SUPPLEMENTARY DATA

A cyclin D1-dependent transcriptional program predicts clinical outcome in mantle cell lymphoma

Santiago Demajo *et al.*

SUPPLEMENTARY DATA INDEX

SUPPLEMENTARY METHODS

Western blot, immunoprecipitation, and qRT-PCR

Western blot (WB) analysis was performed as previously described (1), using cyclin D1 (Santa Cruz Biotechnology, sc-753, RRID:AB_2070433) and tubulin (Sigma-Aldrich, T5168, RRID:AB_477579) antibodies. Co-immunoprecipitation assays were performed as described before (2), using cyclin D1 antibody (Santa Cruz Biotechnology, sc-8396, RRID:AB_627344) or control IgG (Santa Cruz Biotechnology, sc-2025, RRID:AB_737182) followed by protein G-magnetic beads (Invitrogen) incubation and elution with Glycine 100mM pH=2.5. Co-IP experiments were performed within five weeks after cell thawing. Cyclin D1 (Santa Cruz Biotechnology, sc-753), E2F4 (Bethyl, A302-134A, RRID:AB_1720353), FOXM1 (Santa Cruz Biotechnology, sc-502, RRID:AB_631523), and CBP (Santa Cruz Biotechnology, sc-7300, RRID:AB_626817) antibodies were used for WB detection. In figure 1A and supplementary figure S2A, the same blot was probed with cyclin D1 and tubulin antibodies by cutting the membrane. In figure 2H, cyclin D1 and CBP blots correspond to the same membrane while E2F4 and FOXM1 blots correspond to an independent membrane. Image acquisition was performed with ImageQuant LAS 4000 mini (GE Healthcare). Image processing and quantification were performed with Multi Gauge software (Fujifilm).

For qRT-PCR analysis, cDNA was generated from 1 µg RNA with qScript cDNA Synthesis kit (Quantabio). qRT–PCR reaction was performed using SYBR green (Roche). Cyclin D1 amplification was performed with forward primer 5'-GACCTTCGTTGCCCTCTGT-3' and reverse primer 5'-AGCGTGTGAGGCGGTAGTAG-3'. Results were normalized using *PUM1*, amplified with forward primer 5'-CGGTCGTCCTGAGGATAAAA-3' and reverse primer 5'-CGTACGTGAGGCGTGAGTAA-3'.

RNA-sequencing

RNA-seq libraries were prepared using the TruSeq Stranded mRNA Sample Prep Kit v2 (Illumina) according to the manufacturer's protocol. Briefly, 1 µg of total RNA was used for poly(A)-mRNA selection using streptavidin-coated magnetic beads and were fragmented to approximately 300 bp. cDNA was synthesized using reverse transcriptase SuperScript II (Invitrogen) and random primers, incorporating dUTP in place of dTTP in the second strand. dsDNA was subjected to A-tailing and ligation of the barcoded Truseq adapters. All purification steps were performed with AMPure XP beads. Libraries were amplified by PCR using the corresponding primer cocktail. Final libraries were analyzed using Agilent DNA 1000 chip and were quantified by the KAPA Library Quantification Kit (Kapa Biosystems) prior to amplification with Illumina's cBot.

Raw sequence files (.fastq) underwent quality control analysis using FastQC and reads were aligned to the Human Feb. 2009 (GRCh37/hg19) genome using TopHat v2.1.1 allowing for unique alignments and up to two mismatches. The resulting alignments were summarized by Ensembl gene identifiers to evaluate number of uniquely aligned reads per transcript and per sample. The raw read counts were analyzed using the 'limma' package v3.9 from the Bioconductor (http://bioconductor.org) and were used as input to form a DGEList object. Scale normalization was applied and the calculation of normalized signal was performed by voom function of the 'limma' package and converted to RPKM (3). Genes with expression lower than 0.5 RPKM in both shCtrl and shCycD1 cells were considered not expressed and were excluded. Differential gene expression analysis between shCtrl and shCycD1 was performed using 'limma' package, with cut-offs of fold change > 1.5 and adjusted P-value < 0.05 to identify differentially expressed genes.

Gene expression microarray

RNA was extracted from cyclin D1-overexpressing and control JVM13 cells as reported before (1). 150 ng of each RNA sample were processed according to 3'IVT PLUS chemistry using an automated system (Biomek FX System, Beckman Coulter). Biotinylated cRNA were

prepared according to the standard Affymetrix protocol. Following fragmentation, 6.6 ug of fragmented and labeled cRNA were hybridized on Affymetrix Human Genome U219 Array Plate for 16 hr at 45ºC, using the automated GeneTitan System, which includes the hybridization oven, Fluidic Station and Scanner. Raw U219 microarray data were normalized using the robust multiarray average (RMA) method.

Differential gene expression analyses were performed using an adjusted p-value < 0.05 and selecting, for each gene, the probe set with the highest interquartile range. In all the analyses with HG-U133 Plus 2.0 microarrays, cyclin D1 expression was evaluated by the 208711_s_at probe. To assess the status of cyclin D1 RNA (full length or truncated 3'UTR) the ratio between 208712_at and 208711_s_at cyclin D1 probes was evaluated. The different hierarchical clustering analyses and heatmaps were created with 'gplots', 'heatmap.plus' and 'ComplexHeatmap' R packages. Gene expression data corresponding to breast cancer and multiple myeloma patients (4,5) were log transformed and analyzed as described for MCL. Regarding breast cancer, those patients with very low cyclin D1 signal (below 5) were considered cyclin D1-negative and excluded from the analysis, in both ER-positive and ER-negative subsets. ER status was evaluated by *ESR1* gene expression (205225_at probe). Kaplan-Meier curves from breast cancer patients were calculated with the gene expression and clinical data available online (http://kmplot.com/) (5).

NanoString

Digital gene expression quantification by the NanoString platform was performed according to the manufacturer's protocol. Briefly, Probe Mixes A and B and hybridization master mix (including Elements TagSets) were prepared, mixed with the RNA samples and incubated in a PCR machine at 67ºC for 16 hours. Hybridized samples were processed in the NanoString nCounter Prep Station and immobilized in the cartridge for data collection on the nCounter Digital Analyzer using 280 fields of view. To assess the status of cyclin D1 RNA (full length or truncated 3'UTR) the ratio between the two cyclin D1 probes (exonic and 3'UTR) was evaluated.

ChIP-sequencing, genomic analysis, and bioinformatic analysis

Cyclin D1 ChIP-seq data were analyzed as described before (1). Cyclin D1 target genes (n = 8,638) were defined as the genes presenting a peak of cyclin D1 in its proximal promoter, within 1 kilobase upstream of the TSS, in the four MCL cell lines analyzed. Cyclin D1 non-regulated genes were defined as the cyclin D1 ChIP-seq targets that were neither upregulated nor downregulated in JeKo-1 shCycD1 RNA-seq using stringent criteria (adjusted p-value > 0.2 and fold-change < 1.1).

Motif enrichment analysis was performed using the AME tool from the MEME suite tool (http://meme-suite.org/tools/ame), with the CisBP motif database for *Homo sapiens* (6). Analysis of E2F and CHR motif enrichment was performed using the 'Biostrings' R package with the motif definition previously described (7); statistical significance was calculated in comparison to all gene promoters, defined as -736 to +828 from the TSS, which are the mean values corresponding to cyclin D1 peaks in the cyclin D1-dependent gene program; these sequences were obtained from UCSC Table Browser (https://genome-euro.ucsc.edu/). To study the overlap between cyclin D1 and other transcription factors, ChIP-seq data from the ENCODE project (https://www.encodeproject.org) (8), available at the UCSC Genome Browser, were used; colocalization was considered positive when peaks were overlapping in at least one bp. The plots of ChIP-seq enrichment signals around the TSS of cyclin D1-dependent transcriptional program genes were generated by the sitepro script from the CEAS package. The wig files for each transcription factor were generated with MACS (9). For all these analyses, cyclin D1 ChIP-seq data corresponding to JeKo-1 cells were used. E2F4 and FOXM1 ChIP-seq target genes were obtained from ChEA3 database (https://amp.pharm.mssm.edu/chea3/) (10) and corresponded to target genes in GM12878 cells from the ENCODE project. A gene expression score for both E2F4 and FOXM1 was obtained based on the mean expression of all the corresponding target genes in each MCL patient.

Venn diagrams were calculated using 'venneuler' R package and Venny v2.1.0 (http://bioinfogp.cnb.csic.es/tools/venny). The significance of the overlaps was evaluated by

Fisher's test considering as background the genes that could be evaluated by the corresponding techniques and, in the case of the RNA-seq, that in addition could be detected (RPKM>0). Gene ontology analysis was performed using DAVID v6.7 (https://david.ncifcrf.gov) and the GO consortium database (http://geneontology.org).

Identification of a simplified cyclin D1 signature

To simplify the full 295-gene cyclin D1-dependent transcriptional program identified in cell lines, we integrated the gene expression analyses performed in MCL patients. Several filters were applied to select the genes that fulfilled several conditions (supplementary Figure S9). First, we selected the genes that were upregulated in MCL considering the following comparisons: genes upregulated in MCL peripheral blood samples versus normal naïve and memory B-cells, genes upregulated in MCL lymphoid tissues samples versus normal lymphoid tissues, and genes upregulated in MCL peripheral blood samples versus other leukemic cyclin D1 negative B-cell lymphoid neoplasms (analyses corresponding to supplementary Figure S5A). Afterwards, we selected the genes whose expression directly correlated with cyclin D1 expression, in both blood and tissue MCL samples (analysis corresponding to supplementary Figure S5B). Finally, we selected the genes whose expression positively correlated with death risk, in both blood and tissue MCL samples (analysis corresponding to supplementary Figure S8A). Overall, this integrative analysis resulted in 38 genes, which were analyzed by NanoString. One additional gene was discarded due to poor correlation between microarray and NanoString expression data, leading to a final 37-gene simplified signature.

