Supplemental Data

The Immune Microenvironment Shapes Transcriptional and Genetic Heterogeneity in Chronic Lymphocytic Leukemia

Clare Sun, Yun-Ching Chen, Aina Zurita Martinez, Maria Joao Baptista, Stefania Pittaluga, Delong Liu, PhD, Daniel Rosebrock, Satyen Harish Gohil, Nakhle S. Saba, Theresa Davies-Hill, Sarah E. M. Herman, Gad Getz, Mehdi Pirooznia, Catherine J. Wu, Adrian Wiestner

Table of Contents

Supplemental Methods 2
Supplemental Table S1 7
Supplemental Table S2 11
Supplemental Table S3 13
Supplemental Figure S1 14
Supplemental Figure S2 15
Supplemental Figure S4 17
Supplemental Figure S5 18
Supplemental Figure S6 19
Supplemental References 20

Supplemental Methods

Bulk RNA Sequencing and Analysis

RNA was extracted from CD19$^+$ CLL cells or snap frozen LN biopsies using RNeasy kits (Qiagen, Germantown, MD, USA). Total RNA libraries were prepared according to the Illumina TruSeq protocol. Libraries were pooled and sequenced on a HiSeq 2000 (Illumina). Reads were aligned to hg38 using STAR [1], summarized by featureCounts,[2] and normalized across samples.[3] Counts per million (CPM) were used to measure gene expression. Genes with CPM > 1 in ≥ 3 samples were analyzed in downstream analysis. Paired gene-level differential analysis was conducted using the limma package.[3] Gene Set Enrichment Analysis (GSEA)[4] or overrepresentation analysis using ClusterProfiler[5] was used to test for enrichment of a curated set of lymphocyte gene-expression signatures (Supplemental Table S2). Enrichment was defined as a normalized enrichment score (NES) ≥ 1.6 and FDR < 0.05 with ≥ 5 leading-edge genes and > 20% of total leading-edge genes that were unique to each signature. Some signatures (e.g., HIF1α ↓, KLF2 ↓ and nutrient deprivation ↓) were defined by downregulated genes in a pathway and had a positive NES for pathways enriched in PB samples.

CIBERSORT and the LM22 leukocyte signature matrix[6] was used for deconvolution of bulk RNA-seq data from normal LN samples into 22 immune cell types. For CLL LN samples, CIBERSORT was applied using a custom signature matrix generated by replacing gene expression profiles of naïve and memory B cells with that of CD19$^+$ CLL cells purified from LN samples.[7] RNA-seq data was quantified in FPKM (Fragments Per Kilobase Million) as input for CIBERSORT.

Single cell RNA Sequencing and Analysis

Individually barcoded single-cell libraries were created from LN SCS with the Chromium Single Cell 3′ Reagent Kit V2. Libraries were sequenced on NovaSeq (Illumina) and processed by Cell Ranger Single Cell Software Suite 3.0.1.[8] Genes expressed in < 3 cells and cells with < 200 genes detected or > 10% mitochondrial gene content were filtered.

Downstream analyses were performed on Seurat v3.[9] For each sample, gene expression was normalized to count per 10,000 and the top 2,000 variable genes were identified and scaled with effects of total number of detected genes and mitochondrial gene content regressed out using Seurat functions 'FindVariableFeatures' and 'ScaleData' respectively. Samples were integrated using canonical correlation analysis (CCA) with 30 CCA components. Clustering was done by IKAP,[10] an iterative parameter-choosing algorithm implemented on top of Seurat SNN (Shared Nearest Neighbor) clustering. Differentially expressed genes were identified using Seurat function 'FindMarkers'. G2/M and S scores were computed by Seurat function 'CellCycleScoring'. When integrating PB and LN samples, LN samples were used as reference. As expected, results from downstream analyses of the integrated PB and LN dataset, including visualization in UMAP, were not identical to those obtained from LN single cells only.

RNA velocity was performed by first counting unspliced and spliced mRNA for each LN sample using the package 'velocyto' followed by estimating cell transitions using the R packages 'SeuratWrappers' and 'velocyto.R' for the deterministic model and the Python packages 'scVelo' for stochastic and dynamical models.[11,12]

All parameters were set as default unless otherwise specified.

Whole Exome Sequencing and Analysis

DNA was extracted from $CD19^+$ CLL cells using DNeasy kits (Qiagen, Germantown, MD, USA). Libraries were prepared using Agilent SureSelect Human All Exon kit and Nextera DNA Library Prep for Enrichment. Pooled libraries were sequenced on Illumina next-generation sequencers. BAM files were generated from the Picard pipeline then analyzed on the Firehose platform, including quality control, local re-alignment, mutation calling, small insertion and deletion identification, rearrangement detection and coverage calculations.

Somatic mutations were identified in targeted exon data using the MuTect algorithm.[13] Cross-contamination was estimated with ContEst[14] and was used to set the lowest allelic fraction at which somatic mutations could be detected on a per sample basis. Candidate indels were detected using Indelocator. All mutations were filtered using a panel of normal

samples, which removes mutations commonly seen across a large number of sequenced normal (non-cancer) samples. All paired tumor-normal pairs were run through deTiN[15] to estimate the contamination of tumor DNA in the normal sample and keep mutations which would otherwise have been removed by germline filters. For each patient, the union of all point mutations and indels from every sample was created. The mutant and reference allele counts of every mutation in this union were measured in each sample using samtools in a process called force-calling.[16]

Somatic copy number alterations were inferred from the ratio of tumor read depth to the expected read depth derived from a panel of normal samples using the ReCapSeg tool.[17,18] ABSOLUTE[19] was used to estimate sample purity, ploidy, absolute somatic copy number, and the cancer cell fraction (CCF) of mutations. To distinguish subclonal populations, a Bayesian clustering procedure[20,21] was applied to CCFs across PB and LN samples from each patient.

Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
LIVE/DEAD Fixable Aqua Dead Cell Stain Kit	Invitrogen	Cat# L34965
HLA-DR FITC (G46-6)	BD Biosciences	Cat# 555811, RRID:AB_396145
CCR7 PE-CF594 (150503)	BD Biosciences	Cat# 562381, RRID:AB_11153301
CD45RO APC (UCHL1)	BD Biosciences	Cat# 559865, RRID:AB_398673
CD3 APC-H7 (SK7)	BD Biosciences	Cat# 560176, RRID:AB_1645475
CD14 V500 (M5E2)	BD Biosciences	Cat# 561391, RRID:AB_10611856
CD19 V500 (HIB19)	BD Biosciences	Cat# 561121, RRID:AB_10562391
CD8 BV650 (RPA-T8)	BD Biosciences	Cat# 563821, RRID:AB_2744462
CD4 BV786 (SK3)	BD Biosciences	Cat# 563877, RRID:AB_2738462
Critical Commercial Assays		

