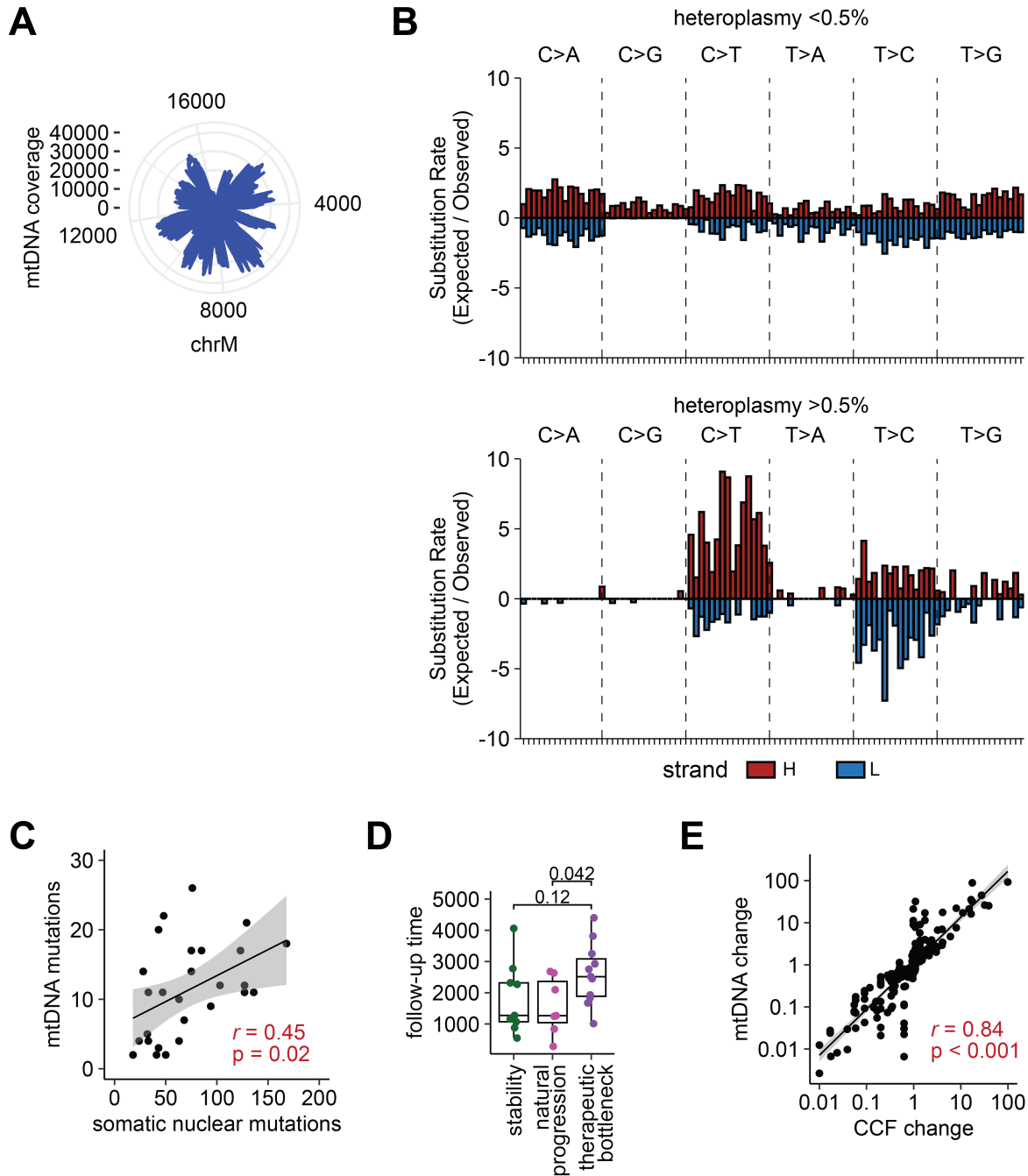


Supplementary Data

Tracking rare single donor and recipient immune and leukemia cells after allogeneic hematopoietic cell transplantation using mitochondrial DNA mutations

Supplementary Figures



Suppl. Fig. 1. Extraction of mitochondrial DNA mutations from bulk RNA-seq data in CLL.

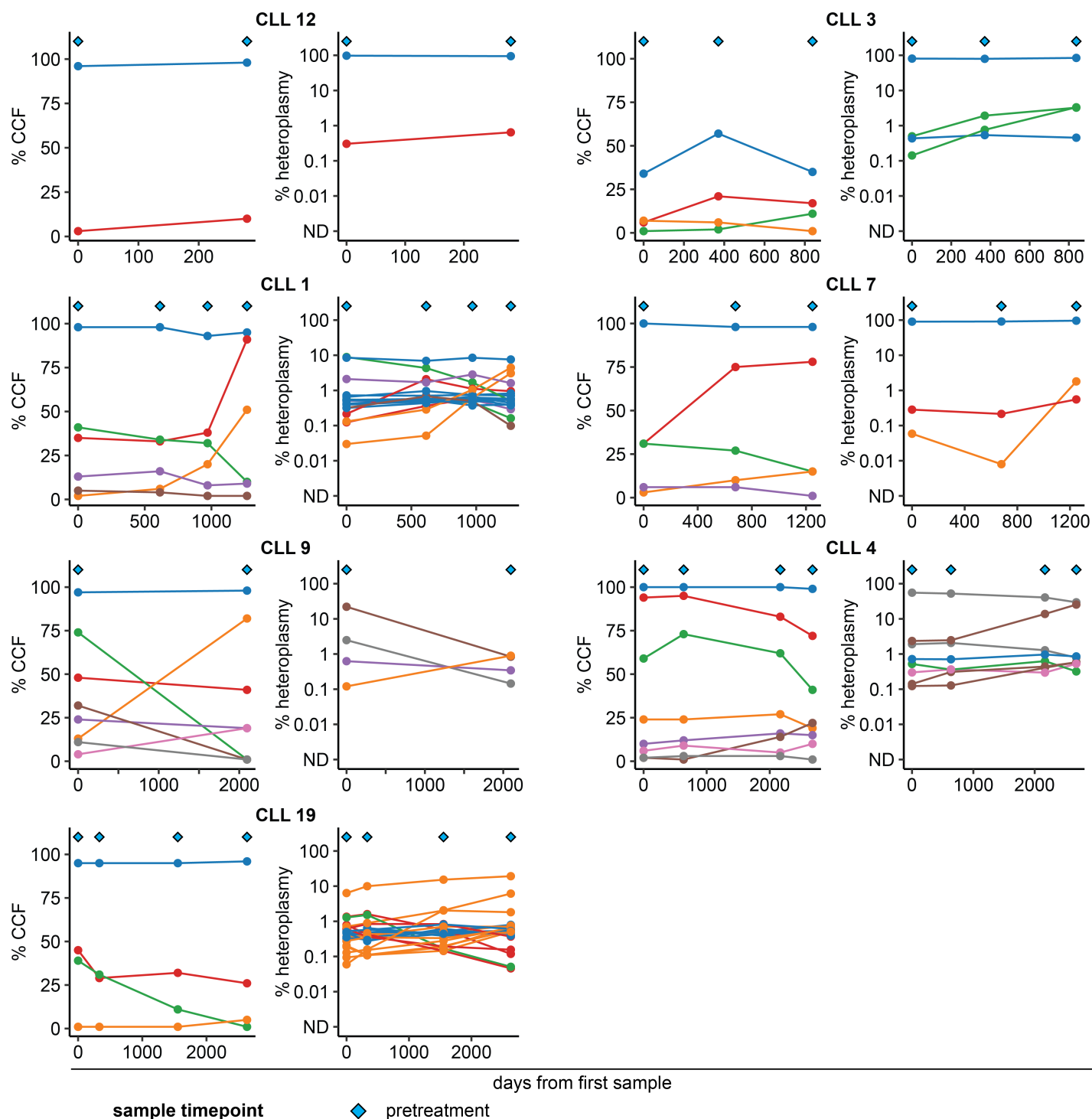
A Coverage of mitochondrial transcripts in bulk RNA-seq data of CLL.

B Mutational profiles in mitochondrial DNA (mtDNA) mutations extracted from bulk RNA-seq data. Those mtDNA mutations with heteroplasmy <0.5% (top) were excluded due to their unspecific mutational profiles, while mtDNA mutations with a heteroplasmy >0.5% in at least one sample (bottom) showed a characteristic profile and were used for further analyses.

C Association of somatic nuclear and mitochondrial DNA mutations. Statistical testing using Pearson's correlation coefficient.

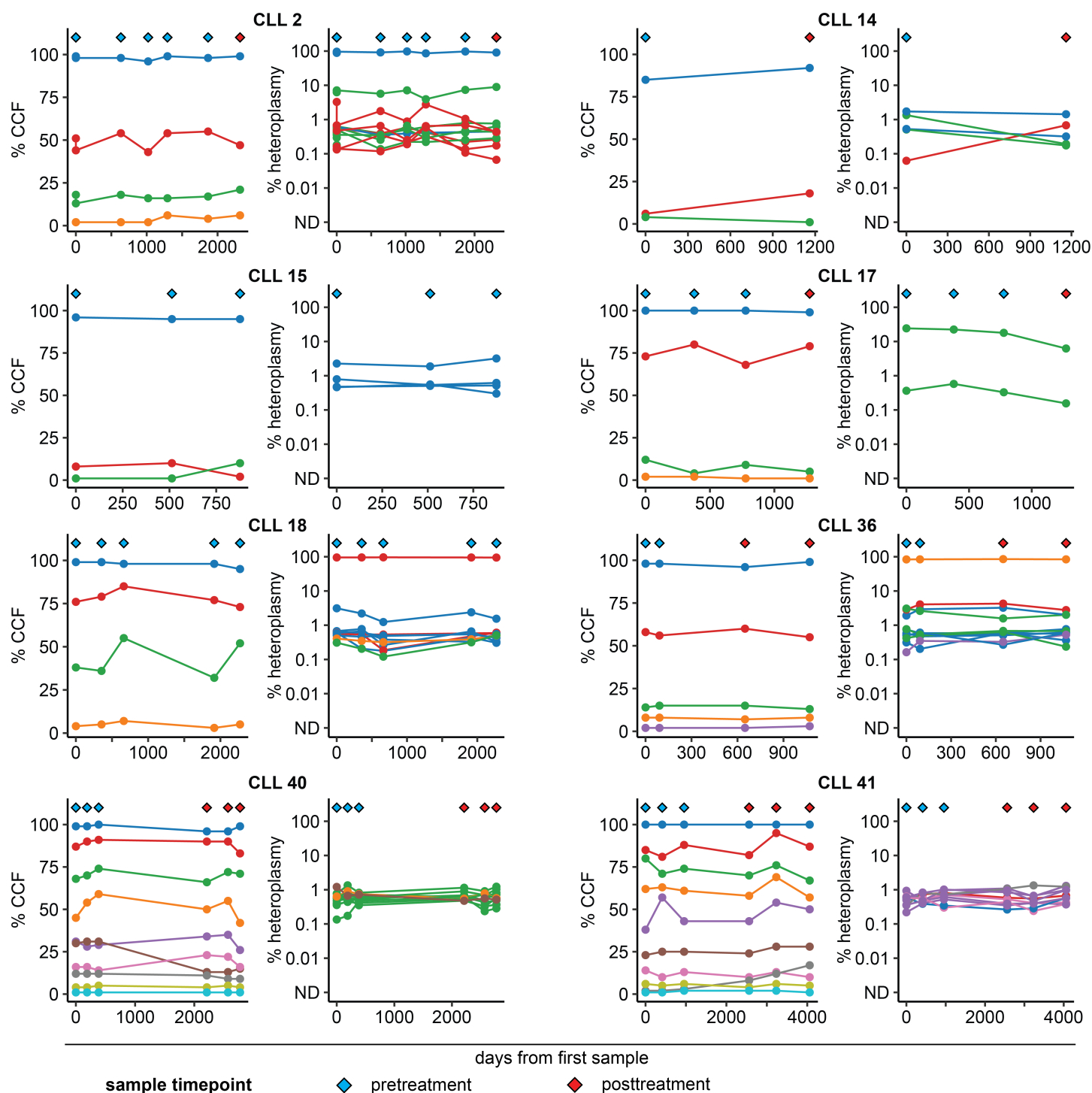
D Follow-up time from first to last sample across genetic CLL subgroups in cohort.

E Correlation of change in mtDNA mutations and CCF values. Statistical testing using Pearson's correlation coefficient.

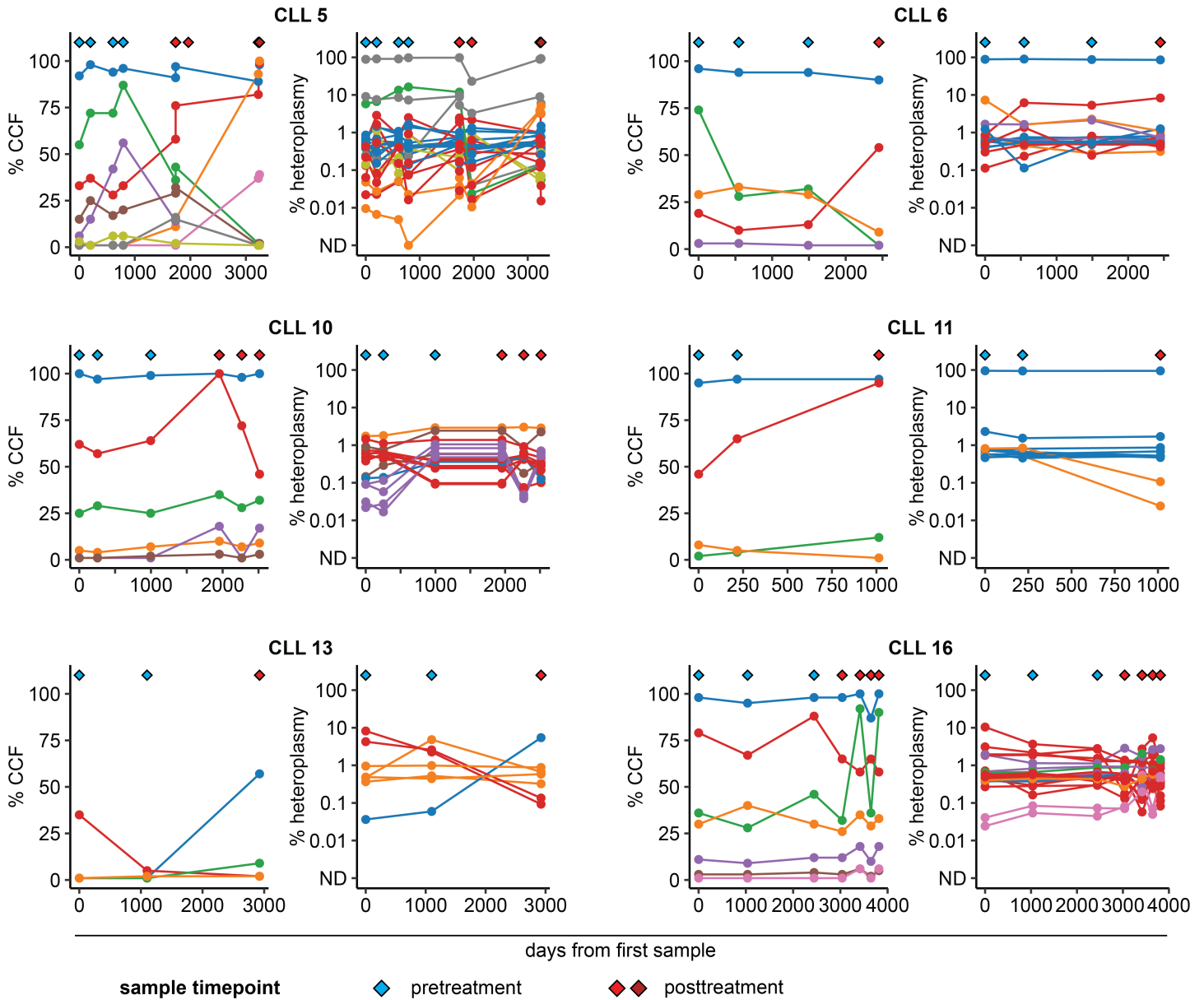


Suppl. Fig. 2. Longitudinal changes of somatic nuclear and mitochondrial DNA mutations in naturally progressing CLL.