SUPPLEMENTARY REFERENCES

1. Albero R, Enjuanes A, Demajo S, Castellano G, Pinyol M, García N, et al. Cyclin D1 overexpression induces global transcriptional downregulation in lymphoid neoplasms. J Clin Invest. 2018;128:4132–47.
2. Demajo S, Uribesalgo I, Gutiérrez A, Ballaré C, Capdevila S, Roth M, et al. ZRF1 controls the retinoic acid pathway and regulates leukemogenic potential in acute myeloid leukemia. Oncogene. 2014;33:5501–10.
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:R29.
4. Broyl A, Hose D, Lokhorst H, de Knegt Y, Peeters J, Jauch A, et al. Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. Blood. 2010;116:2543–53.
5. Győrffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Res Treat. 2010;123:725–31.
6. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME Suite: Tools for motif discovery and searching. Nucleic Acids Res. 2009;37:W202–8.
7. Müller GA, Stangner K, Schmitt T, Wintsche A, Engeland K. Timing of transcription during the cell cycle: Protein complexes binding to E2F, E2F/CLE, CDE/CHR, or CHR promoter elements define early and late cell cycle gene expression. Oncotarget. 2017;8:97736–48.
8. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489:57–74.
9. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based Analysis of ChIP-Seq (MACS). Genome Biol. 2008;9:R137.
10. Keenan AB, Torre D, Lachmann A, Leong AK, Wojciechowicz ML, Utti V, et al. ChEA3: transcription factor enrichment analysis by orthogonal omics integration. Nucleic Acids Res. Oxford University Press; 2019;47:W212–24.

Supplementary Table S1. The 295 genes from the cyclin D1-dependent transcriptional program.

AMOTL1	*CDKN3*	*FAM72A*	*LBR*	*POLA2*	*SNRPB*
ANLN	*CDT1*	*FAM72B*	*LIG1*	*POLD1*	*SP1*
ANP32E	*CENPA*	*FAM72D*	*LIN54*	*POLE*	*SPAG5*
ARHGAP11B	*CENPF*	*FAM83D*	*LIN9*	*POLE2*	*SPC24*
ARID1A	*CENPH*	*FANCB*	*LMNB1*	*PPIH*	*SPC25*
ARL6IP6	*CENPK*	*FANCD2*	*LMNB2*	*PRC1*	*SPDL1*
ASF1B	*CENPM*	*FANCG*	*LRRCC1*	*PRIM1*	*SPIN4*
ASPM	*CENPQ*	*FANCI*	*MAD2L1*	*PRR11*	*SRSF1*
ATAD2	*CENPW*	*FBXO5*	*MAGOHB*	*PSRC1*	*SSRP1*
ATAD5	*CEP152*	*FEN1*	*MASTL*	*PXMP2*	*STIL*
AUNIP	*CEP55*	*FIGNL1*	*MCM10*	*RACGAP1*	*STMN1*
AURKA	*CHAF1A*	*FMR1*	*MCM2*	*RAD1*	*SUV39H1*
BIRC5	*CHAF1B*	*FOXM1*	*MCM3*	*RAD51*	*SUV39H2*
BLM	*CHEK1*	*GINS1*	*MCM4*	*RAD51AP1*	*TACC3*
BRCA1	*CKAP2*	*GINS2*	*MCM6*	*RAD51C*	*TCF19*
BRCA2	*CKAP2L*	*GINS3*	*MCM7*	*RAD54L*	*TICRR*
BRIP1	*CKAP5*	*GINS4*	*MCM8*	*RANBP1*	*TIMELESS*
BUB1	*CKS1B*	*GMNN*	*MCMBP*	*RBL1*	*TIPIN*
BUB1B	*CLSPN*	*GSG2*	*MELK*	*RBM14*	*TK1*
C16orf59	*CTDSPL2*	*GTSE1*	*MIS18A*	*RCCD1*	*TMEM106C*
C17orf53	*CTNNAL1*	*H2AFV*	*MMS22L*	*RECQL4*	*TMPO*
C1orf112	*DARS2*	*HAT1*	*MND1*	*RFC2*	*TMSB15B*
C4orf46	*DBF4*	*HAUS4*	*MSH2*	*RFC3*	*TOP2A*
C9orf40	*DCK*	*HELLS*	*MSH6*	*RFC4*	*TPGS2*
CASP6	*DCLRE1B*	*HIRIP3*	*MTBP*	*RFC5*	*TPX2*
CASP8AP2	*DCP2*	*HJURP*	*MTFR2*	*RFWD3*	*TRAIP*
CBFB	*DEK*	*HMGA1*	*MTHFD1*	*RFXAP*	*TRIP13*
CBX2	*DEPDC1B*	*HMGB2*	*MYBL2*	*RHNO1*	*TROAP*
CBX5	*DLGAP5*	*HMGB3*	*NAP1L1*	*RIF1*	*TTK*
CCDC167	*DNA2*	*HMGN2*	*NCAPD2*	*RMI1*	*TUBB4B*
CCDC34	*DNMT1*	*HMGXB4*	*NCAPG*	*RMI2*	*UBE2C*
CCNA2	*DONSON*	*HMMR*	*NCAPG2*	*RNASEH2A*	*UBE2T*
CCNB1	*DSCC1*	*HNRNPD*	*NCAPH*	*RNASEH2B*	*UHRF1*
CCNB2	*DTYMK*	*HP1BP3*	*NEK2*	*RPA2*	*USP1*
CCNE2	*DUT*	*IFI27L1*	*NSD2*	*RRM1*	*USP13*
CCNF	*E2F1*	*INCENP*	*NUF2*	*RRM2*	*USP37*
CDC20	*E2F8*	*ITPR3*	*NUP153*	*SAE1*	*VRK1*
CDC25A	*EBP*	*KANK2*	*NUP85*	*SAP30*	*WDHD1*
CDC25B	*ECT2*	*KBTBD2*	*NUSAP1*	*SAPCD2*	*WDR5*
CDC25C	*ERCC6L*	*KIAA1524*	*OIP5*	*SASS6*	*WDR62*
CDC45	*ERI2*	*KIF11*	*ORC6*	*SFR1*	*WDR76*
CDC6	*ERLIN1*	*KIF14*	*PARP2*	*SHCBP1*	*WEE1*
CDC7	*ESCO2*	*KIF15*	*PARPBP*	*SIVA1*	*XRCC2*
CDCA2	*ESPL1*	*KIF18B*	*PBK*	*SKA2*	*ZNF367*
CDCA3	*EXO1*	*KIF20A*	*PCNA*	*SKA3*	*ZWILCH*
CDCA4	*EXOSC8*	*KIF20B*	*PHF19*	*SLBP*	
CDCA5	*EZH2*	*KIF22*	*PITX1*	*SMC2*	
CDCA7	*FADS1*	*KIF23*	*PLK1*	*SMC3*	
CDCA8	*FAM111A*	*KIFC1*	*PLK4*	*SMC4*	
CDK1	*FAM111B*	*KNTC1*	*POC1A*	*SNRNP48*	

Supplementary Table S2. Gene Ontology (GO) analysis of the cyclin D1-dependent transcriptional program.

GO Term	Count	%	P-value	FDR
GO:0007049 cell cycle	182	61.7	3.4E-116	6.1E-113
GO:0022402 cell cycle process	165	55.9	7.8E-110	1.4E-106
GO:0000278 mitotic cell cycle	142	48.1	1.4E-102	2.5E-99
GO:1903047 mitotic cell cycle process	137	46.4	3.6E-101	6.4E-98
GO:0051276 chromosome organization	128	43.4	6.41E-74	1.15E-70
GO:0000280 nuclear division	97	32.9	2.29E-71	4.11E-68
GO:0048285 organelle fission	98	33.2	7.73E-70	1.39E-66
GO:0007067 mitotic nuclear division	85	28.8	3.79E-68	6.80E-65
GO:0006260 DNA replication	72	24.4	1.33E-64	2.39E-61
GO:0006259 DNA metabolic process	110	37.3	7.24E-63	1.30E-59
GO:0051301 cell division	88	29.8	6.72E-61	1.20E-57
GO:0007059 chromosome segregation	69	23.4	6.44E-56	1.16E-52
GO:0006996 organelle organization	178	60.3	1.40E-52	2.50E-49
GO:0044770 cell cycle phase transition	79	26.8	3.50E-52	6.28E-49
GO:0000819 sister chromatid segregation	56	19.0	9.32E-50	1.67E-46
GO:0044772 mitotic cell cycle phase transition	75	25.4	1.16E-49	2.08E-46
GO:0006974 cellular response to DNA damage stimulus	87	29.5	2.39E-47	4.29E-44
GO:0098813 nuclear chromosome segregation	59	20.0	2.85E-47	5.12E-44
GO:0006261 DNA-dependent DNA replication	45	15.3	2.97E-44	5.32E-41
GO:0006281 DNA repair	70	23.7	3.49E-43	6.26E-40
GO:0051726 regulation of cell cycle	89	30.2	1.39E-42	2.50E-39
GO:0010564 regulation of cell cycle process	71	24.1	7.28E-41	1.31E-37
GO:0006310 DNA recombination	48	16.3	1.07E-35	1.92E-32
GO:0000070 mitotic sister chromatid segregation	36	12.2	9.13E-32	1.64E-28
GO:0016043 cellular component organization	193	65.4	2.71E-31	4.86E-28
GO:0000075 cell cycle checkpoint	42	14.2	3.95E-31	7.08E-28
GO:0071840 cellular component organization or biog.	195	66.1	6.42E-31	1.15E-27
GO:0007346 regulation of mitotic cell cycle	56	19.0	9.16E-31	1.64E-27
GO:0071103 DNA conformation change	45	15.3	1.34E-30	2.41E-27
GO:0033554 cellular response to stress	99	33.6	9.50E-30	1.70E-26
GO:0007062 sister chromatid cohesion	33	11.2	2.55E-29	4.57E-26
GO:0044843 cell cycle G1/S phase transition	39	13.2	1.43E-26	2.56E-23
GO:1902589 single-organism organelle organization	89	30.2	5.57E-26	9.98E-23
GO:0000082 G1/S transition of mitotic cell cycle	37	12.5	1.07E-25	1.92E-22
GO:0006302 double-strand break repair	35	11.9	1.05E-24	1.89E-21

