CR	3	2	
PR	7		
NR	5		
Not available	9	3	
AA severity			
Moderate	13		
Severe	9		
Very severe	1		
Not available	1		
PNH clone in flow cytometry			
Yes	4		
No	9		
Subclinical (<10%)	6		
Not available	5	5	
Concomitant LGL diagnosis			
Yes	1		
No	12	5	
Not available	11		
Progression to MDS/leukemia			
Yes	1		
No	16		
Not available	11		
MDS subtype			
RA		1	
RCMD		2	
MDS-U		1	
MDS del(5q)		1	
IPSS-R classification			
Low		1	
Very low		4	

Table S2 Somatic variants in CD4+ or CD8+ T cells with immunogene panel sequencing

sample_ID	group	Chr	Start	Ref	Alt	Variant type	Gene	Amino acid change	VAF (%)
AA-1_CD4	AA CD4	19	41761486	C	T	nonsynonymous SNV	CEACAM6	p.A221V	6,1
AA-14_CD4	AA CD4	X	16869605	A	G	nonsynonymous SNV	RBBP7	p.S13P	3,6
AA-14_CD4	AA CD4	1	19657029	A	C	nonsynonymous SNV	MINOS1-NBL	p.H149P	5,2
AA-14_CD4	AA CD4	15	74433809	AGC	-	nonframeshift deletion	SEMA7A	p.36_37del	4,9
AA-14_CD4	AA CD4	5	160354783	C	-	frameshift deletion	C1QTNF2	p.D122fs	2,4
AA-14_CD4	AA CD4	4	186076660	T	C	nonsynonymous SNV	TLR3	p.L14P	2,1
AA-15_CD4	AA CD4	X	3322815	G	C	nonsynonymous SNV	MXRA5	p.S957C	3,1
AA-15_CD4	AA CD4	19	48137655	G	A	nonsynonymous SNV	LIG1	p.A306V	4,5
AA-15_CD4	AA CD4	7	48403724	A	G	nonsynonymous SNV	ABCA13	p.Q3972R	4,4
AA-15_CD4	AA CD4	19	54338199	G	T	nonsynonymous SNV	LILRA4	p.P131H	3,1
AA-15_CD4	AA CD4	2	61498915	-	AAA	splicing	XPO1		7,3
AA-15_CD4	AA CD4	1	64873427	-	T	frameshift insertion	JAK1	p.K142fs	2,4
AA-15_CD4	AA CD4	2	159944930	-	T	frameshift insertion	PLA2R1	p.G1372fs	2,4
AA-15_CD4	AA CD4	5	180620630	GAG	-	nonframeshift deletion	FLT4	p.795_795del	2,7
AA-19_CD4	AA CD4	20	1578678	G	C	nonsynonymous SNV	SIRPB1	p.D31E	3,7
AA-19_CD4	AA CD4	16	2172589	T	C	nonsynonymous SNV	TRAF7	p.S262P	2,4
AA-19_CD4	AA CD4	1	2629382	G	C	nonsynonymous SNV	MMEL1	p.L35V	2,5
AA-19_CD4	AA CD4	20	23085355	C	-	frameshift deletion	CD93	p.D280fs	1,4
AA-19_CD4	AA CD4	22	39521946	C	-	frameshift deletion	ATF4	p.P134fs	1,4
AA-19_CD4	AA CD4	10	67884929	GCG	-	nonframeshift deletion	SIRT1	p.70_70del	3,7
AA-19_CD4	AA CD4	10	79307537	T	C	nonsynonymous SNV	ZMIZ1	p.L934P	3,0
AA-19_CD4	AA CD4	3	123649196	-	GG	splicing	MYLK		4,6
AA-19_CD4	AA CD4	2	218270043	G	-	frameshift deletion	AAMP	p.P15fs	1,2
AA-19_CD4	AA CD4	2	226795115	AT	GG	nonframeshift substitut	IRS1		9,4
AA-2_3_CD4	AA CD4	11	48121032	G	T	nonsynonymous SNV	PTPRJ	p.V128F	2,3
AA-2_3_CD4	AA CD4	19	53807550	C	-	stopgain	NLRP12	p.V731X	2,1
AA-2_3_CD4	AA CD4	12	68158223	G	A	nonsynonymous SNV	IFNG	p.L51F	2,9
AA-2_3_CD4	AA CD4	7	108123139	A	-	frameshift deletion	LAMB4	p.L9fs	1,9
AA-2_3_CD4	AA CD4	1	114716123	C	T	nonsynonymous SNV	NRAS	p.G13D	4,5
AA-23_2_CD4	AA CD4	12	56349033	G	-	frameshift deletion	STAT2	p.P489fs	3,0
AA-23_2_CD4	AA CD4	8	130943463	C	A	splicing	ADCY8		2,2
AA-23_2_CD4	AA CD4	5	180620630	GAG	-	nonframeshift deletion	FLT4	p.795_795del	2,1
AA-23_2_CD4	AA CD4	3	192170653	C	A	nonsynonymous SNV	FGF12	p.V140F	4,8
AA-23_2_CD4	AA CD4	2	238329189	G	-	frameshift deletion	TRAF3IP1	p.E254fs	2,2
AA-24_CD4	AA CD4	20	1578678	G	C	nonsynonymous SNV	SIRPB1	p.D31E	3,2
AA-24_CD4	AA CD4	2	25313957	C	G	nonsynonymous SNV	DNMT3A	p.G10R	6,3
AA-24_CD4	AA CD4	17	31966158	A	T	nonsynonymous SNV	SUZ12	p.H133L	4,7
AA-24_CD4	AA CD4	3	38581154	-	A	frameshift insertion	SCN5A	p.P1002fs	2,5
AA-26_CD4	AA CD4	17	4716525	C	T	nonsynonymous SNV	ARRB2	p.P77S	4,6
AA-26_CD4	AA CD4	3	16885140	A	G	nonsynonymous SNV	PLCL2	p.E34G	3,3
AA-26_CD4	AA CD4	17	19783757	G	-	frameshift deletion	ULK2	p.P800fs	3,1
AA-26_CD4	AA CD4	17	31993864	G	T	splicing	SUZ12		2,2
AA-26_CD4	AA CD4	2	64104699	ACC	TTT	frameshift substitution	PELI1		6,8
AA-26_CD4	AA CD4	8	143932459	CCT	-	nonframeshift deletion	PLEC	p.625_626del	3,1
AA-26_CD4	AA CD4	1	181755252	G	T	nonsynonymous SNV	CACNA1E	p.V1263L	2,4
AA-3_2_CD4	AA CD4	6	26156595	G	C	nonsynonymous SNV	HIST1H1E	p.A69P	2,5
AA-3_2_CD4	AA CD4	9	34568906	T	C	nonsynonymous SNV	CNTFR	p.S26G	4,0
AA-3_2_CD4	AA CD4	3	53179768	G	T	stopgain	PRKCD	p.E103X	2,1
AA-3_2_CD4	AA CD4	15	74799309	C	T	nonsynonymous SNV	CSK	p.R94W	2,1
AA-3_2_CD4	AA CD4	4	105272621	C	A	nonsynonymous SNV	TET2	p.Q1414K	2,6
AA-3_2_CD4	AA CD4	4	182680597	T	C	nonsynonymous SNV	TENM3	p.L565P	2,2
AA-3_3_BM_CD4	AA CD4	19	40256925	C	A	splicing	AKT2		4,2
AA-3_3_BM_CD4	AA CD4	19	42096209	A	G	nonsynonymous SNV	POU2F2	p.L179P	4,2
AA-3_3_BM_CD4	AA CD4	2	61498915	-	AA	splicing	XPO1		6,6
AA-3_3_BM_CD4	AA CD4	20	62911482	T	G	nonsynonymous SNV	DIDO1	p.E44A	3,5
AA-3_3_BM_CD4	AA CD4	4	105272621	C	A	nonsynonymous SNV	TET2	p.Q1414K	2,0
AA-3_3_BM_CD4	AA CD4	X	124045368	G	A	nonsynonymous SNV	STAG2	p.A223T	8,7
AA-3_3_BM_CD4	AA CD4	1	206474382	C	T	nonsynonymous SNV	IKBKE	p.R47W	2,4
AA-3_3_PB_CD4	AA CD4	9	2039793	AG	-	frameshift deletion	SMARCA2	p.Q228fs	4,2
AA-3_3_PB_CD4	AA CD4	9	2039796	A	-	frameshift deletion	SMARCA2	p.Q229fs	4,2
AA-3_3_PB_CD4	AA CD4	16	10907447	-	C	frameshift insertion	CIITA	p.S652fs	3,4

AA-3_3_PB_CD4	AA CD4	15	52336478	T	-	frameshift deletion	MYO5A	p.I1413fs	3,3
AA-3_3_PB_CD4	AA CD4	19	54360981	C	T	nonsynonymous SNV	LAIR1	p.R82H	2,1
AA-3_3_PB_CD4	AA CD4	20	62911482	T	G	nonsynonymous SNV	DIDO1	p.E44A	2,7
AA-3_3_PB_CD4	AA CD4	14	105473335	-	G	frameshift insertion	CRIP2	p.R30fs	15,3
AA-3_3_PB_CD4	AA CD4	11	128772868	T	-	frameshift deletion	FLI1	p.F158fs	2,2
AA-3_3_PB_CD4	AA CD4	5	180620630	GAG	-	nonframeshift deletion	FLT4	p.795_795del	2,3
AA-3_3_PB_CD4	AA CD4	1	200992344	GCC	-	nonframeshift deletion	KIF21B	p.774_775del	2,9
AA-4_2_CD4	AA CD4	12	49032991	T	A	nonsynonymous SNV	KMT2D	p.Q3905L	2,0
AA-4_2_CD4	AA CD4	12	124363682	G	-	frameshift deletion	NCOR2	p.P957fs	2,7
AA-40_CD4	AA CD4	19	10288183	G	T	stoploss	ICAM4	p.X273L	3,0
AA-40_CD4	AA CD4	10	13197702	A	G	nonsynonymous SNV	MCM10	p.K685R	2,0
AA-40_CD4	AA CD4	19	41242951	G	A	nonsynonymous SNV	AXL	p.A193T	2,3
AA-40_CD4	AA CD4	13	48452991	A	G	splicing	RB1		3,6
AA-40_CD4	AA CD4	19	53807550	C	-	stopgain	NLRP12	p.V731X	3,9
AA-40_CD4	AA CD4	12	57162440	C	-	frameshift deletion	LRP1	p.P776fs	2,4
AA-40_CD4	AA CD4	15	58611061	G	A	nonsynonymous SNV	ADAM10	p.T550M	3,6
AA-40_CD4	AA CD4	12	112489068	C	T	nonsynonymous SNV	PTPN11	p.R502W	5,9
AA-40_CD4	AA CD4	12	124426673	C	T	nonsynonymous SNV	NCOR2	p.R425H	3,5
AA-40_CD4	AA CD4	3	169112846	G	T	nonsynonymous SNV	MECOM	p.L652I	3,4
AA-40_CD4	AA CD4	5	171436135	G	T	stopgain	FGF18	p.E38X	2,9
AA-41_CD4	AA CD4	16	2181305	G	A	nonsynonymous SNV	CASKIN1	p.A688V	2,2
AA-41_CD4	AA CD4	17	28042952	CAC	-	nonframeshift deletion	NLK	p.27_27del	2,1
AA-41_CD4	AA CD4	11	47667658	-	T	frameshift insertion	AGBL2	p.K751fs	3,2
AA-41_CD4	AA CD4	16	57382345	C	-	frameshift deletion	CX3CL1	p.L84fs	2,3
AA-41_CD4	AA CD4	16	66570026	-	T	frameshift insertion	CMTM1	p.V175fs	2,3
AA-42_CD4	AA CD4	16	10957804	A	-	frameshift deletion	CLEC16A	p.K35fs	2,5
AA-42_CD4	AA CD4	8	22527512	A	-	frameshift deletion	PPP3CC	p.E355fs	2,3
AA-42_CD4	AA CD4	17	28042952	CAC	-	nonframeshift deletion	NLK	p.27_27del	2,2
AA-42_CD4	AA CD4	17	28385267	C	T	nonsynonymous SNV	SARM1	p.A541V	7,3
AA-42_CD4	AA CD4	19	43261911	G	A	nonsynonymous SNV	PSG9	p.R220W	2,6
AA-42_CD4	AA CD4	7	44106848	C	-	frameshift deletion	AEBP1	p.P186fs	4,5
AA-42_CD4	AA CD4	X	53409123	C	T	nonsynonymous SNV	SMC1A	p.S495N	5,0
AA-42_CD4	AA CD4	3	53770457	T	-	frameshift deletion	CACNA1D	p.F1302fs	2,0
AA-42_CD4	AA CD4	10	71811331	G	A	nonsynonymous SNV	CDH23	p.D792N	3,6
AA-42_CD4	AA CD4	10	79307505	C	-	frameshift deletion	ZMIZ1	p.H923fs	2,3
AA-42_CD4	AA CD4	9	136503188	C	T	nonsynonymous SNV	NOTCH1	p.V1721M	3,4
AA-42_CD4	AA CD4	3	196472595	C	T	nonsynonymous SNV	RNF168	p.V314I	3,0
AA-43_CD4	AA CD4	19	13286719	C	T	nonsynonymous SNV	CACNA1A	p.A1114T	4,9
AA-43_CD4	AA CD4	1	36466179	C	T	nonsynonymous SNV	CSF3R	p.A757T	2,3
AA-43_CD4	AA CD4	8	39918846	G	A	nonsynonymous SNV	IDO1	p.C112Y	3,8
AA-43_CD4	AA CD4	1	89116994	G	T	nonsynonymous SNV	GBP2	p.P289H	2,6
AA-43_CD4	AA CD4	3	105776477	G	T	nonsynonymous SNV	CBLB	p.P190H	3,4
AA-43_CD4	AA CD4	11	129882181	G	T	nonsynonymous SNV	NFRKB	p.L391I	2,9
AA-44_CD4	AA CD4	16	11679278	G	-	frameshift deletion	TXNDC11	p.Q480fs	4,3
AA-44_CD4	AA CD4	3	13496768	C	-	frameshift deletion	HDAC11	p.I95fs	2,8
AA-44_CD4	AA CD4	20	44622613	G	A	nonsynonymous SNV	ADA	p.P139S	2,4
AA-44_CD4	AA CD4	6	47609224	-	A	frameshift insertion	CD2AP	p.V578fs	2,0
AA-44_CD4	AA CD4	6	47609225	A	-	frameshift deletion	CD2AP	p.K579fs	2,4
AA-44_CD4	AA CD4	19	54847969	G	C	nonsynonymous SNV	KIR2DS2	p.R292T	2,2
AA-44_CD4	AA CD4	19	54847980	CG	TC	nonframeshift substitut	KIR2DS2		2,5
AA-44_CD4	AA CD4	16	66565991	-	A	frameshift insertion	CKLF	p.E62fs	2,8
AA-44_CD4	AA CD4	15	93002204	A	-	frameshift deletion	CHD2	p.K1389fs	2,3
AA-44_CD4	AA CD4	11	119206451	G	-	frameshift deletion	CBL	p.G12fs	3,9
AA-44_CD4	AA CD4	5	148120312	AA	-	frameshift deletion	SPINK5	p.E820fs	2,5
AA-44_CD4	AA CD4	1	153390239	T	C	nonsynonymous SNV	S100A8	p.K73R	1,1
AA-44_CD4	AA CD4	2	159944930	-	T	frameshift insertion	PLA2R1	p.G1372fs	3,1
AA-45_CD4	AA CD4	17	819471	C	T	nonsynonymous SNV	NXN	p.R155Q	2,6
AA-45_CD4	AA CD4	5	1294550	G	-	frameshift deletion	TERT	p.P112fs	5,6
AA-45_CD4	AA CD4	11	7062467	C	G	nonsynonymous SNV	NLRP14	p.A980G	7,4
AA-45_CD4	AA CD4	19	13286541	G	-	frameshift deletion	CACNA1A	p.P1173fs	5,1
AA-45_CD4	AA CD4	19	18162265	C	A	nonsynonymous SNV	PIK3R2	p.P322Q	8,1
AA-45_CD4	AA CD4	22	21867421	G	A	nonsynonymous SNV	MAPK1	p.A7V	6,3
AA-45_CD4	AA CD4	3	27721976	A	G	nonsynonymous SNV	EOMES	p.S107P	4,1
AA-45_CD4	AA CD4	21	31120412	G	A	nonsynonymous SNV	TIAM1	p.R1578C	9,4