CD19 MicroBeads	MACS Miltenyi	Cat# 130-097-055
LS Columns	MACS Miltenyi	Cat# 130-042-401
LD Columns	MACS Miltenyi	Cat# 130-042-901
TruSeq® Stranded Total RNA Library Prep	Illumina	Cat# 20020597
Chromium Single Cell 3' Library and Gel Bead Kit v2	10X Genomics	Cat# PN-120237
Chromium Single Cell A Chip Kit	10X Genomics	Cat# PN-1000009
Agilent SureSelect Human All Exon kit	Illumina	Cat# 20020188
Nextera DNA Library Prep for Enrichment	Illumina	Cat# 20020188
Software and Algorithms		
STAR	Dobin et al., 2013	https://github.com/alexdobin/STAR
featureCounts	Liao et al., 2014	http://subread.sourceforge.net
limma	Law et al., 2014	https://github.com/cran/limma
GSEA	Subramanian et al., 2005	http://software.brobroadinstit.org/gsea/index.jsp
clusterProfiler	Yu et al., 2012	http://bioconductor.org/packages/release/bioc/html/clusterProfiler.html
CEMiTool	Russo et al., 2018	https://bioconductob.org/packages/relerel/bioc/html/CEMiCEMi.html
Cell Ranger	10X Genomics	https://bioconductob.org/packages/relerel/bioc/html/CEMiCEMi.html
Seurat	Butler et al., 2018	https://satijalab.org/seurat
Picard	Broad Institute, Cambridge, MA	http://picard.sourceforge.net
Firehose	Broad Institute, Cambridge, MA	http://www.broadinstitute.org/cancer/cga/Firehose
MuTect	Broad Institute, Cambridge, MA	http://www.broadinstitute.org/cancer/cga/mutect
ContEst	Broad Institute, Cambridge, MA	https://software.broadinstitute.org/cancer/cga/contest
Indelocator	Broad Institute, Cambridge, MA	https://software.broadinstitute.org/cancer/cga/indelocator
ReCapSeg	Broad Institute, Cambridge, MA	http://gatkforums.bbroadinstitut.org/gaga/categories/recapsre-documentation

FlowJo v10.0.7	Tree Star, Inc., Ashland, OR	http://www.flowjo.com
Prism 7	GraphPad Software, Inc., San Diego, CA	

Supplemental Table S1. Differentially expressed genes between lymph node and peripheral blood.

Gene	Log_2 FC	*p* value	FDR
MKI67	4.63	2.41E-29	5.29E-25
FOS	4.56	3.15E-18	3.01E-15
FOSB	4.48	2.20E-16	8.79E-14
RGS1	3.95	1.18E-19	1.52E-16
HIST1H3B	3.81	2.41E-22	1.06E-18
ASPM	3.28	1.40E-21	5.12E-18
TOP2A	3.19	1.30E-22	9.56E-19
RRM2	2.92	3.46E-20	7.62E-17
CENPF	2.90	3.89E-19	4.76E-16
CCL3	2.84	2.41E-21	7.59E-18
HIST2H3C	2.77	3.33E-17	1.88E-14
HIST2H3A	2.77	2.85E-17	1.65E-14
TPX2	2.76	5.71E-20	1.12E-16
AC069363.1	2.74	4.88E-21	1.34E-17
CCL4	2.59	7.64E-20	1.18E-16
DLGAP5	2.59	4.95E-18	4.03E-15
CDK1	2.46	3.35E-18	3.07E-15
NR4A2	2.46	3.52E-14	6.73E-12
BUB1	2.41	1.05E-19	1.44E-16
HJURP	2.41	6.99E-20	1.18E-16
KIF14	2.32	8.04E-17	3.77E-14
BIRC5	2.29	3.85E-18	3.39E-15
CEP55	2.27	1.35E-17	9.29E-15
DTL	2.22	1.47E-16	6.09E-14
ANLN	2.21	1.43E-16	6.08E-14
MCM10	2.16	8.27E-18	6.06E-15
POLQ	2.16	4.21E-18	3.56E-15
KLF4	2.15	2.61E-10	1.05E-08
KIF23	2.07	5.41E-17	2.83E-14
ESPL1	2.07	7.28E-16	2.43E-13
CDCA5	2.04	1.21E-16	5.33E-14
NCAPH	2.03	1.70E-17	1.13E-14
CR2	2.03	1.21E-27	1.33E-23
HIST1H3G	2.02	2.36E-10	9.66E-09
KIF4A	2.02	5.58E-17	2.85E-14
DUSP1	2.02	5.67E-12	4.22E-10
ZWINT	2.01	2.58E-17	1.53E-14
CDC45	2.00	7.60E-18	5.77E-15
MELK	2.00	2.15E-18	2.15E-15

Gene	Log_2 FC	*p* value	FDR
CSF1	1.26	1.56E-10	6.76E-09
MYC	1.26	3.83E-12	3.06E-10
HIST1H4A	1.25	1.08E-10	5.02E-09
KIFC1	1.25	2.66E-13	3.49E-11
COL6A3	1.25	1.16E-07	1.73E-06
LDLRAD4	1.25	1.28E-12	1.20E-10
MMP9	1.25	1.11E-07	1.67E-06
HIST1H3F	1.24	1.04E-07	1.57E-06
FABP5	1.24	2.11E-12	1.86E-10
SAPCD2	1.24	2.32E-14	4.78E-12
ASB2	1.24	3.37E-15	9.26E-13
PIF1	1.24	2.79E-15	7.86E-13
CACNA1D	1.23	4.11E-11	2.20E-09
C3	1.23	1.86E-06	1.81E-05
SUSD1	1.23	1.54E-16	6.27E-14
CD4	1.22	7.30E-11	3.55E-09
TULP2	1.22	2.78E-09	7.52E-08
PYCR1	1.21	8.26E-14	1.30E-11
CPNE7	1.21	4.64E-12	3.61E-10
PAG1	1.21	3.68E-08	6.52E-07
PDCD1	1.20	1.75E-17	1.13E-14
GIMAP5	1.20	6.73E-09	1.57E-07
JUN	1.19	8.97E-06	6.92E-05
MCM4	1.19	4.24E-11	2.24E-09
PRKCH	1.19	1.14E-07	1.70E-06
MRC2	1.19	3.62E-11	1.98E-09
ASF1B	1.19	3.96E-11	2.13E-09
LPIN1	1.19	3.13E-14	6.15E-12
EGR3	1.18	5.89E-08	9.78E-07
KIF20A	1.18	3.77E-11	2.05E-09
LHFP	1.18	2.69E-12	2.25E-10
DSCC1	1.18	1.83E-12	1.65E-10
MRC1	1.18	1.16E-08	2.47E-07
H2AFX	1.18	3.24E-12	2.65E-10
UHRF1	1.17	9.10E-11	4.31E-09
CD3E	1.17	5.91E-10	2.04E-08
PTGER4	1.16	7.93E-16	2.60E-13
CXCL12	1.16	3.90E-07	4.77E-06
ARHGAP11A	1.16	2.14E-12	1.88E-10