Cancer cell fractions (CCF) of CLL subclones were calculated from whole-exome sequencing (WES) data. Heteroplasmy was calculated from mitochondrial DNA mutations extracted from matched bulk RNA-seq data. Colors in CCF plots indicate CLL subclones identified from WES data, while colors in heteroplasmy plots indicate the closest-matching CLL subclone for each mitochondrial DNA mutation.

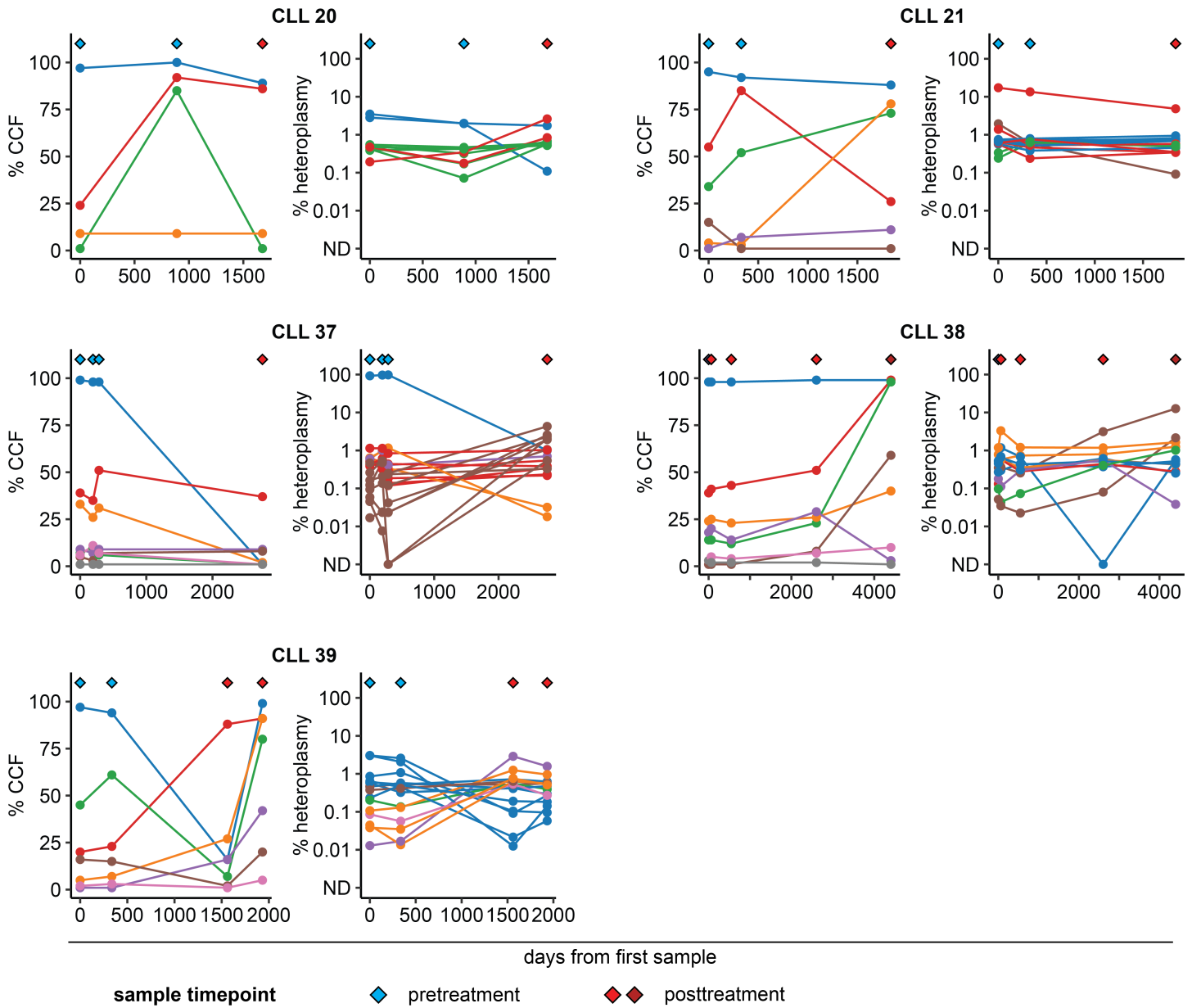


Suppl. Fig. 3. Longitudinal changes of somatic nuclear and mitochondrial DNA mutations in genetically stable CLL. Cancer cell fractions (CCF) of CLL subclones were calculated from whole-exome sequencing (WES) data. Heteroplasmy was calculated from mitochondrial DNA mutations extracted from matched bulk RNA-seq data. Colors in CCF plots indicate CLL subclones identified from WES data, while colors in heteroplasmy plots indicate the closest-matching CLL subclone for each mitochondrial DNA mutation.



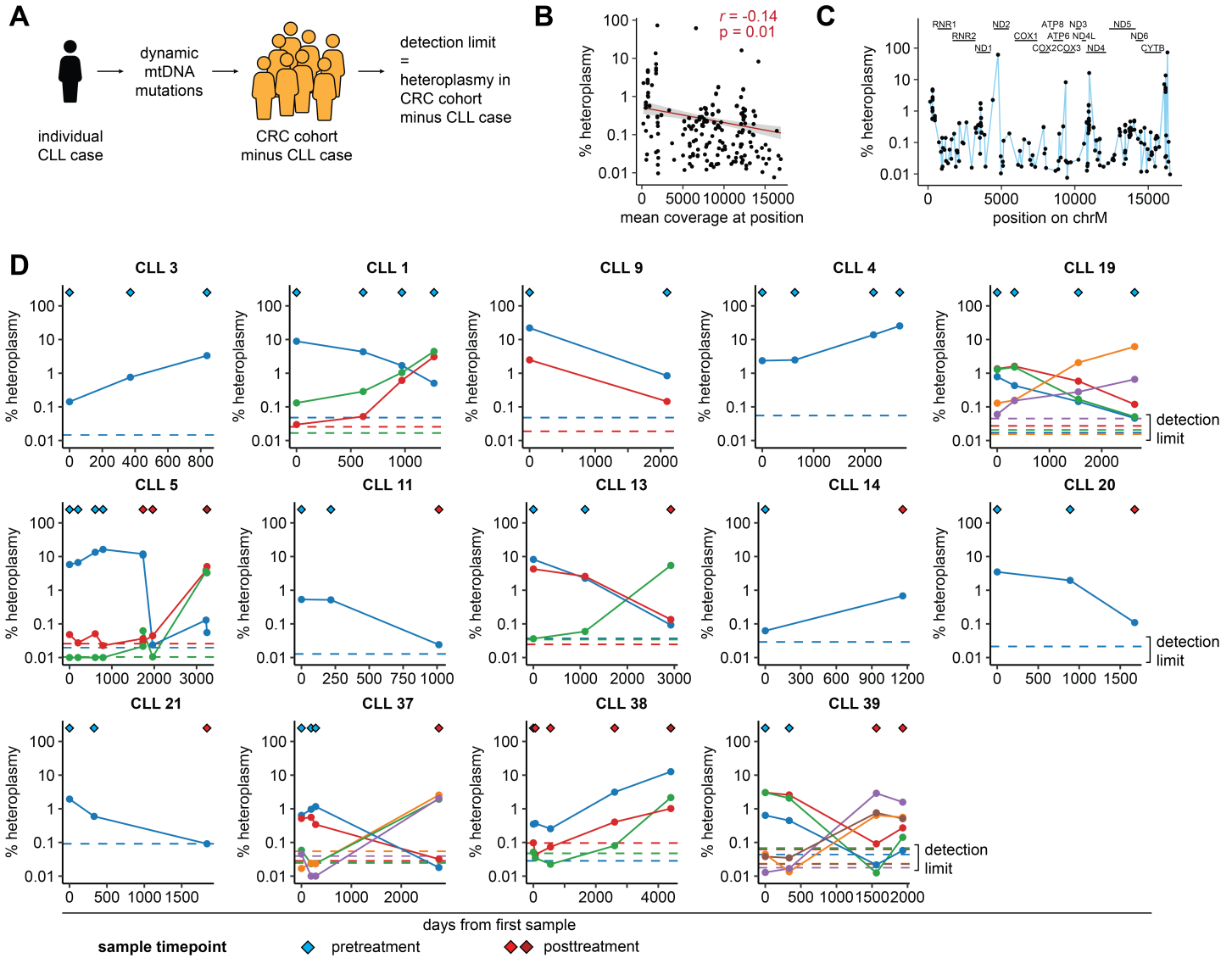
Suppl. Fig. 4. Longitudinal changes of somatic nuclear and mitochondrial DNA mutations in evolving CLL.

Cancer cell fractions (CCF) of CLL subclones were calculated from whole-exome sequencing (WES) data. Heteroplasmy was calculated from mitochondrial DNA mutations extracted from matched bulk RNA-seq data. Colors in CCF plots indicate CLL subclones identified from WES data, while colors in heteroplasmy plots indicate the closest-matching CLL subclone for each mitochondrial DNA mutation.



Suppl. Fig. 5. Longitudinal changes of somatic nuclear and mitochondrial DNA mutations in evolving CLL.

Cancer cell fractions (CCF) of CLL subclones were calculated from whole-exome sequencing (WES) data. Heteroplasmy was calculated from mitochondrial DNA mutations extracted from matched bulk RNA-seq data. Colors in CCF plots indicate CLL subclones identified from WES data, while colors in heteroplasmy plots indicate the closest-matching CLL subclone for each mitochondrial DNA mutation.



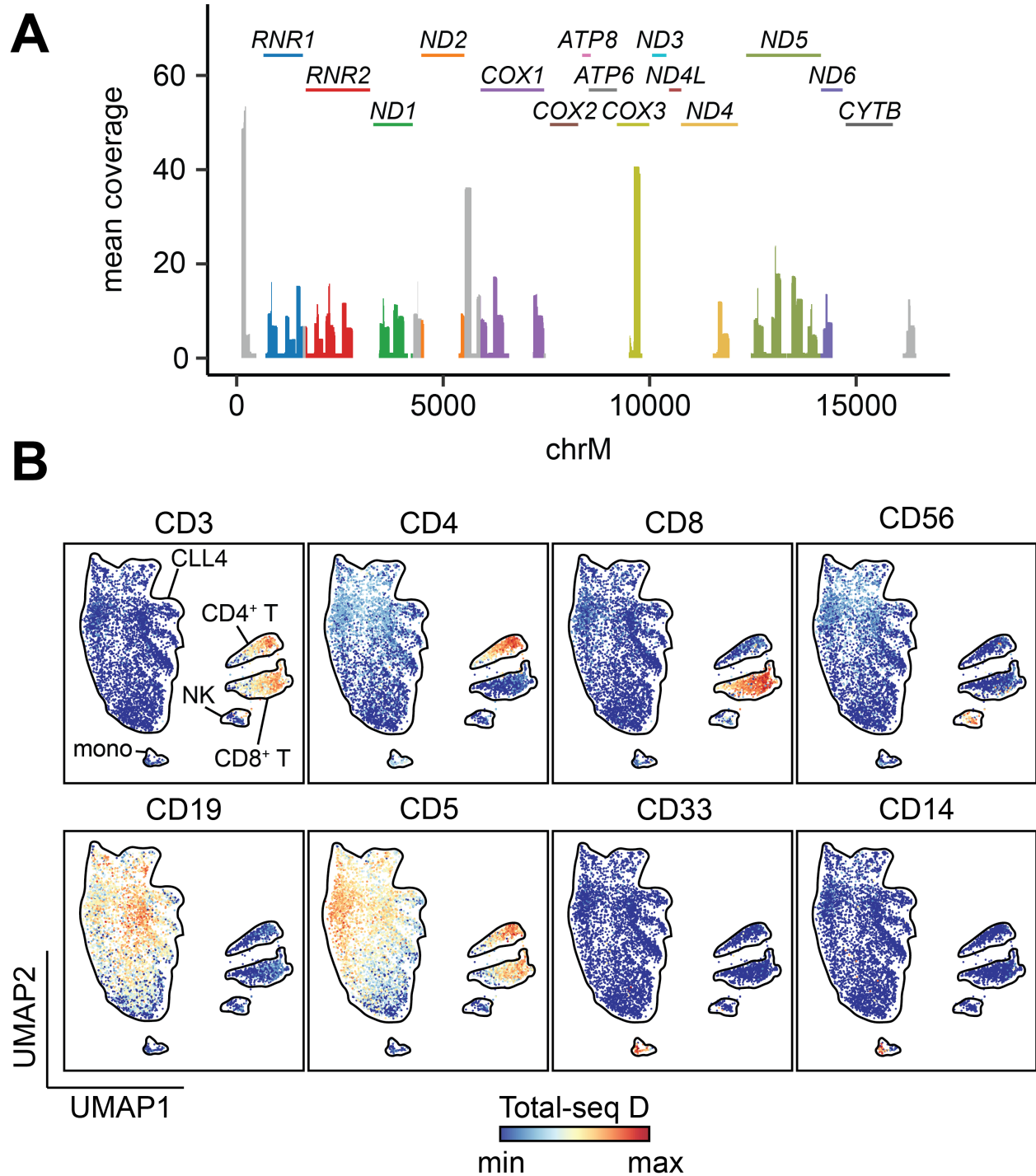
Suppl. Fig. 6. Limit of detection of mitochondrial DNA mutations extracted from bulk RNA-seq data.