AA-45_CD4	AA CD4	10	34131566	C	A	nonsynonymous SNV	PARD3	p.S1037I	5,8
AA-45_CD4	AA CD4	20	34438603	ACC	-	nonframeshift deletion	ITCH	p.107_108del	2,2
AA-45_CD4	AA CD4	22	41178675	C	-	frameshift deletion	EP300	p.P2322fs	1,6
AA-45_CD4	AA CD4	21	42111218	G	A	nonsynonymous SNV	UMODL1	p.E594K	7,9
AA-45_CD4	AA CD4	15	42142593	G	A	nonsynonymous SNV	PLA2G4F	p.P755L	6,7
AA-45_CD4	AA CD4	21	44901575	C	G	nonsynonymous SNV	ITGB2	p.E220Q	10,4
AA-45_CD4	AA CD4	19	49061396	C	T	nonsynonymous SNV	NTF4	p.C201Y	4,1
AA-45_CD4	AA CD4	16	68122002	A	T	nonsynonymous SNV	NFATC3	p.D40V	2,6
AA-45_CD4	AA CD4	12	68158261	-	AA	splicing	IFNG		12,9
AA-45_CD4	AA CD4	10	71791196	GCT	-	nonframeshift deletion	CDH23	p.2038_2039del	2,3
AA-45_CD4	AA CD4	11	119285565	T	C	nonsynonymous SNV	CBL	p.M647T	8,8
AA-45_CD4	AA CD4	X	154366454	G	A	nonsynonymous SNV	FLNA	p.A361V	3,2
AA-45_CD4	AA CD4	1	155324488	T	C	nonsynonymous SNV	RUSC1	p.L24P	5,4
AA-45_CD4	AA CD4	3	184372727	G	A	nonsynonymous SNV	THPO	p.S283F	10,2
AA-45_CD4	AA CD4	2	216162010	G	A	nonsynonymous SNV	XRCC5	p.R599H	8,2
AA-46_CD4	AA CD4	11	299396	G	-	frameshift deletion	IFITM5	p.P32fs	2,6
AA-46_CD4	AA CD4	1	2629382	G	C	nonsynonymous SNV	MMEL1	p.L35V	7,1
AA-46_CD4	AA CD4	11	14644838	G	A	nonsynonymous SNV	PDE3B	p.A255T	2,6
AA-46_CD4	AA CD4	9	14788920	T	-	frameshift deletion	FREM1	p.K1392fs	2,4
AA-46_CD4	AA CD4	2	39250640	G	A	nonsynonymous SNV	MAP4K3	p.A867V	2,1
AA-46_CD4	AA CD4	14	45176751	-	A	frameshift insertion	FANCM	p.Q1307fs	2,5
AA-46_CD4	AA CD4	13	48459775	T	-	frameshift deletion	RB1	p.L683fs	2,3
AA-46_CD4	AA CD4	1	64841313	-	T	frameshift insertion	JAK1	p.P861fs	3,5
AA-46_CD4	AA CD4	7	98224843	C	-	frameshift deletion	TECPR1	p.G883fs	2,7
AA-47_CD4	AA CD4	11	14867704	G	T	nonsynonymous SNV	PDE3B	p.D1029Y	6,3
AA-48_CD4	AA CD4	19	53807550	C	-	stopgain	NLRP12	p.V731X	2,9
AA-48_CD4	AA CD4	17	58271727	C	T	nonsynonymous SNV	MPO	p.R653H	2,7
AA-48_CD4	AA CD4	1	67395689	A	-	frameshift deletion	IL12RB2	p.E644fs	3,1
AA-48_CD4	AA CD4	4	69957571	A	-	frameshift deletion	CSN2	p.F125fs	2,1
AA-48_CD4	AA CD4	11	102843458	C	T	nonsynonymous SNV	MMP3	p.S30N	2,9
AA-48_CD4	AA CD4	1	146019198	C	T	nonsynonymous SNV	HFE2	p.A99T	2,2
AA-48_CD4	AA CD4	5	148120312	AA	-	frameshift deletion	SPINK5	p.E820fs	3,4
AA-48_CD4	AA CD4	1	231209016	C	T	nonsynonymous SNV	TRIM67	p.A630V	2,0
AA-49_CD4	AA CD4	16	3738616	C	T	nonsynonymous SNV	CREBBP	p.R1408H	1,1
AA-49_CD4	AA CD4	17	5521756	AT	GC	nonframeshift substitut	NLRP1	NM_033007	15,3
AA-49_CD4	AA CD4	17	7439675	G	C	nonsynonymous SNV	FGF11	p.G19R	4,1
AA-49_CD4	AA CD4	10	8058529	A	C	nonsynonymous SNV	GATA3	p.T156P	6,2
AA-49_CD4	AA CD4	12	8838361	C	T	nonsynonymous SNV	A2ML1	p.A294V	14,3
AA-49_CD4	AA CD4	22	17738219	C	T	nonsynonymous SNV	BID	p.R29K	15,1
AA-49_CD4	AA CD4	20	32083930	G	A	nonsynonymous SNV	HCK	p.R169Q	12,1
AA-49_CD4	AA CD4	6	32183926	A	C	nonsynonymous SNV	AGER	p.C38W	15,8
AA-49_CD4	AA CD4	3	39266083	C	T	nonsynonymous SNV	CX3CR1	p.V143M	15,3
AA-49_CD4	AA CD4	19	42099726	C	G	nonsynonymous SNV	POU2F2	p.Q133H	21,6
AA-49_CD4	AA CD4	3	50302418	C	T	nonsynonymous SNV	HYAL1	p.R180Q	13,2
AA-49_CD4	AA CD4	3	58094867	G	T	nonsynonymous SNV	FLNB	p.Q273H	13,1
AA-49_CD4	AA CD4	17	58199623	T	G	nonsynonymous SNV	EPX	p.S456A	15,0
AA-49_CD4	AA CD4	11	61875938	A	C	nonsynonymous SNV	FADS3	p.V400G	11,7
AA-49_CD4	AA CD4	16	67174367	G	T	nonsynonymous SNV	NOL3	p.Q66H	5,0
AA-49_CD4	AA CD4	7	76302959	A	T	nonsynonymous SNV	HSPB1	p.S83C	30,0
AA-49_CD4	AA CD4	11	108250753	TG	-	frameshift deletion	ATM	p.C430fs	3,5
AA-49_CD4	AA CD4	7	142960945	C	T	nonsynonymous SNV	KEL	p.R128Q	15,1
AA-49_CD4	AA CD4	1	155290622	G	A	stopgain	PKLR	p.R559X	17,2
AA-49_CD4	AA CD4	6	162201135	G	T	nonsynonymous SNV	PRKN	p.T177N	16,9
AA-49_CD4	AA CD4	4	184635393	T	C	nonsynonymous SNV	CASP3	p.M27V	13,2
AA-5_2_CD4	AA CD4	19	17841498	C	T	nonsynonymous SNV	JAK3	p.V345M	2,1
AA-5_2_CD4	AA CD4	12	22659061	A	G	nonsynonymous SNV	ETNK1	p.N244S	5,7
AA-5_2_CD4	AA CD4	21	39183309	T	C	nonsynonymous SNV	PSMG1	p.E26G	7,4
AA-5_2_CD4	AA CD4	19	40667975	C	-	frameshift deletion	NUMBL	p.Q400fs	5,6
AA-5_2_CD4	AA CD4	19	40667977	GC	-	frameshift deletion	NUMBL	p.Q399fs	5,3
AA-5_2_CD4	AA CD4	22	41150125	G	A	nonsynonymous SNV	EP300	p.R915H	2,2
AA-5_2_CD4	AA CD4	14	100262307	A	-	frameshift deletion	YY1	p.E228fs	2,2
AA-50_CD4	AA CD4	7	2731590	G	A	nonsynonymous SNV	GNA12	p.T170M	2,6
AA-50_CD4	AA CD4	6	26156910	A	-	frameshift deletion	HIST1H1E	p.K174fs	2,1
AA-50_CD4	AA CD4	19	51454239	TG	CA	nonframeshift substitut	SIGLEC8		28,1