HIST1H3C	1.98	8.92E-13	8.92E-11
SLC40A1	1.95	1.04E-10	4.86E-09
CLSPN	1.95	7.28E-16	2.43E-13
EGR1	1.93	6.08E-09	1.44E-07
DIAPH3	1.91	1.74E-18	1.83E-15
FYB	1.90	3.65E-11	2.00E-09
HSPA1B	1.90	1.04E-08	2.25E-07
DUSP4	1.90	1.74E-12	1.57E-10
CENPE	1.89	2.37E-17	1.45E-14
DST	1.89	8.10E-10	2.66E-08
AC145110.1	1.87	5.67E-16	2.12E-13
ANKRD13B	1.87	1.32E-17	9.29E-15
VCAM1	1.86	5.57E-11	2.85E-09
DUSP2	1.84	1.64E-13	2.37E-11
HMMR	1.84	5.00E-14	8.80E-12
KIF18B	1.84	8.84E-16	2.86E-13
TICRR	1.84	1.60E-15	4.75E-13
CDCA2	1.83	4.58E-15	1.23E-12
CKAP2L	1.83	2.31E-15	6.61E-13
UBE2C	1.81	8.09E-14	1.29E-11
SEMA7A	1.80	3.96E-17	2.18E-14
SLC2A3	1.77	1.09E-14	2.47E-12
RGS2	1.72	9.20E-08	1.42E-06
DUSP6	1.72	9.26E-15	2.15E-12
TROAP	1.71	3.92E-14	7.43E-12
TTK	1.70	4.17E-14	7.64E-12
CDC6	1.70	7.25E-13	7.59E-11
GTSE1	1.67	4.12E-14	7.62E-12
RP11-452K12.6	1.67	3.71E-13	4.44E-11
EXO1	1.67	3.99E-15	1.08E-12
NUSAP1	1.66	5.91E-16	2.17E-13
PPP1R15A	1.66	1.65E-12	1.52E-10
CDC25A	1.64	1.24E-14	2.76E-12
IPCEF1	1.63	9.16E-17	4.20E-14
EGR2	1.62	1.05E-11	7.15E-10
SIGLEC1	1.62	3.41E-11	1.87E-09
CDK5R1	1.60	6.41E-18	5.04E-15
AURKB	1.59	1.17E-15	3.64E-13
DCAF12	1.58	3.96E-14	7.44E-12
CD1C	1.58	1.44E-10	6.33E-09

RP11-452K12.7	1.15	4.47E-12	3.50E-10
TNFRSF9	1.15	1.92E-10	8.12E-09
XRCC2	1.14	3.51E-12	2.83E-10
ITK	1.14	7.36E-08	1.18E-06
CCL4L2	1.14	1.35E-14	2.98E-12
NUF2	1.14	8.15E-10	2.67E-08
PAQR4	1.13	6.13E-14	1.04E-11
FAM72C	1.13	1.48E-11	9.32E-10
BRCA1	1.12	1.78E-13	2.55E-11
CCR6	1.12	5.81E-09	1.39E-07
NME1	1.12	1.50E-11	9.36E-10
E2F1	1.11	2.67E-13	3.49E-11
IL2RB	1.11	1.70E-11	1.04E-09
DMXL2	1.11	5.10E-07	6.01E-06
BCAR3	1.11	4.08E-14	7.61E-12
IL21R	1.11	5.32E-12	4.02E-10
RP11-564A8.4	1.10	1.03E-11	7.03E-10
CCDC152	1.10	1.52E-06	1.54E-05
CD2	1.09	6.81E-09	1.59E-07
CTD-2313F11.1	1.09	8.89E-08	1.38E-06
RP11-564A8.8	1.09	5.31E-11	2.73E-09
UBE2T	1.09	3.10E-13	3.86E-11
FAM72D	1.09	7.49E-12	5.35E-10
FMNL3	1.09	2.87E-15	8.00E-13
ABCG1	1.08	5.98E-09	1.42E-07
HIST2H4B	1.08	3.10E-09	8.22E-08
SLC2A14	1.08	1.13E-08	2.41E-07
HIST2H4A	1.08	3.05E-09	8.13E-08
SLC29A1	1.08	2.43E-11	1.40E-09
FAM95B1	1.08	6.62E-11	3.28E-09
FABP5P7	1.07	3.29E-11	1.82E-09
HIST1H1B	1.07	4.69E-06	4.01E-05
FAM95B1	1.07	5.94E-11	3.00E-09
MIR4435-2HG	1.07	3.02E-12	2.49E-10
IL21R-AS1	1.07	2.11E-11	1.24E-09
C1QB	1.07	1.06E-08	2.29E-07
PLK3	1.06	6.44E-16	2.25E-13
GFI1	1.06	3.82E-13	4.54E-11
RP11-146D12.2	1.05	5.91E-11	3.00E-09
PTGDS	1.05	1.75E-08	3.46E-07

RP11-293M10.2	1.54	5.21E-12	3.97E-10
SLCO2B1	1.54	8.71E-10	2.83E-08
CCND2	1.52	2.17E-22	1.06E-18
PRR11	1.52	1.17E-14	2.63E-12
TYMS	1.52	4.53E-14	8.17E-12
KIF2C	1.52	6.84E-16	2.35E-13
DAB2	1.51	8.73E-10	2.83E-08
CD69	1.50	8.95E-12	6.18E-10
TNFAIP2	1.49	9.78E-09	2.13E-07
CCNB1	1.49	2.45E-13	3.26E-11
CKS2	1.47	4.17E-16	1.61E-13
DEPDC1B	1.47	9.96E-15	2.28E-12
RGCC	1.46	4.43E-10	1.64E-08
KIAA0101	1.46	2.92E-12	2.42E-10
IQGAP3	1.45	1.52E-14	3.21E-12
KLF10	1.45	4.55E-10	1.67E-08
ETV5	1.44	1.41E-14	3.08E-12
SHCBP1	1.44	2.57E-12	2.16E-10
BUB1B	1.43	5.83E-12	4.31E-10
RP11-4O1.2	1.43	6.33E-17	3.17E-14
CD83	1.43	6.20E-16	2.20E-13
ESCO2	1.43	5.52E-13	6.07E-11
PKMYT1	1.42	4.97E-15	1.29E-12
NR4A3	1.42	5.41E-12	4.08E-10
TBC1D4	1.42	3.03E-08	5.50E-07
NEK2	1.40	8.60E-13	8.68E-11
SEPP1	1.40	8.08E-08	1.27E-06
ZMIZ1	1.39	7.92E-14	1.27E-11
RP11-823E8.3	1.39	9.00E-13	8.96E-11
MFSD2A	1.38	8.04E-20	1.18E-16
SPC24	1.37	1.18E-12	1.13E-10
BTBD19	1.37	3.43E-16	1.35E-13
STUM	1.36	6.48E-15	1.63E-12
INPP4B	1.36	1.26E-10	5.61E-09
CDCA7	1.35	1.98E-12	1.76E-10
CDCA3	1.35	1.35E-12	1.26E-10
MAF	1.33	5.81E-11	2.96E-09
DTHD1	1.33	1.14E-08	2.44E-07
RP11-167N4.4	1.33	2.24E-09	6.28E-08
HSPA1A	1.32	3.12E-07	3.97E-06

