Cancer Cell, Volume *30*

Supplemental Information

A Druggable TCF4- and BRD4-Dependent Transcriptional Network Sustains Malignancy in Blastic Plasmacytoid Dendritic Cell Neoplasm

Michele Ceribelli, Zhiying Esther Hou, Priscilla N. Kelly, Da Wei Huang, George Wright, Karthik Ganapathi, Moses O. Evbuomwan, Stefania Pittaluga, Arthur L. Shaffer, Guido Marcucci, Stephen J. Forman, Wenming Xiao, Rajarshi Guha, Xiaohu Zhang, Marc Ferrer, Laurence Chaperot, Joel Plumas, Elaine S. Jaffe, Craig J. Thomas, Boris Reizis, and Louis M. Staudt

SUPPLEMENTAL DATA

Table S1, related to Figure 1. Provided as a separate Excel file.

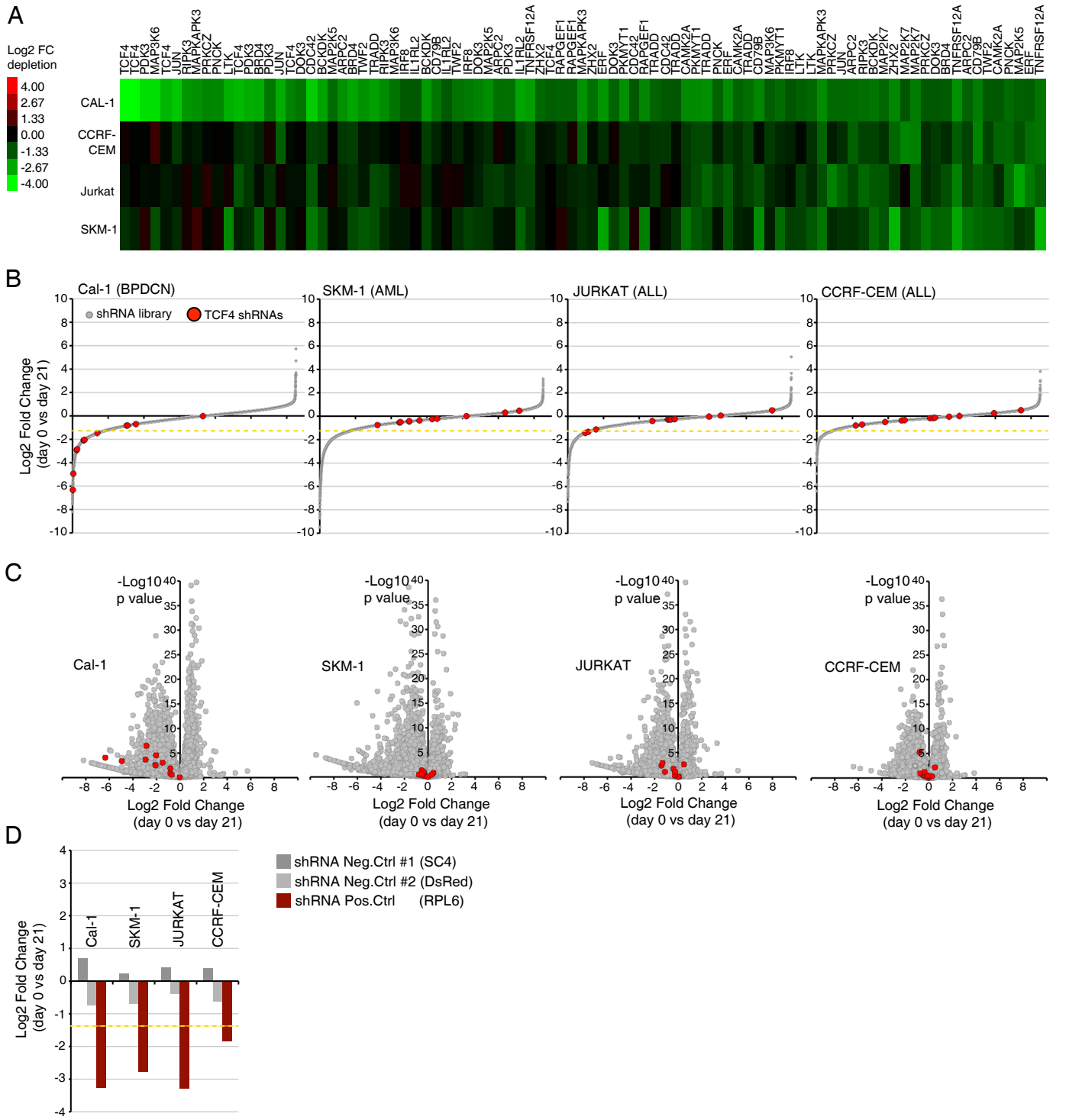

Figure S1, related to Figure 1: **Genetic dependencies of BPDCN identified by shRNA screen.**
A) The shRNAs toxic to Cal-1 cells were ranked based on their differential toxicity with respect to the average of 3 control leukemia cell lines: CCRF-CEM, T-cell acute lymphoblastic leukemia (T-ALL); Jurkat, T-ALL; SKM-1, acute myeloid leukemia (AML). **B)** The average depletion (Log2 fold change) of every shRNA present in the library is shown for each of the indicated cell lines. The 12 TCF4 shRNAs present in the library are highlighted in red. **C)** For each cell line, 4 independent biological replicates were