A The limit of detection (LOD) for mtDNA mutations for each CLL case in bulk RNA-seq data were defined as the heteroplasmy for each mutation within the rest of the study cohort (i.e. background heteroplasmy), excluding samples from the respective CLL case.

B LOD of mtDNA mutations is weakly correlated with the average sequencing coverage at that position on chrM.

C LOD of mtDNA mutations is driven by their location on chrM.

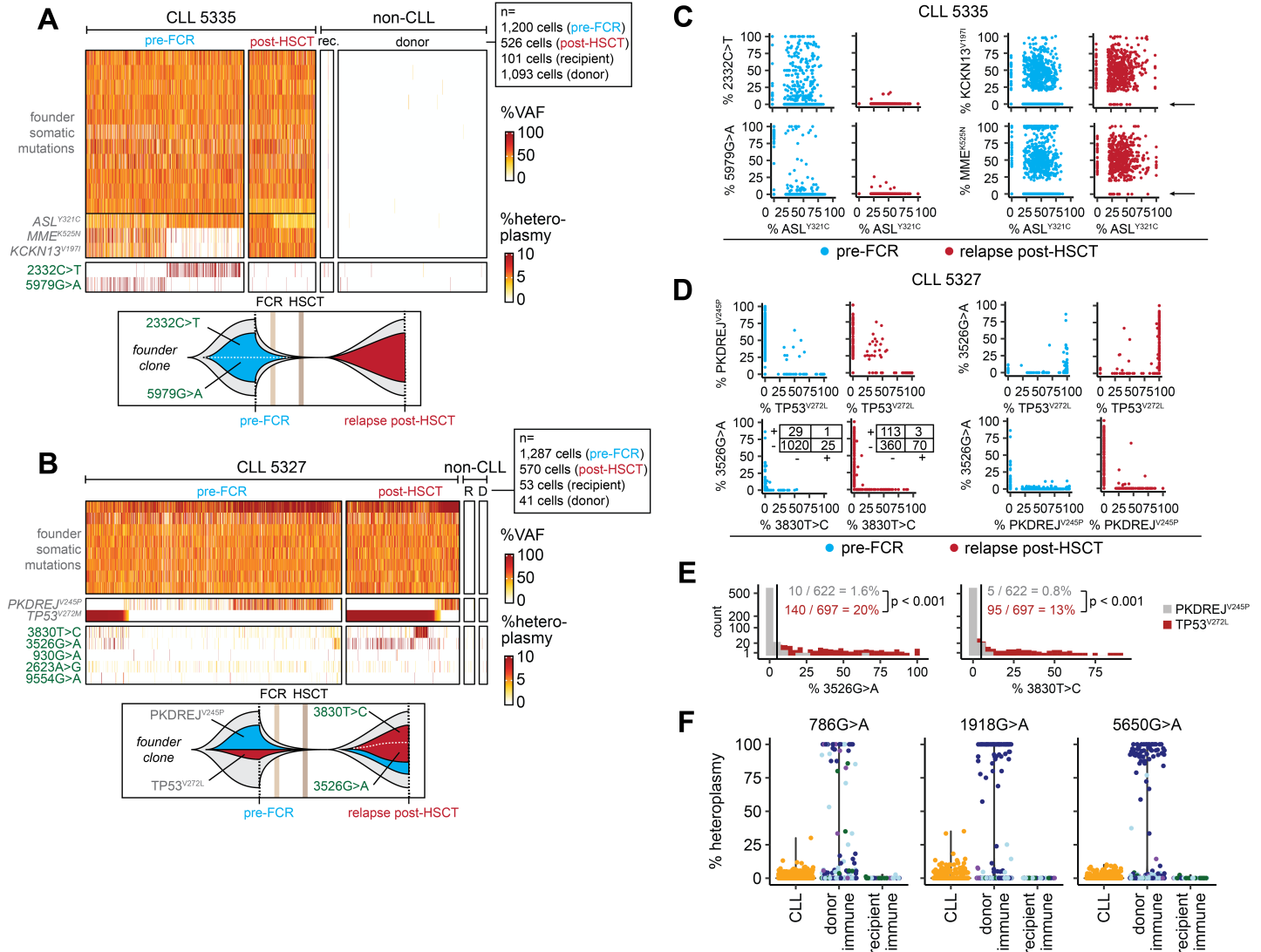
D Longitudinal tracking of mtDNA mutations alongside their respective LOD. Each line corresponds to one individual mtDNA mutation.



Suppl. Fig. 7. Tracking co-evolution of somatic nuclear and mitochondrial DNA mutations in CLL and AML using single cell DNA and protein sequencing (Tapestri).

A Pseudobulk coverage profile of the mitochondrial genome using mitochondrial DNA (mtDNA) panel for scDNA-seq.

B Expression of CD3, CD4 and CD8 (T cells), CD56 (NK cells), CD19 and CD5 (CLL), CD33 and CD14 (monocytes) in CLL4 (n=5,596).



Suppl. Fig. 8. Co-segregation of mitochondrial and somatic nuclear DNA mutations in CLL.

A Single-cell variant allele frequencies (VAFs) and heteroplasmy of mitochondrial DNA (mtDNA) mutations in CLL5335 pre-FCR and post-HSCT, in recipient-derived (rec.) and donor-derived (donor) immune cells. The fish plot (bottom) summarizes the results, with outgrowth of a resistant clone (purple) that co-existed beside two subclones marked by 2332C>T and 5979G>A.

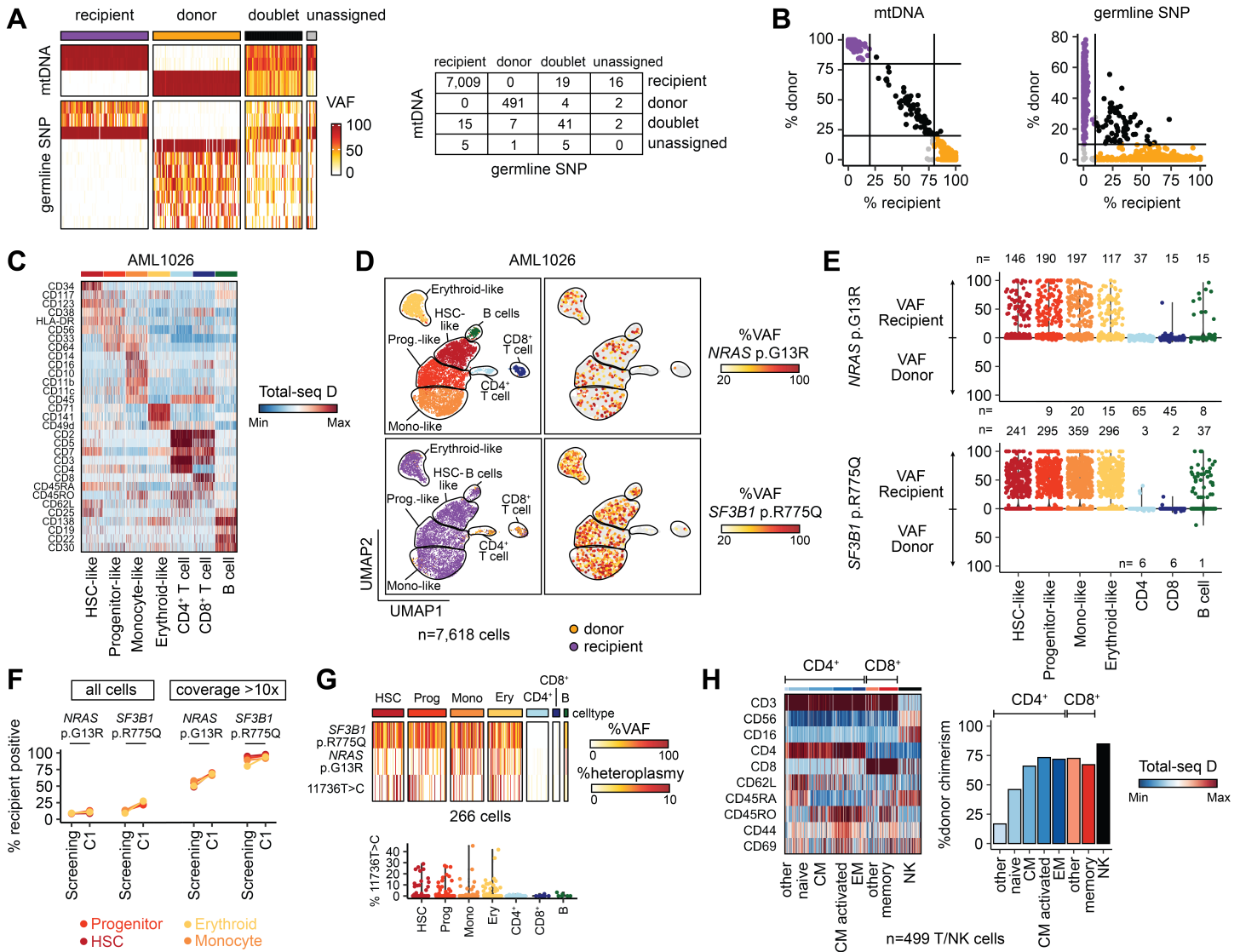
B Single-cell VAFs and heteroplasmy of mtDNA mutations in CLL5327 pre-FCR and post-HSCT, in recipient-derived (R) and donor-derived (D) immune cells. CLL clones relapse at a ratio which is shifted towards the subclone marked by *TP53*^{V272L}, 3830T>C and 3526G>A (fish plot bottom).

C Co-occurrence of 2332C>T, 5979G>A, *ASL*^{Y321C}, *MME*^{K525N} and *KCKN*^{V197I} pre-FCR (blue) and post-HSCT (red) in CLL5335. Both mtDNA mutations are lost at relapse post-HSCT. The arrows indicate that *ASL*^{Y321C} and *MME*^{K525N} do not always co-occur with *KCKN*^{V197I} pre-FCR but almost always co-occur at relapse post-HSCT, indicating outgrowth of CLL5335 from a subclone that harbors all three mutations at relapse post-HSCT.

D Mutual exclusivity of a subclone in CLL5327 marked by *PKDREJ*^{V245P} versus a subclone marked by *TP53*^{V272L}, 3830T>C and 3526G>A. Further, 3526G>A and 3830T>C are mutually exclusive, as demonstrated by the insets that show the number of cells that are positive (heteroplasmy >5%) for 3526G>A (top left), 3830T>C (bottom right), both (top right) or none (bottom left).

E Distribution of 3526G>A (left) or 3830T>C (right) in cells that are *PKDREJ*^{V245P} (grey) or *TP53*^{V272L} (red), illustrating that 3526G>A and 3830T>C cooccur with *TP53* mutated cells in CLL5327.

F Heteroplasmy of three mtDNA mutations with expansion in donor-derived immune cells of CLL5328.



Suppl. Fig. 9. Co-segregation of mitochondrial and somatic nuclear DNA mutations in AML bone marrow.

A Donor-recipient deconvolution of AML1026 using mtDNA mutations or germline single nucleotide polymorphisms are highly consistent.

B Donor-recipient deconvolution with mitochondrial DNA (mtDNA) (left) and single nucleotide polymorphisms (SNPs) (right).

C Identification of hematopoietic stem cell (HSC)-like, progenitor-like, monocyte-like, erythroid, CD4⁺ T cell, CD8⁺ T cell and B cell populations with 45 Total-seq D antibodies in AML1026.

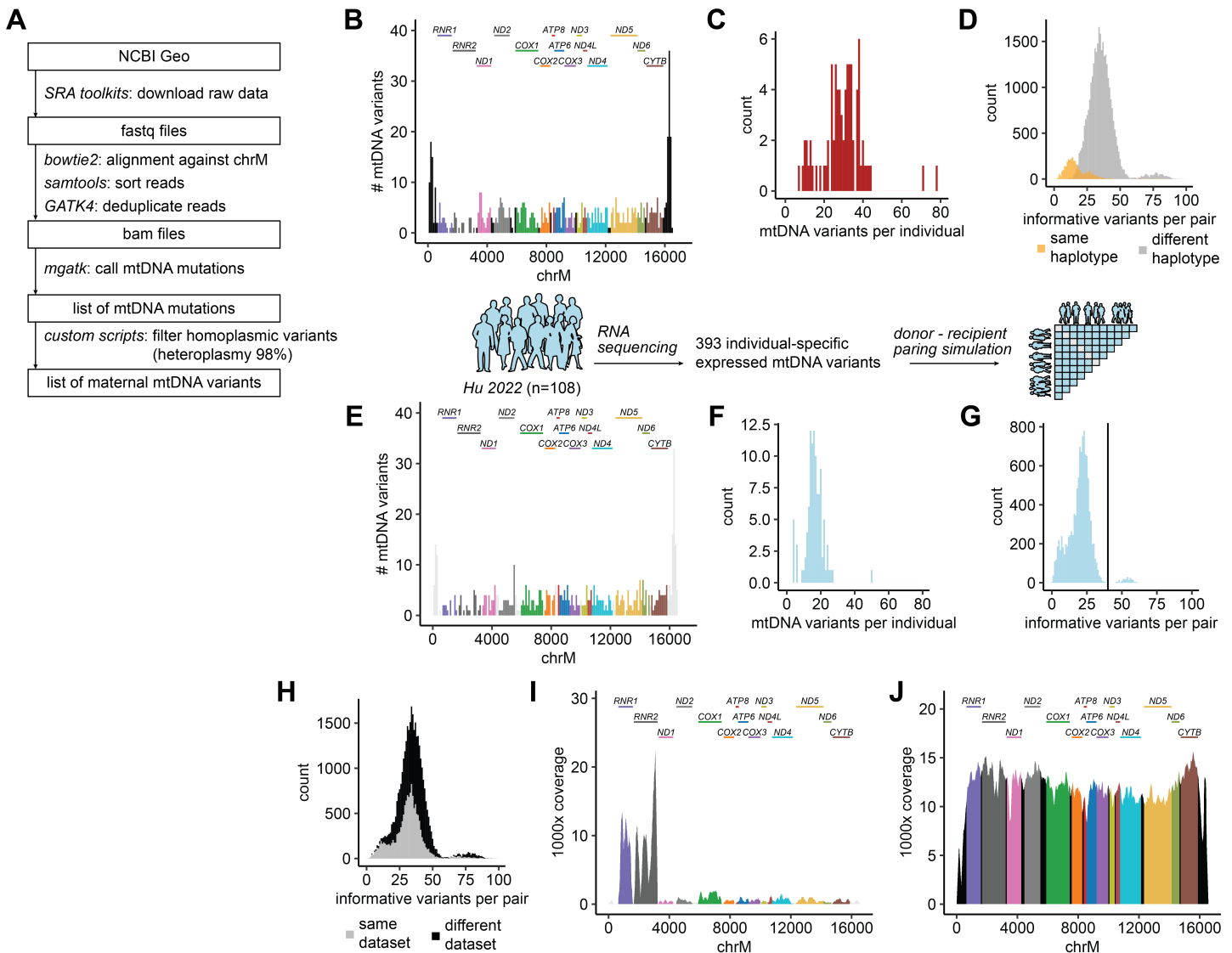
D UMAP projections of surface marker expression data obtained from AML1026 annotated by cell types (top left), donor-recipient annotation (bottom left), apparent variant allele frequencies (VAFs) of *NRAS*^{G13R} (top right) and *SF3B1*^{P775Q} (bottom right).