THBS1	1.05	1.25E-06	1.31E-05
RP11-347P5.1	1.05	1.94E-08	3.76E-07
NELL2	1.05	1.04E-08	2.25E-07
UBE2S	1.04	8.50E-12	5.95E-10
TIMD4	1.04	1.65E-10	7.06E-09
CHI3L2	1.04	6.27E-11	3.15E-09
HK2	1.04	7.36E-15	1.80E-12
LINC00996	1.03	1.54E-11	9.56E-10
ST8SIA1	1.03	9.18E-08	1.42E-06
ECT2	1.03	5.15E-10	1.85E-08
WDFY3	1.03	1.63E-07	2.29E-06
BRIP1	1.03	6.31E-11	3.16E-09
LMNA	1.03	9.25E-05	5.06E-04
APOE	1.03	2.76E-07	3.57E-06
KIF11	1.03	1.20E-12	1.14E-10
FAM72B	1.02	1.77E-11	1.07E-09
PTPN7	1.02	7.78E-14	1.27E-11
SPRED1	1.02	8.33E-09	1.88E-07
GINS1	1.01	1.22E-10	5.48E-09
IL7R	1.01	1.27E-05	9.35E-05
METTL1	1.01	1.27E-15	3.87E-13
SH2D1A	1.01	1.13E-08	2.42E-07
PLXNB2	1.01	9.51E-10	3.06E-08
MIR222HG	1.01	6.97E-13	7.38E-11
CCL5	1.00	1.71E-08	3.41E-07
SGK1	1.00	7.86E-11	3.82E-09
DNASE1L3	1.00	5.57E-12	4.18E-10
RASSF4	1.00	1.53E-09	4.58E-08
RP11-876N24.1	1.00	4.27E-11	2.25E-09
ITGAD	1.00	5.75E-10	2.01E-08
LPL	1.00	1.43E-11	9.08E-10
MAP1B	-1.00	5.34E-10	1.90E-08
CTD-3224K15.2	-1.00	4.14E-11	2.20E-09
RP5-1042I8.7	-1.01	8.08E-15	1.95E-12
STMN3	-1.01	4.65E-13	5.33E-11
PLAC8	-1.02	7.81E-09	1.78E-07
LINC00264	-1.02	8.06E-13	8.25E-11
AC092620.3	-1.02	1.60E-13	2.33E-11
RXFP4	-1.04	6.92E-10	2.34E-08
TLE1	-1.05	3.17E-10	1.25E-08

CDCA8	1.32	1.84E-12	1.66E-10
RP11-424C20.2	1.32	5.54E-15	1.42E-12
CD163L1	1.32	3.87E-09	9.82E-08
SOCS3	1.32	4.56E-10	1.67E-08
ICAM1	1.32	6.58E-15	1.63E-12
DOK3	1.32	7.29E-17	3.55E-14
RP5-1028K7.2	1.32	1.51E-11	9.44E-10
CD163	1.31	1.28E-08	2.69E-07
CD28	1.31	9.09E-11	4.31E-09
CCND2-AS1	1.31	1.86E-17	1.17E-14
BCL11B	1.31	2.89E-10	1.15E-08
PBK	1.30	8.11E-13	8.26E-11
PHACTR1	1.30	1.39E-20	3.39E-17
THEMIS	1.30	4.63E-09	1.15E-07
CIT	1.29	1.05E-12	1.02E-10
CDC20	1.29	4.61E-10	1.68E-08
SPAG5	1.29	1.23E-12	1.15E-10
MPEG1	1.29	4.17E-11	2.22E-09
AURKA	1.29	1.16E-10	5.24E-09
GZMK	1.29	6.88E-08	1.11E-06
SKA3	1.28	5.06E-13	5.73E-11
AICDA	1.27	4.88E-08	8.34E-07
CXCL9	1.27	1.34E-08	2.78E-07
RRP12	1.26	5.32E-13	5.91E-11

SSBP2	-1.06	4.05E-11	2.17E-09
KLF11	-1.06	5.87E-12	4.32E-10
KLF7	-1.06	8.56E-08	1.34E-06
RP5-1068E13.7	-1.07	4.93E-16	1.87E-13
CRIP2	-1.10	7.58E-07	8.46E-06
RP11-231E19.1	-1.11	1.19E-11	7.79E-10
KCNH2	-1.11	1.22E-10	5.46E-09
VCAN	-1.13	1.58E-06	1.59E-05
EMP3	-1.13	1.46E-14	3.15E-12
HRK	-1.15	1.04E-10	4.84E-09
CA14	-1.16	4.21E-12	3.32E-10
RP11-128B16.5	-1.20	2.01E-13	2.80E-11
BCL7A	-1.26	6.68E-11	3.30E-09
RP1-140C12.2	-1.27	6.09E-20	1.12E-16
RP11-96K19.4	-1.30	4.74E-14	8.41E-12
KLF3	-1.32	5.61E-13	6.15E-11
CSRP2	-1.34	1.50E-13	2.23E-11
ARRDC3	-1.35	7.30E-13	7.59E-11
PDK4	-1.40	5.13E-07	6.04E-06
TXNIP	-1.65	7.12E-19	7.85E-16
HBA1	-2.06	2.84E-08	5.21E-07
HBA2	-2.12	1.82E-08	3.57E-07
HBB	-2.50	9.69E-10	3.10E-08

FC: fold-change; FDR: false discovery rate

Supplemental Table S2. Curated list of gene expression signatures.