performed, and Volcano scatter plots were generated to compare the significance (p value) vs the fold-change depletion observed for each shRNA present in the library. The 12 TCF4 shRNAs present in the library are highlighted in red. **D)** The average depletion (Log2 fold change) of a positive control shRNA (ribosomal protein RPL6) and of two negative control shRNA is shown for the indicated cell lines.

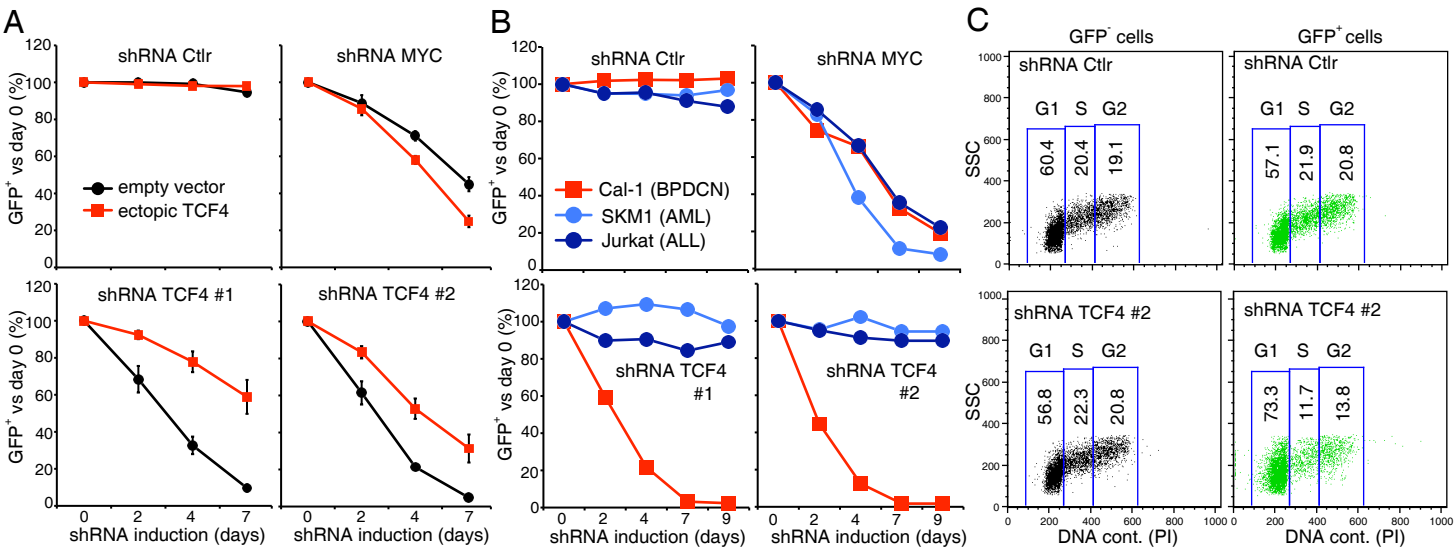

Figure S2, related to Figure 2: **The E-box transcription factor TCF4 is required for BPDCN viability.**
A) Cal-1 cells were first infected witheither an empty vector or a TCF4 rescue construct. After puromycin selection, cells were superinfected with the indicated shRNAs, and shRNA toxicity was assessed by depletion of the GFP$^+$ population over time, as described in Fig. 2A. **B)** Cal-1, SKM-1 and Jurkatcells were infected with Ctrl, MYC or TCF4 shRNAs. Shown is the fraction of live, shRNA expressing (GFP$^+$) cells over time after shRNA induction, compared to the day 0 un-induced value. **C)** A representative flow cytometry cell cycle analysis, based on DNA content (Propidium Iodide, PI staining), is shown for Ctrl and TCF4 shRNAs at day 4 post induction in Cal-1 cells. Error bars represent SEM of triplicates.