AA-50_CD4	AA CD4	1	64841313	-	T	frameshift insertion	JAK1	p.P861fs	2,1
AA-50_CD4	AA CD4	7	102283025	G	A	nonsynonymous SNV	CUX1	p.A642T	2,2
AA-6_CD4	AA CD4	16	3729220	G	A	nonsynonymous SNV	CREBBP	p.P1905S	2,0
AA-6_CD4	AA CD4	20	23085355	C	-	frameshift deletion	CD93	p.D280fs	1,4
AA-6_CD4	AA CD4	8	47954350	TT	-	frameshift deletion	PRKDC	p.K165fs	2,4
AA-6_CD4	AA CD4	1	64841313	-	T	frameshift insertion	JAK1	p.P861fs	2,0
AA-6_CD4	AA CD4	14	95116697	T	A	splicing	DICER1		2,3
AA-6_CD4	AA CD4	14	105466451	TG	CC	nonframeshift substitut	MTA1		4,5
AA-6_CD4	AA CD4	6	109467736	-	A	splicing	ZBTB24		3,5
AA-6_CD4	AA CD4	3	194396880	C	-	frameshift deletion	GP5	p.G468fs	2,4
AA-7_1_CD4	AA CD4	17	4716534	C	T	nonsynonymous SNV	ARRB2	p.P80S	5,4
AA-7_1_CD4	AA CD4	1	16760814	G	C	nonsynonymous SNV	MST1L	p.S92W	13,0
AA-7_1_CD4	AA CD4	3	38575422	C	T	nonsynonymous SNV	SCN5A	p.V1127M	7,7
AA-7_1_CD4	AA CD4	15	52179798	C	G	nonsynonymous SNV	GNB5	p.E28Q	4,7
AA-7_1_CD4	AA CD4	16	88716376	G	A	nonsynonymous SNV	PIEZO1	p.T2345I	10,0
AA-7_1_CD4	AA CD4	16	88731725	C	T	nonsynonymous SNV	PIEZO1	p.M1059I	4,1
AA-7_2_CD4	AA CD4	17	4716534	C	T	nonsynonymous SNV	ARRB2	p.P80S	7,0
AA-7_2_CD4	AA CD4	16	15718400	C	T	nonsynonymous SNV	MYH11	p.R1737Q	4,1
AA-7_2_CD4	AA CD4	20	23085825	G	A	nonsynonymous SNV	CD93	p.T123M	11,1
AA-7_2_CD4	AA CD4	6	26156815	C	T	nonsynonymous SNV	HIST1H1E	p.T142M	2,0
AA-7_2_CD4	AA CD4	17	31993864	G	T	splicing	SUZ12		2,9
AA-7_2_CD4	AA CD4	2	58226728	C	A	nonsynonymous SNV	FANCL	p.L91F	3,2
AA-7_2_CD4	AA CD4	17	59064371	C	A	nonsynonymous SNV	TRIM37	p.V248F	2,1
AA-7_2_CD4	AA CD4	11	61816898	A	G	nonsynonymous SNV	FADS1	p.L11P	4,2
AA-7_2_CD4	AA CD4	15	66703941	G	A	nonsynonymous SNV	SMAD6	p.R228H	4,8
AA-1_CD8	AA CD8	17	4716515	C	-	frameshift deletion	ARRB2	p.F73fs	4,5
AA-1_CD8	AA CD8	3	53186637	G	-	frameshift deletion	PRKCD	p.G432fs	1,6
AA-1_CD8	AA CD8	19	54575980	A	T	nonsynonymous SNV	LILRA2	p.N364Y	3,1
AA-1_CD8	AA CD8	8	66154051	A	G	nonsynonymous SNV	TRIM55	p.E414G	2,8
AA-14_CD8	AA CD8	1	2629382	G	C	nonsynonymous SNV	MMEL1	p.L35V	4,7
AA-14_CD8	AA CD8	21	14965960	CTCT	-	frameshift deletion	NRIP1	p.R744fs	10,3
AA-14_CD8	AA CD8	13	36819536	C	-	frameshift deletion	RFXAP	p.A60fs	2,5
AA-14_CD8	AA CD8	1	44801718	C	G	nonsynonymous SNV	PLK3	p.Q178E	9,2
AA-14_CD8	AA CD8	X	45110146	G	A	nonsynonymous SNV	KDM6A	p.C1279Y	2,6
AA-14_CD8	AA CD8	12	53482752	C	A	nonsynonymous SNV	MAP3K12	p.G684V	10,1
AA-14_CD8	AA CD8	19	53807550	C	-	stopgain	NLRP12	p.V731X	2,6
AA-14_CD8	AA CD8	18	63987234	A	-	frameshift deletion	SERPINB8	p.K181fs	4,8
AA-14_CD8	AA CD8	17	67030763	G	T	nonsynonymous SNV	CACNG4	p.R248M	10,8
AA-14_CD8	AA CD8	7	116559262	G	A	nonsynonymous SNV	CAV1	p.R140H	3,4
AA-14_CD8	AA CD8	11	119154858	C	T	nonsynonymous SNV	ABCG4	p.A210V	5,7
AA-14_CD8	AA CD8	8	144312189	T	-	frameshift deletion	HSF1	p.S363fs	6,0
AA-14_CD8	AA CD8	5	148127023	C	G	nonsynonymous SNV	SPINK5	p.P970A	2,4
AA-15_CD8	AA CD8	12	5985057	C	T	nonsynonymous SNV	VWF	p.E2322K	3,5
AA-15_CD8	AA CD8	12	40278178	A	G	nonsynonymous SNV	LRRK2	p.N720D	9,9
AA-15_CD8	AA CD8	22	44845292	A	G	nonsynonymous SNV	ARHGAP8,PR	p.E207G	5,0
AA-15_CD8	AA CD8	12	49631428	TT	CC	nonframeshift substitut	PRPF40B	NM_001031698	4,5
AA-15_CD8	AA CD8	13	51964981	G	A	nonsynonymous SNV	ATP7B	p.T587M	3,6
AA-15_CD8	AA CD8	12	70656729	T	C	nonsynonymous SNV	PTPRR	p.I374V	5,8
AA-15_CD8	AA CD8	7	74338976	A	G	nonsynonymous SNV	CLIP2	p.K217R	4,5
AA-15_CD8	AA CD8	11	104037388	A	-	frameshift deletion	DDI1	p.E189fs	3,0
AA-15_CD8	AA CD8	X	130013925	A	C	nonsynonymous SNV	BCORL1	p.T385P	10,3
AA-15_CD8	AA CD8	5	151119386	G	A	nonsynonymous SNV	ANXA6	p.A419V	2,5
AA-15_CD8	AA CD8	X	153872212	A	C	nonsynonymous SNV	L1CAM	p.C109G	10,2
AA-15_CD8	AA CD8	4	159352799	C	T	nonsynonymous SNV	RAPGEF2	p.S1166L	10,5
AA-15_CD8	AA CD8	2	161223975	C	T	nonsynonymous SNV	TANK	p.P130S	7,3
AA-19_CD8	AA CD8	X	21426662	T	C	splicing	CNKSR2		2,9
AA-19_CD8	AA CD8	21	46003711	G	A	nonsynonymous SNV	COL6A1	p.E929K	2,9
AA-19_CD8	AA CD8	3	57104352	T	A	nonsynonymous SNV	IL17RD	p.Y268F	2,1
AA-2_2_CD8	AA CD8	18	63318543	C	-	frameshift deletion	BCL2	p.A42fs	3,0
AA-2_2_CD8	AA CD8	8	88074670	C	T	nonsynonymous SNV	MMP16	p.R386Q	2,7
AA-2_2_CD8	AA CD8	15	89306164	G	A	nonsynonymous SNV	FANCI	p.M1109I	3,9
AA-2_2_CD8	AA CD8	11	89340169	T	G	nonsynonymous SNV	NOX4	p.D407A	2,5
AA-2_2_CD8	AA CD8	3	93893055	C	T	nonsynonymous SNV	PROS1	p.D345N	3,5
AA-2_2_CD8	AA CD8	2	101842667	G	T	nonsynonymous SNV	MAP4K4	p.Q336H	2,6

AA-2_2_CD8	AA CD8	8	143923856	G	A	nonsynonymous SNV	PLEC	p.R2011W	7,7
AA-2_2_CD8	AA CD8	1	182843294	G	T	nonsynonymous SNV	DHX9	p.V38F	3,2
AA-2_3_CD8	AA CD8	17	31993864	G	T	splicing	SUZ12		3,8
AA-2_3_CD8	AA CD8	17	39410039	T	C	nonsynonymous SNV	MED1	p.M728V	2,8
AA-2_3_CD8	AA CD8	5	40762977	G	A	nonsynonymous SNV	PRKAA1	p.S494L	2,4
AA-2_3_CD8	AA CD8	11	48121014	G	T	nonsynonymous SNV	PTPRJ	p.V122L	2,2
AA-2_3_CD8	AA CD8	2	60918414	G	T	nonsynonymous SNV	REL	p.V221L	2,3
AA-2_3_CD8	AA CD8	14	87941401	C	A	nonsynonymous SNV	GALC	p.D587Y	2,0
AA-2_3_CD8	AA CD8	11	89340169	T	G	nonsynonymous SNV	NOX4	p.D407A	2,9
AA-2_3_CD8	AA CD8	3	93893055	C	T	nonsynonymous SNV	PROS1	p.D345N	3,4
AA-2_3_CD8	AA CD8	2	101842667	G	T	nonsynonymous SNV	MAP4K4	p.Q336H	3,4
AA-2_3_CD8	AA CD8	7	108048044	C	T	nonsynonymous SNV	LAMB4	p.R1397Q	2,6
AA-2_3_CD8	AA CD8	X	138857661	G	-	frameshift deletion	FGF13	p.P23fs	1,3
AA-2_3_CD8	AA CD8	6	161098343	C	-	frameshift deletion	MAP3K4	p.A1193fs	2,2
AA-2_3_CD8	AA CD8	6	161098345	GC	-	frameshift deletion	MAP3K4	p.A1194fs	2,2
AA-2_3_CD8	AA CD8	1	197040526	G	T	nonsynonymous SNV	F13B	p.Q650K	3,9
AA-23_2_CD8	AA CD8	16	11255329	G	-	frameshift deletion	SOCS1	p.P50fs	4,6
AA-23_2_CD8	AA CD8	3	30650380	AA	-	frameshift deletion	TGFBR2	p.E125fs	1,8
AA-23_2_CD8	AA CD8	19	53807550	C	-	stopgain	NLRP12	p.V731X	3,2
AA-24_CD8	AA CD8	9	34659772	G	C	nonsynonymous SNV	IL11RA	p.G275A	5,0
AA-24_CD8	AA CD8	15	79058383	C	T	nonsynonymous SNV	RASGRF1	p.R161Q	5,4
AA-24_CD8	AA CD8	8	101631668	G	A	nonsynonymous SNV	GRHL2	p.R414Q	2,8
AA-24_CD8	AA CD8	2	102022186	-	A	splicing	IL1R2		2,3
AA-24_CD8	AA CD8	12	121274443	GCT	-	nonframeshift deletion	CAMKK2	p.28_28del	2,2
AA-26_CD8	AA CD8	20	428972	G	A	nonsynonymous SNV	RBCK1	p.A274T	4,9
AA-26_CD8	AA CD8	10	8058529	A	C	nonsynonymous SNV	GATA3	p.T156P	3,8
AA-26_CD8	AA CD8	4	15838141	G	A	nonsynonymous SNV	CD38	p.R212H	2,4
AA-26_CD8	AA CD8	17	19958502	A	C	nonsynonymous SNV	AKAP10	p.L130W	3,0
AA-26_CD8	AA CD8	13	35156093	A	C	nonsynonymous SNV	NBEA	p.K846N	2,3
AA-26_CD8	AA CD8	15	41985865	T	C	nonsynonymous SNV	PLA2G4E	p.N726D	3,7
AA-26_CD8	AA CD8	15	42150123	T	A	nonsynonymous SNV	PLA2G4F	p.S302C	5,6
AA-26_CD8	AA CD8	19	55912123	T	C	nonsynonymous SNV	NLRP13	p.E565G	6,1
AA-26_CD8	AA CD8	11	66564557	C	T	splicing	CTSF		4,1
AA-26_CD8	AA CD8	16	75235818	C	T	nonsynonymous SNV	BCAR1	p.V151M	3,5
AA-26_CD8	AA CD8	5	78002972	A	G	nonsynonymous SNV	AP3B1	p.I1023T	8,7
AA-26_CD8	AA CD8	X	78273075	-	T	frameshift insertion	CYSLTR1	p.N224fs	2,2
AA-26_CD8	AA CD8	11	104037183	G	A	nonsynonymous SNV	DDI1	p.A121T	4,5
AA-26_CD8	AA CD8	12	120215218	G	A	nonsynonymous SNV	PXN	p.P344L	3,7
AA-26_CD8	AA CD8	11	126292936	T	C	nonsynonymous SNV	TIRAP	p.I176T	2,4
AA-26_CD8	AA CD8	5	136054753	G	T	nonsynonymous SNV	TGFBI	p.L434F	2,1
AA-26_CD8	AA CD8	6	137203516	A	-	frameshift deletion	IFNGR1	p.F239fs	4,8
AA-26_CD8	AA CD8	X	153866826	C	T	nonsynonymous SNV	L1CAM	p.V747M	4,2
AA-26_CD8	AA CD8	3	190609032	G	A	nonsynonymous SNV	IL1RAP	p.E130K	5,4
AA-26_CD8	AA CD8	1	207144685	G	A	nonsynonymous SNV	C4BPA	p.A588T	2,3
AA-26_CD8	AA CD8	2	218387897	G	A	nonsynonymous SNV	SLC11A1	p.G246D	2,9
AA-3_1_CD8	AA CD8	16	30083679	T	C	nonsynonymous SNV	PPP4C	p.S134P	3,1
AA-3_1_CD8	AA CD8	3	46265851	-	A	frameshift insertion	CCR3	p.S231fs	5,1
AA-3_1_CD8	AA CD8	14	50204360	T	C	nonsynonymous SNV	SOS2	p.Y46C	3,3
AA-3_1_CD8	AA CD8	10	69368560	C	T	stopgain	HK1	p.R174X	2,9
AA-3_1_CD8	AA CD8	11	102724768	T	C	nonsynonymous SNV	MMP8	p.T30A	3,7
AA-3_1_CD8	AA CD8	3	134610310	A	G	nonsynonymous SNV	KY	p.S220P	4,1
AA-3_1_CD8	AA CD8	8	143932459	CCT	-	nonframeshift deletion	PLEC	p.625_626del	3,7
AA-3_1_CD8	AA CD8	8	143934880	AGC	-	nonframeshift deletion	PLEC	p.277_278del	4,1
AA-3_1_CD8	AA CD8	6	161098346	C	-	frameshift deletion	MAP3K4	p.A1194fs	3,6
AA-3_1_CD8	AA CD8	6	161098348	GC	-	frameshift deletion	MAP3K4	p.A1195fs	3,5
AA-3_1_CD8	AA CD8	1	203772091	C	T	nonsynonymous SNV	LAX1	p.R96C	3,1
AA-3_2_CD8	AA CD8	19	7174723	C	T	nonsynonymous SNV	INSR	p.C328Y	3,4
AA-3_2_CD8	AA CD8	3	9893298	T	A	nonsynonymous SNV	JAGN1	p.V158E	3,0
AA-3_2_CD8	AA CD8	17	19382248	A	C	nonsynonymous SNV	MAPK7	p.I510L	4,0
AA-3_2_CD8	AA CD8	12	25225628	C	G	nonsynonymous SNV	KRAS	p.A146P	3,0
AA-3_2_CD8	AA CD8	9	32526074	G	-	frameshift deletion	DDX58	p.P31fs	2,3
AA-3_2_CD8	AA CD8	20	51523337	C	T	nonsynonymous SNV	NFATC2	p.V282M	2,1
AA-3_2_CD8	AA CD8	18	62384865	C	T	nonsynonymous SNV	TNFRSF11A	p.A244V	2,9
AA-3_2_CD8	AA CD8	1	200174158	A	G	nonsynonymous SNV	NR5A2	p.D453G	4,1