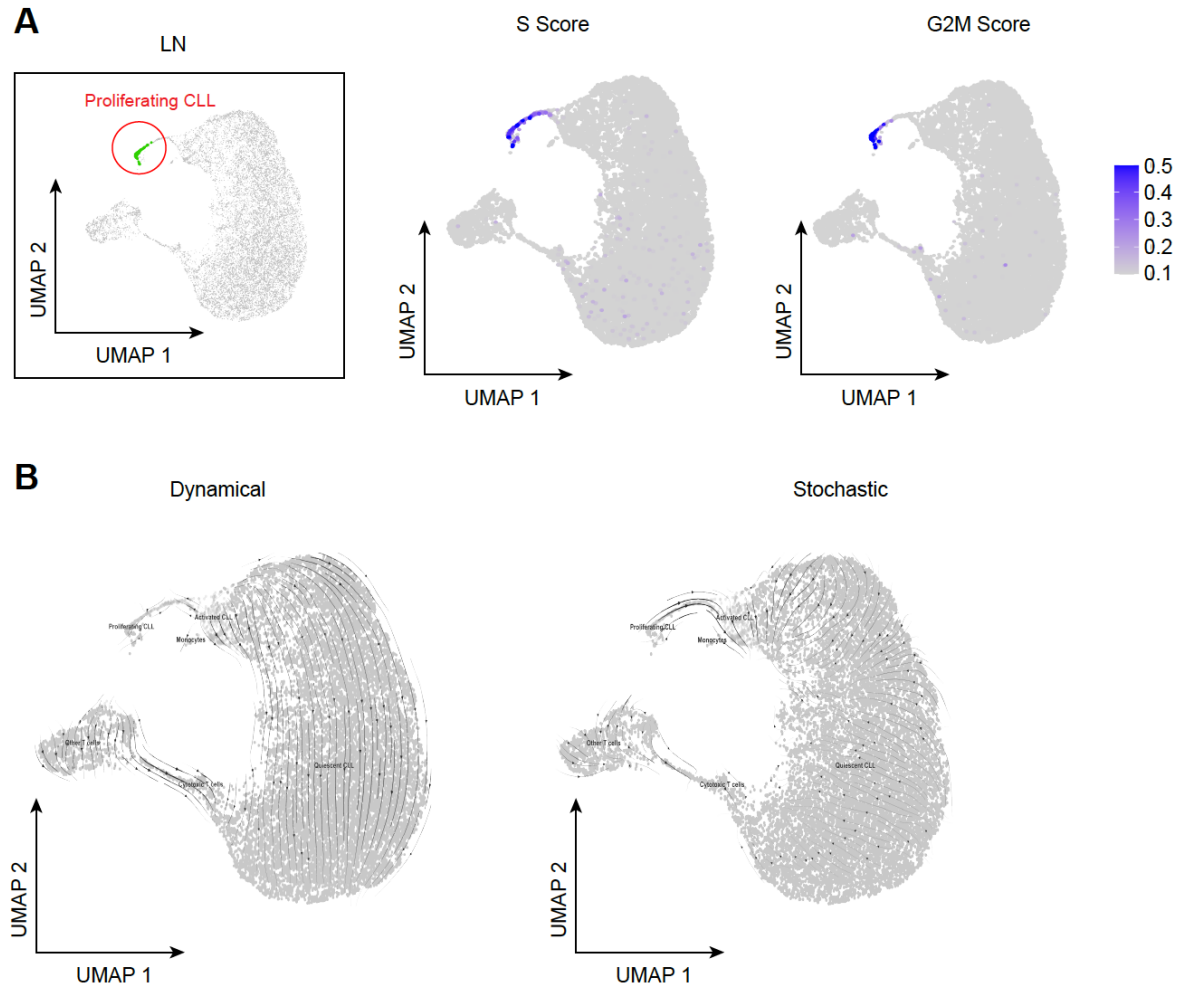
Signature	Description	Reference
activation_BCR	BCR target genes_LN UP	Herishanu et al. Blood 117:563 (2011)
BAnergyUp-1	B_cell_up_anergy	Glynne et al. Nature 403:672 (2000)
BCL6-1	BCL6_repressed	Shaffer et al. Immunity 13:199 (2000)
Blimp-1	Blimp_Bcell_repressed	Shaffer et al. Immunity 17:51 (2002)
CD40Dn-1	CD40_downregulated_Burkitt_lymphoma	Basso et al. Blood 104:4088 (2004)
CD40Up-1	CD40_upregulated_Burkitt_lymphoma	Basso et al. Blood 104:4088 (2004)
E2F3Up-1	E2F3_overexpression_2x_up	Bild et al. Nature 439:353 (2006)
HIF1aDn-1	HIF1alpha_1.5x_down	Manalo et al. Blood 105:659 (2005)
HIF1aUp-1	HIF1alpha_1.5x_Up	Manalo et al. Blood 105:659 (2005)
HRASDn-1	HRAS_overexpression_2x_down	Bild et al. Nature 439:353 (2006)
HRASUp-1	HRAS_overexpression_2x_up	Bild et al. Nature 439:353 (2006)
IFN-1	IFN_PMBC_2x_Up	Baechler et al. PNAS 100:2610 (2003)
Ig-1	Immunoglobulin_Node2868	Su et al. PNAS 101:6062 (2004)
IL10Up-1	IL10_OCILy3_Up	Lam et al. Blood. 111:3701 (2008)
IL-4_reg	IL-4 target genes_CLL	Ruiz-Lafuente et al. PLOSone 9 (2014)
IL6Up-4	IL6_OCILy10_Up_all	Lam et al. Blood. 111:3701 (2008)
IRF3-1	IRF3_target_gene	Ogawa et al. Cell 122:707 (2005)
IRF4Up-3	IRF4_myeloma_induced_all	Shaffer et al. Nature. 454:226 (2008)
JAKUp-1	JAK_IL10_Ly10_Up	Lam et al. Blood. 111:3701 (2008)
KLF2Dn-2	KLF2_repressed	Haaland et al. Mol. Immunol. 42:627 (2005)
KLF2Up-1	KLF2_induced	Haaland et al. Mol. Immunol. 42:627 (2005)
KRASDn-1	KRAS_Down	Sweet-Cordero et al. Nat Genet 37:48 (2005)
KRASUp-1	KRAS_Up	Sweet-Cordero et al. Nat Genet 37:48 (2005)
MCL_proliferation	Proliferation_MCL	Rosenwald et al. Cancer Cell. 3:185 (2003)
MYCDn-1	Myc_overexpression_2x_down	Bild et al. Nature 439:353 (2006)
MYCUp-2	Myc_overexpression_1.5x_up	Bild et al. Nature 439:353 (2006)
NFkB-1-3-10	NFkB_Up_all_OCILy3_Ly10_K1106_HBL1	Lam et al. Clinical Cancer Research 11:1 (2005) & Staudt Lab unpublished data
NotchDN1-2-3	Notch_T-ALL_down_Weng_Sharma_Palomero	Weng et al. Genes Dev 20:2096 (2006) Sharma et al. Mol Cell Biol 26:8022 (2006) Palomerao et al. PNAS, 103:18261 (2006)
NotchUP1-2-4	Notch_T-ALL_up_Weng_Sharma_Palomero	Weng et al. Genes Dev 20:2096 (2006) Sharma et al. Mol Cell Biol 26:8022 (2006) Palomerao et al. PNAS, 103:18261 (2006)
p53up-1	p53_up_Xray	Rosenwald et al. Blood. 104:1428 (2004)
PAX5-1	PAX5_repressed	Delogu et al. Immunity 24:269 (2006)
PGC1up-1	PGC-1alpha_overepxression_up	Mootha et al. Nature Genetics 34:267 (2003)
Quiesce-1	Quiescence_heme_all	Su et al. PNAS 101:6062 (2004)
SerumDown	Serum_response_Fb_down	Chang et al. PLOS Biol 2:206 (2004)
SerumUp	Serum_response_Fb_up	Chang et al. PLOS Biol 2:206 (2004)
SREBPUp-1	SREBP1a&2_up_Scap_dep	Horton et al. PNAS 100:12027 (2003)
StarveDn-4	Glutamine_Glucose_starve_both_down	Peng et al. MCB 22:5575 (2002)

StarveUp-4	Glutamine_Glucose_starve_both_up	Peng et al. MCB 22:5575 (2002)
STAT3Up-1-2	STAT3_up_OCILy10_high_ABC DLBCL_subgroup	Lam et al. Blood. 111:3701 (2008)
TActDn-5	Tcell_PIrep_CsAup4x	Feske et al. Nat Immunol 2:316 (2001)
TActUp-5	Tcell_PIind_CsAdown4x	Feske et al. Nat Immunol 2:316 (2001)
TAnergy-1	T_cell_up_ionomycin_anergy	Macian et al. Cell 109:719 (2002)
TcytDn-1	Tcell_cytokine_repressed	Kovanen et al. J Biol Chem 278:5205 (2003)
TcytUp-6	Tcell_cytokine_induced_PMBC_Bcell_induced	Kovanen et al. J Biol Chem 278:5205 (2003)
TGFBDn-1-2	TGFbeta_down_epithelial_small_large	Kang et al. Mol Cell 11:915-26 (2003)
TGFBUp-5-6	TGFbeta_up_epithelial_small_large	Kang et al. Mol Cell 11:915-26 (2003)
TLR-1	DC_TLR4_TLR8_synergy	Napolitani et al. Nat Immunol 6:769 (2005)
TLR-CpG	TLR_Target_genes_CpG_stim	Bomben et al. Leukemia 26:1584 (2012)
XBP-1	XBP1_target_all	Shaffer et al. Immunity 2004 21:81-93