E Apparent VAFs per cell of *NRAS*^{G13R} (top) and *SF3B1*^{P775Q} (bottom) cell types for recipient- and donor-derived cells.

F Apparent single cell VAF of *NRAS*^{G13R} and *SF3B1*^{P775Q} in AML1026 across all scDNA-seq profiles (left) or only cells with coverage of both loci with at least 10x (right), demonstrating the relevance of drop-out for analysis of scDNA-seq data.

G Co-occurrence of 11736T>C with *NRAS*^{G13R} and *SF3B1*^{P775Q} in AML1026 (top). Heteroplasmy of 11736T>C across cell populations in AML1026 (bottom).

H Identification of T cell subpopulations with surface marker expression (Total-seq D) (left). Donor chimerism across T cell subpopulations (right)



Suppl. Fig. 10. Applicability of expressed mitochondrial DNA (mtDNA) mutations to distinguish unrelated individuals.

A Workflow for extraction of maternal mtDNA variants from public sequencing data (1,2).

B Position of 624 homoplasmic mitochondrial DNA mutations representing maternal DNA variants that allow to identify individuals across the mitochondrial genome chrM (hg38)

C A median of 30 homoplasmic mtDNA mutations (heteroplasmy >98%) per individual (range 7 – 78) were detectable in bulk mtDNA-seq datasets.

D Distribution of the number of mtDNA mutations that distinguish unrelated individuals from the same mtDNA haplotype (orange) or from different mtDNA haplotypes (grey).

E Distribution of expressed mtDNA mutations across the mitochondrial chromosome (hg38) identified in 108 bulk RNA-seq profiles. (2) Grey – non-expressed variants.

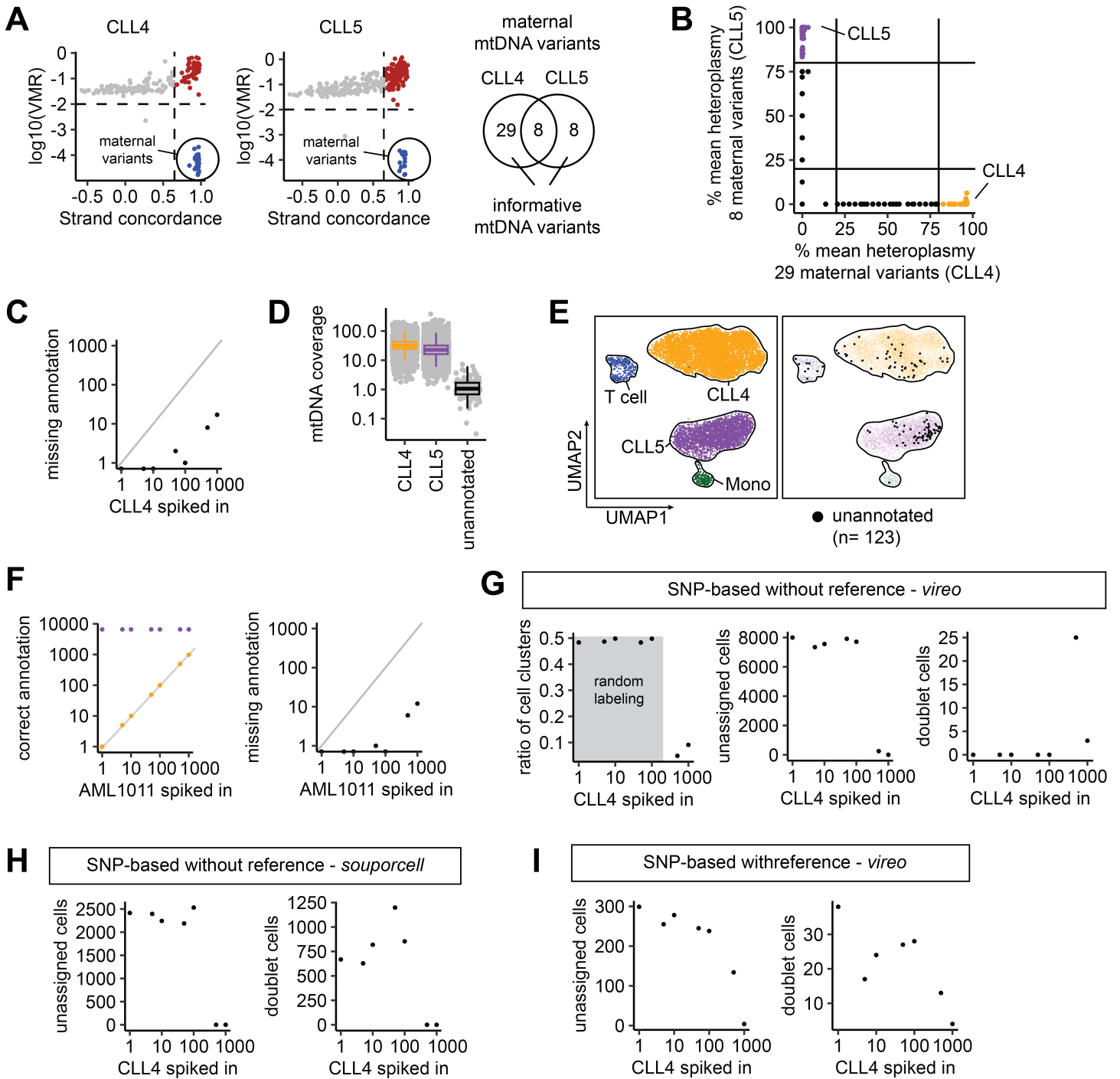
F Number of mtDNA mutations expressed per individual.

G Number of expressed mtDNA mutations that distinguish any simulated pair of two individuals. Outlier pairs with >40 mtDNA mutations are indicated by the vertical line.

H Distribution of the number of mtDNA mutations that distinguish unrelated individuals from the same dataset (grey) or from different published datasets (black), indicating that the number of mtDNA mutations are not related to the underlying data source.

I Coverage of transcripts across the mitochondrial chromosome obtained from bulk RNA-seq profiles.

J Coverage of DNA fragments across the mitochondrial chromosome obtained from bulk mtDNA-seq profiles.



Suppl. Figure 11. Assessment of deconvolution performance of synthetic chimeric cell populations in *in-silico* mixing experiments.

A Identification of maternally inherited mtDNA variants that distinguish single cell chromatin profiles of CLL4 and CLL5 reanalyzed from *Penter & Gohil et al., Cancer Discovery 2021 (3)* (informative mtDNA variants) using variance-to-mean ratio (VMR) and strand concordance.

B Deconvolution of single cells from CLL4 (orange) and CLL5 (purple) based on mean heteroplasmy of 29 (CLL4) and 8 (CLL5) maternal mtDNA variants. Black dots indicate cells which cannot be assigned to either individual.

C Number of unannotated cells from CLL4 cells after mtDNA-based deconvolution of mixed data for conditions from 1 to 1,000 CLL4 cells spiked into 7,579 CLL5 cells.

D Single cell mtDNA coverage in CLL4 (orange), CLL5 (purple) and cells that could not be annotated.

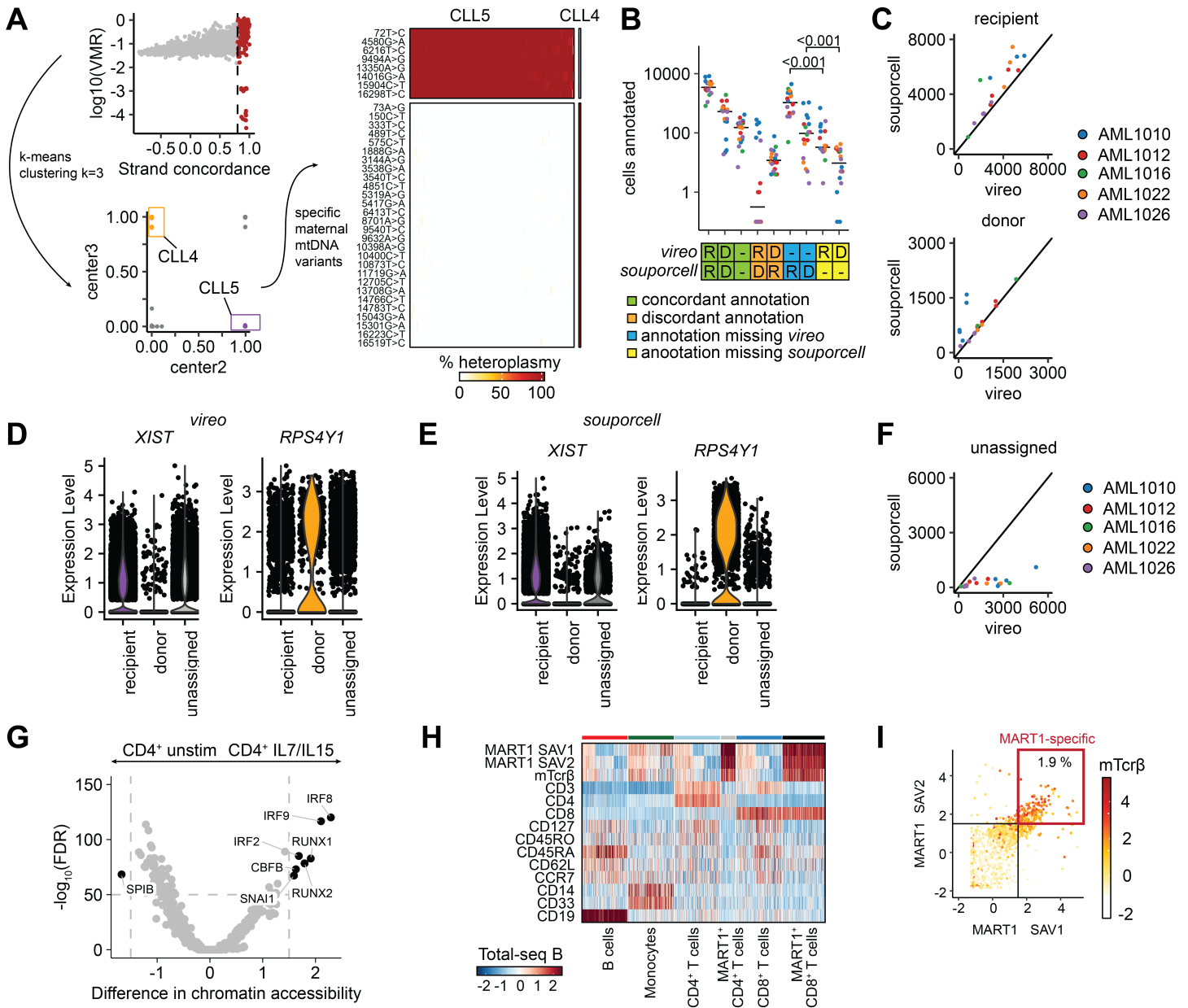
E UMAP projection of single cell chromatin accessibility profiles of CLL4 and CLL5 annotated by cell types (left) and by cells that could not be annotated using maternal mtDNA variants (right).

F Spike-in experiment with single cells from AML1011 spiked into AML1012.

G Germline-free deconvolution of mixed single cell RNA sequencing (scRNA-seq) profiles with *vireo* leads to random annotation for fewer than 300 (3%) CLL4 cells into 10,000 CLL5 cells (left). Number of unassigned cells (middle) and cells detected as doublets (right) for spike-in experiment of scRNA-seq profiles of CLL4 and CLL5 reanalyzed from *Penter & Gohil et al., Cancer Discovery 2021* using *vireo* without a germline reference.

H Number of unassigned cells (left) and cells detected as doublets (right) for spike-in data using *souporcell* without a germline reference.

I Number of unassigned cells (left) and cells detected as doublets (right) for spike-in data using *vireo* with a germline reference.



Suppl. Fig. 12. Deconvolution of chimeric populations using mitochondrial DNA mutations and germline single nucleotide polymorphisms.

A Identification of maternal mitochondrial DNA (mtDNA) variants from *in-silico* mixing experiment of CLL4 and CLL5 mtscATAC-seq profiles reanalyzed from Penter & Gohil et al., Cancer Discovery 2021 (3) at a ratio of 10 (CLL4) to 7,579 (CLL5). After filtering of potential maternal mtDNA variants based on strand-concordance >0.8, variants are clustered using k-means clustering with k=3 and filtered for mtDNA variants that are mutually exclusive across both CLL populations.

B-C Concordance of single nucleotide polymorphisms (SNP)-based germline-free donor-recipient deconvolution with *souporcell* and SNP-based whole-exome sequencing (WES) germline-based donor-recipient deconvolution with *vireo* across single cell RNA sequencing (scRNA-seq) profiles of 5 AML patients (4) for recipient- (B) and donor-derived cells (C).