GO:0090304	nucleic acid metabolic process	164	55.6	4.72E-24	8.47E-21
GO:0000226	microtubule cytoskeleton organization	46	15.6	5.10E-24	9.15E-21
GO:1901990	regulation of mitotic cell cycle phase tran.	39	13.2	1.24E-22	2.22E-19
GO:1901987	regulation of cell cycle phase transition	40	13.6	1.92E-22	3.45E-19
GO:0007017	microtubule-based process	52	17.6	2.32E-22	4.15E-19
GO:0044839	cell cycle G2/M phase transition	32	10.8	5.85E-22	1.05E-18
GO:0006312	mitotic recombination	19	6.4	5.75E-21	1.03E-17
GO:0006271	DNA strand elongate. in DNA replication	16	5.4	5.81E-21	1.04E-17
GO:0051052	regulation of DNA metabolic process	40	13.6	6.34E-21	1.14E-17
GO:0022616	DNA strand elongation	17	5.8	6.39E-21	1.14E-17
GO:0051783	regulation of nuclear division	29	9.8	6.40E-21	1.15E-17
GO:0000722	telomere maintenance via recombination	17	5.8	1.22E-20	2.19E-17
GO:0010948	negative regulation of cell cycle process	33	11.2	2.18E-20	3.91E-17
GO:0006323	DNA packaging	31	10.5	2.38E-20	4.27E-17
GO:0006270	DNA replication initiation	17	5.8	4.12E-20	7.39E-17
GO:0006139	nucleobase-containing comp. metabolic p.	167	56.6	9.04E-20	1.62E-16
GO:0000086	G2/M transition of mitotic cell cycle	29	9.8	1.36E-19	2.44E-16
GO:0000731	DNA synthesis involved in DNA repair	21	7.1	1.85E-19	3.31E-16
GO:0007088	regulation of mitotic nuclear division	26	8.8	2.75E-19	4.94E-16
GO:0031570	DNA integrity checkpoint	27	9.2	4.38E-19	7.85E-16
GO:0090068	positive regulation of cell cycle process	32	10.8	8.50E-19	1.52E-15
GO:0046483	heterocycle metabolic process	167	56.6	9.85E-19	1.77E-15
GO:0006725	cellular aromatic compound metabolic p.	167	56.6	2.32E-18	4.16E-15
GO:0045786	negative regulation of cell cycle	41	13.9	4.19E-18	7.52E-15
GO:0007093	mitotic cell cycle checkpoint	26	8.8	6.67E-18	1.20E-14
GO:1901360	organic cyclic compound metabolic p.	169	57.3	7.92E-18	1.42E-14
GO:0045787	positive regulation of cell cycle	35	11.9	1.13E-17	2.02E-14
GO:0045930	negative regulation of mitotic cell cycle	29	9.8	1.78E-17	3.19E-14
GO:0071897	DNA biosynthetic process	28	9.5	1.91E-17	3.42E-14
GO:0000725	recombinational repair	19	6.4	1.20E-15	2.19E-12
GO:0000910	cytokinesis	21	7.1	4.93E-15	8.76E-12
GO:0051321	meiotic cell cycle	27	9.2	5.97E-15	1.07E-11
GO:0006325	chromatin organization	47	15.9	9.71E-15	1.75E-11

GOTERM_BP_ALL category was used. The gene count, the percentage from the 295-gene transcriptional program, the p-value, and the FDR adjusted p-value are shown for each GO term. Only GO terms with $p < 10^{-14}$ are shown.

Supplementary Table S3. Colocalization analysis between cyclin D1 peaks from the cyclin D1-dependent program genes and transcriptional regulators from the ENCODE database.

Protein	Peaks	% peaks	Odds ratio	P-value	Adj. p-value
E2F4	264	89.5	4.61	<1E-10	<1E-10
NFYA	97	32.9	3.39	<1E-10	<1E-10
cFos	111	37.6	3.22	<1E-10	<1E-10
FOXM1	196	66.4	2.09	<1E-10	<1E-10
NFYB	193	65.4	2.09	<1E-10	<1E-10
SIN3A	253	85.8	1.83	1E-10	7.4E-09
RFX5	115	39.0	2.15	7E-10	5.11E-08
SP1	251	85.1	1.67	2.19E-08	1.58E-06
CHD2	243	82.4	1.66	3.51E-08	2.49E-06
WHIP	142	48.1	1.77	2.88E-07	2.01E-05
Mxi1	264	89.5	1.56	5.31E-07	3.67E-05
Max	216	73.2	1.5	1.40E-05	0.0009
ETS1	114	38.6	1.55	0.0002	0.0124
RXRA	37	12.5	2.08	0.0002	0.0156
PML	218	73.9	1.4	0.0003	0.0186
TAF1	242	82.0	1.37	0.0004	0.0287
NFIC	102	34.6	1.52	0.0005	0.0325
ELK1	130	44.1	1.45	0.0007	0.0419
TBP	174	59.0	1.4	0.0007	0.0424
Pol24H8	234	79.3	1.35	0.0007	0.0425
POU2F2	191	64.7	1.37	0.0009	0.0530
BRCA1	27	9.2	2.12	0.0011	0.0639
CHD1	97	32.9	1.45	0.0021	0.1181
Pol2	288	97.6	1.29	0.0022	0.1214
ZBTB33	54	18.3	1.58	0.0034	0.1866
RUNX3	238	80.7	1.28	0.0043	0.2327
Nrf1	140	47.5	1.34	0.0044	0.2349
STAT5A	58	19.7	1.51	0.0061	0.3184
ELF1	266	90.2	1.25	0.0073	0.3700
YY1	237	80.3	1.23	0.0155	0.7754
Pbx3	79	26.8	1.34	0.0196	0.9589
ATF2	110	37.3	1.28	0.0215	1
GABP	142	48.1	1.25	0.0226	1
cMyc	84	28.5	1.32	0.0231	1
MAZ	219	74.2	1.21	0.0248	1
TCF3	128	43.4	1.21	0.0533	1
NFKB	74	25.1	1.25	0.0642	1
MTA3	66	22.4	1.25	0.0720	1
STAT1	25	8.5	1.41	0.0798	1
TR4	19	6.4	1.48	0.0854	1

ATF3	32	10.8	1.33	0.0927	1
STAT3	32	10.8	1.31	0.1034	1
PAX5N19	104	35.3	1.16	0.1154	1
Pol2s2	139	47.1	1.14	0.1183	1
SRF	100	33.9	1.15	0.1400	1
NFE2	13	4.4	1.43	0.1575	1
TBLR1	97	32.9	1.14	0.1578	1
SIX5	100	33.9	1.14	0.1582	1
p300	64	21.7	1.14	0.1978	1
BCLAF1	77	26.1	1.13	0.1990	1
USF2	76	25.8	1.13	0.2012	1
Egr1	184	62.4	1.08	0.2305	1
EBF1	117	39.7	1.08	0.2596	1
PAX5C20	142	48.1	1.08	0.2669	1
TCF12	122	41.4	1.06	0.3144	1
COREST	20	6.8	1.15	0.3181	1
PU.1	94	31.9	1.06	0.3326	1
NFATC1	35	11.9	1.1	0.3430	1
BCL3	46	15.6	1.08	0.3523	1
IRF4	52	17.6	1.07	0.3579	1
Znf143	123	41.7	1.04	0.3858	1
BHLHE40	98	33.2	1.04	0.4023	1
USF1	86	29.2	1.04	0.4044	1
JunD	1	0.3	1.55	0.5105	1
IKZF1	6	2.0	1.06	0.5152	1
CEBPB	21	7.1	1	0.5371	1
ZZZ3	5	1.7	0.99	0.5761	1
MEF2C	29	9.8	0.88	0.7607	1
NRSF	46	15.6	0.9	0.7611	1
CTCF	127	43.1	0.92	0.7788	1
MEF2A	48	16.3	0.89	0.7843	1
ZEB1	47	15.9	0.85	0.8612	1
BATF	21	7.1	0.78	0.8762	1
BCL11A	14	4.7	0.68	0.9374	1
EZH2	4	1.4	0.45	0.9729	1
SMC3	38	12.9	0.65	0.9957	1
Rad21	47	15.9	0.61	0.9994	1
Pol3	0	0.0	0	1	1
ZNF274	0	0.0	0	1	1

ENCODE ChIP-seq data correspond to GM12878 cells. Number and percentage of the 295 cyclin D1-activated genes containing a peak of the corresponding transcription factor overlapping with the cyclin D1 peak are indicated. Odds ratios, p-values, and adjusted (Adj.) p-values are also indicated.

Supplementary Table S4. Motif enrichment analysis using the AME tool.