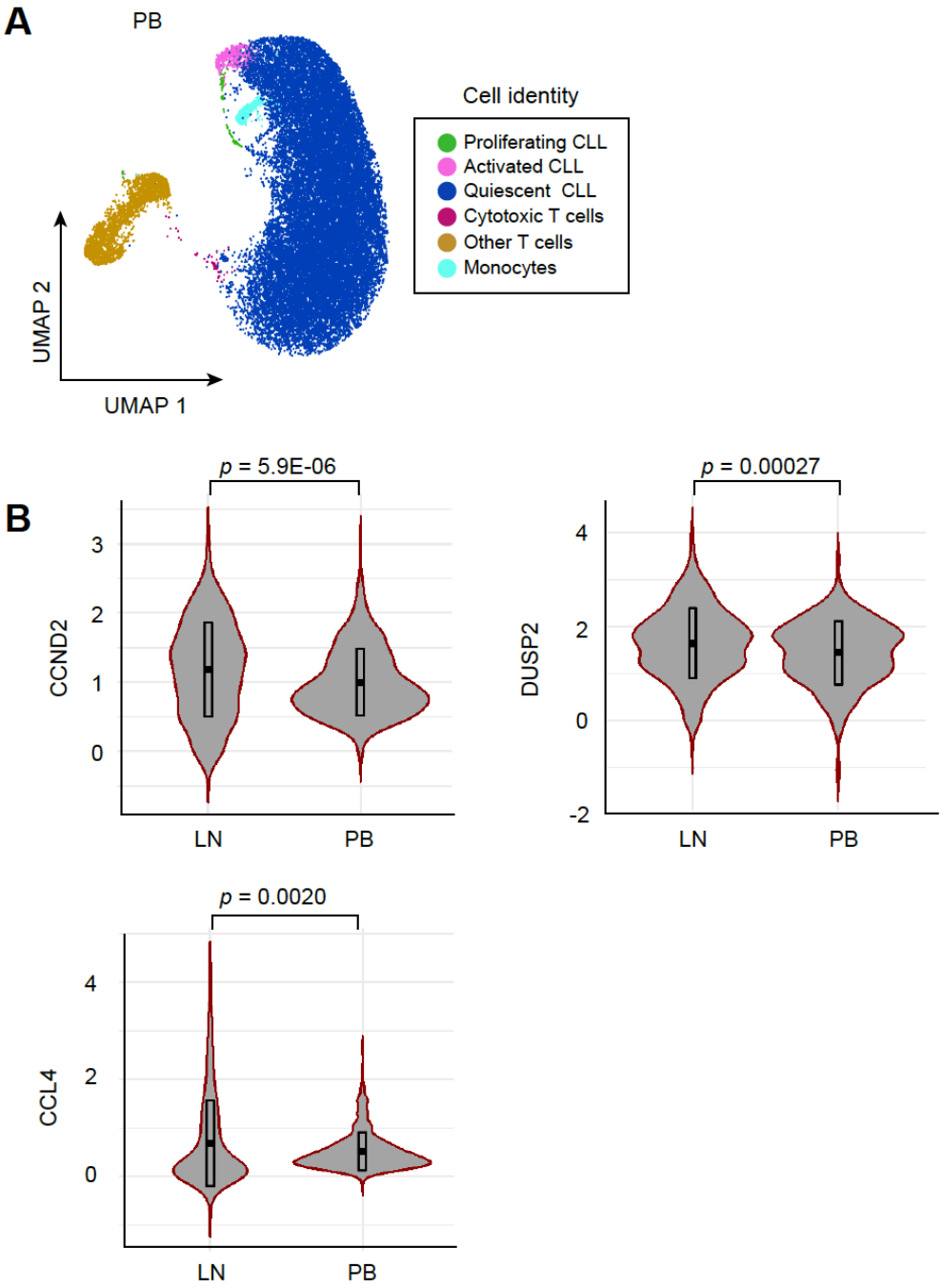
Supplemental Table S3. Genes overexpressed in lymph node relative to peripheral blood in shifted and stable groups of patients.

Shifted Group	
CENPF	MTFP1
HIST1H3B	FAM72A
HIST2H3C	ABCG1
HIST2H3A	AC009237.8
CPNE7	TMEM75
ESPL1	KIAA0101
PIF1	RP11-146D12.2
PKMYT1	FAM95B1
UHRF1	FAM95B1
SPC24	CDT1
CDCA3	E2F2
CDC25A	PLPP3
HIST1H3C	DTX1
ESCO2	RASAL1
CDC20	BNIP3P41
ARHGAP11A	TNS3
RP4-569M23.4	C3
CIT	VASH2
RECQL4	GLYATL2
CENPI	HIST1H2AB
RP4-569M23.2	IQGAP2
CDCA8	

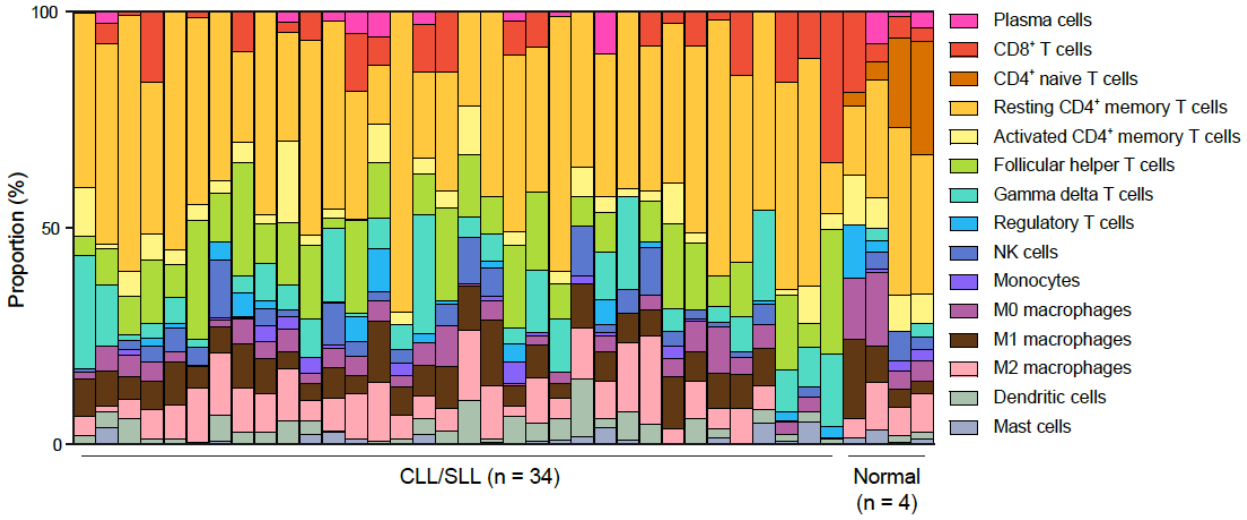
Stable Group	
SLC2A3	SGK1
RGS1	ANK2
FOS	DUSP5
RP11-823E8.3	CD2
FOSB	LGMN
SLC40A1	GZMK
DST	RP11-347P5.1
DUSP6	A2M
CD4	SORL1
NR4A3	IL7R
FYB	SIK1
CD83	MIR24-2
NR4A2	TC2N
APOE	RP11-52J3.3
INPP4B	CTLA4
DUSP1	LMNA
CD69	CD96
SOCS3	RGS2
RGCC	
RP11-293M10.2	
TBC1D4	
FRMD4B	