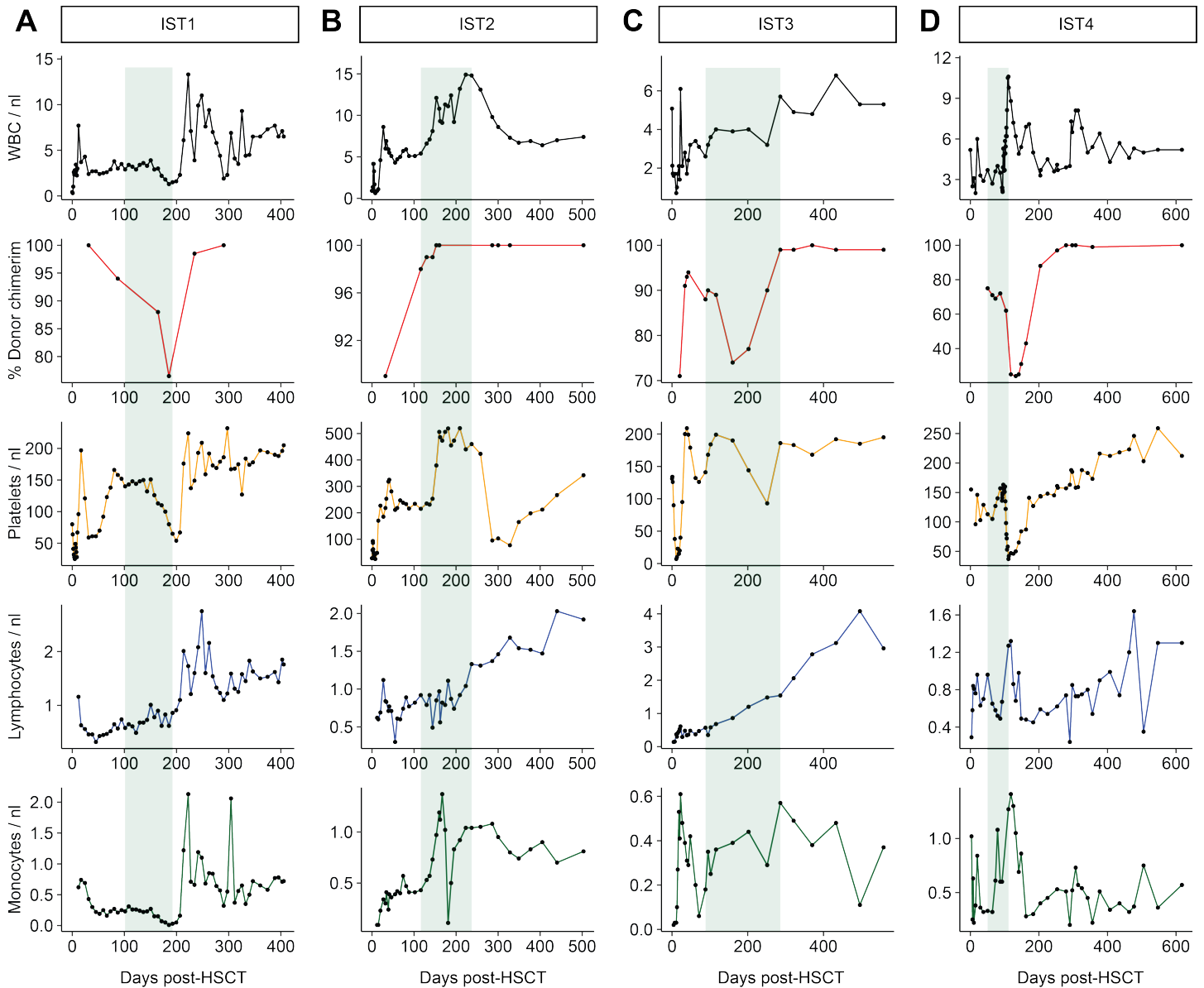
D-E Mutual exclusive detection of sex-specific genes *XIST* and *RPS4Y1* in AML1010 following donor-recipient deconvolution with *vireo* (D) and *souporcell* (E) confirms the donor-recipient annotations.

F Number of unassigned cells following SNP-based germline-free donor-recipient deconvolution with *souporcell* and SNP-based whole-exome sequencing germline-based donor-recipient deconvolution with *vireo* across scRNA-seq profiles of 5 AML patients.

G Effect of T cell expansion protocol on chromatin profiles of CD4⁺ (G) and CD8⁺ T cells (H) demonstrated by differential chromatin accessibility of transcription factor binding motifs.

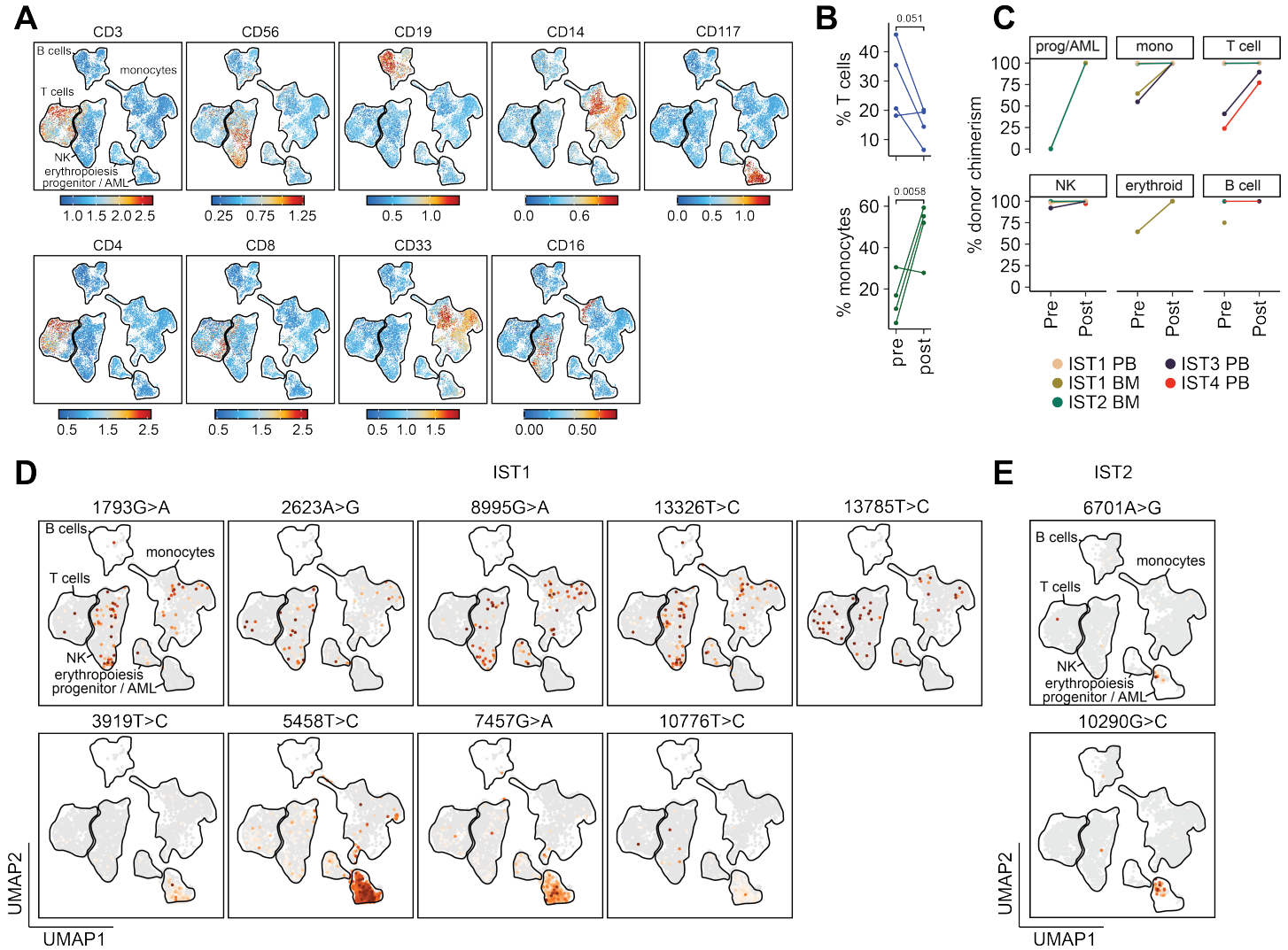
H Surface marker expression of TSA and TSB oligotags across cell types. SAV: streptavidin, mTcrβ: murinized Tcrβ

I Detection of MART1-specific T cells using TSA SAV:HLA-monomer conjugates. The color indicates the expression of the murinized Tcrβ sequence.



Suppl. Fig. 13. Blood and bone marrow counts of immunosuppression tapering (IST) cohort from routine clinical measurements.

A-D White blood cell (WBC) count, % bulk bone marrow donor chimerism using cytogenetics (X/Y chromosome), platelets, lymphocytes and monocytes measured in peripheral blood of patients IST1-4. The period of immunosuppression tapering (IST) is highlighted in green.



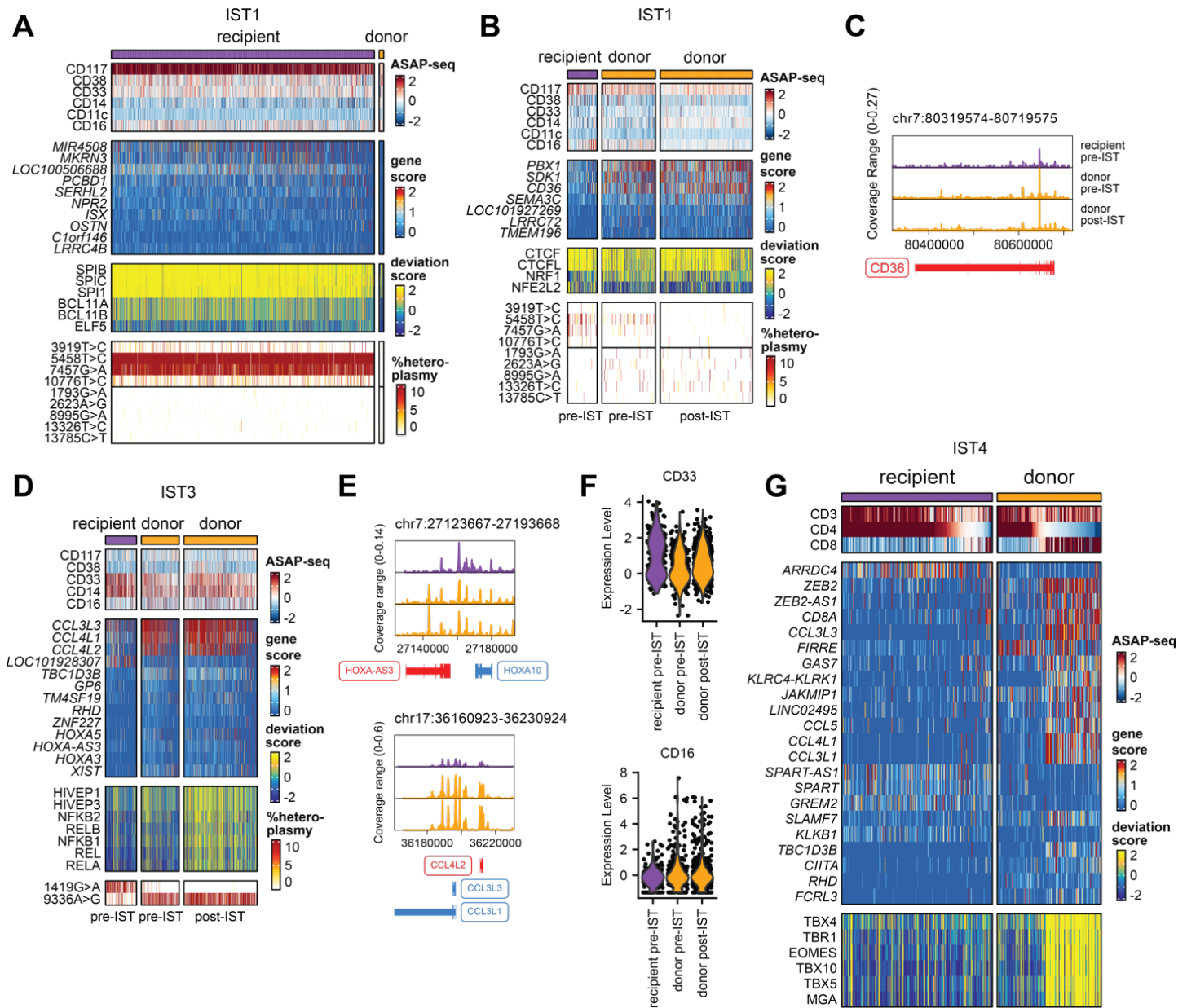
Suppl. Fig. 14. Identification of cell types and tracking of mitochondrial DNA mutations in immunosuppression tapering (IST) cohort using ASAP-seq.

A UMAP projection of single cell chromatin accessibility profiles annotated by surface marker expression of CD3, CD4 and CD8 (T cells), CD56 (NK cells), and CD14, CD16, CD33 and CD117 (myeloid populations) measured using ASAP-seq and Total-seq B antibodies.

B Percentage of T cells and monocytes in patients IST1-4 before (pre) and after (post) IST.

C Donor chimerism across progenitor/AML, monocytic (mono), T cell, NK, erythroid and B cell populations before (Pre) and after (Post) IST.

D-E UMAP projection of single cell chromatin accessibility profiles annotated by detection of mitochondrial DNA (mtDNA) mutations in IST1 (D) and IST2 (E).



Suppl. Fig. 15. Differential phenotypes of physiologic (donor-) and malignant hematopoietic (recipient-derived) cells.

A Surface marker expression (ASAP-seq), gene scores, transcription factor motif activity and mitochondrial DNA (mtDNA) mutations in AML (recipient-) and physiologic donor-derived hematopoietic stem cell (HSC)-like cells in IST1.

B Surface marker expression (ASAP-seq), gene scores, transcription factor motif activity and mitochondrial DNA (mtDNA) mutations in AML (recipient-) and donor-derived erythroid cells in IST1.