AA-3_3_BM_CD8	AA CD8	18	334818	G	-	frameshift deletion	COLEC12	p.P580fs	3,3
AA-3_3_BM_CD8	AA CD8	12	25225628	C	G	nonsynonymous SNV	KRAS	p.A146P	4,4
AA-3_3_BM_CD8	AA CD8	11	27369145	-	A	splicing	LGR4		2,8
AA-3_3_BM_CD8	AA CD8	20	51523337	C	T	nonsynonymous SNV	NFATC2	p.V282M	7,6
AA-3_3_BM_CD8	AA CD8	19	51824563	T	C	nonsynonymous SNV	FPR3	p.L272S	3,2
AA-3_3_BM_CD8	AA CD8	2	130075335	C	T	nonsynonymous SNV	POTEF	p.G713S	16,1
AA-3_3_BM_CD8	AA CD8	2	130656786	G	A	nonsynonymous SNV	POTEJ	p.G676S	17,6
AA-3_3_BM_CD8	AA CD8	6	137875835	G	-	frameshift deletion	TNFAIP3	p.D212fs	3,9
AA-3_3_BM_CD8	AA CD8	6	161548996	C	T	nonsynonymous SNV	PRKN	p.R165Q	2,2
AA-3_3_BM_CD8	AA CD8	3	189880121	A	T	nonsynonymous SNV	TP63	p.K375N	2,9
AA-3_3_PB_CD8	AA CD8	9	2039793	A	-	frameshift deletion	SMARCA2	p.Q228fs	3,6
AA-3_3_PB_CD8	AA CD8	6	6318597	G	A	nonsynonymous SNV	F13A1	p.A23V	3,0
AA-3_3_PB_CD8	AA CD8	3	9893298	T	A	nonsynonymous SNV	JAGN1	p.V158E	3,5
AA-3_3_PB_CD8	AA CD8	12	25225628	C	G	nonsynonymous SNV	KRAS	p.A146P	7,8
AA-3_3_PB_CD8	AA CD8	17	31966179	G	T	nonsynonymous SNV	SUZ12	p.G140V	2,7
AA-3_3_PB_CD8	AA CD8	21	44366853	G	-	frameshift deletion	TRPM2	p.G175fs	2,6
AA-3_3_PB_CD8	AA CD8	20	51523337	C	T	nonsynonymous SNV	NFATC2	p.V282M	7,4
AA-3_3_PB_CD8	AA CD8	19	51824563	T	C	nonsynonymous SNV	FPR3	p.L272S	3,1
AA-3_3_PB_CD8	AA CD8	16	57026672	C	T	nonsynonymous SNV	NLRC5	p.L577F	2,2
AA-3_3_PB_CD8	AA CD8	10	71791196	GCT	-	nonframeshift deletion	CDH23	p.2038_2039del	2,8
AA-3_3_PB_CD8	AA CD8	7	82005473	C	A	nonsynonymous SNV	CACNA2D1	p.A514S	3,3
AA-3_3_PB_CD8	AA CD8	1	113859412	T	C	nonsynonymous SNV	PTPN22	p.T46A	3,5
AA-3_3_PB_CD8	AA CD8	X	124045368	G	A	nonsynonymous SNV	STAG2	p.A223T	2,0
AA-3_3_PB_CD8	AA CD8	6	137875835	G	-	frameshift deletion	TNFAIP3	p.D212fs	6,6
AA-3_3_PB_CD8	AA CD8	3	189880121	A	T	nonsynonymous SNV	TP63	p.K375N	4,2
AA-4_1_CD8	AA CD8	6	26234866	T	-	frameshift deletion	HIST1H1D	p.K23fs	2,2
AA-4_1_CD8	AA CD8	13	26676562	T	C	nonsynonymous SNV	WASF3	p.I185T	9,8
AA-4_1_CD8	AA CD8	11	27369077	A	G	nonsynonymous SNV	LGR4	p.F525S	9,4
AA-4_1_CD8	AA CD8	17	39791445	G	A	nonsynonymous SNV	IKZF3	p.A154V	16,8
AA-4_1_CD8	AA CD8	17	42322464	T	A	nonsynonymous SNV	STAT3	p.Y640F	17,7
AA-4_1_CD8	AA CD8	2	43824020	C	T	nonsynonymous SNV	ABCG5	p.R406Q	12,7
AA-4_1_CD8	AA CD8	8	47788994	T	C	nonsynonymous SNV	PRKDC	p.N3605S	6,8
AA-4_1_CD8	AA CD8	10	71791196	GCT	-	nonframeshift deletion	CDH23	p.2038_2039del	2,1
AA-4_1_CD8	AA CD8	1	119734662	C	G	nonsynonymous SNV	PHGDH	p.P180R	10,2
AA-4_1_CD8	AA CD8	1	158645571	C	A	nonsynonymous SNV	SPTA1	p.S1307I	4,7
AA-4_2_CD8	AA CD8	19	7899712	A	T	nonsynonymous SNV	LRRC8E	p.N268I	3,7
AA-4_2_CD8	AA CD8	16	24032142	C	T	nonsynonymous SNV	PRKCB	p.R99C	3,7
AA-4_2_CD8	AA CD8	15	41095776	T	G	nonsynonymous SNV	INO80	p.N99T	2,4
AA-4_2_CD8	AA CD8	12	49631475	C	-	frameshift deletion	PRPF40B	p.I25fs	1,9
AA-4_2_CD8	AA CD8	6	135463146	T	-	frameshift deletion	AHI1	p.T304fs	4,5
AA-4_2_CD8	AA CD8	2	159944930	-	T	frameshift insertion	PLA2R1	p.G1372fs	2,1
AA-4_2_CD8	AA CD8	1	174449175	G	C	nonsynonymous SNV	GPR52	p.R355P	3,4
AA-4_2_CD8	AA CD8	1	200173962	G	A	splicing	NR5A2		2,7
AA-4_2_CD8	AA CD8	2	227815495	A	T	nonsynonymous SNV	CCL20	p.I39F	2,4
AA-40_CD8	AA CD8	8	3162250	G	T	nonsynonymous SNV	CSMD1	p.P1918Q	7,7
AA-40_CD8	AA CD8	6	30899296	A	C	stoploss	DDR1	p.X768C	9,3
AA-40_CD8	AA CD8	6	31577176	C	T	nonsynonymous SNV	TNF	p.A114V	2,2
AA-40_CD8	AA CD8	21	44391350	G	A	nonsynonymous SNV	TRPM2	p.G507R	6,2
AA-40_CD8	AA CD8	10	102655202	C	G	nonsynonymous SNV	TRIM8	p.S231R	14,3
AA-40_CD8	AA CD8	3	120002204	CC	-	frameshift deletion	GSK3B	p.V41fs	7,1
AA-40_CD8	AA CD8	3	120002219	-	A	frameshift insertion	GSK3B	p.V37fs	8,0
AA-40_CD8	AA CD8	12	122989858	T	G	nonsynonymous SNV	PITPNM2	p.H887P	12,5
AA-40_CD8	AA CD8	1	155904795	T	C	nonsynonymous SNV	RIT1	p.Y22C	2,3
AA-40_CD8	AA CD8	2	213007409	C	T	nonsynonymous SNV	IKZF2	p.R511H	6,6
AA-41_CD8	AA CD8	16	3058483	C	-	frameshift deletion	MMP25	p.P411fs	4,0
AA-41_CD8	AA CD8	16	11967657	C	T	nonsynonymous SNV	TNFRSF17	p.T122M	2,3
AA-41_CD8	AA CD8	2	24918346	CTG	-	nonframeshift deletion	ADCY3	p.214_214del	5,7
AA-41_CD8	AA CD8	10	79307505	C	-	frameshift deletion	ZMIZ1	p.H923fs	2,7
AA-41_CD8	AA CD8	8	143918777	C	T	nonsynonymous SNV	PLEC	p.A3668T	3,7
AA-41_CD8	AA CD8	8	143934811	T	C	nonsynonymous SNV	PLEC	p.E301G	3,0
AA-41_CD8	AA CD8	5	148120312	AA	-	frameshift deletion	SPINK5	p.E820fs	3,0
AA-42_CD8	AA CD8	21	44321793	G	A	nonsynonymous SNV	PFKL	p.R419H	10,4
AA-42_CD8	AA CD8	3	58008763	C	A	nonsynonymous SNV	FLNB	p.H67N	4,3
AA-43_CD8	AA CD8	20	23084886	C	-	frameshift deletion	CD93	p.G436fs	2,6