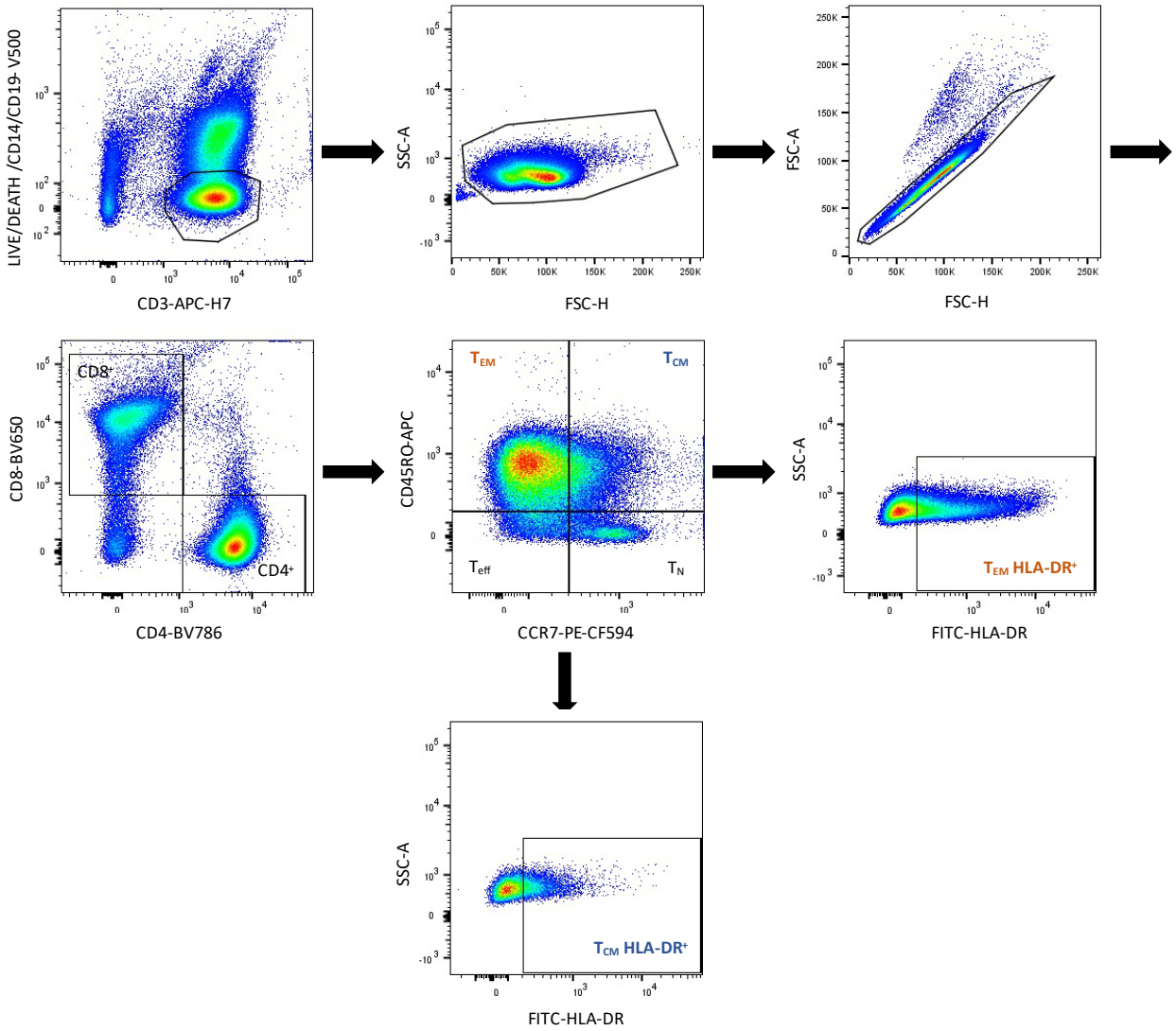
Supplemental Figure S1. Characteristics of activated and proliferating CLL cells. (A) Uniform Manifold Approximation and Projection (UMAP) of G2/M and S phase markers show overexpression in proliferating CLL cluster in LN samples. (B) RNA velocities derived from dynamical and stochastic modeling projected on a UMAP of single cells.



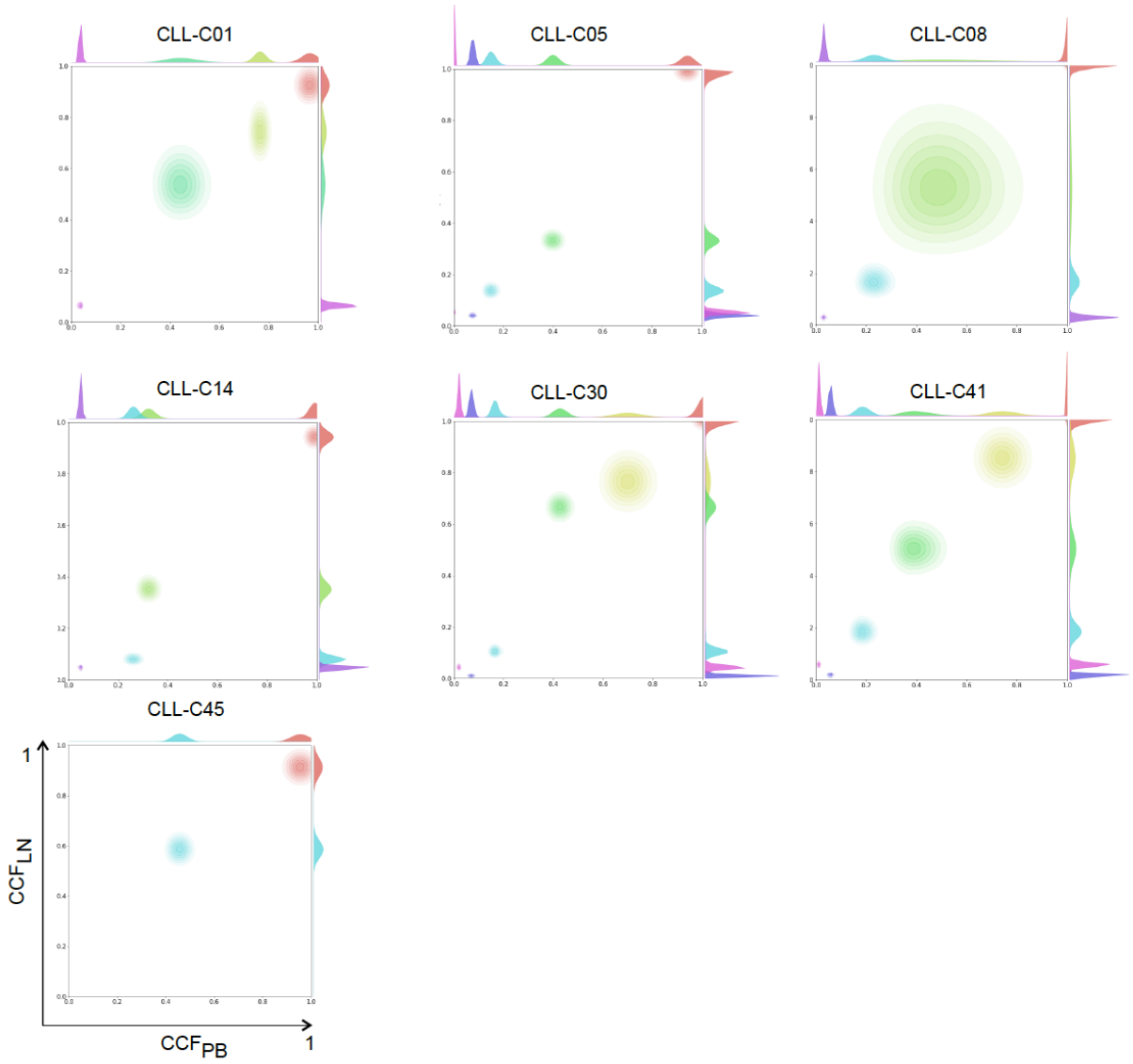
Supplemental Figure S2. Differences in activated CLL cells between LN and PB. (A) UMAP and clustering of cell identities of peripheral blood (PB) single cells from the integrated dataset of PB and lymph node (LN) samples. (B) Violin plots of *CCND2*, *DUSP2*, and *CCL4* expression of single cells in LN and PB. Bars indicate mean ± standard deviation.