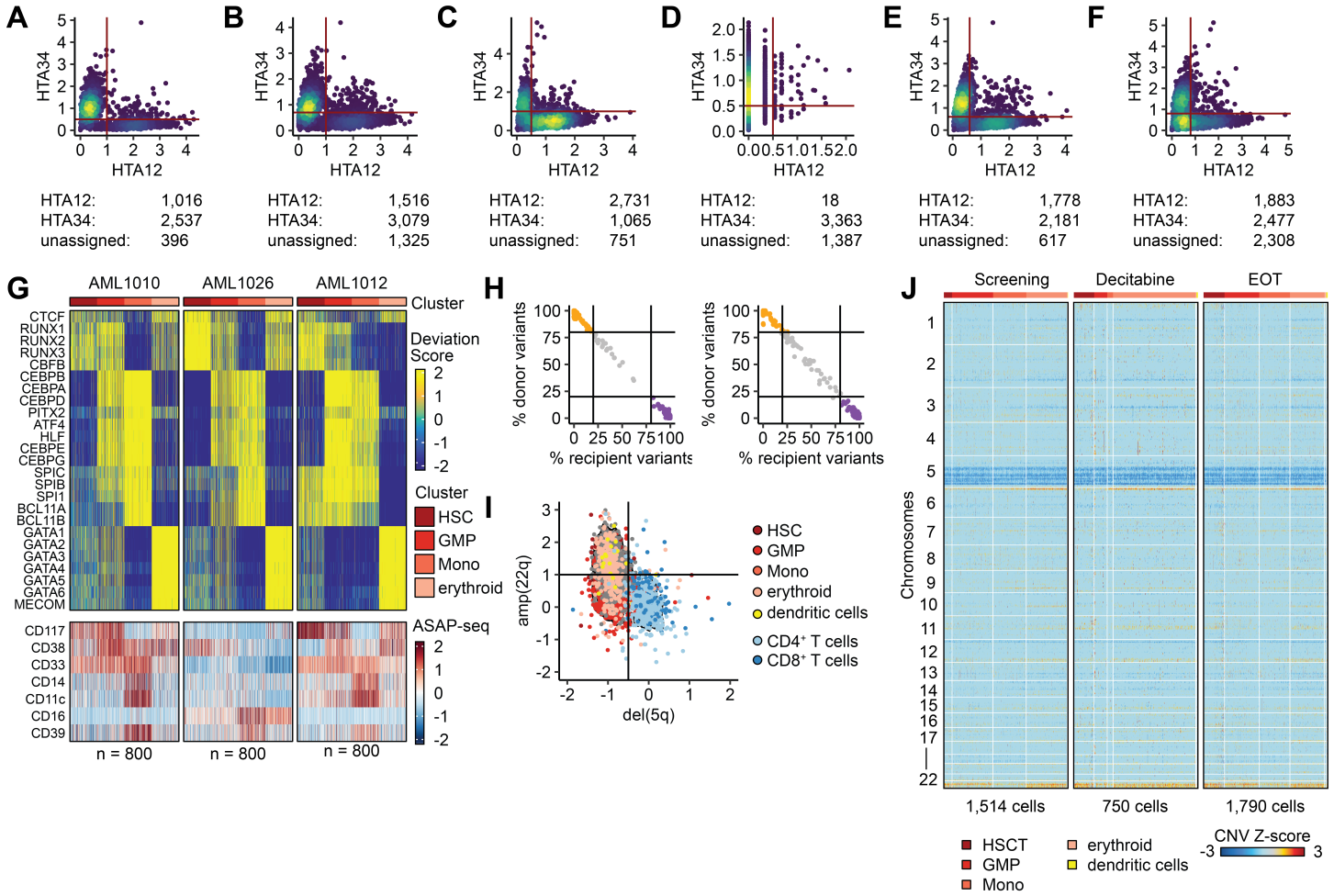
C Chromatin accessibility of *CD36* in AML / recipient- and donor-derived erythroid cells in IST1.

D Surface marker expression (ASAP-seq), gene scores, transcription factor motif activity and mitochondrial DNA (mtDNA) mutations in AML (recipient-) and donor-derived monocytic cells in IST3.

E Chromatin accessibility of *HOXA-AS3* and *HOXA10* (top) and *CCL4L2*, *CCL3L3* and *CCL3L1* (bottom) in AML (recipient-) and donor-derived erythroid cells in IST3.

F Surface marker expression of *CD33* and *CD16* in AML (recipient-) and donor-derived monocytic cells in IST3.

G Surface marker expression (ASAP-seq), gene scores and transcription factor motif activity in AML (recipient-) and donor-derived T cells in IST4.



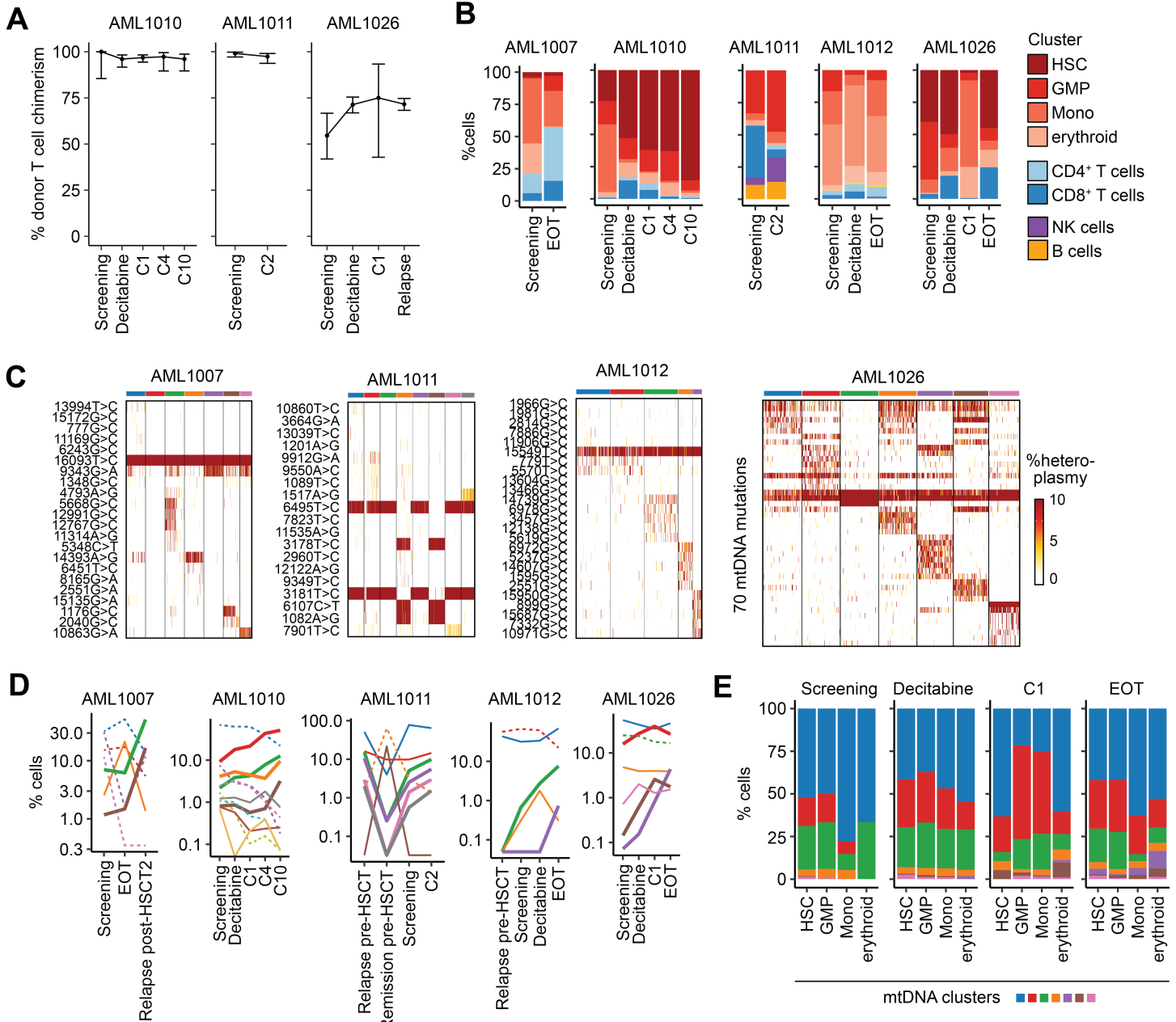
Suppl. Fig. 16. Deconvolution and identification of AML phenotypes in bone marrow using ASAP-seq.

A-F Deconvolution of pooled samples using mean expression Total-seq A hashtags (HTA) measured with ASAP-seq. Numbers indicate the number of cells that could be assigned to HTA1 and HTA2 (HTA12) and HTA3 and HTA4 (HTA34), or that could not be deconvoluted (unassigned).

G Chromatin accessibility profiles (top) and surface marker expression (bottom) across hematopoietic stem cell (HSC)-like, granulocyte macrophage progenitor (GMP)-like, monocytic (Mono) and erythroid cells in AML1010, AML1026 and AML1012.

H Donor-recipient deconvolution using maternal mtDNA variants in AML1010 (left) and AML1026 (right). Orange and purple indicate donor and recipient-derived cells, while cells without clear annotation (i.e. doublets or low-quality cells) are indicated in grey.

I, J Identification of copy number changes *del(5q)* and *amp(22q)* in AML1012 across cell types indicated by the color key. T cells do not harbor any copy number change, while myeloid and progenitor cells have *del(5q)* (I – bottom left) or *del(5q)* and *amp(22q)* (I – top left). *Amp(22q)* is not detected in isolation (I – top right).



Suppl. Fig. 17. Identification of AML subclones using mitochondrial DNA (mtDNA) mutations.

A Longitudinal mtDNA-based single cell T cell chimerism in AML1010, AML1011 and AML1026 throughout treatment with decitabine and ipilimumab.

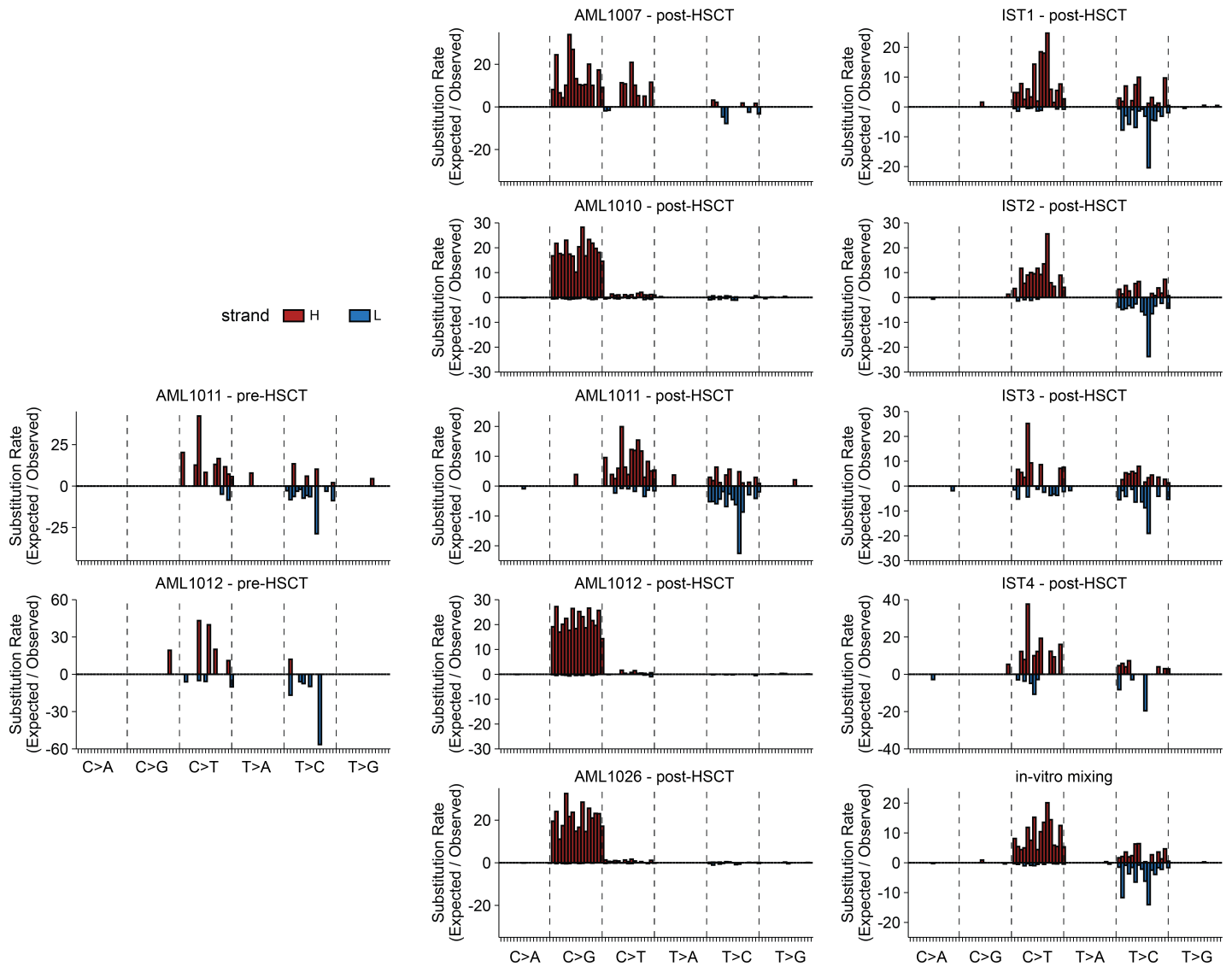
B Longitudinal dynamics of cell types throughout the ETCTN/CTEP 10026 study in the five studied AML cases. In AML1026 blue indicates all T cells.

C Identification of mtDNA-derived subclones in four studied AML cases.

D Longitudinal dynamics of AML subclones through treatment on ETCTN/CTEP 10026. Dotted lines represent clones that diminish over time, while solid lines indicate clones with increase in representation.

E Longitudinal dynamics of mtDNA-based subclusters in AML1026 across different phenotypic compartments, which suggests an increase of subclonal complexity throughout treatment.

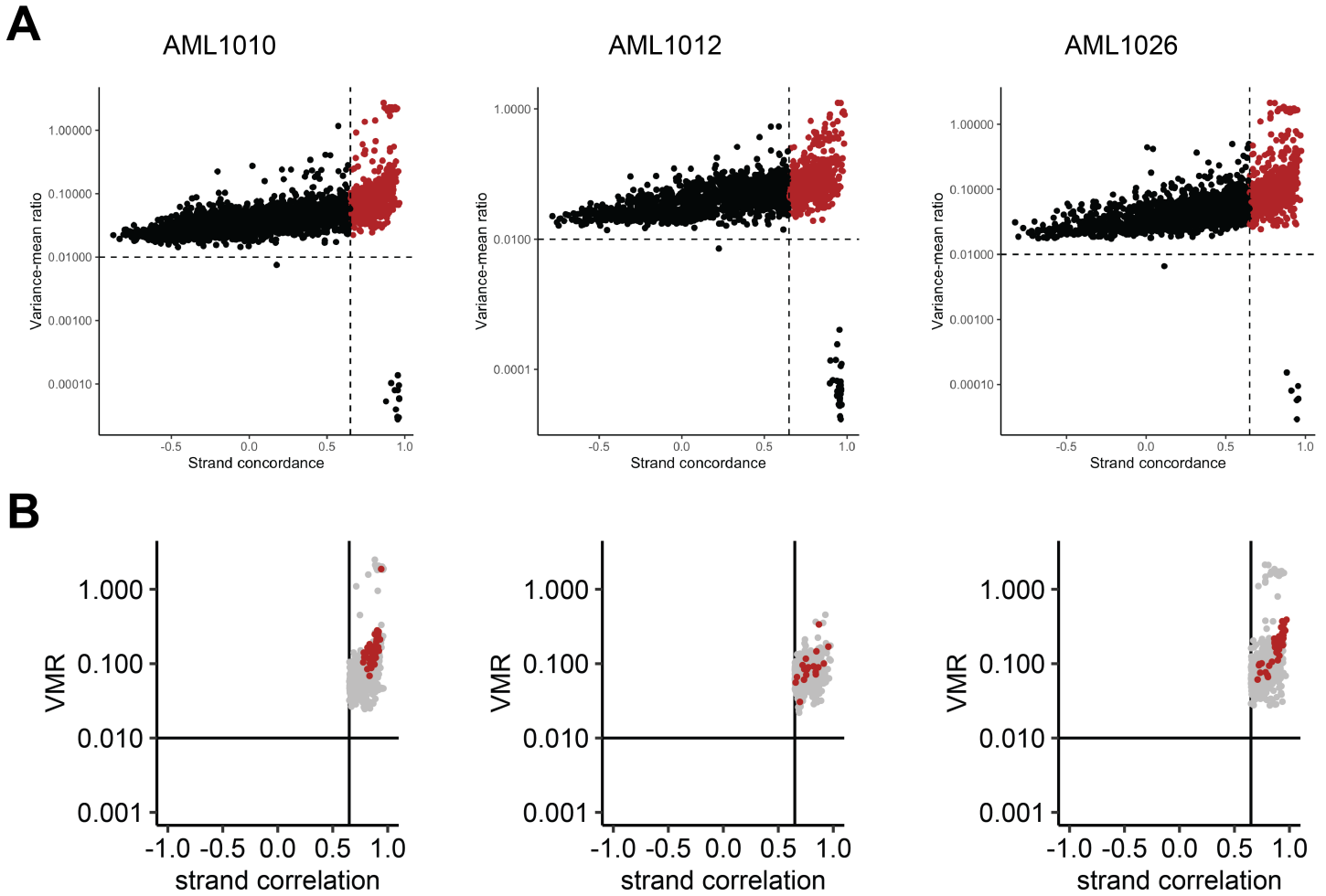
Screening – baseline sample prior to decitabine/ipilimumab treatment. Decitabine – sample after decitabine monotherapy priming cycle. C1/C2/C4/C10 – number of treatment cycles of decitabine/ipilimumab. EOT – end of treatment.