Motif ID	Protein	P-vaule	Adj. p-value
M6482_1.02	SP3	2.73E-50	2.00E-47
M2314_1.02	SP2	4.28E-49	3.14E-46
M1906_1.02	SP1	1.15E-46	8.40E-44
M6147_1.02	ARID3A	1.10E-43	8.09E-41
M4459_1.02	EGR1	2.55E-42	1.87E-39
M2391_1.02	KLF5	5.77E-41	4.23E-38
M6442_1.02	PURA	1.32E-40	9.68E-38
M6535_1.02	WT1	2.28E-39	1.67E-36
M6539_1.02	ZBTB7B	4.82E-39	3.54E-36
M4604_1.02	ZNF263	1.40E-37	1.03E-34
M2273_1.02	E2F6	9.95E-37	7.29E-34
M2283_1.02	FOXP1	8.17E-36	5.99E-33
M4537_1.02	E2F4	1.61E-34	1.18E-31
M6399_1.02	ONECUT2	2.73E-34	2.00E-31
M6324_1.02	KLF4	3.85E-34	2.82E-31
M6378_1.02	NKX3-1	9.97E-34	7.30E-31
M6245_1.02	FOXO1	1.40E-33	1.03E-30
M5856_1.02	SP8	1.76E-33	1.29E-30
M5593_1.02	KLF16	2.13E-33	1.56E-30
M6483_1.02	SP4	3.45E-33	2.53E-30
M6336_1.02	MAZ	7.95E-33	5.83E-30
M6552_1.02	ZNF148	8.47E-33	6.21E-30
M6242_1.02	FOXJ3	8.98E-33	6.58E-30
M6241_1.02	FOXJ2	2.39E-32	1.75E-29
M6238_1.02	FOXF1	4.50E-32	3.30E-29
M6299_1.02	HOXC6	2.55E-28	1.87E-25
M6237_1.02	FOXD3	2.68E-28	1.96E-25
M6244_1.02	FOXM1	4.94E-28	3.62E-25
M5322_1.02	CPEB1	1.61E-26	1.18E-23
M5697_1.02	ONECUT3	1.70E-26	1.25E-23
M6546_1.02	ZFHX3	1.75E-26	1.28E-23
M4536_1.02	E2F1	1.87E-26	1.37E-23
M1882_1.02	IRF1	3.20E-26	2.35E-23
M6201_1.02	EGR4	7.72E-26	5.66E-23
M6251_1.02	FUBP1	4.34E-25	3.18E-22
M4453_1.02	BCL11A	9.42E-25	6.90E-22
M6221_1.02	ETS2	9.46E-25	6.94E-22
M6321_1.02	KLF15	1.31E-24	9.63E-22
M6114_1.02	FOXA1	1.76E-24	1.29E-21
M0633_1.02	DMRT2	3.37E-24	2.47E-21
M6490_1.02	SRY	5.43E-24	3.98E-21
M6199_1.02	EGR2	5.93E-24	4.35E-21

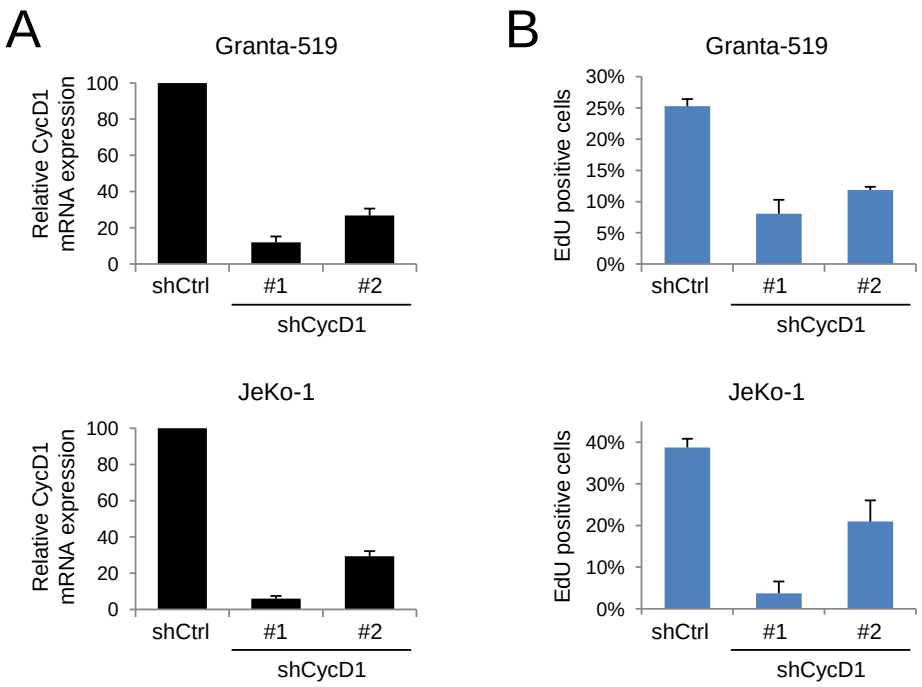
M6329_1.02	LHX3	3.81E-23	2.79E-20
M6250_1.02	FOXQ1	5.46E-23	4.00E-20
M6456_1.02	RREB1	1.02E-22	7.47E-20
M6212_1.02	EPAS1	1.38E-22	1.01E-19
M6239_1.02	FOXF2	1.58E-22	1.16E-19
M5445_1.02	FOXD2	4.06E-22	2.98E-19
M6246_1.02	FOXO3	5.14E-22	3.76E-19
M2305_1.02	NRF1	1.27E-21	9.32E-19
M6438_1.02	PROP1	1.34E-21	9.81E-19
M6249_1.02	FOXP3	1.50E-21	1.10E-18
M2385_1.02	FOXP2	2.11E-21	1.55E-18
M5592_1.02	KLF14	4.39E-20	3.22E-17
M5740_1.02	POU4F1	6.28E-20	4.61E-17
M6234_1.02	FOXA3	9.78E-20	7.17E-17
M6547_1.02	ZFX	1.41E-19	1.03E-16
M6426_1.02	POU3F2	1.50E-19	1.10E-16
M5735_1.02	POU3F3	1.75E-19	1.29E-16
M5743_1.02	POU4F3	2.35E-19	1.72E-16
M6279_1.02	HMGA1	4.39E-19	3.22E-16
M5291_1.02	ARX	1.37E-18	1.00E-15
M2277_1.02	FLI1	1.37E-18	1.01E-15
M6301_1.02	HOXD10	1.73E-18	1.27E-15
M6150_1.02	ARNT2	1.74E-18	1.28E-15
M2307_1.02	PRDM1	3.53E-18	2.59E-15
M6417_1.02	POU1F1	3.71E-18	2.72E-15
M6119_1.02	SPI1	5.51E-18	4.04E-15
M6311_1.02	IRF5	5.92E-18	4.34E-15
M6325_1.02	KLF6	8.03E-18	5.89E-15
M6313_1.02	IRF8	1.14E-17	8.39E-15
M4640_1.02	ZBTB7A	1.25E-17	9.15E-15

Only motifs with adjusted (Adj.) p-value < 10^{-14} are shown.

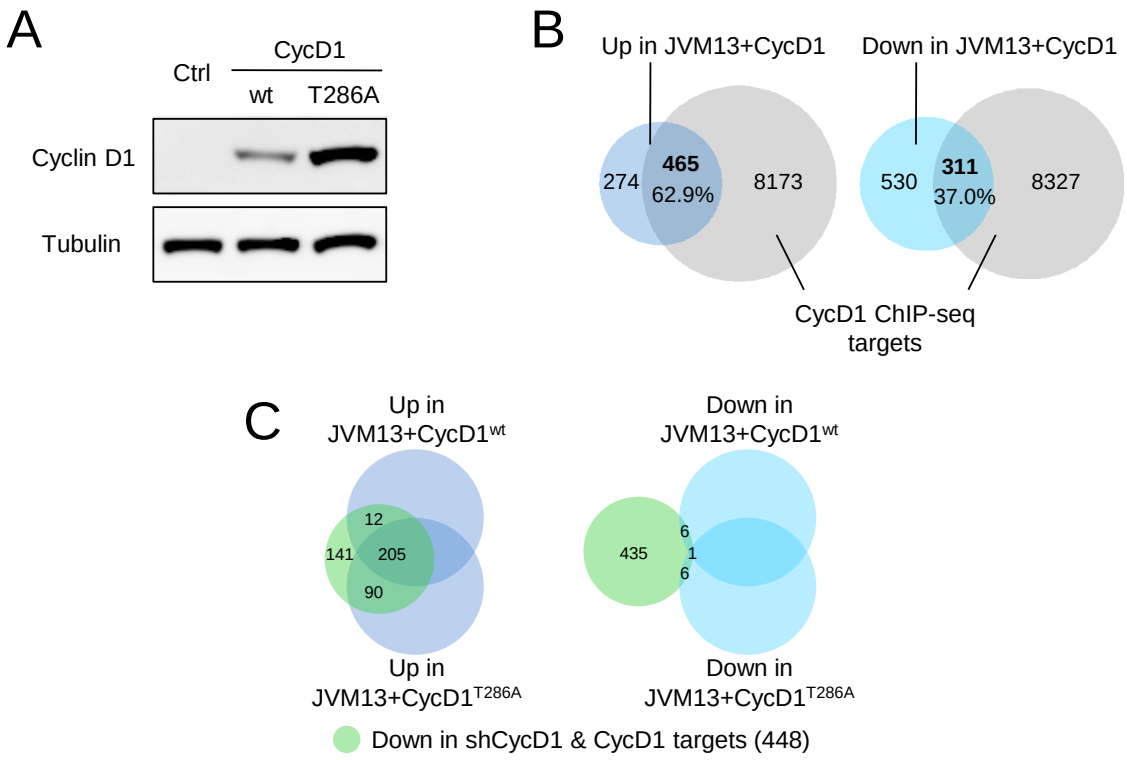
Supplementary Table S5. Two-variable Cox regression models in the 53 leukemic MCL cases from the validation series, considering the 37-gene cyclin D1 signature evaluated by NanoString and each of the different molecular factors analyzed.

	HR	95% CI	P-value
Cyclin D1 Signature	2.39	1.17-4.89	0.017
Cyclin D1 expression	0.95	0.47-1.95	0.908
Cyclin D1 Signature	2.35	1.41-3.91	< 0.001
cMCL vs. nnMCL	0.85	0.28-2.55	0.767
Cyclin D1 Signature	2.49	1.42-4.38	0.001
17p/*TP53*	3.14	1.10-8.91	0.032
Cyclin D1 Signature	2.35	1.18-4.67	0.015
9p/*CDKN2A*	0.92	0.21-4.07	0.908
Cyclin D1 Signature	2.39	1.43-3.97	< 0.001
11q/*ATM*	0.38	0.11-1.34	0.132
Cyclin D1 Signature	1.88	1.07-3.32	0.028
CNA (cont.)	1.08	1.03-1.13	0.008
Cyclin D1 Signature	2.52	1.33-4.77	0.004
Cyclin D1 Full vs. Truncated	0.75	0.20-2.82	0.669
Cyclin D1 Signature	2.82	0.90-8.87	0.076
Proliferation Signature (cont.)	1.00	0.99-1.01	0.704
Cyclin D1 Signature	6.4	1.67-24.74	0.007
Proliferation Signature (cat.)	4.5	0.73-27.66	0.104

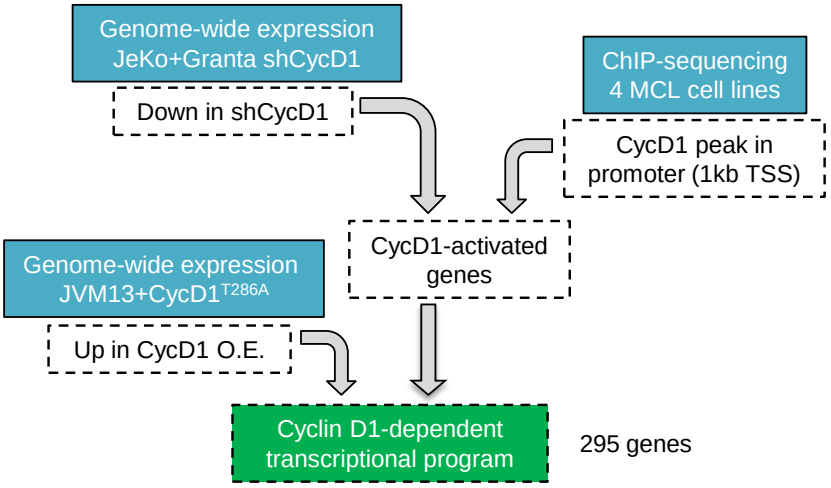
Abbreviations: HR: hazard ratio; CI: confidence interval; CNA: copy number alterations; cont.: continuous variable; cat.: categorical variable. The proliferation signature was analyzed both as a continuous and as a categorical (fixing same single level shift) variable.