AA-43_CD8	AA CD8	17	41586598	GCT	-	nonframeshift deletion	KRT14	p.79_79del	5,8
AA-43_CD8	AA CD8	17	42322464	T	A	nonsynonymous SNV	STAT3	p.Y640F	10,2
AA-43_CD8	AA CD8	21	44366853	G	-	frameshift deletion	TRPM2	p.G175fs	4,3
AA-43_CD8	AA CD8	11	64267482	G	T	nonsynonymous SNV	PLCB3	p.G1144C	5,9
AA-43_CD8	AA CD8	15	89801149	C	T	nonsynonymous SNV	ANPEP	p.G594D	3,4
AA-43_CD8	AA CD8	9	95922401	A	C	nonsynonymous SNV	ERCC6L2	p.S477R	14,1
AA-43_CD8	AA CD8	10	121551382	G	A	nonsynonymous SNV	FGFR2	p.R63C	4,2
AA-43_CD8	AA CD8	12	131915951	C	-	frameshift deletion	ULK1	p.A557fs	3,0
AA-43_CD8	AA CD8	9	136507992	G	A	nonsynonymous SNV	NOTCH1	p.A1158V	3,5
AA-43_CD8	AA CD8	8	143917124	C	T	nonsynonymous SNV	PLEC	p.A4219T	2,6
AA-43_CD8	AA CD8	8	143932459	CCT	-	nonframeshift deletion	PLEC	p.625_626del	2,3
AA-43_CD8	AA CD8	2	159944931	T	-	frameshift deletion	PLA2R1	p.K1371fs	2,6
AA-44_CD8	AA CD8	16	2180708	G	A	nonsynonymous SNV	CASKIN1	p.A887V	4,5
AA-44_CD8	AA CD8	16	3590712	G	T	nonsynonymous SNV	SLX4	p.H976N	2,3
AA-44_CD8	AA CD8	16	29664695	G	T	nonsynonymous SNV	SPN	p.G323W	3,0
AA-44_CD8	AA CD8	21	44318587	G	A	nonsynonymous SNV	PFKL	p.V352M	2,9
AA-44_CD8	AA CD8	X	78273664	G	T	nonsynonymous SNV	CYSLTR1	p.T28N	3,4
AA-44_CD8	AA CD8	9	88991879	T	-	frameshift deletion	C9orf47	p.L167fs	2,4
AA-44_CD8	AA CD8	12	112486532	G	A	nonsynonymous SNV	PTPN11	p.V432M	7,9
AA-44_CD8	AA CD8	9	131192225	A	T	nonsynonymous SNV	NUP214	p.T24S	6,3
AA-44_CD8	AA CD8	7	132179838	G	T	nonsynonymous SNV	PLXNA4	p.S1241R	2,9
AA-44_CD8	AA CD8	8	143932459	CCT	-	nonframeshift deletion	PLEC	p.625_626del	2,6
AA-44_CD8	AA CD8	2	216141226	G	T	nonsynonymous SNV	XRCC5	p.M461I	3,1
AA-45_CD8	AA CD8	19	4816453	G	-	frameshift deletion	TICAM1	p.P642fs	5,0
AA-45_CD8	AA CD8	9	14857747	G	T	nonsynonymous SNV	FREM1	p.L212M	2,3
AA-45_CD8	AA CD8	1	48736294	-	TT	frameshift insertion	BEND5	p.K182fs	3,0
AA-45_CD8	AA CD8	11	60468274	-	A	frameshift insertion	MS4A1	p.E234fs	3,0
AA-45_CD8	AA CD8	1	64841313	-	T	frameshift insertion	JAK1	p.P861fs	2,3
AA-45_CD8	AA CD8	10	79307505	C	-	frameshift deletion	ZMIZ1	p.H923fs	2,3
AA-45_CD8	AA CD8	11	102331110	C	G	nonsynonymous SNV	BIRC3	p.T398R	3,7
AA-45_CD8	AA CD8	3	138402184	A	-	frameshift deletion	MRAS	p.E105fs	2,2
AA-45_CD8	AA CD8	5	177508533	G	A	nonsynonymous SNV	DOK3	p.R26W	5,8
AA-45_CD8	AA CD8	2	230294309	-	A	frameshift insertion	SP140	p.R555fs	4,4
AA-46_CD8	AA CD8	17	4796222	G	T	stopgain	PSMB6	p.G10X	2,7
AA-46_CD8	AA CD8	1	67395689	A	-	frameshift deletion	IL12RB2	p.E644fs	2,7
AA-46_CD8	AA CD8	14	100239722	G	A	nonsynonymous SNV	YY1	p.G160S	2,1
AA-46_CD8	AA CD8	5	150054169	CTG	-	nonframeshift deletion	CSF1R	p.939_940del	2,5
AA-46_CD8	AA CD8	1	200174077	A	T	nonsynonymous SNV	NR5A2	p.Q426L	6,8
AA-47_CD8	AA CD8	19	13286524	G	-	frameshift deletion	CACNA1A	p.L1179fs	3,8
AA-47_CD8	AA CD8	9	14788920	T	-	frameshift deletion	FREM1	p.K1392fs	4,3
AA-47_CD8	AA CD8	X	21843270	G	A	nonsynonymous SNV	MBTPS2	p.R59H	2,5
AA-47_CD8	AA CD8	5	40764540	T	C	nonsynonymous SNV	PRKAA1	p.Y470C	6,0
AA-47_CD8	AA CD8	19	43071931	G	A	nonsynonymous SNV	PSG2	p.H245Y	2,8
AA-47_CD8	AA CD8	11	47579105	C	T	nonsynonymous SNV	NDUFS3	p.A5V	5,5
AA-47_CD8	AA CD8	16	66588031	A	-	frameshift deletion	CMTM2	p.E167fs	3,7
AA-47_CD8	AA CD8	10	79307505	C	-	frameshift deletion	ZMIZ1	p.H923fs	2,5
AA-47_CD8	AA CD8	10	89338944	-	G	frameshift insertion	IFIT3	p.W97fs	2,5
AA-47_CD8	AA CD8	1	186675998	T	C	nonsynonymous SNV	PTGS2	p.Q386R	4,5
AA-48_CD8	AA CD8	11	9749928	AG	-	frameshift deletion	SWAP70	p.E514fs	4,2
AA-48_CD8	AA CD8	8	22527512	A	-	frameshift deletion	PPP3CC	p.E355fs	3,1
AA-48_CD8	AA CD8	6	47609225	A	-	frameshift deletion	CD2AP	p.K579fs	5,7
AA-48_CD8	AA CD8	19	48234413	C	-	frameshift deletion	CARD8	p.D114fs	4,0
AA-48_CD8	AA CD8	12	103941642	G	A	nonsynonymous SNV	HSP90B1	p.R415H	3,6
AA-48_CD8	AA CD8	3	180947921	GAT	-	frameshift deletion	FXR1		3,5
AA-48_CD8	AA CD8	5	181242277	G	A	nonsynonymous SNV	RACK1	p.R60W	4,1
AA-48_CD8	AA CD8	2	197396092	G	A	nonsynonymous SNV	SF3B1	p.A1168V	2,2
AA-49_CD8	AA CD8	11	1759600	G	-	frameshift deletion	CTSD	p.Q90fs	5,2
AA-49_CD8	AA CD8	1	6462047	G	A	nonsynonymous SNV	TNFRSF25	p.A108V	7,8
AA-49_CD8	AA CD8	17	9022590	C	-	frameshift deletion	NTN1	p.P73fs	4,0
AA-49_CD8	AA CD8	14	24372470	C	T	nonsynonymous SNV	NFATC4	p.P409L	6,3
AA-49_CD8	AA CD8	2	24918346	CTG	-	nonframeshift deletion	ADCY3	p.214_214del	4,0
AA-49_CD8	AA CD8	13	35232515	C	T	nonsynonymous SNV	NBEA	p.A1891V	3,4
AA-49_CD8	AA CD8	1	41001077	G	A	nonsynonymous SNV	CTPS1	p.V196M	6,0
AA-49_CD8	AA CD8	19	45408613	-	A	frameshift insertion	CD3EAP	p.K217fs	4,5

AA-49_CD8	AA CD8	17	46876521	C	A	nonsynonymous SNV	WNT9B	p.P293T	6,5
AA-49_CD8	AA CD8	X	53405385	C	T	nonsynonymous SNV	SMC1A	p.A640T	3,7
AA-49_CD8	AA CD8	5	83541675	C	T	nonsynonymous SNV	VCAN	p.S1904F	3,0
AA-49_CD8	AA CD8	15	89805396	C	T	nonsynonymous SNV	ANPEP	p.A228T	4,2
AA-49_CD8	AA CD8	15	92967428	A	-	frameshift deletion	CHD2	p.K702fs	2,1
AA-49_CD8	AA CD8	X	108076203	C	T	nonsynonymous SNV	VSIG1	p.T272I	3,2
AA-49_CD8	AA CD8	1	116764568	A	G	nonsynonymous SNV	CD2	p.Y259C	6,3
AA-49_CD8	AA CD8	4	122456323	C	G	nonsynonymous SNV	IL2	p.D40H	2,6
AA-49_CD8	AA CD8	1	186965457	A	G	nonsynonymous SNV	PLA2G4A	p.K483R	2,1
AA-5_2_CD8	AA CD8	11	6391632	-	C	frameshift insertion	SMPD1	p.K189fs	20,6
AA-5_2_CD8	AA CD8	19	16425252	A	G	nonsynonymous SNV	EPS15L1	p.L208P	9,3
AA-5_2_CD8	AA CD8	2	25244613	CGT	-	nonframeshift deletion	DNMT3A	p.379_380del	2,1
AA-5_2_CD8	AA CD8	4	26415503	C	T	nonsynonymous SNV	RBPJ	p.L62F	6,2
AA-5_2_CD8	AA CD8	2	46941624	G	A	nonsynonymous SNV	MCFD2,TTC7	p.R14W	5,7
AA-5_2_CD8	AA CD8	2	65334634	A	-	frameshift deletion	SPRED2	p.V112fs	3,3
AA-5_2_CD8	AA CD8	11	108250708	A	T	nonsynonymous SNV	ATM	p.I415F	5,2
AA-5_2_CD8	AA CD8	6	160027223	C	-	frameshift deletion	IGF2R	p.P229fs	2,1
AA-50_CD8	AA CD8	11	9713523	A	-	frameshift deletion	SWAP70	p.K100fs	3,2
AA-50_CD8	AA CD8	16	10957804	A	-	frameshift deletion	CLEC16A	p.K35fs	2,2
AA-50_CD8	AA CD8	19	16889968	G	A	nonsynonymous SNV	F2RL3	p.A169T	3,8
AA-50_CD8	AA CD8	2	24827975	C	T	nonsynonymous SNV	ADCY3	p.V787M	2,3
AA-50_CD8	AA CD8	16	66566839	C	T	nonsynonymous SNV	CMTM1	p.A109V	2,3
AA-50_CD8	AA CD8	11	113205544	C	T	nonsynonymous SNV	NCAM1	p.A123V	1,7
AA-50_CD8	AA CD8	5	132073941	C	T	nonsynonymous SNV	CSF2	p.R40W	25,8
AA-50_CD8	AA CD8	6	160585167	G	A	stopgain	LPA	p.R1390X	25,6
AA-6_CD8	AA CD8	16	2967559	TGC	-	nonframeshift deletion	KREMEN2	p.339_340del	5,6
AA-6_CD8	AA CD8	20	3698027	G	-	frameshift deletion	SIGLEC1	p.P631fs	2,9
AA-6_CD8	AA CD8	12	9095695	G	T	stopgain	A2M	p.Y469X	9,8
AA-6_CD8	AA CD8	16	10957804	A	-	frameshift deletion	CLEC16A	p.K35fs	2,1
AA-6_CD8	AA CD8	1	32280202	C	G	nonsynonymous SNV	LCK	p.P389R	2,5
AA-6_CD8	AA CD8	6	33320958	G	A	nonsynonymous SNV	DAXX	p.R198W	2,1
AA-6_CD8	AA CD8	21	44366853	G	-	frameshift deletion	TRPM2	p.G175fs	2,8
AA-6_CD8	AA CD8	12	57471342	C	T	nonsynonymous SNV	GLI1	p.P740S	2,0
AA-6_CD8	AA CD8	15	66703338	GCG	-	nonframeshift deletion	SMAD6	p.27_28del	3,3
AA-6_CD8	AA CD8	11	72434313	C	-	frameshift deletion	CLPB	p.G54fs	3,6
AA-6_CD8	AA CD8	11	75284234	C	T	splicing	ARRB1		5,8
AA-6_CD8	AA CD8	9	99828485	T	C	nonsynonymous SNV	NR4A3	p.F148S	3,7
AA-6_CD8	AA CD8	12	111447460	-	G	frameshift insertion	SH2B3	p.D182fs	29,9
AA-7_1_CD8	AA CD8	12	12908969	G	A	stopgain	GPRC5A	p.W240X	2,6
AA-7_1_CD8	AA CD8	17	19958119	G	C	nonsynonymous SNV	AKAP10	p.H258D	3,0
AA-7_1_CD8	AA CD8	16	66565992	A	-	frameshift deletion	CKLF	p.E62fs	1,6
AA-7_1_CD8	AA CD8	17	75737568	T	C	nonsynonymous SNV	ITGB4	p.I715T	2,2
AA-7_1_CD8	AA CD8	X	77698600	C	A	nonsynonymous SNV	ATRX	p.D55Y	2,4
AA-7_1_CD8	AA CD8	11	108271285	G	T	nonsynonymous SNV	ATM	p.V986F	3,0
AA-7_2_CD8	AA CD8	1	2629382	G	C	nonsynonymous SNV	MMEL1	p.L35V	2,9
AA-7_2_CD8	AA CD8	18	12836879	C	G	nonsynonymous SNV	PTPN2	p.R29P	5,8
AA-7_2_CD8	AA CD8	17	19958119	G	C	nonsynonymous SNV	AKAP10	p.H258D	5,0
AA-7_2_CD8	AA CD8	13	26685699	A	-	frameshift deletion	WASF3	p.K452fs	1,7
AA-7_2_CD8	AA CD8	21	44366853	G	-	frameshift deletion	TRPM2	p.G175fs	2,2
AA-7_2_CD8	AA CD8	7	48192971	G	T	nonsynonymous SNV	ABCA13	p.A28S	3,5
AA-7_2_CD8	AA CD8	13	48303979	CCG	-	nonframeshift deletion	RB1	p.23_23del	2,4
AA-7_2_CD8	AA CD8	1	64873428	T	-	frameshift deletion	JAK1	p.K142fs	2,2
AA-7_2_CD8	AA CD8	X	77698600	C	A	nonsynonymous SNV	ATRX	p.D55Y	5,5
AA-7_2_CD8	AA CD8	14	105470183	G	A	nonsynonymous SNV	MTA1	p.V706I	4,8
HC-1_CD4	healthy CD4	19	55711924	C	T	nonsynonymous SNV	NLRP9	p.A907T	2,2
HC-1_CD4	healthy CD4	8	143925208	G	A	nonsynonymous SNV	PLEC	p.A1560V	4,3
HC-1_CD4	healthy CD4	3	180948439	-	A	frameshift insertion	FXR1	p.V121fs	1,3
HC-10_CD4	healthy CD4	16	2189459	A	G	nonsynonymous SNV	CASKIN1	p.I117T	3,1
HC-10_CD4	healthy CD4	2	39120337	T	C	nonsynonymous SNV	SOS1	p.K29R	3,0
HC-10_CD4	healthy CD4	3	141487133	C	T	nonsynonymous SNV	RASA2	p.A17V	4,2
HC-10_CD4	healthy CD4	1	156298799	C	A	nonsynonymous SNV	VHLL	p.A131S	2,4
HC-11_CD4	healthy CD4	19	17844303	G	-	frameshift deletion	JAK3	p.Q39fs	3,4
HC-11_CD4	healthy CD4	15	90476692	-	A	frameshift insertion	IQGAP1	p.T938fs	2,6
HC-12_CD4	healthy CD4	18	334818	G	-	frameshift deletion	COLEC12	p.P580fs	3,0