Suppl. Fig. 18. Mutational signatures of mitochondrial DNA (mtDNA) identified using ASAP-seq.

Substitution rate (observed over expected) of mtDNA mutations (y-axis) shown in each category of mononucleotide and trinucleotide modifications are color-coded by the mtDNA heavy (red; upwards) and light (blue; downwards) strand. While post-transplant AML relapse cases AML1007, AML1010, AML1012 and AML1026 show an unexpected pattern of C>G transversion mutations, other samples have the typical signature with transition mutations.

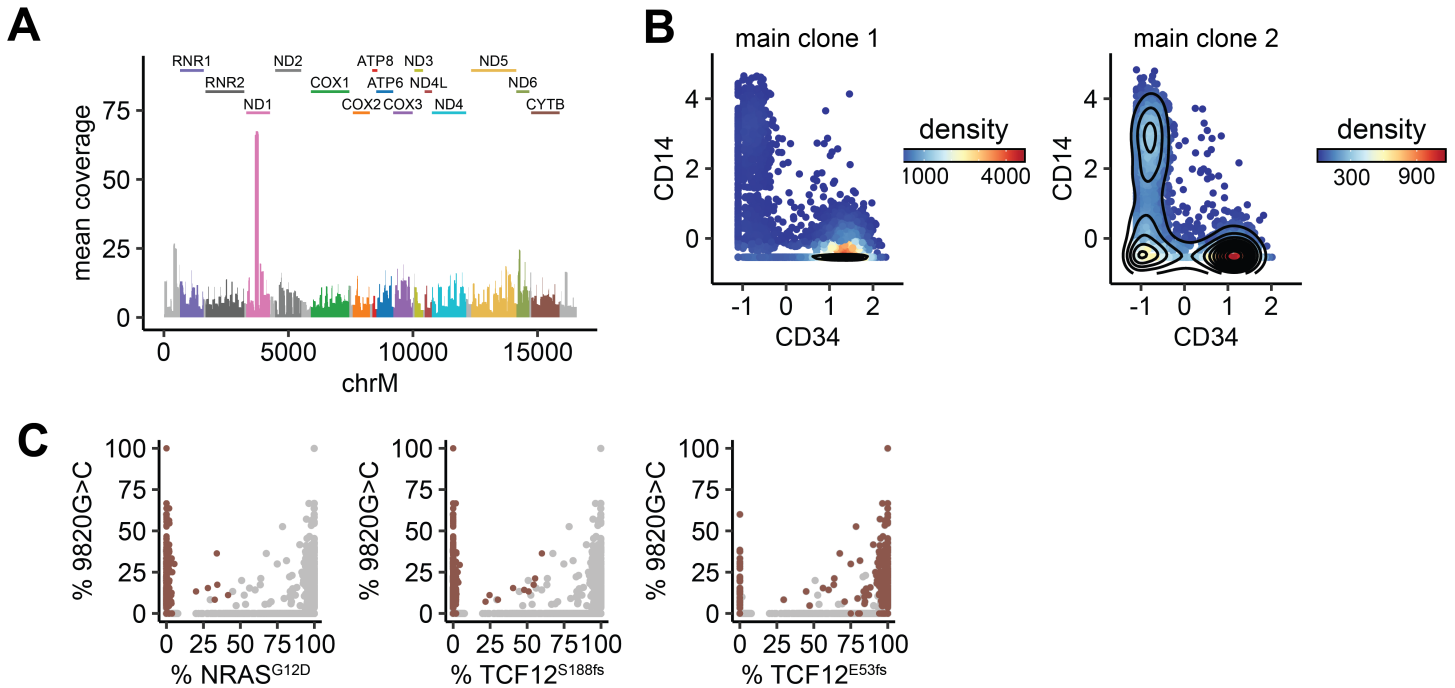
Remarkably, C>G transversion mutations are absent in AML1012 prior to stem cell transplantation.



Suppl. Fig. 19. Identification of high-confidence mitochondrial DNA mutations for Tapestry analysis.

A High-confidence mitochondrial DNA (mtDNA) mutations (red) identified based on strand concordance and variance-mean ratio (VMR) from ASAP-seq data in AML1010, AML1012 and AML1026 served as a comparator for selection of mtDNA variants from Tapestry data.

B As Tapestry amplicons do not have overlapping reads which renders analysis of strand concordance infeasible, mitochondrial DNA mutations used for analysis of Tapestry data (red) were correlated with high-confidence mtDNA mutations found using ASAP-seq (grey).



Suppl. Fig. 20. Co-evolution of somatic nuclear and mitochondrial DNA mutations in AML.

A Mean coverage of the pan mitochondrial DNA (pan-mtDNA) panel used for single cell DNA sequencing with the Tapestry platform (Mission Bio).

B Expression of CD34 and CD14 in main clone 1 (left) and main clone 2 (right) in AML1010.

C Identification of *NRAS*^{wt} cells (absence of *NRAS*^{G12D}, highlighted in brown) also marked by 9820G>A and *TCF12*^{S53fs}.

Supplementary Tables

Suppl. Table 1. Clinical data of CLL cohort reanalyzed from *Gruber et al.*

CLL	IGHV mutation status	Gender	Age at diagnosis in years	Treatment*	Response to therapy
-----	----------------------	--------	---------------------------	------------	---------------------

Genetically stable CLL

2	mutated	male	55	FR	PD
14	unmutated	male	47	FR	CR
15	unmutated	male	58	Alemtuzumab, FCR	CR
17	mutated	female	58	FR	PD
18	mutated	male	47	Bendamustine, Lenalidomide, Rituximab	No Data
36	unmutated	male	52	FR	PR
40	mutated	male	72	FR	No Data
41	mutated	male	65	FR	PR

Clonal evolution (prior to treatment)

1	unmutated	female	43	(FCR)	SD
3	unmutated	female	63	(FR)	CR
4	mutated	male	51	(FR)	No Data
7	mutated	female	56	(FR)	No Data
9	mutated	male	54	(Lenalidomide, FCR)	CR
12	unmutated	male	53	(FCR)	CR
19	mutated	female	53	(FCR)	PR

Clonal evolution (following treatment)

5	unmutated	male	37	FCR	CR
6	unmutated	male	58	FR	No Data
10	mutated	female	58	FR	No Data
11	unmutated	male	54	FR	PD
13	unmutated	male	57	F	PR
16	mutated	male	81	FR	PR
20	mutated	male	56	FR	No Data
21	unmutated	male	35	Lenalidomide, FR	No Data
37	unmutated	female	49	FCR	No Data
38	unmutated	male	37	F	No Data
39	unmutated	male	58	FR	CR

*CLL cases with clonal evolution during natural disease progression were sampled prior to treatment, which is provided in the table only for completeness.

F: fludarabine, C: cyclophosphamide, R: rituximab

PD: progressive disease, CR: complete remission, PR: partial response, SD: stable disease

Suppl. Table 2. Mitochondrial and somatic nuclear DNA mutations identified in CLL using single cell DNA sequencing.

Identifier Penter & Gohil et al.	Identifier Bachireddy et al.	Mutation detection timepoint	Clonal somatic mutations in hg19	Amino acid change	Clonal mtDNA mutations in hg38
CLL 4	CLL 5328	post-HSCT	chr2:233323457	ALPI p.S433S	3538G>A
		founder	chr19:45317870	BCAM p.E311K	4344T>C
		pre-FCR	chr8:86360275	CA3 p.R226W	13916G>A
		post-HSCT	chr10:128841379	DOCK1 p.V697I	16247A>G
		post-HSCT	chr10:103354473	DPCD p.D41E	6426G>A
		post-HSCT	chr16:67709195	GFOD2 p.G341R	16290C>T
		pre-FCR	chr22:23235972	IGLL5 p.G100E	
		pre-FCR	chr5:5464776	KIAA0947 p.R1777*	
		post-HSCT	chr12:4920442	KCNA6 p.A412V	
		pre-FCR	chr20:60912714	LAMA5 p.R699L	
		post-HSCT	chr6:136682295	MAP7 p.Q517*	
		post-HSCT	chr15:42019401	MGA p.Q1152*	
		pre-FCR	chr11:63721451	NAA40 p.M147L	
		pre-FCR	chr8:54163591	OPRK1 p.S3T	
		founder	chr17:57043202	PPM1E p.M244T	
		founder	chr2:198267371	SF3B1 p.H662Q	
		founder	chr19:4363776	SH3GL1 p.K189*	
		post-HSCT	chr4:170043260	SH3RF1 p.R446H	
		pre-FCR	chr1:19439249	UBR4 p.R3857H	
		founder	chr12:966369	WNK1 p.V452L	
		post-HSCT	chr11:6953628	ZNF215 p.H42P	
CLL 5	CLL 5335	founder	chr7:20782612	ABCB5 p.V1046A	2332C>T
		founder	chr9:75543910	ALDH1A1 p.L114F	5979G>A
		subclonal	chr7:65554306	ASL p.Y321C	
		founder	chr2:168726231	B3GALT1 p.P228S	
		founder	chr1:207791577	CR1 p.D2351N	
		founder	chr4:155163826	DCHS2 p.F1892Y	
		founder	chr19:42546856	GRIK5 p.E441K	
		subclonal	chr14:90650709	KCNK13 p.V197I	
		subclonal	chr3:154866416	MME p.K525N	
		founder	chr14:56645144	PELI2 p.S57C	
		founder	chr11:65429479	RELA p.E39Q	
		founder	chr6:54191662	TINAG p.R191H	
		founder	chr19:56172500	U2AF2 p.M144K	
		founder	chr2:158974339	UPP2 p.G115S	
CLL 6	CLL 5327	founder	chr20:31761967	BPIFA2 p.V129L	3830T>C
		founder	chr1:75102115	C1orf173 p.P151R	3526G>A
		founder	chr8:61765460	CHD7 p.2061_2062insL	930G>A
		founder	chr3:98600507	DCBLD2 p.R104C	2623A>G
		founder	chr22:41257549	DNAJB7 p.F150L	9554G>A
		subclonal	chr22:46658486	PKDREJ p.V245G	
		founder	chr8:144994679	PLEC p.G3241fs	
		subclonal	chr17:7577124	TP53 p.V272M	
		founder	chr13:27255288	WASF3 p.V272L	
		founder	chr6:87967798	ZNF292 p.Q1484fs	

The samples were previously analyzed by mtscATAC-seq (Penter & Gohil et al.)(3) and whole-exome sequencing (Bachireddy et al.)(5). The relevant mutations for this study are indicated.

Suppl. Table 3. Total-seq A oligotags.

	Antigen	Clone	Isotype	Total-seq B	Catalogue number	Barcode	Marker purpose
1		LNH-94	IgG1	A0251	394601	GTCAACTCTTTAGCG	Hashtag 1
2		LNH-94	IgG1	A0252	394603	TGATGGCCTATTGGG	Hashtag 2
3		LNH-94	IgG1	A0253	394605	TTCCGCCTCTCTTTG	Hashtag 3
4		LNH-94	IgG1	A0254	394607	AGTAAGTTCAGCGTA	Hashtag 4
5				A0951	405251	AACCTTTGCCACTGC	SAV1
6				A0952	405253	GTCCGACTAATAGCT	SAV2
7				A0953	405255	CAGGTTGTTGTCATT	SAV3
8				A0954	405257	TATTTCCACCCGGTC	SAV4
9				A0955	405259	GTTGTGAGCACGAGA	SAV5
10				A0961	405189	ACCCGGTATCAATTA	SAV6
11	CD34	581	IgG1	A0054	343537	GCAGAAATCTCCCTT	Stem cells

Suppl. Table 4. Total-seq B oligotags.