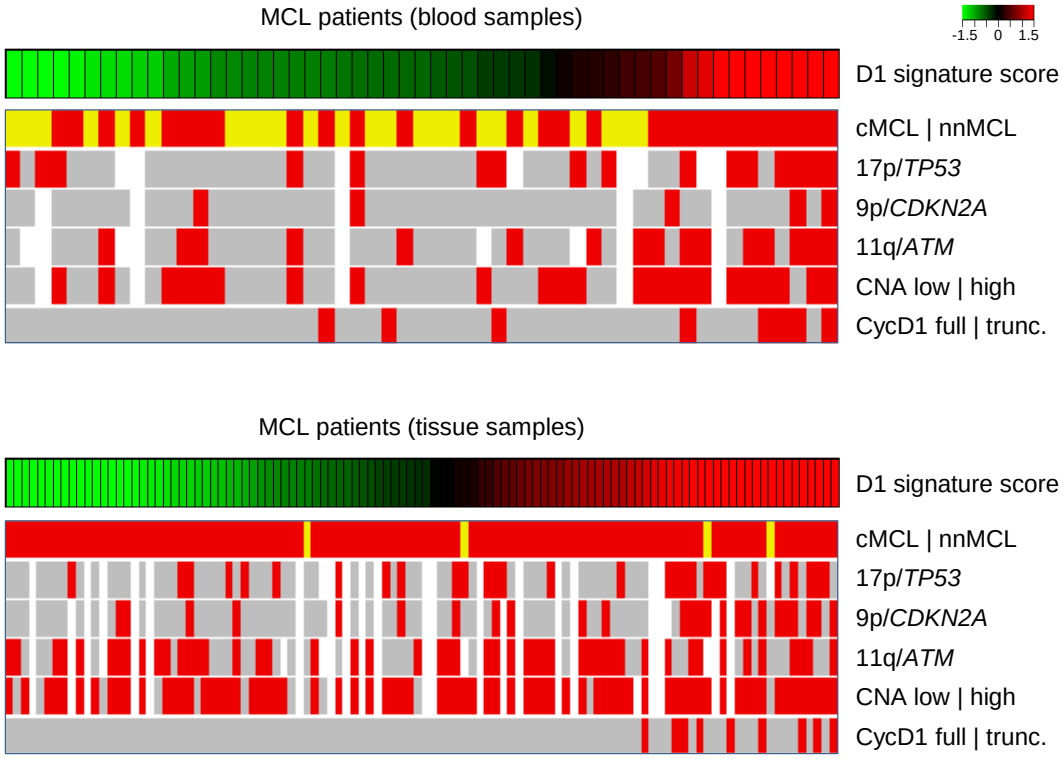
Supplementary Figure S1. (A) Cyclin D1 (CycD1) mRNA expression by qRT-PCR analysis in cyclin D1 silenced (shCycD1 #1 and #2) relative to control (shCtrl) MCL cell lines. Cyclin D1 expression was normalized to the *PUM1* housekeeping gene. Data are the means ± s.e.m. of three independent experiments. (B) Percentage of EdU positive cells in shCtrl and shCycD1 MCL cell lines. Data are the means ± s.e.m. of three independent experiments.



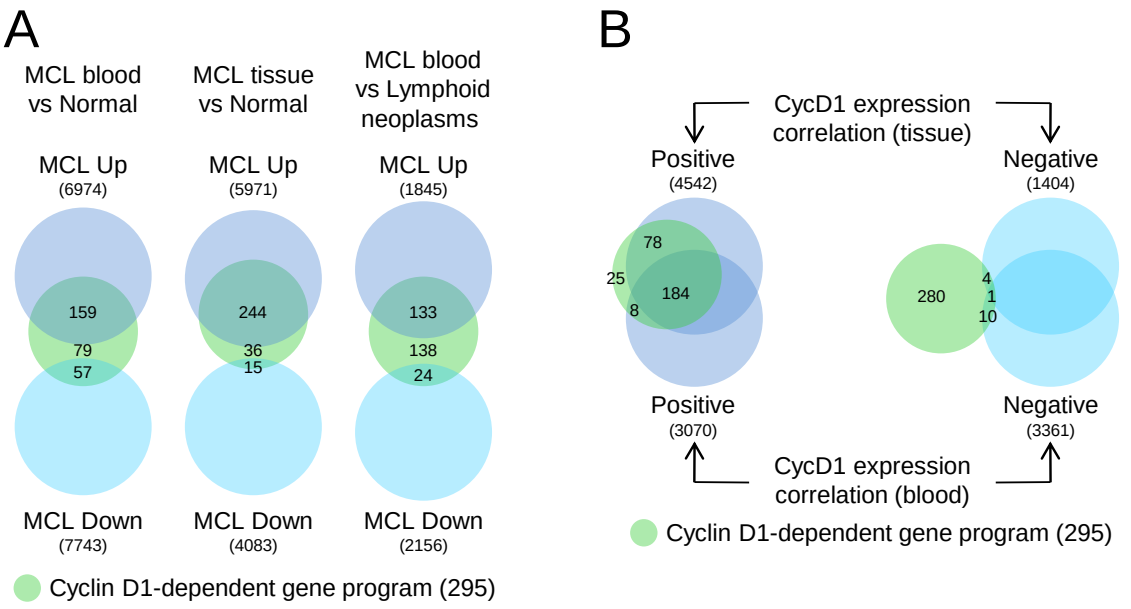
Supplementary Figure S2. (A) Western blot analysis of cyclin D1 in control (Ctrl) and cyclin D1-overexpressing (CycD1^{wt} and CycD1^{T286A}) JVM13 cells. Tubulin was used as loading control. (B) Venn diagrams showing the overlap between differentially expressed genes in cyclin D1-overexpressing JVM13 cells and the cyclin D1 target genes by ChIP-seq in four MCL cell lines (n = 8,638). Upregulated (dark blue) and downregulated (light blue) genes were selected by overlapping results from CycD1^{wt} and CycD1^{T286A} cell models, selecting only the genes dysregulated in both. Statistical significance was assessed by one-tailed Fisher's test. (C) Venn diagrams showing the overlap between the cyclin D1-activated genes identified by RNA-seq and ChIP-seq in MCL cell lines (in green, n = 448) and differential gene expression analysis in cyclin D1-overexpressing JVM13 cells. Genes either upregulated (dark blue) or downregulated (light blue) in CycD1^{wt} or CycD1^{T286A} cell models are shown.



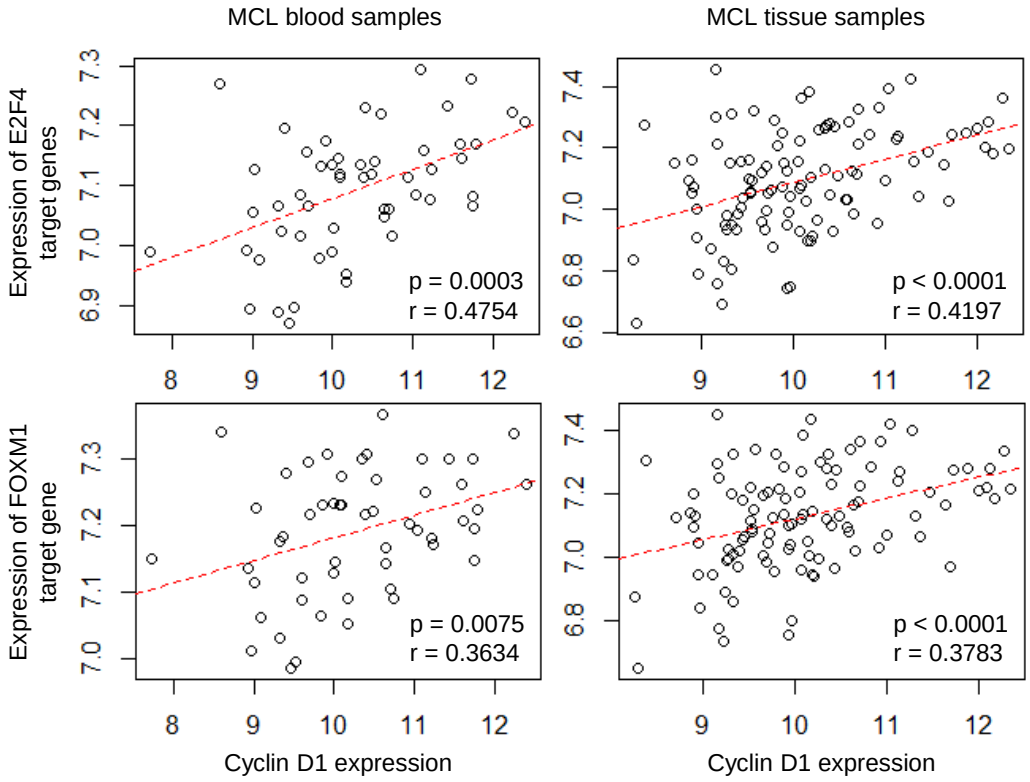
Supplementary Figure S3. Diagram showing the definition of the cyclin D1-dependent transcriptional program through the integration of different experiments. The 295 genes included in the cyclin D1 program fulfilled the following conditions: 1) they were downregulated in cyclin D1-silenced MCL cell lines (in both Granta-519 and JeKo-1); 2) they had a cyclin D1 peak in their proximal promoter in four MCL cell lines, as observed by ChIP-seq; and 3) they were upregulated in CycD1^{T286A} overexpressing (O.E.) JVM13 cells.