HC-12_CD4	healthy CD4	16	1766530	G	T	stopgain	MAPK8IP3	p.E934X	2,7
HC-12_CD4	healthy CD4	15	92967428	A	-	frameshift deletion	CHD2	p.K702fs	2,7
HC-12_CD4	healthy CD4	11	116836148	C	T	nonsynonymous SNV	APOA1	p.R155H	3,1
HC-13_CD4	healthy CD4	1	1212061	G	-	frameshift deletion	TNFRSF4	p.P172fs	6,1
HC-13_CD4	healthy CD4	8	39917899	G	T	nonsynonymous SNV	IDO1	p.D38Y	3,5
HC-13_CD4	healthy CD4	5	150054169	CTG	-	nonframeshift deletion	CSF1R	p.939_940del	3,2
HC-13_CD4	healthy CD4	2	178450415	C	A	nonsynonymous SNV	PRKRA	p.C10F	2,7
HC-14_CD4	healthy CD4	3	32502661	A	G	nonsynonymous SNV	CMTM6	p.F29L	5,0
HC-14_CD4	healthy CD4	19	55738356	A	-	frameshift deletion	NLRP9	p.S7fs	2,3
HC-14_CD4	healthy CD4	12	94149380	C	T	nonsynonymous SNV	PLXNC1	p.R137W	8,3
HC-14_CD4	healthy CD4	9	95111535	G	-	frameshift deletion	FANCC	p.P419fs	3,2
HC-14_CD4	healthy CD4	3	123286692	C	T	nonsynonymous SNV	ADCY5	p.R867H	2,4
HC-15_CD4	healthy CD4	6	160532553	G	A	nonsynonymous SNV	LPA	p.A1980V	4,8
HC-16_CD4	healthy CD4	2	15613226	A	G	nonsynonymous SNV	DDX1	p.E320G	2,1
HC-16_CD4	healthy CD4	7	82005473	C	A	nonsynonymous SNV	CACNA2D1	p.A514S	3,8
HC-16_CD4	healthy CD4	6	89411928	C	A	nonsynonymous SNV	RRAGD	p.E22D	7,1
HC-16_CD4	healthy CD4	12	107319300	C	-	frameshift deletion	BTBD11	p.V120fs	2,3
HC-16_CD4	healthy CD4	3	154315189	T	-	frameshift deletion	DHX36	p.M154fs	2,0
HC-17_CD4	healthy CD4	7	2002123	T	A	splicing	MAD1L1		4,7
HC-17_CD4	healthy CD4	11	6391624	C	-	frameshift deletion	SMPD1	p.P187fs	4,1
HC-17_CD4	healthy CD4	10	18140783	C	T	nonsynonymous SNV	CACNB2	p.A16V	6,6
HC-17_CD4	healthy CD4	7	83977195	G	A	nonsynonymous SNV	SEMA3A	p.R552C	2,6
HC-17_CD4	healthy CD4	11	89490908	G	-	frameshift deletion	NOX4	p.P22fs	3,1
HC-18_CD4	healthy CD4	19	43071931	G	A	nonsynonymous SNV	PSG2	p.H245Y	3,1
HC-18_CD4	healthy CD4	12	57185564	G	A	nonsynonymous SNV	LRP1	p.C2166Y	2,8
HC-18_CD4	healthy CD4	11	60468274	-	A	frameshift insertion	MS4A1	p.E234fs	3,1
HC-18_CD4	healthy CD4	9	88991636	G	A	nonsynonymous SNV	C9orf47	p.R138Q	3,7
HC-18_CD4	healthy CD4	14	99231443	A	G	nonsynonymous SNV	BCL11B	p.L180P	3,1
HC-19_CD4	healthy CD4	3	9911574	C	-	frameshift deletion	IL17RE	p.P286fs	2,8
HC-19_CD4	healthy CD4	16	10907448	C	-	frameshift deletion	CIITA	p.S652fs	2,4
HC-2_CD4	healthy CD4	12	7065935	A	-	frameshift deletion	C1S	p.Q112fs	1,1
HC-2_CD4	healthy CD4	17	29110525	G	A	nonsynonymous SNV	MYO18A	p.R1000C	2,2
HC-2_CD4	healthy CD4	3	38581154	-	A	frameshift insertion	SCN5A	p.P1002fs	4,2
HC-2_CD4	healthy CD4	19	46320484	C	T	nonsynonymous SNV	HIF3A	p.T287M	2,1
HC-2_CD4	healthy CD4	20	48650031	C	T	nonsynonymous SNV	PREX1	p.R998H	2,2
HC-2_CD4	healthy CD4	16	67154780	G	A	nonsynonymous SNV	TRADD	p.R270W	3,7
HC-2_CD4	healthy CD4	10	110498419	C	T	stopgain	DUSP5	p.R100X	3,0
HC-2_CD4	healthy CD4	6	136792005	CGG	-	nonframeshift deletion	MAP3K5	p.51_51del	2,8
HC-20_CD4	healthy CD4	16	11345611	C	T	nonsynonymous SNV	RMI2	p.A47V	4,5
HC-20_CD4	healthy CD4	19	14117529	C	T	nonsynonymous SNV	PRKACA	p.A7T	12,9
HC-20_CD4	healthy CD4	20	31666016	C	T	nonsynonymous SNV	BCL2L1	p.R149H	2,0
HC-20_CD4	healthy CD4	2	60918408	C	T	nonsynonymous SNV	REL	p.R219C	2,6
HC-20_CD4	healthy CD4	11	86245255	C	T	nonsynonymous SNV	EED	p.A9V	2,5
HC-21_CD4	healthy CD4	18	62354427	C	T	nonsynonymous SNV	TNFRSF11A	p.T107M	4,3
HC-21_CD4	healthy CD4	6	100540378	C	T	nonsynonymous SNV	ASCC3	p.E1854K	3,5
HC-21_CD4	healthy CD4	1	114716123	C	T	nonsynonymous SNV	NRAS	p.G13D	2,0
HC-3_CD4	healthy CD4	2	11192156	G	T	nonsynonymous SNV	ROCK2	p.N1299K	4,5
HC-3_CD4	healthy CD4	16	31410790	C	-	frameshift deletion	ITGAD	p.A423fs	3,5
HC-3_CD4	healthy CD4	X	53421893	A	C	splicing	SMC1A		2,7
HC-3_CD4	healthy CD4	3	57165208	C	T	nonsynonymous SNV	IL17RD	p.A27T	5,3
HC-3_CD4	healthy CD4	7	140783127	G	-	frameshift deletion	BRAF	p.P403fs	2,0
HC-4_CD4	healthy CD4	3	9933285	G	T	nonsynonymous SNV	IL17RC	p.D589Y	3,8
HC-4_CD4	healthy CD4	16	29985697	G	A	nonsynonymous SNV	TAOK2	p.A610T	4,2
HC-4_CD4	healthy CD4	21	44427098	G	A	nonsynonymous SNV	TRPM2	p.A1321T	4,0
HC-4_CD4	healthy CD4	5	74636373	A	G	nonsynonymous SNV	ENC1	p.L38P	2,4
HC-4_CD4	healthy CD4	7	108066501	C	T	nonsynonymous SNV	LAMB4	p.R849H	2,3
HC-4_CD4	healthy CD4	12	112141000	A	G	nonsynonymous SNV	TRAFD1	p.E140G	2,7
HC-4_CD4	healthy CD4	3	140683138	G	A	nonsynonymous SNV	TRIM42	p.A340T	2,3
HC-4_CD4	healthy CD4	3	189867903	G	A	nonsynonymous SNV	TP63	p.R139H	2,3
HC-4_CD4	healthy CD4	3	194396974	G	A	nonsynonymous SNV	GP5	p.R437W	3,4
HC-4_CD4	healthy CD4	3	197010748	G	A	nonsynonymous SNV	MELTF	p.T427M	3,7
HC-4_CD4	healthy CD4	2	219224395	C	-	frameshift deletion	ATG9A	p.A326fs	2,9
HC-5_CD4	healthy CD4	17	1643883	G	T	nonsynonymous SNV	SCARF1	p.A117E	6,7
HC-5_CD4	healthy CD4	22	23580104	CAG	-	nonframeshift deletion	IGLL1	p.29_29del	6,1

HC-5_CD4	healthy CD4	7	27184450	T	C	nonsynonymous SNV	HOXA11	p.K232R	4,2
HC-5_CD4	healthy CD4	8	47954350	TT	-	frameshift deletion	PRKDC	p.K165fs	2,4
HC-6_CD4	healthy CD4	3	30650380	AA	-	frameshift deletion	TGFBR2	p.E125fs	1,5
HC-6_CD4	healthy CD4	19	40667975	C	-	frameshift deletion	NUMBL	p.Q400fs	4,3
HC-6_CD4	healthy CD4	19	40667977	GC	-	frameshift deletion	NUMBL	p.Q399fs	4,1
HC-6_CD4	healthy CD4	5	177404622	C	T	nonsynonymous SNV	F12	p.R226H	3,4
HC-8_CD4	healthy CD4	12	8129340	-	A	frameshift insertion	CLEC4A	p.V53fs	2,2
HC-8_CD4	healthy CD4	12	49033106	G	A	stopgain	KMT2D	p.Q3867X	3,1
HC-8_CD4	healthy CD4	15	98948617	C	T	stopgain	IGF1R	p.Q1211X	2,6
HC-1_CD8	healthy CD8	18	48949961	A	G	nonsynonymous SNV	SMAD7	p.L155P	2,4
HC-1_CD8	healthy CD8	1	64142566	A	C	nonsynonymous SNV	ROR1	p.N364H	2,4
HC-10_CD8	healthy CD8	16	3728370	G	A	nonsynonymous SNV	CREBBP	p.A2188V	2,4
HC-10_CD8	healthy CD8	4	152347051	C	A	nonsynonymous SNV	FBXW7	p.C84F	8,0
HC-11_CD8	healthy CD8	20	34445308	C	A	nonsynonymous SNV	ITCH	p.N219K	2,3
HC-12_CD8	healthy CD8	19	45885548	G	T	nonsynonymous SNV	IRF2BP1	p.A76D	2,9
HC-12_CD8	healthy CD8	11	112961661	G	T	nonsynonymous SNV	NCAM1	p.A17S	2,1
HC-13_CD8	healthy CD8	19	16890260	C	T	nonsynonymous SNV	F2RL3	p.A266V	2,5
HC-13_CD8	healthy CD8	15	41056702	G	A	nonsynonymous SNV	INO80	p.R664C	6,3
HC-13_CD8	healthy CD8	4	54285866	G	A	nonsynonymous SNV	PDGFRA	p.R822H	1,7
HC-13_CD8	healthy CD8	12	107581159	C	T	nonsynonymous SNV	BTBD11	p.L37F	2,7
HC-13_CD8	healthy CD8	4	159350173	C	T	nonsynonymous SNV	RAPGEF2	p.S1089F	2,1
HC-14_CD8	healthy CD8	9	377066	C	A	nonsynonymous SNV	DOCK8	p.S697R	4,3
HC-14_CD8	healthy CD8	16	2173540	G	A	nonsynonymous SNV	TRAF7	p.A358T	2,5
HC-14_CD8	healthy CD8	17	17823608	C	T	splicing	SREBF1		3,4
HC-14_CD8	healthy CD8	12	40320067	-	A	frameshift insertion	LRRK2	p.S1636fs	2,0
HC-14_CD8	healthy CD8	12	40320068	A	-	frameshift deletion	LRRK2	p.S1636fs	2,0
HC-14_CD8	healthy CD8	17	42310529	G	A	nonsynonymous SNV	STAT5A	p.D719N	2,0
HC-14_CD8	healthy CD8	17	42844820	C	T	nonsynonymous SNV	AOC2	p.T65I	4,3
HC-14_CD8	healthy CD8	10	71791196	GCT	-	nonframeshift deletion	CDH23	p.2038_2039del	2,0
HC-15_CD8	healthy CD8	20	3801047	G	A	nonsynonymous SNV	CDC25B	p.R28H	2,5
HC-15_CD8	healthy CD8	6	36373580	T	G	splicing	ETV7		8,0
HC-15_CD8	healthy CD8	22	41149902	A	C	nonsynonymous SNV	EP300	p.T841P	3,2
HC-15_CD8	healthy CD8	17	42217243	C	G	nonsynonymous SNV	STAT5B	p.E433Q	2,4
HC-15_CD8	healthy CD8	19	43184803	A	G	nonsynonymous SNV	PSG5	p.Y137H	4,7
HC-15_CD8	healthy CD8	3	49302514	A	G	nonsynonymous SNV	USP4	p.F339S	4,4
HC-15_CD8	healthy CD8	5	111072933	G	A	splicing	TSLP		5,8
HC-16_CD8	healthy CD8	11	14644608	A	G	nonsynonymous SNV	PDE3B	p.D178G	2,1
HC-16_CD8	healthy CD8	12	68158261	T	A	splicing	IFNG		10,3
HC-16_CD8	healthy CD8	15	74433798	C	T	nonsynonymous SNV	SEMA7A	p.A41T	6,3
HC-16_CD8	healthy CD8	1	100731268	C	G	nonsynonymous SNV	VCAM1	p.C333W	2,3
HC-16_CD8	healthy CD8	3	184365148	T	G	nonsynonymous SNV	POLR2H	p.L58W	3,5
HC-17_CD8	healthy CD8	7	6446094	T	G	nonsynonymous SNV	DAGLB	p.I36L	3,1
HC-17_CD8	healthy CD8	11	18245442	T	C	nonsynonymous SNV	SAA2	p.K102E	2,1
HC-17_CD8	healthy CD8	6	26156572	CTCTC	-	frameshift deletion	HIST1H1E	p.A61fs	3,7
HC-17_CD8	healthy CD8	21	31251838	C	G	nonsynonymous SNV	TIAM1	p.A439P	2,6
HC-17_CD8	healthy CD8	16	31266064	A	T	nonsynonymous SNV	ITGAM	p.E115V	5,2
HC-17_CD8	healthy CD8	21	31266560	T	A	nonsynonymous SNV	TIAM1	p.D138V	4,4
HC-17_CD8	healthy CD8	20	33676813	G	C	stopgain	E2F1	p.Y411X	3,3
HC-17_CD8	healthy CD8	22	50603951	G	-	frameshift deletion	MAPK8IP2	p.G218fs	5,7
HC-17_CD8	healthy CD8	19	54633060	A	T	nonsynonymous SNV	LILRB1	p.T335S	8,0
HC-17_CD8	healthy CD8	7	108123139	A	-	frameshift deletion	LAMB4	p.L9fs	1,5
HC-17_CD8	healthy CD8	6	135463145	-	T	frameshift insertion	AHI1	p.T304fs	2,2
HC-17_CD8	healthy CD8	5	148827319	T	G	nonsynonymous SNV	ADRB2	p.L163R	5,6
HC-18_CD8	healthy CD8	X	1341780	G	A	stopgain	IL3RA	p.W5X	2,7
HC-18_CD8	healthy CD8	18	11881049	G	A	nonsynonymous SNV	GNAL	p.V147M	9,1
HC-18_CD8	healthy CD8	7	84011182	C	T	splicing	SEMA3A		3,7
HC-18_CD8	healthy CD8	9	91409924	A	G	nonsynonymous SNV	NFIL3	p.S271P	2,2
HC-18_CD8	healthy CD8	10	119430388	A	-	frameshift deletion	GRK5	p.K183fs	2,1
HC-18_CD8	healthy CD8	8	143932459	CCT	-	nonframeshift deletion	PLEC	p.625_626del	3,7
HC-18_CD8	healthy CD8	5	180269371	A	G	nonsynonymous SNV	MAPK9	p.V54A	2,3
HC-19_CD8	healthy CD8	5	1293640	G	A	stopgain	TERT	p.R416X	4,1
HC-19_CD8	healthy CD8	17	1362010	T	A	splicing	YWHAE		14,5
HC-19_CD8	healthy CD8	15	66703750	GCT	-	nonframeshift deletion	SMAD6	p.164_165del	3,6
HC-2_CD8	healthy CD8	17	75737568	T	C	nonsynonymous SNV	ITGB4	p.I715T	2,5