	Antigen	Clone	Isotype	Total-seq B	Catalogue number	Barcode	Marker purpose
1	CD3	UCHT1	IgG1	B0034	300477	CTCATTGTAACCTCT	T cell
2	CD4	RPA-T4	IgG1	B0072	300565	TGTTCCCGCTCAACT	CD4 T cell
3	CD8	RPA-T8	IgG1	B0080	301069	GCTGCGCTTTCCATT	CD8 T cell
4	CD45RA	HI100	IgG2b	B0063	304161	TCAATCCTTCCGCTT	T cell differentiation
5	CD45RO	UCHL1	IgG2a	B0087	304257	CTCCGAATCATGTTG	T cell differentiation
6	CD62L	DREG-56	IgG1	B0147	304849	GTCCCTGCAACTTGA	T cell differentiation
7	CCR7	G043H7	IgG2a	B0148	353249	AGTTCAGTCAACCGA	T cell differentiation
8	CD28	CD28.2	IgG1	B0386	302961	TGAGAACGACCCTAA	T cell differentiation
9	CD38	HIT2	IgG1	B0389	303547	TGTACCCGCTTGTGA	T cell state / AML blast
10	CD56	QA17A16	IgG1	B0084	392423	TTCGCCGCATTGAGT	NK cell
11	CD57	QA17A04	IgG1	B0168	393323	AACTCCCTATGGAGG	NK cell / T cell state
12	CD127	A019D5	IgG1	B0390	351354	GTGTGTTGTCCTATG	memory T cell
13	CD25	BC96	IgG1	B0085	302647	TTTGTCTGTACGCC	Treg
14	CD39	A1	IgG1	B0176	328241	TTACCTGGTATCCGT	T cell exhaustion
15	PD-1	EH12.2H7	IgG1	B0088	329961	ACAGCGCCGTATTTA	T cell state
16	CD19	HIB19	IgG1	B0050	302263	CTGGGCAATTACTCG	B cell
17	CD138	DL-101	IgG1	B0831	352329	GTATAGACCAAAGCC	Plasma cell
18	CD33	P67.6	IgG1	B0052	366635	TAACTCAGGGCCTAT	AML blast
19	CD117	104D2	IgG1	B0061	313247	AGACTAATAGCTGAC	AML blast
20	CD11c	S-HCL-3	IgG2b	B0053	371523	TACGCCTATAACTTG	Dendritic cell
21	CD14	M5E2	IgG2a	B0081	301857	TCTCAGACCTCCGTA	Monocyte
22	CD16	3G8	IgG1	B0083	302063	AAGTTCACTCTTTGC	Monocyte
23	IgG1	MOPC-21	IgG1	B0090	400185	GCCGGACGACATTAA	Isotype control
24	IgG2b	MPC-11	IgG2b	B0092	400379	ATATGTATCACGCGA	Isotype control
25	IgG2a	MOPC-173	IgG2a	B0091	400291	CTCCTACCTAAACTG	Isotype control

Suppl. Table 5. Immunosuppression tapering cohort.

Patient	Genotype	Conditioning	IS	Days to relapse	Days to remission	Toxicity	Notes
IST1	normal karyotype	Flu/Bu RIC	Rap/Tac/MTX/rituximab	185	105	skin aGvHD	
IST2	<i>inv(11)</i>	Flu/Bu RIC	Tac/MTX	116	184	pancreatitis	participation in GM-CSF vaccination study (6)
IST3	normal karyotype	Flu/Bu RIC	Rap/Tac/MTX	159	126	skin cGvHD	
IST4	<i>tri(8)</i>	Flu/Bu RIC	Rap/Tac/MTX	87	166	nephritis with acute kidney injury and dialysis; oral cGvHD	

Flu: fludarabine; Bu: busulfan; RIC: reduced-intensity conditioning; Rap: sirolimus; Tac: tacrolimus; MTX: methotrexate; aGvHD: acute graft-versus-host disease; cGvHD: chronic graft-versus-host disease; GM-CSF: vaccination with autologous myeloblasts admixed with GM-K562 cells

Suppl. Table 6. Participants of ETCTN/CTEP 10026 studied with ASAP-seq.

Patient	Clinical karyotype	Donor	%blasts screening aspirate	%blasts screening histology	Best response	Time on treatment
AML1007	normal karyotype	related	13	5-10	marrow CR without HI	7 months
AML1010	normal karyotype	unrelated	44	30	stable disease	10 months
AML1011	normal karyotype	unrelated	28	50	stable disease	7 months
AML1012	46,XX,t(1;12)(p12;p13), inv(3)(p?21q?25), del(5)(q11.2q34), add(21)(q22)[15]	related	32	50	stable disease	6 months
AML1026	46,XX,t(9;12)(q22;q13)	unrelated	26	5	stable disease	3 months

CR: complete remission, HI: hematologic improvement

Suppl. Table 7. Mutations targeted using single cell DNA sequencing in AML1010.

Chromosome	Genomic location	Gene	REF	ALT
chr1	114716126	NRAS	C	T
chr1	158547374	OR6Y1	G	A
chr2	166405668	SCN7A	A	G
chr2	240572515	RNPEPL1	G	C
chr3	46021532	XCR1	C	T
chr3	128481288	GATA2	CTTCCTTCTTCAT	C
chr4	129102637	C4orf33	G	A
chr5	157285369	CYFIP2	C	T
chr6	20402431	E2F3	A	G
chr6	168653159	SMOC2	G	A
chr7	21739637	DNAH11	G	T
chr7	100097939	MCM7	A	C
chr8	102560782	ODF1	CTGCAACCCCTGCAGCCCCTGCAACCCG	C
chr10	86900110	BMPR1A	A	C
chr13	52724084	CNMD	T	C
chr13	110480306	COL4A2	G	A
chr14	23080323	ACIN1	G	A
chr15	40903001	VPS18	T	C
chr15	57063758	TCF12	G	GA
chr15	57243509	TCF12	C	CA
chr16	82068237	HSD17B2	C	T
chr17	1659937	PRPF8	C	T
chr19	9186097	OR7D2	T	A
chr19	30443831	ZNF536	G	A
chr19	38473571	RYS1	C	T
chr20	34072083	RALY	G	A
chr22	21000146	THAP7	C	T
chr22	50266979	MAPK11	G	C

REF: reference allele, ALT: alternate allele

Suppl. Table 8. Mutations targeted using single cell DNA sequencing in AML1012.

Chromosome	Genomic location	Gene	REF	ALT
chr3	47916994	MAP4	A	G
chr4	152328233	FBXW7	G	A
chr4	173292151	GALNT7	G	A
chr5	476296	SLC9A3	C	T
chr7	122702400	CADPS2	G	A
chr8	21909554	DOK2	G	T
chr8	109486749	PKHD1L1	G	A
chr8	143818227	PUF60	A	C
chr10	43557531	ZNF239	G	T
chr11	120420764	ARHGEF12	C	T
chr12	2606651	CACNA1C	A	G
chr16	22915253	HS3ST2	G	A
chr17	7675095	TP53	C	T
chr19	49290212	SLC6A16	G	T
chr22	17727240	BCL2L13	AACAAC	A

REF: reference allele, ALT: alternate allele

Suppl. Table 9. Mutations targeted using single cell DNA sequencing in AML1026.

Chromosome	Genomic location	Gene	REF	ALT
chr1	1815790	GNB1	T	C
chr1	114716124	NRAS	C	G
chr1	211829270	LPGAT1	C	T
chr2	124764024	CNTNAP5	C	A
chr2	197401788	SF3B1	C	T
chr2	219491794	SPEG	C	T
chr3	39119933	TTC21A	C	T
chr3	128481291	GATA2	CCTTCTT	C
chr4	663811	PDE6B	G	C
chr4	896481	GAK	G	A
chr5	16701042	MYO10	G	C
chr6	70294211	COL9A1	C	A
chr6	121310781	TBC1D32	C	G
chr8	93786227	TMEM67	C	G
chr9	34726502	FAM205A	C	G
chr9	135106798	OLFM1	C	G
chr10	21517360	SKIDA1	A	T
chr11	5545344	OR52H1	C	G
chr11	5611199	TRIM5	G	C
chr11	64343825	CCDC88B	G	C
chr12	6938704	ATN1	G	C
chr12	53482711	MAP3K12	G	A
chr12	57489913	MARS1	C	T
chr12	106376431	POLR3B	C	G
chr12	123001068	PITPNM2	G	C
chr13	113160062	PROZ	G	A
chr15	40294323	PLCB2	G	A
chr15	41866163	SPTBN5	G	A
chr16	4600149	C16orf96	G	C
chr17	76736877	SRSF2	G	T
chr17	82253039	CSNK1D	G	C
chr19	8950131	MUC16	G	C
chr19	10581776	AP1M2	G	A
chr19	41880917	CD79A	C	G
chr19	49864337	PNKP	G	C
chr19	49909475	IL4I1	G	C
chr21	30592756	KRTAP6-3	C	T
chr21	44637679	TSPEAR	T	A
chrX	49880642	USP27X	G	A
chrX	101041322	TRMT2B	G	A
chrX	130243603	ZNF280C	G	A

REF: reference allele, ALT: alternate allele

Suppl. Table 10. Samples analyzed for this work using single cell sequencing platforms.

Cohort	Patient	Sample	Tissue	ASAP-seq	scRNA-seq	scDNA-seq
<i>ETCTN/CTEP 10026</i>	AML1007	Screening	BM	yes	reanalysis	no
		EOT	BM	yes	reanalysis	no
		relapse post-HSCT2	BM	yes	no	no
	AML1010	Screening	BM	yes	reanalysis	yes
		EOLN	BM	yes	reanalysis	no
		C1	BM	yes	reanalysis	no
		C4	BM	yes	reanalysis	no
		C10	BM	yes	reanalysis	yes
	AML1011	relapse pre-HSCT	BM	yes	no	no
		remission pre-HSCT	BM	yes	no	no
		Screening	BM	yes	no	no
		C2	BM	yes	no	no
	AML1012	relapse pre-HSCT	BM	yes	no	no
		Screening	BM	yes	reanalysis	yes
		Decitabine	BM	yes	reanalysis	no
		C2	BM	yes	reanalysis	no
		EOT	BM	yes	reanalysis	yes
	AML1016	Screening	BM	no	reanalysis	no
		C2	BM	no	reanalysis	no
	AML1022	Screening	BM	no	reanalysis	no
		EOLN	BM	no	reanalysis	no
		C1	BM	no	reanalysis	no
	AML1026	Screening	BM	yes	reanalysis	yes
		EOLN	BM	yes	reanalysis	no
		C1	BM	yes	reanalysis	yes
		Relapse	BM	yes	reanalysis	no
<i>CLL relapse post-HSCT</i>	CLL5327 (CLL6)	pre-FCR	PB	no	no	yes
		post-HSCT	PB	no	no	yes
	CLL5328 (CLL4)	pre-FCR	PB	no	no	yes
		post-HSCT	PB	no	no	yes
	CLL5335 (CLL5)	pre-FCR	PB	no	no	yes
		post-HSCT	PB	no	no	yes
<i>Immuno-suppression tapering</i>	IST1	pre-IST	PB, BM	yes	no	no
		response	PB, BM	yes	no	no
	IST2	pre-IST	BM	yes	no	no
		response	BM	yes	no	no
	IST3	pre-IST	PB	yes	no	no
		response	PB	yes	no	no
	IST4	pre-IST	PB	yes	no	no
		response	PB	yes	no	no
<i>In-vitro mixing</i>	mix	ratio 1:3	T cells, PB	yes	no	no
	mix	ratio 1:30	T cells, PB	yes	no	no
	mix	ratio 1:300	T cells, PB	yes	no	no

EOT – end of treatment, C1/C2/C4/C10 – cycle 1/2/4/10 of decitabine/ipilimumab, pre-FCR – prior to immunochemotherapy with fludarabine, cyclophosphamide and rituximab (FCR), pre-/post-HSCT – prior to/following allogeneic hematopoietic stem cell transplantation (HSCT), PB – peripheral blood, BM – bone marrow, IST – immunosuppression tapering

Supplementary Notes

Analysis of maternal DNA variants extracted from bulk RNA-seq data.

To address the question of how many maternally inherited mitochondrial DNA mutations would likely be able to distinguish between donor and recipient in the setting of unrelated transplants, we simulated unrelated transplants and compared the maternal mtDNA variants of these simulated pairs using bulk mitochondrial DNA profiles of 81 individuals extracted from two public datasets (PRJNA486215 and PRJNA563929) (**Suppl. Fig. 10A**). We identified the homoplasmic mitochondrial DNA variants per individual (cut-off: heteroplasmy >98%), as those are the ones that are maternally inherited and thus can be used for distinguishing individuals. We then counted the number of mutually exclusive (= diverging) mitochondrial variants in each of the 6,480 (i.e. 81 x 80) possible pairings.

Mutational profiles of mitochondrial DNA mutations in AML relapse cases

Compared to the expected mutational signatures with C>T and T>C transition mutations previously identified using mtscATAC-seq(3,7), post-transplant AML relapse cases contained an usually high number of G>C transversion mutations (**Suppl. Fig. 18**). These mtDNA mutations passed filtering for high-confidence variants (strand concordance >0.65, variance-mean ratio <0.01 and detection in more than 5 cells) (**Suppl. Fig. 19A**) and were also detectable using single cell DNA sequencing (**Suppl. Fig. 19B**), indicating that they do not constitute an obvious sequencing artifact. Although we can only speculate on their origin, we think it likely that they may arise due to oxidative stress(8), either as a result of AML disease biology or the context of allogeneic stem cell transplantation.

Interestingly, in AML1012 G>C mutations were absent prior to transplantation, potentially suggesting an effect induced by the conditioning therapy or graft-versus-leukemia responses (GvL).

Supplementary references

1. Hjelm BE, Rollins B, Morgan L, Sequeira A, Mamdani F, Pereira F, et al. Splice-Break: exploiting an RNA-seq splice junction algorithm to discover mitochondrial DNA deletion breakpoints and analyses of psychiatric disorders. *Nucleic Acids Res.* 2019;47:e59.
2. Hu Z, van der Ploeg K, Chakraborty S, Arunachalam PS, Mori DAM, Jacobson KB, et al. Early immune markers of clinical, virological, and immunological outcomes in patients with COVID-19: a multi-omics study. *Elife.* 2022;11:e77943.
3. Penter L, Gohil SH, Lareau C, Ludwig LS, Parry EM, Huang T, et al. Longitudinal single-cell dynamics of chromatin accessibility and mitochondrial mutations in chronic lymphocytic leukemia mirror disease history. *Cancer Discov. American Association for Cancer Research;* 2021;11:3048–63.
4. Penter L, Liu Y, Wolff JO, Yang L, Taing L, Jhaveri A, et al. Mechanisms of response and resistance to combined decitabine and ipilimumab for advanced myeloid disease. *Blood.* 2023;141:1817–30.
5. Bachireddy P, Ennis C, Nguyen VN, Gohil SH, Clement K, Shukla SA, et al. Distinct evolutionary paths in chronic lymphocytic leukemia during resistance to the graft-versus-leukemia effect. *Science Translational Medicine.* 2020;12:eabb7661.
6. Ho VT, Kim HT, Bavli N, Mihm M, Pozdnyakova O, Piesche M, et al. Vaccination with autologous myeloblasts admixed with GM-K562 cells in patients with advanced MDS or AML after allogeneic HSCT. *Blood Adv.* 2017;1:2269–79.
7. Penter L, ten Hacken E, Southard J, Lareau CA, Ludwig LS, Li S, et al. Mitochondrial DNA Mutations as Natural Barcodes for Lineage Tracing of Murine Tumor Models. *Cancer Research.* 2022;OF1–6.
8. Kino K, Sugiyama H. Possible cause of G-C→C-G transversion mutation by guanine oxidation product, imidazolone. *Chem Biol.* 2001;8:369–78.