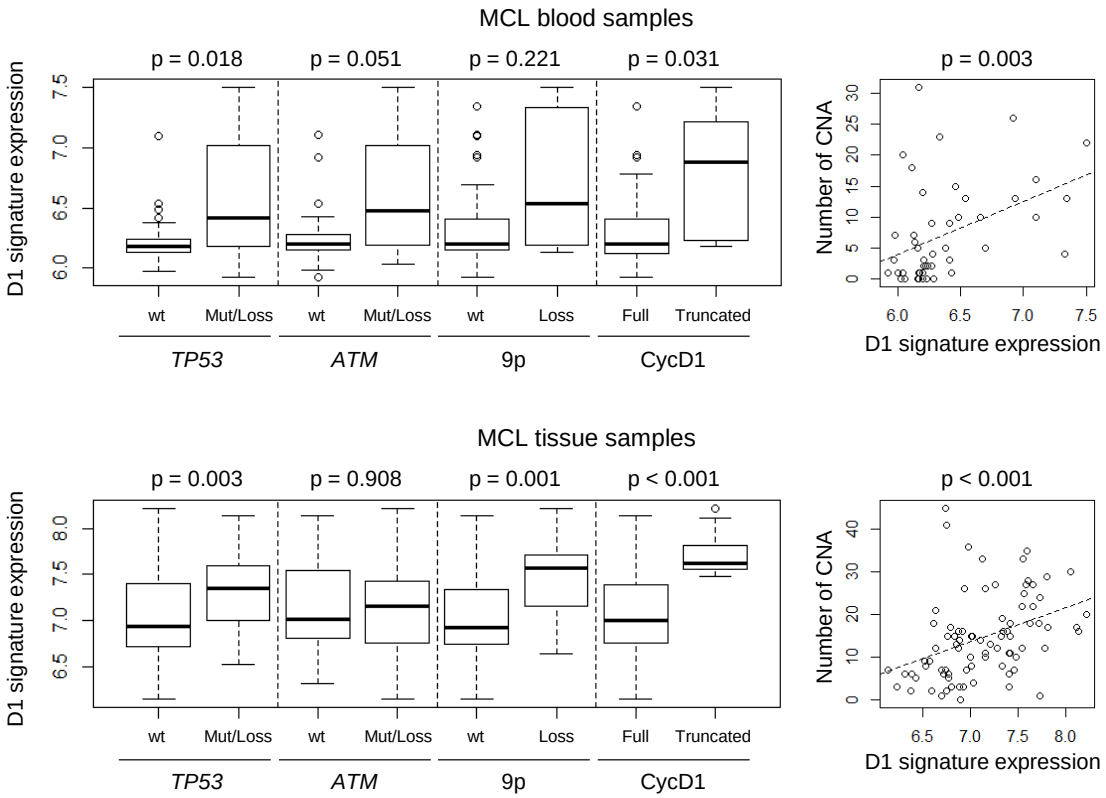
Supplementary Figure S4. Heatmaps of the cyclin D1 signature and several molecular features in MCL primary cases from both peripheral blood (n = 53, top) and lymphoid tissue (n = 106, bottom) samples. MCL patients are shown in columns ordered by cyclin D1 signature score, and are classified in cMCL (red) and nnMCL (yellow). 17p/*TP53*, 9p/*CDKN2A*, and 11q/*ATM* genetic alterations are represented in red. Patients with high number (≥5) of copy number alterations (CNA) are shown in red. Patients with full length and truncated 3'UTR cyclin D1 RNA are represented in grey and red, respectively. White: data not available.



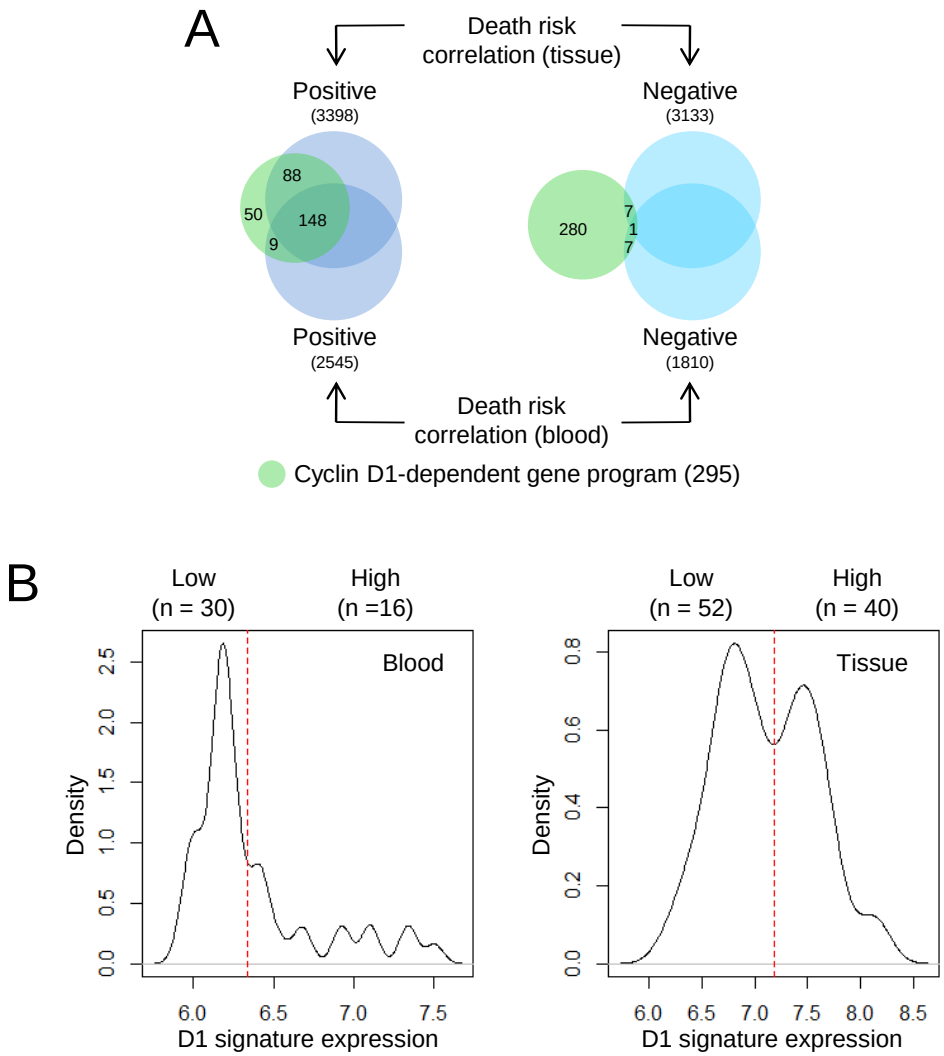
Supplementary Figure S5. (A) Venn diagrams corresponding to Figure 3C. Overlap of the cyclin D1-dependent gene program with genes upregulated (Up) and downregulated (Down) in MCL in the three differential expression analyses: 1) MCL peripheral blood samples versus normal naïve and memory B-cells; 2) MCL lymphoid tissues samples versus normal lymphoid tissues; and 3) MCL peripheral blood samples versus cyclin D1-negative leukemic B-cell chronic lymphoid neoplasms. (B) Venn diagrams representing the overlap between the cyclin D1-dependent gene program and the genes whose expression either positively (dark blue) or negatively (light blue) correlated to cyclin D1 expression in primary MCL. Correlation was assessed by Pearson's r independently in blood and tissue MCL samples.



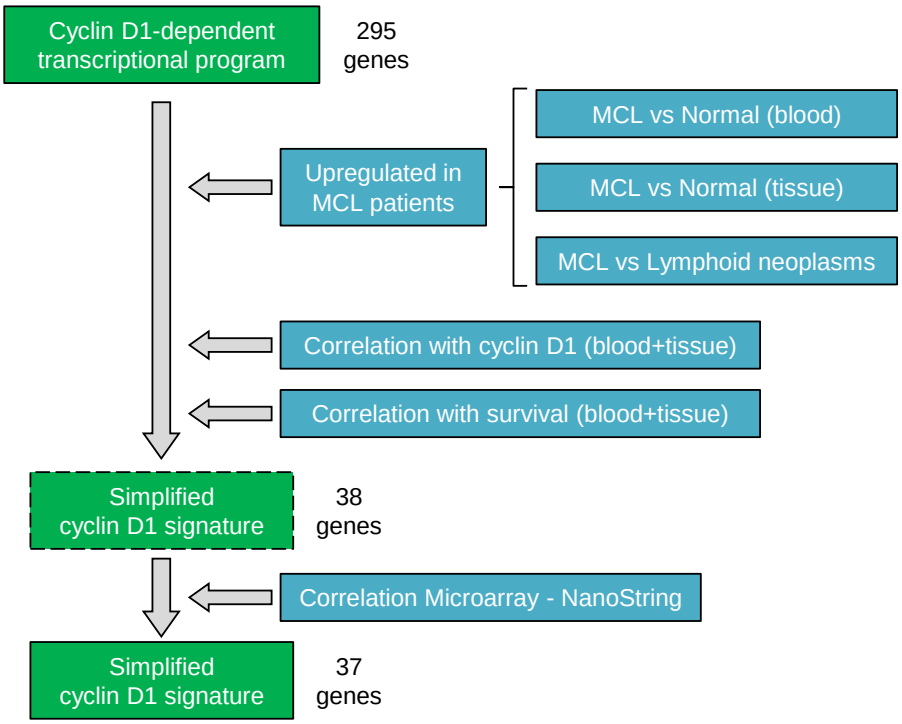
Supplementary Figure S6. Correlation between cyclin D1 expression and the mean expression of either E2F4 target genes (up) or FOXM1 target genes (bottom), in the MCL primary cases from peripheral blood (left) and lymphoid tissue (right) samples. E2F4 and FOXM1 target genes correspond to ChIP-seq analyses in GM12878 cells from the ENCODE database. Correlation was assessed by Pearson's r.