HC-2_CD8	healthy CD8	3	197675055	G	A	nonsynonymous SNV	RUBCN	p.A961V	3,2
HC-2_CD8	healthy CD8	1	207468699	C	-	frameshift deletion	CR2	p.V206fs	1,0
HC-20_CD8	healthy CD8	11	320713	G	A	nonsynonymous SNV	IFITM3	p.A34V	3,1
HC-20_CD8	healthy CD8	11	47667658	-	T	frameshift insertion	AGBL2	p.K751fs	2,3
HC-20_CD8	healthy CD8	7	48508005	C	A	nonsynonymous SNV	ABCA13	p.L4494I	2,7
HC-20_CD8	healthy CD8	1	89133386	GT	AC	nonframeshift substitut	GBP7		6,3
HC-20_CD8	healthy CD8	10	122576623	C	T	nonsynonymous SNV	DMBT1	p.R170C	4,1
HC-20_CD8	healthy CD8	11	129870009	T	C	nonsynonymous SNV	NFRKB	p.I1031V	5,7
HC-20_CD8	healthy CD8	5	181204737	G	A	nonsynonymous SNV	TRIM7	p.A125V	3,7
HC-21_CD8	healthy CD8	17	4716539	-	C	frameshift insertion	ARRB2	p.R81fs	6,7
HC-21_CD8	healthy CD8	20	46013347	G	C	nonsynonymous SNV	MMP9	p.V475L	3,0
HC-21_CD8	healthy CD8	5	53062931	T	A	splicing	ITGA2		7,1
HC-21_CD8	healthy CD8	15	90441520	A	T	nonsynonymous SNV	IQGAP1	p.I222F	4,0
HC-3_CD8	healthy CD8	20	1578678	G	C	nonsynonymous SNV	SIRPB1	p.D31E	4,7
HC-3_CD8	healthy CD8	16	2076509	C	T	nonsynonymous SNV	TSC2	p.L721F	3,3
HC-3_CD8	healthy CD8	9	95941479	G	T	nonsynonymous SNV	ERCC6L2	p.V604F	3,7
HC-3_CD8	healthy CD8	1	156873710	G	A	nonsynonymous SNV	NTRK1	p.A310T	2,1
HC-4_CD8	healthy CD8	16	2179268	GGC	-	nonframeshift deletion	CASKIN1	p.1277_1278del	6,3
HC-4_CD8	healthy CD8	1	11046966	G	-	frameshift deletion	MASP2	p.P53fs	2,6
HC-4_CD8	healthy CD8	19	54990027	T	G	nonsynonymous SNV	NLRP2	p.V769G	2,7
HC-4_CD8	healthy CD8	19	56032615	C	T	nonsynonymous SNV	NLRP5	p.R761W	2,3
HC-4_CD8	healthy CD8	11	64369566	G	A	nonsynonymous SNV	RPS6KA4	p.V511M	2,3
HC-4_CD8	healthy CD8	1	64841313	-	T	frameshift insertion	JAK1	p.P861fs	2,6
HC-4_CD8	healthy CD8	2	190995094	C	T	nonsynonymous SNV	STAT1	p.R304H	2,5
HC-4_CD8	healthy CD8	1	206685319	C	-	frameshift deletion	MAPKAPK2	p.H30fs	4,3
HC-5_CD8	healthy CD8	17	28537368	C	-	frameshift deletion	FOXN1	p.P627fs	2,5
HC-5_CD8	healthy CD8	13	32746085	T	G	nonsynonymous SNV	PDS5B	p.C907W	2,4
HC-5_CD8	healthy CD8	3	58104061	G	-	frameshift deletion	FLNB	p.W529fs	2,6
HC-5_CD8	healthy CD8	11	128758118	G	A	nonsynonymous SNV	FLI1	p.A8T	2,0
HC-6_CD8	healthy CD8	1	41362186	A	G	nonsynonymous SNV	FOXO6	p.N86D	5,4
HC-6_CD8	healthy CD8	16	67174734	G	-	frameshift deletion	NOL3	p.P133fs	2,0
HC-6_CD8	healthy CD8	15	74433809	AGC	-	nonframeshift deletion	SEMA7A	p.36_37del	4,7
HC-7_CD8	healthy CD8	6	109467736	-	A	splicing	ZBTB24		3,3
HC-7_CD8	healthy CD8	11	116837116	G	-	frameshift deletion	APOA1	p.Q29fs	3,7
HC-7_CD8	healthy CD8	9	128635189	A	G	nonsynonymous SNV	WDR34	p.V295A	4,6
HC-7_CD8	healthy CD8	5	173234846	C	T	nonsynonymous SNV	NKX2-5	p.A80T	2,3
HC-7_CD8	healthy CD8	2	187467754	T	-	frameshift deletion	TFPI	p.K269fs	2,2
HC-7_CD8	healthy CD8	2	230814047	AG	TT	splicing	CAB39		3,3
HC-8_CD8	healthy CD8	20	408739	A	G	nonsynonymous SNV	RBCK1	p.M1V	2,1
HC-8_CD8	healthy CD8	2	25240306	A	G	nonsynonymous SNV	DNMT3A	p.L621P	4,2
HC-8_CD8	healthy CD8	22	39521946	C	-	frameshift deletion	ATF4	p.P134fs	1,2
HC-8_CD8	healthy CD8	9	130880097	T	C	nonsynonymous SNV	ABL1	p.S485P	5,5
HC-8_CD8	healthy CD8	9	130884637	C	-	frameshift deletion	ABL1	p.P783fs	2,0
HC-8_CD8	healthy CD8	5	170270842	C	G	nonsynonymous SNV	LCP2	p.E134Q	5,1
hMDS-1_CD4	hMDS CD4	3	390997	C	A	nonsynonymous SNV	CHL1	p.H877N	2,0
hMDS-1_CD4	hMDS CD4	8	27610498	G	A	nonsynonymous SNV	CLU	p.T25M	3,4
hMDS-1_CD4	hMDS CD4	3	30650380	AA	-	frameshift deletion	TGFBR2	p.E125fs	1,9
hMDS-1_CD4	hMDS CD4	18	41970416	G	T	nonsynonymous SNV	PIK3C3	p.S101I	2,4
hMDS-1_CD4	hMDS CD4	2	43828025	G	A	nonsynonymous SNV	ABCG5	p.R198W	3,4
hMDS-1_CD4	hMDS CD4	20	48651578	G	T	nonsynonymous SNV	PREX1	p.P825T	3,6
hMDS-1_CD4	hMDS CD4	3	52505683	T	C	nonsynonymous SNV	STAB1	p.S533P	4,2
hMDS-1_CD4	hMDS CD4	17	60600712	T	-	frameshift deletion	PPM1D	p.F100fs	3,4
hMDS-1_CD4	hMDS CD4	11	66468160	C	-	frameshift deletion	PELI3	p.S11fs	5,1
hMDS-1_CD4	hMDS CD4	2	203870733	C	T	nonsynonymous SNV	CTLA4	p.A86V	2,2
hMDS-10_CD4	hMDS CD4	20	1578678	G	C	nonsynonymous SNV	SIRPB1	p.D31E	2,1
hMDS-10_CD4	hMDS CD4	17	19382248	A	C	nonsynonymous SNV	MAPK7	p.I510L	4,7
hMDS-10_CD4	hMDS CD4	9	27524902	A	-	frameshift deletion	IFNK	p.E189fs	2,0
hMDS-10_CD4	hMDS CD4	22	41129965	T	C	nonsynonymous SNV	EP300	p.L415P	1,0
hMDS-10_CD4	hMDS CD4	12	49631475	C	-	frameshift deletion	PRPF40B	p.I25fs	2,5
hMDS-10_CD4	hMDS CD4	2	60918414	G	T	nonsynonymous SNV	REL	p.V221L	2,7
hMDS-10_CD4	hMDS CD4	10	79307505	C	-	frameshift deletion	ZMIZ1	p.H923fs	2,2
hMDS-10_CD4	hMDS CD4	2	99601551	T	C	nonsynonymous SNV	AFF3	p.S444G	2,7
hMDS-10_CD4	hMDS CD4	8	102262044	A	-	frameshift deletion	UBR5	p.F2570fs	1,1
hMDS-11_CD4	hMDS CD4	21	44441747	G	A	nonsynonymous SNV	TRPM2	p.R162H	4,1