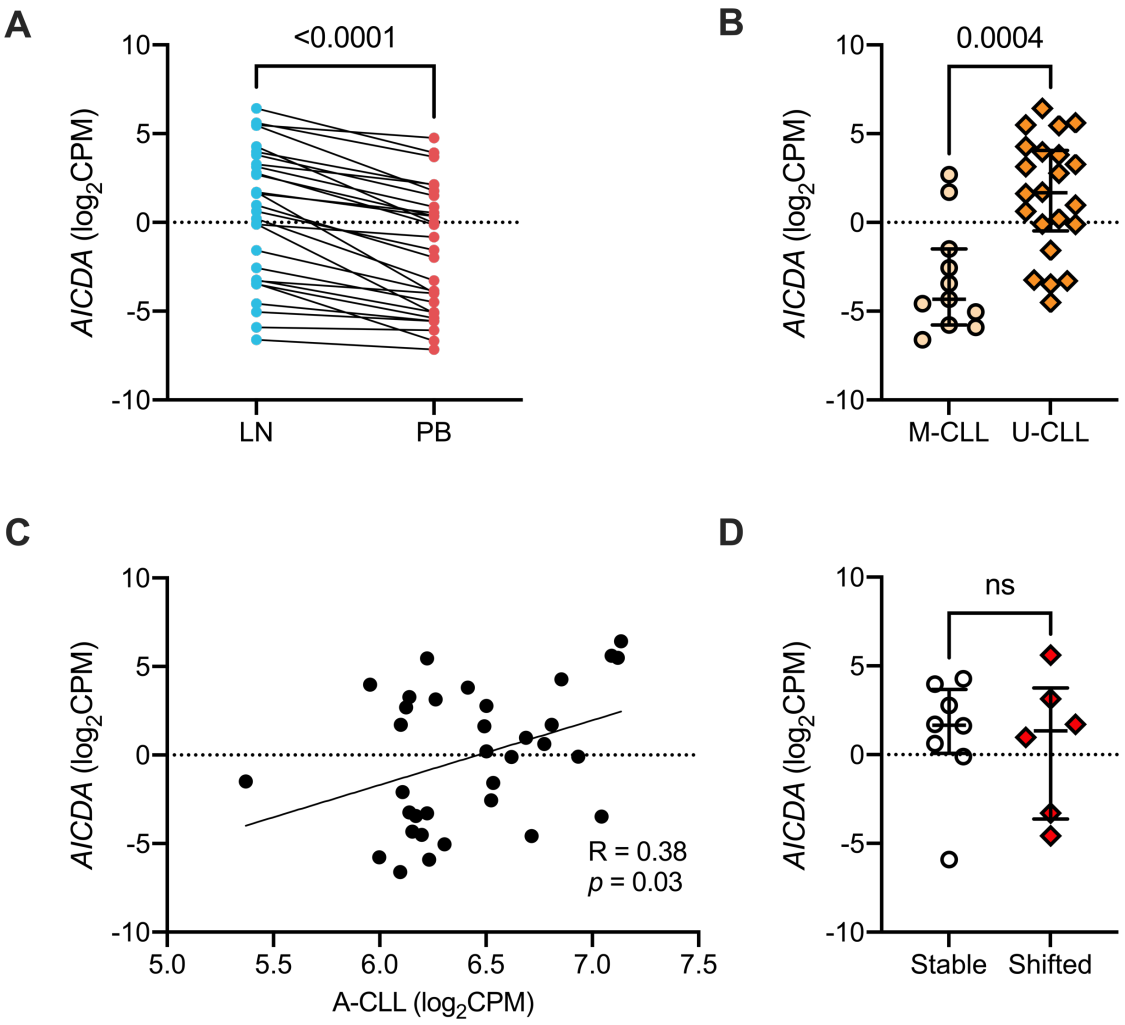
Supplemental Figure S3. Deconvolution of LN transcriptomes. Abundance of the indicated non-CLL/non-B cell types as estimated by CIBERSORT deconvolution of bulk RNA-seq data in individual LN samples.



Supplemental Figure S4. Gating strategy for activated CD4+ memory T-cell subsets. Live $CD3^+$ cells were first selected, then cell debris and doublets were excluded. T cells were subdivided based on the expression of CD4 and CD8. $CD4^+$ effector memory T cells (T_{EM}) and central memory T cells (T_{CM}) were defined by $CD45RO^+CCR7^-$ and $CD45RO^+CCR7^+$, respectively. $HLA\text{-}DR^+$ cells in each subset were gated in the final step.



Supplemental Figure S5. Patients without genetic compartmentalization between PB and LN. Density plots of CCF in PB and LN in patients with no subclonal expansion in either PB or LN.



Supplemental Figure S6. *AICDA* gene expression in purified CD19⁺ tumor cells. (A) *AICDA* expression in paired LN and PB samples. Dot plot, median and interquartile range of *AICDA* expression in LN samples in (B) M-CLL and U-CLL cases, and (D) stable and shifted groups. (C) Correlation of *AICDA* expression and A-CLL signature in LN samples.

Supplemental References

1. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.
2. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30(7):923-930.
3. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15(2):R29.
4. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
5. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284-287.
6. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453-457.
7. Herishanu Y, Perez-Galan P, Liu D, et al. The lymph node microenvironment promotes B-cell receptor signaling, NF-kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood*. 2011;117(2):563-574.
8. Zheng GX, Terry JM, Belgrader P, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun*. 2017;8:14049.
9. Stuart T, Butler A, Hoffman P, et al. Comprehensive Integration of Single-Cell Data. *Cell*. 2019;177(7):1888-1902 e1821.
10. Chen YC, Suresh A, Underbayev C, et al. IKAP-Identifying K mAjor cell Population groups in single-cell RNA-sequencing analysis. *Gigascience*. 2019;8(10).
11. La Manno G, Soldatov R, Zeisel A, et al. RNA velocity of single cells. *Nature*. 2018;560(7719):494-498.
12. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol*. 2020;38(12):1408-1414.
13. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013;31(3):213-219.

14. Cibulskis K, McKenna A, Fennell T, Banks E, DePristo M, Getz G. ContEst: estimating cross-contamination of human samples in next-generation sequencing data. *Bioinformatics*. 2011;27(18):2601-2602.
15. Taylor-Weiner A, Stewart C, Giordano T, et al. DeTiN: overcoming tumor-in-normal contamination. *Nat Methods*. 2018;15(7):531-534.
16. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.
17. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*. 2004;5(4):557-572.
18. Ramos AH, Lichtenstein L, Gupta M, et al. Oncotator: cancer variant annotation tool. *Hum Mutat*. 2015;36(4):E2423-2429.
19. Carter SL, Cibulskis K, Helman E, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol*. 2012;30(5):413-421.
20. Escobar MD, West M. Bayesian Density-Estimation and Inference Using Mixtures. *Journal of the American Statistical Association*. 1995;90(430):577-588.
21. Landau DA, Carter SL, Stojanov P, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013;152(4):714-726.