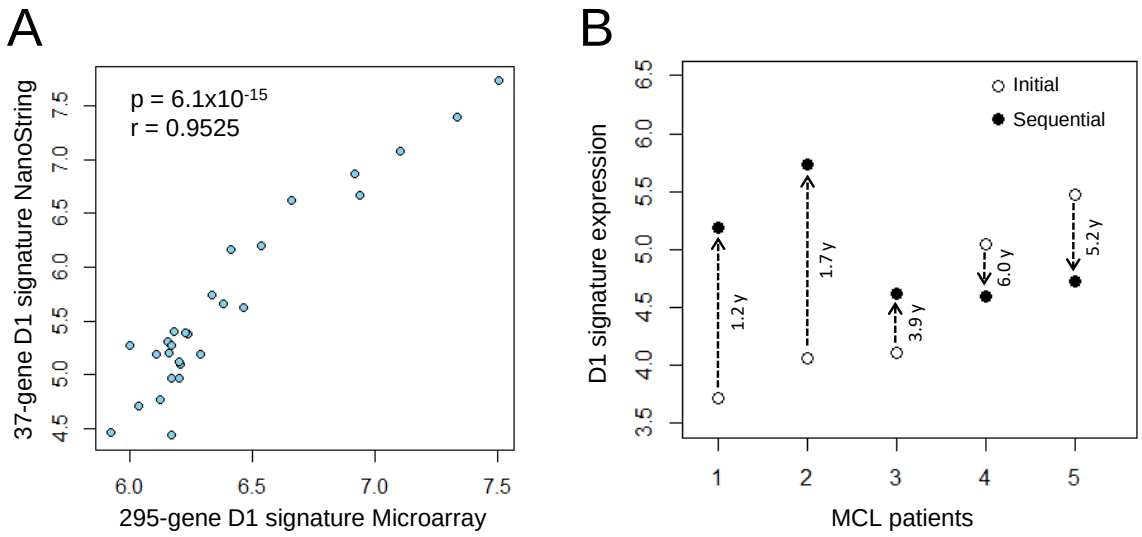
Supplementary Figure S7. Left, boxplots showing cyclin D1 signature expression in the MCL primary cases from peripheral blood (up) and lymphoid tissue (bottom) samples, classified by different molecular features. Statistical significance was assessed by two-tailed Student's t-test. Right, correlation between cyclin D1 signature expression and the number of copy number alterations (CNA) in the two MCL series. Statistical significance was assessed by Pearson's r.



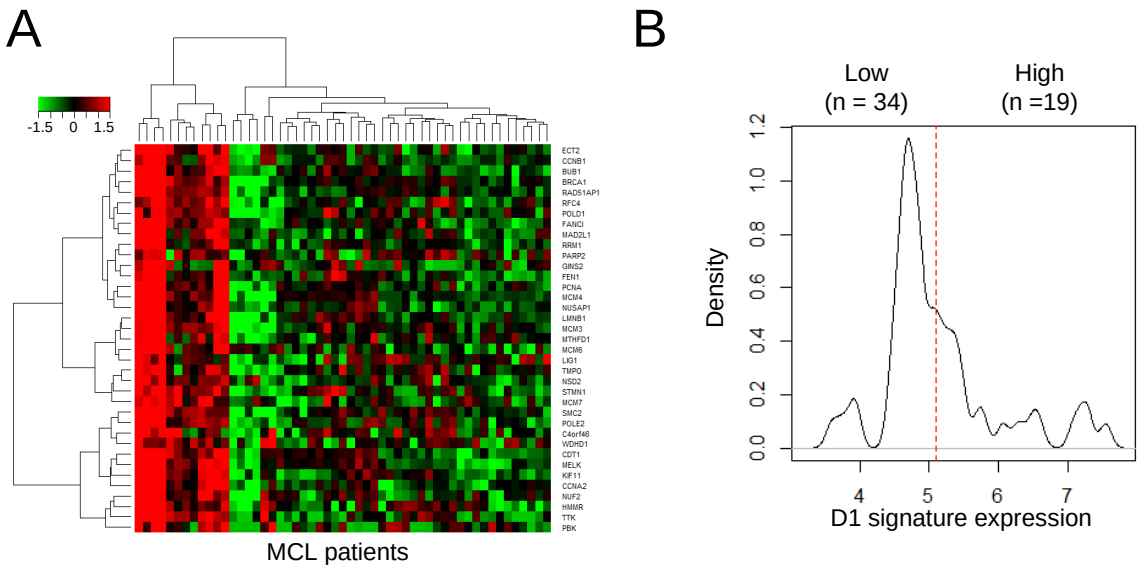
Supplementary Figure S8. (A) Venn diagrams representing the overlap between the cyclin D1-dependent gene program and the genes whose expression positively (dark blue) or negatively (light blue) correlated to death risk in MCL patients. Survival analysis was performed by Cox regression independently in blood and tissue MCL samples. (B) Distribution of cyclin D1 signature score in MCL primary samples from peripheral blood and lymphoid tissues. The thresholds that were used to divide the patients in "low" and "high" groups and the number of cases corresponding to each group are represented.



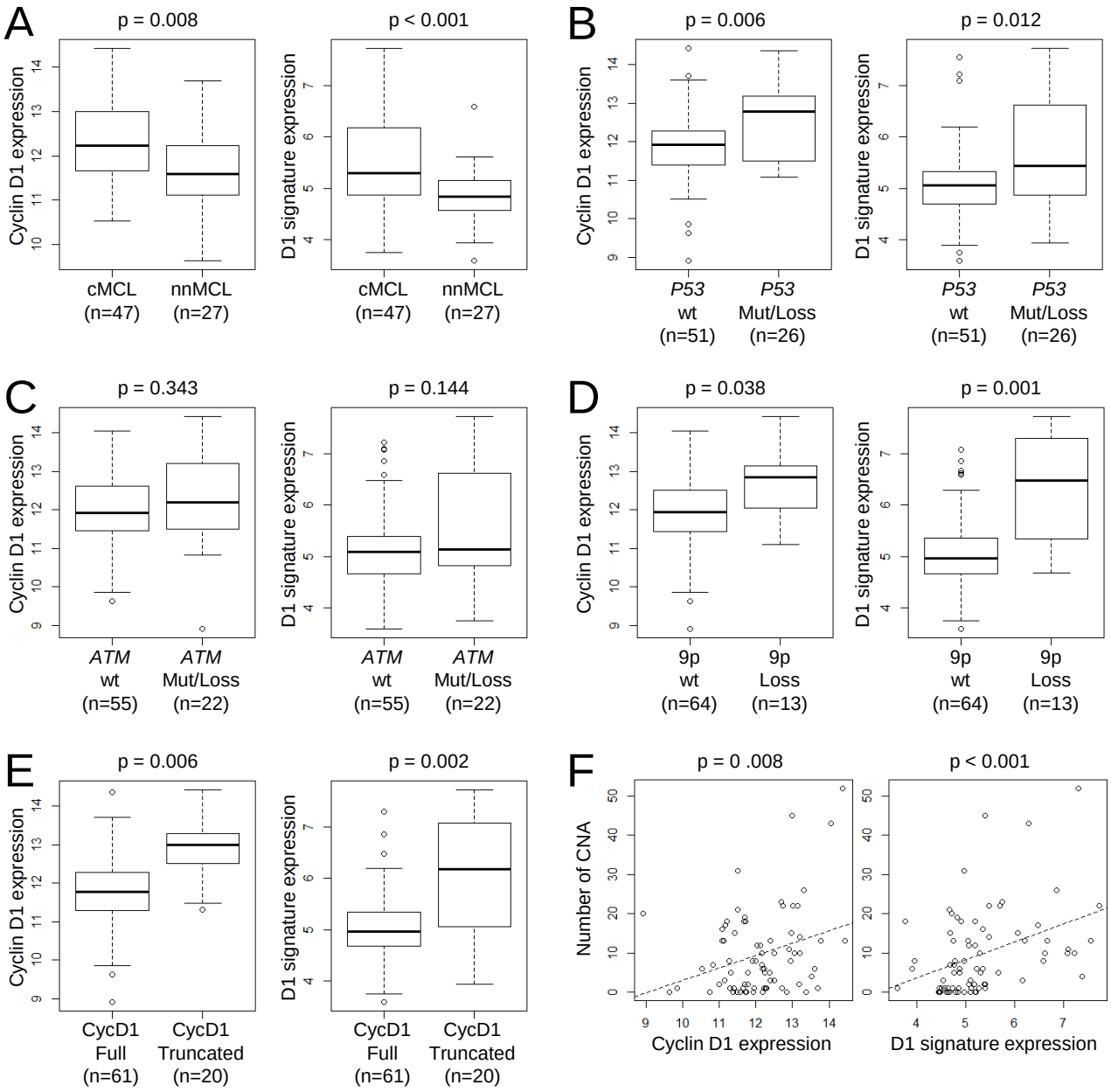
Supplementary Figure S9. Diagram representing the integration of the different analyses used to simplify the full 295-gene cyclin D1-dependent program identified in cell lines into a 37-gene cyclin D1 signature; see supplementary Methods for further details.