hMDS-11_CD4	hMDS CD4	X	47637935	G	A	nonsynonymous SNV	ELK1	p.A301V	4,8
hMDS-11_CD4	hMDS CD4	X	53413102	G	A	nonsynonymous SNV	SMC1A	p.R218W	2,2
hMDS-11_CD4	hMDS CD4	16	67154780	G	A	nonsynonymous SNV	TRADD	p.R270W	9,5
hMDS-11_CD4	hMDS CD4	X	78273075	-	T	frameshift insertion	CYSLTR1	p.N224fs	2,8
hMDS-11_CD4	hMDS CD4	12	103941642	G	A	nonsynonymous SNV	HSP90B1	p.R415H	2,4
hMDS-11_CD4	hMDS CD4	9	108911154	G	T	nonsynonymous SNV	ELP1	p.Q57K	2,4
hMDS-11_CD4	hMDS CD4	1	181796706	C	T	nonsynonymous SNV	CACNA1E	p.R2021W	3,5
hMDS-13_1_CD4	hMDS CD4	20	3796680	C	T	nonsynonymous SNV	CDC25B	p.S50L	4,1
hMDS-13_1_CD4	hMDS CD4	4	15935996	A	T	nonsynonymous SNV	FGFBP1	p.F213I	9,2
hMDS-13_1_CD4	hMDS CD4	8	16993295	A	T	nonsynonymous SNV	FGF20	p.I138N	8,2
hMDS-13_1_CD4	hMDS CD4	19	35332634	G	A	stopgain	CD22	p.W41X	4,4
hMDS-13_1_CD4	hMDS CD4	17	39408782	A	G	nonsynonymous SNV	MED1	p.S1147P	3,7
hMDS-13_1_CD4	hMDS CD4	22	39521946	C	-	frameshift deletion	ATF4	p.P134fs	2,5
hMDS-13_1_CD4	hMDS CD4	X	49217788	G	A	nonsynonymous SNV	CACNA1F	p.T1019M	2,2
hMDS-13_1_CD4	hMDS CD4	14	61721609	G	T	nonsynonymous SNV	HIF1A	p.M133I	6,0
hMDS-13_1_CD4	hMDS CD4	1	67395689	A	-	frameshift deletion	IL12RB2	p.E644fs	2,3
hMDS-13_1_CD4	hMDS CD4	7	91265979	G	A	nonsynonymous SNV	FZD1	p.V367M	2,6
hMDS-13_1_CD4	hMDS CD4	6	111661739	T	A	stoploss	FYN	p.X535Y	5,4
hMDS-13_1_CD4	hMDS CD4	X	124068618	C	-	frameshift deletion	STAG2	p.H774fs	2,9
hMDS-13_1_CD4	hMDS CD4	3	136398759	G	C	nonsynonymous SNV	STAG1	p.S756C	2,8
hMDS-13_1_CD4	hMDS CD4	X	139548468	A	T	nonsynonymous SNV	F9	p.N128I	7,1
hMDS-13_1_CD4	hMDS CD4	5	158707989	G	C	nonsynonymous SNV	EBF1	p.N519K	6,8
hMDS-13_1_CD4	hMDS CD4	3	183152487	C	T	nonsynonymous SNV	LAMP3	p.R259K	2,2
hMDS-13_1_CD4	hMDS CD4	1	196831856	G	A	nonsynonymous SNV	CFHR1	p.A284T	5,5
hMDS-13_1_CD4	hMDS CD4	2	230202686	C	A	nonsynonymous SNV	SP110	p.R314M	4,0
hMDS-13_2_CD4	hMDS CD4	20	1578678	G	C	nonsynonymous SNV	SIRPB1	p.D31E	5,3
hMDS-13_2_CD4	hMDS CD4	19	1622414	A	C	nonsynonymous SNV	TCF3	p.V184G	7,3
hMDS-13_2_CD4	hMDS CD4	8	16993295	A	T	nonsynonymous SNV	FGF20	p.I138N	2,1
hMDS-13_2_CD4	hMDS CD4	17	40088308	G	A	nonsynonymous SNV	THRA	p.A264T	6,7
hMDS-13_2_CD4	hMDS CD4	22	41149041	A	T	nonsynonymous SNV	EP300	p.M749L	4,3
hMDS-13_2_CD4	hMDS CD4	3	49903542	G	A	nonsynonymous SNV	MST1R	p.A23V	3,3
hMDS-13_2_CD4	hMDS CD4	1	207468699	C	-	frameshift deletion	CR2	p.V206fs	1,4
hMDS-5_CD4	hMDS CD4	16	10907448	C	-	frameshift deletion	CIITA	p.S652fs	2,3
hMDS-5_CD4	hMDS CD4	15	41565128	C	T	nonsynonymous SNV	TYRO3	p.S212F	11,5
hMDS-5_CD4	hMDS CD4	1	46055910	G	A	nonsynonymous SNV	LOC110117498	p.R120C	3,0
hMDS-5_CD4	hMDS CD4	19	49971868	C	T	nonsynonymous SNV	SIGLEC16	p.T357I	3,1
hMDS-5_CD4	hMDS CD4	16	68684601	A	-	frameshift deletion	CDH3	p.K346fs	2,2
hMDS-5_CD4	hMDS CD4	15	78937393	G	-	frameshift deletion	CTSH	p.H52fs	4,1
hMDS-5_CD4	hMDS CD4	14	95535198	GGC	-	nonframeshift deletion	GLRX5	p.37_37del	3,9
hMDS-1_CD8	hMDS CD8	19	4816453	G	-	frameshift deletion	TICAM1	p.P642fs	2,7
hMDS-1_CD8	hMDS CD8	6	26156933	A	-	frameshift deletion	HIST1H1E	p.P181fs	2,2
hMDS-1_CD8	hMDS CD8	11	27369145	-	A	splicing	LGR4		2,6
hMDS-1_CD8	hMDS CD8	17	31993876	C	A	nonsynonymous SNV	SUZ12	p.N412K	3,1
hMDS-1_CD8	hMDS CD8	9	34568930	C	T	nonsynonymous SNV	CNTFR	p.A18T	2,7
hMDS-1_CD8	hMDS CD8	7	41699996	C	G	nonsynonymous SNV	INHBA	p.A127P	2,7
hMDS-1_CD8	hMDS CD8	21	44366853	G	-	frameshift deletion	TRPM2	p.G175fs	2,9
hMDS-1_CD8	hMDS CD8	3	46265851	-	A	frameshift insertion	CCR3	p.S231fs	2,5
hMDS-1_CD8	hMDS CD8	3	49028282	G	T	nonsynonymous SNV	IMPDH2	p.P97H	3,4
hMDS-1_CD8	hMDS CD8	X	150987843	A	-	frameshift deletion	HMGB3	p.K178fs	4,2
hMDS-1_CD8	hMDS CD8	2	159944930	-	T	frameshift insertion	PLA2R1	p.G1372fs	2,0
hMDS-1_CD8	hMDS CD8	1	160325636	-	T	frameshift insertion	COPA	p.N171fs	2,1
hMDS-1_CD8	hMDS CD8	1	207476276	A	C	nonsynonymous SNV	CR2	p.N861T	15,6
hMDS-10_CD8	hMDS CD8	20	391401	T	-	frameshift deletion	TRIB3	p.F136fs	2,4
hMDS-10_CD8	hMDS CD8	19	13506007	G	A	nonsynonymous SNV	CACNA1A	p.T73M	5,4
hMDS-10_CD8	hMDS CD8	6	30571513	C	T	nonsynonymous SNV	ABCF1	p.P9L	1,3
hMDS-10_CD8	hMDS CD8	6	31945182	G	T	nonsynonymous SNV	C2	p.G481V	8,1
hMDS-10_CD8	hMDS CD8	20	40688447	TGG	-	nonframeshift deletion	MAFB	p.134_135del	2,0
hMDS-10_CD8	hMDS CD8	11	47579105	C	T	nonsynonymous SNV	NDUFS3	p.A5V	2,7
hMDS-10_CD8	hMDS CD8	11	69775006	C	G	nonsynonymous SNV	FGF4	p.G27R	6,8
hMDS-10_CD8	hMDS CD8	10	95756250	C	A	nonsynonymous SNV	ENTPD1	p.T4K	2,6
hMDS-10_CD8	hMDS CD8	2	159944930	-	T	frameshift insertion	PLA2R1	p.G1372fs	2,1
hMDS-11_CD8	hMDS CD8	X	3322857	C	A	nonsynonymous SNV	MXRA5	p.G943V	2,3
hMDS-11_CD8	hMDS CD8	17	16382471	C	-	frameshift deletion	UBB	p.I188fs	3,2
hMDS-11_CD8	hMDS CD8	11	18245442	T	C	nonsynonymous SNV	SAA2	p.K102E	3,4

hMDS-11_CD8	hMDS CD8	17	66965065	C	-	frameshift deletion	CACNG4	p.P52fs	3,7
hMDS-11_CD8	hMDS CD8	7	102178640	A	-	frameshift deletion	CUX1	p.K329fs	2,6
hMDS-11_CD8	hMDS CD8	12	112938619	G	T	nonsynonymous SNV	OAS3	p.R30L	4,5
hMDS-13_1_CD8	hMDS CD8	11	8618758	G	A	stopgain	TRIM66	p.R1195X	4,9
hMDS-13_1_CD8	hMDS CD8	17	44349703	C	T	nonsynonymous SNV	GRN	p.R101W	5,1
hMDS-13_1_CD8	hMDS CD8	8	56113888	G	A	nonsynonymous SNV	MOS	p.A32V	5,9
hMDS-13_1_CD8	hMDS CD8	11	75283447	C	T	nonsynonymous SNV	ARRB1	p.R65Q	4,4
hMDS-13_1_CD8	hMDS CD8	15	78932440	A	G	nonsynonymous SNV	CTSH	p.W142R	7,5
hMDS-13_1_CD8	hMDS CD8	2	97724429	G	T	nonsynonymous SNV	ZAP70	p.W131C	14,7
hMDS-13_1_CD8	hMDS CD8	4	109980905	T	A	stopgain	EGF	p.C725X	10,0
hMDS-13_1_CD8	hMDS CD8	4	165040203	G	T	nonsynonymous SNV	TRIM60	p.W44L	11,4
hMDS-13_2_CD8	hMDS CD8	17	38364849	CC	-	frameshift deletion	SOCS7	p.S317fs	6,5
hMDS-13_2_CD8	hMDS CD8	22	40420532	C	T	nonsynonymous SNV	MKL1	p.R309H	2,0
hMDS-13_2_CD8	hMDS CD8	7	50376560	G	C	nonsynonymous SNV	IKZF1	p.S63T	7,1
hMDS-13_2_CD8	hMDS CD8	17	75876776	C	T	nonsynonymous SNV	TRIM47	p.R238H	2,3
hMDS-13_2_CD8	hMDS CD8	8	88186549	A	T	nonsynonymous SNV	MMP16	p.S111T	2,9
hMDS-5_CD8	hMDS CD8	4	1003409	T	A	nonsynonymous SNV	IDUA	p.L530Q	6,8
hMDS-5_CD8	hMDS CD8	17	9221166	AG	CC	splicing	NTN1		4,7
hMDS-5_CD8	hMDS CD8	1	47219898	C	T	nonsynonymous SNV	TAL1	p.G114D	8,4
hMDS-5_CD8	hMDS CD8	12	48778209	A	T	nonsynonymous SNV	ADCY6	p.C305S	4,6
hMDS-5_CD8	hMDS CD8	14	50159964	T	C	nonsynonymous SNV	SOS2	p.N440S	2,2
hMDS-5_CD8	hMDS CD8	17	64053308	T	C	nonsynonymous SNV	ERN1	p.T673A	5,3
hMDS-5_CD8	hMDS CD8	1	64841313	-	T	frameshift insertion	JAK1	p.P861fs	2,2
hMDS-5_CD8	hMDS CD8	16	66577197	G	-	frameshift deletion	CKLF-CMTM	p.G58fs	2,1
hMDS-5_CD8	hMDS CD8	7	81982624	G	A	nonsynonymous SNV	CACNA2D1	p.S633L	3,7
hMDS-5_CD8	hMDS CD8	4	88494197	T	G	nonsynonymous SNV	HERC5	p.F770L	3,2
hMDS-5_CD8	hMDS CD8	5	98903832	TGC	-	nonframeshift deletion	CHD1	p.110_111del	10,0
hMDS-5_CD8	hMDS CD8	3	108755842	G	C	nonsynonymous SNV	RETNLB	p.T91S	2,1
hMDS-5_CD8	hMDS CD8	1	160679075	T	C	nonsynonymous SNV	CD48	p.I237V	4,1

Table S3 Amplicon validation summary

Validation	n
TP	15
FP	1
Amplicon did not work	2

Variants validated with amplicon (TP)

sample_ID	Chr	Start	Ref	Alt	Variant type	Gene	Amino acid change	VAF (%)
AA-3_2_CD8	3	9893298	T	A	nonsynonymous SNV	JAGN1	p.V158E	3,0
AA-7_2_CD8	18	12836879	C	G	nonsynonymous SNV	PTPN2	p.R29P	5,8
AA-5_1_MNC	X	15324664	C	-	splicing	PIGA		14,4
AA-3_3_BM_CD8	12	25225628	C	G	nonsynonymous SNV	KRAS	p.A146P	4,4
AA-3_2_CD8	12	25225628	C	G	nonsynonymous SNV	KRAS	p.A146P	3,0
AA-3_3_PB_CD8	12	25225628	C	G	nonsynonymous SNV	KRAS	p.A146P	7,8
AA-5_1_MNC	2	25234304	A	T	nonsynonymous SNV	DNMT3A	p.L753Q	12,9
HC-8_CD8	2	25240306	A	G	nonsynonymous SNV	DNMT3A	p.L621P	4,2
hMDS-13_2_CD8	17	38364849	CC	-	frameshift deletion	SOCS7	p.S317fs	6,5
AA-5_2_CD8	2	65334634	A	-	frameshift deletion	SPRED2	p.V112fs	3,3
hMDS-5_CD4	15	78937393	G	-	frameshift deletion	CTSH	p.H52fs	4,1
AA-2_2_CD8	15	89306164	G	A	nonsynonymous SNV	FANCI	p.M1109I	3,9
HC-8_CD8	9	130880097	T	C	nonsynonymous SNV	ABL1	p.S485P	5,5
AA-3_3_BM_CD8	6	137875835	G	-	frameshift deletion	TNFAIP3	p.D212fs	3,9
AA-3_3_PB_CD8	6	137875835	G	-	frameshift deletion	TNFAIP3	p.D212fs	6,6
AA-3_3_BM_CD8	20	51523337	C	T	nonsynonymous SNV	NFATC2	p.V282M	7,6
AA-3_3_PB_CD8	20	51523337	C	T	nonsynonymous SNV	NFATC2	p.V282M	7,4
AA-4_1_CD8	17	42322464	T	A	nonsynonymous SNV	STAT3	p.Y640F	1,8
AA-3_3_PB_CD8	1	113859412	T	C	nonsynonymous SNV	PTPN22	p.T46A	3,5

Variants not validated with amplicon (FP)

sample_ID	Chr	Start	Ref	Alt	Variant type	Gene	Amino acid change	VAF (%)
AA-45_CD4	11	119285565	T	C	nonsynonymous SNV	CBL	p.M647T	8,8

Amplicon did not work

sample_ID	Chr	Start	Ref	Alt	Variant type	Gene	Amino acid change	VAF (%)
AA-45_CD8	11	102331110	C	G	nonsynonymous SNV	BIRC3	p.T398R	3,7
AA-2_3_CD4	1	114716123	C	T	nonsynonymous SNV	NRAS	p.G13D	4,5