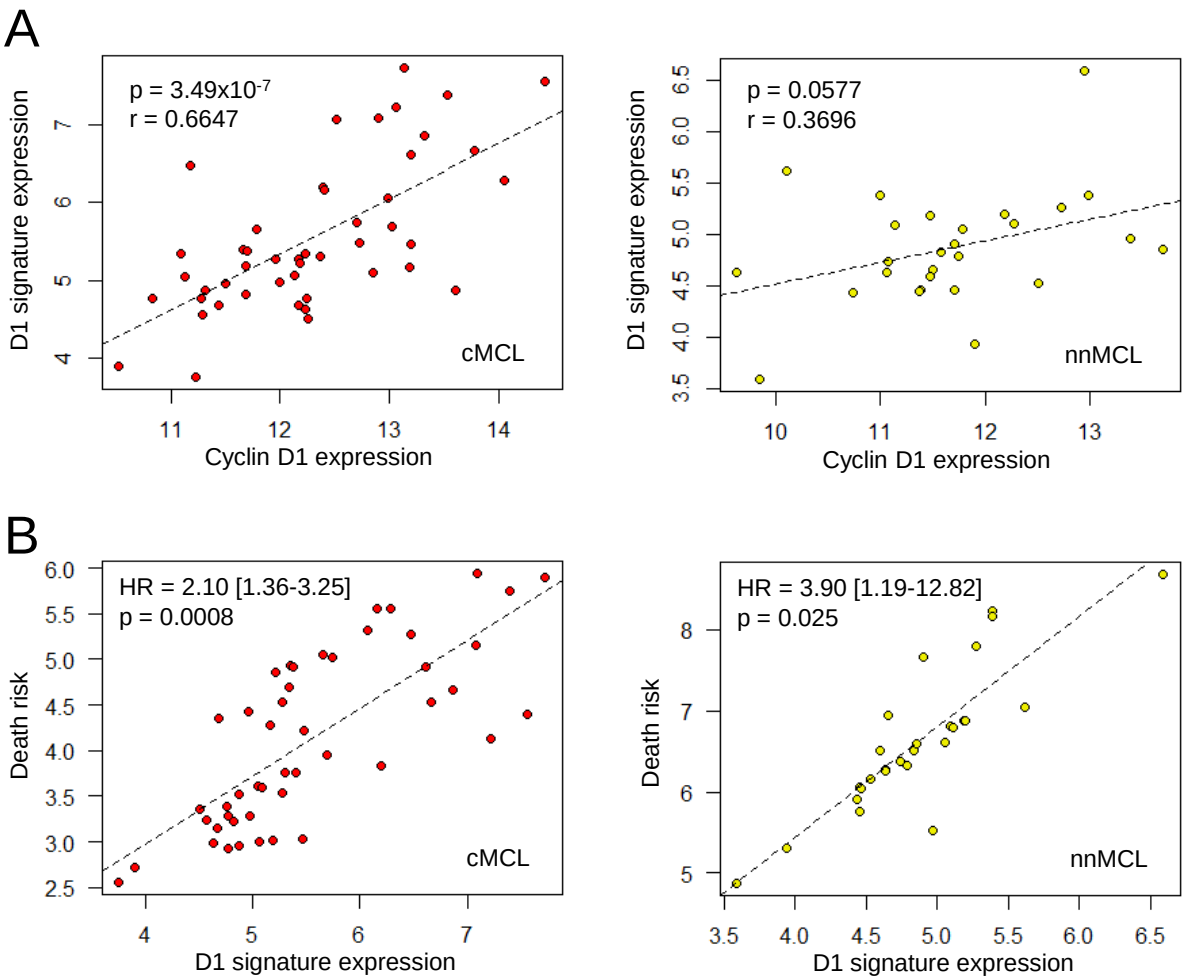
Supplementary Figure S10. (A) Correlation of the 295-gene cyclin D1 signature score by microarray with the 37-gene cyclin D1 signature score by NanoString. Statistical significance was assessed by Pearson's r. (B) 37-gene cyclin D1 signature score (by NanoString) in five MCL cases analyzed in two different time points each: "initial" and "sequential". The elapsed time between the initial and the sequential samples (in years) is indicated for each MCL patient.



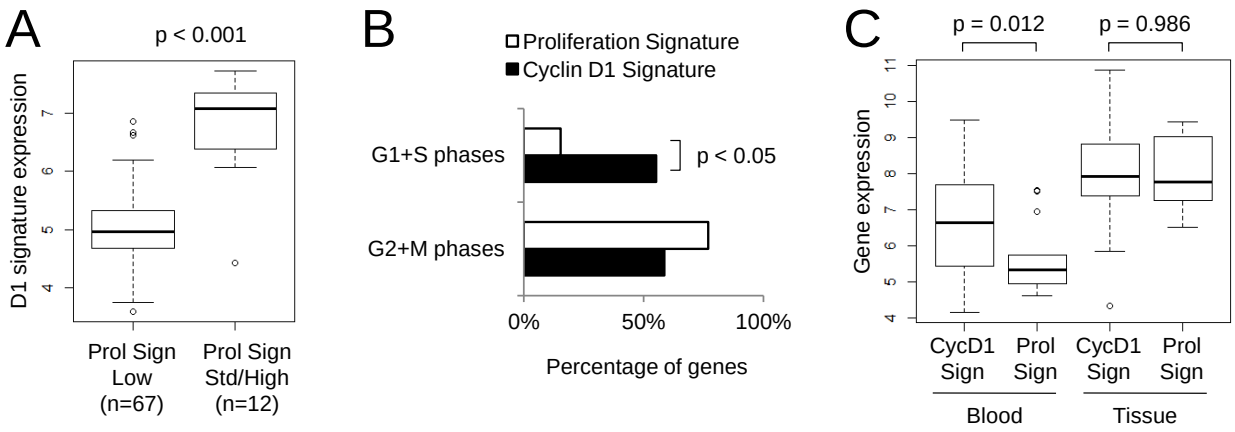
Supplementary Figure S11. (A) Heatmap and hierarchical clustering analysis of the simplified 37-gene cyclin D1 signature analyzed by NanoString in peripheral blood samples from the validation series (n = 53). (B) Distribution of the 37-gene cyclin D1 signature score in MCL samples from the validation series, indicating the threshold that was used to divide the patients in “low” and “high” groups, and the number of cases corresponding to each group.



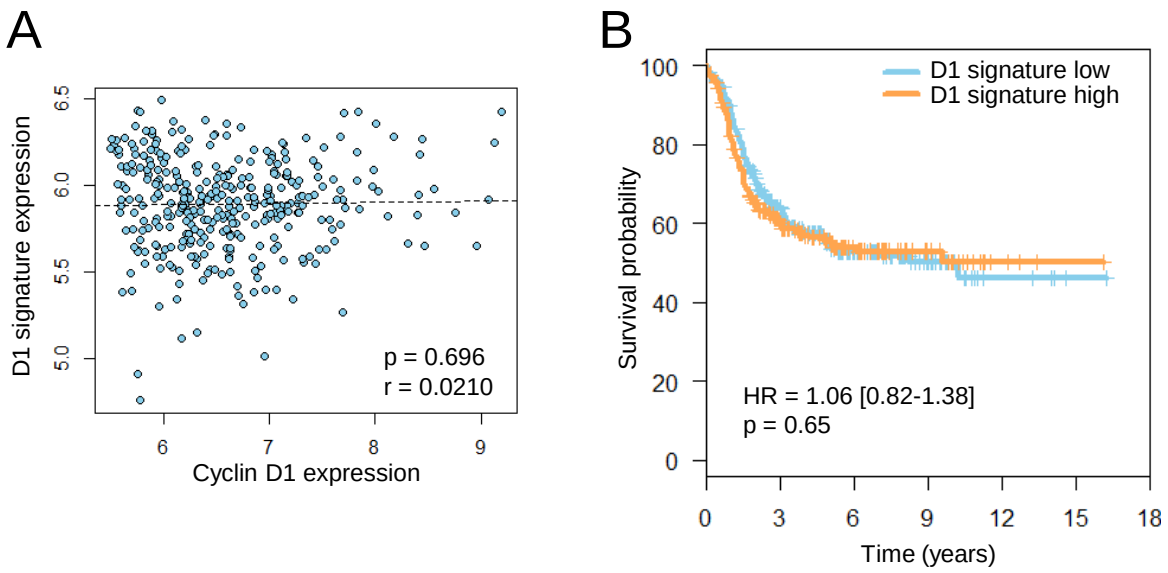
Supplementary Figure S12. (A-E) Boxplots showing either cyclin D1 expression or the 37-gene cyclin D1 signature score in the 81 leukemic MCL patients analyzed by NanoString, classified by different molecular features. Statistical significance was assessed by two-tailed Student's t-test. (F) Correlation between either cyclin D1 expression or the 37-gene cyclin D1 signature score and the number of CNA. Statistical significance was assessed by Pearson's r.



Supplementary Figure S13. (A) Correlation between cyclin D1 expression and the 37-gene cyclin D1 signature score in cMCL (n = 47) and nnMCL (n = 27) patients. Correlation was assessed by Pearson's r. (B) Association between the 37-gene cyclin D1 signature score and the death risk in cMCL and nnMCL. The death risk (y-axis) corresponds to the sum of the martingale residuals and the linear predictors of the fitted OS Cox model; HR with 95% confidence interval and p-value are shown. Survival data were calculated from sampling time.



Supplementary Figure S14. (A) Boxplot of the 37-gene cyclin D1 signature score in MCL patients divided based on their MCL proliferation signature category according to the MCL35 assay: "low" versus "standard"/"high". Statistical significance was assessed by two-tailed Student's t-test. (B) GO analysis of cyclin D1 signature and MCL proliferation signature genes divided in G1/S and G2/M cell cycle phases as described in Figure 2B. Only the genes belonging to the "cell cycle" GO category were considered. Statistical significance was assessed by Fisher's test. (C) Boxplot showing gene expression levels of cyclin D1 and proliferation signatures in blood and tissue samples from the same patients. In the case of the proliferation signature, only the pro-proliferation genes were considered. Statistical significance was assessed by two-tailed Student's t-test.



Supplementary Figure S15. (A) Correlation between cyclin D1 expression and the 295-gene cyclin D1 signature score in ER-negative breast cancer. Correlation was assessed by Pearson's r. (B) Kaplan-Meier curves of the progression free survival in ER-negative breast cancer patients splitted in "high" and "low" groups by the median 37-gene cyclin D1 signature levels.