Table S2, related to Figure 3. Provided as a separate Excel file.
Table S3, related to Figure 3. Provided as a separate Excel file.

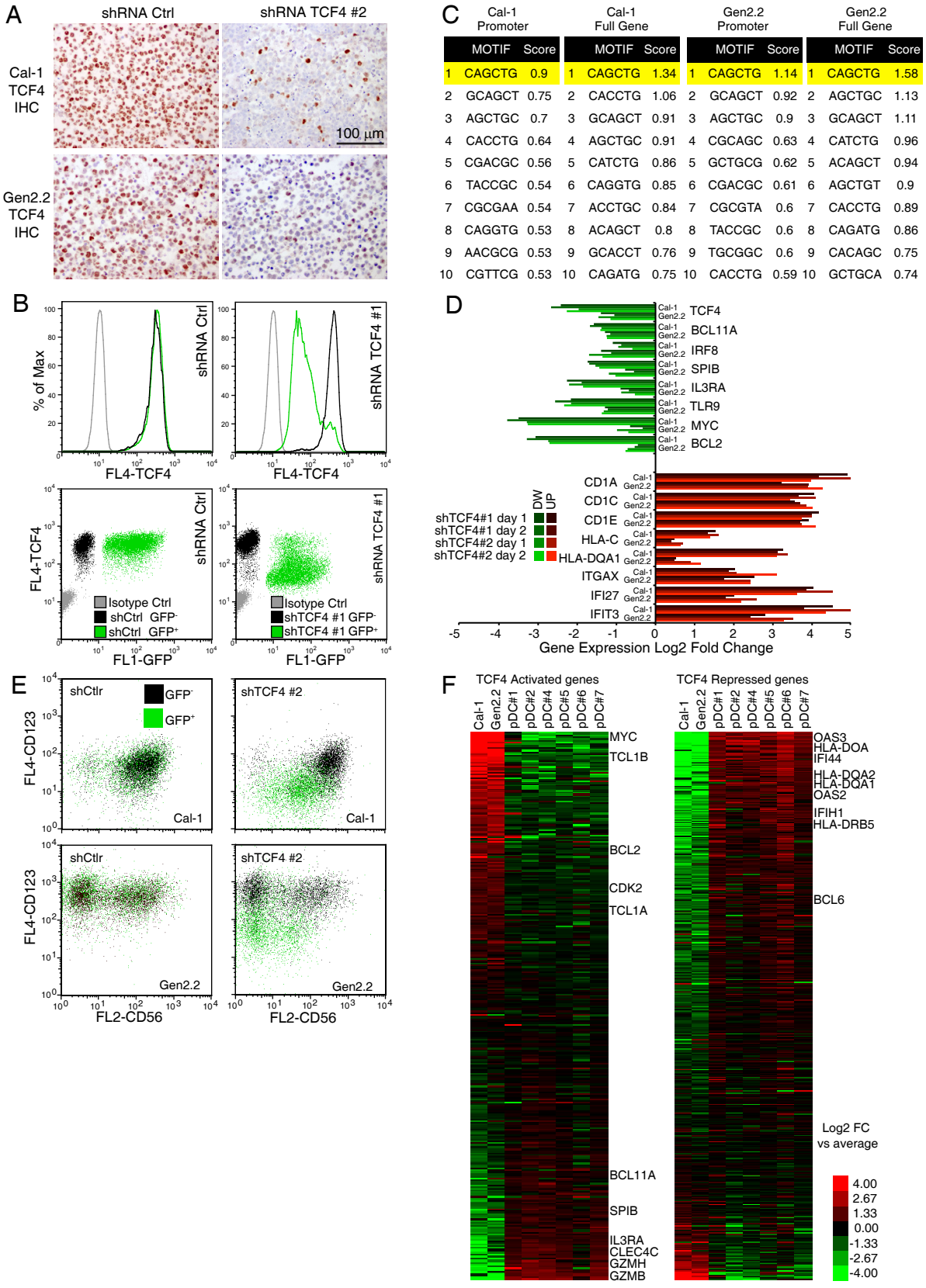

Figure S3, related to Figure 3: **The TCF4 dependent transcriptional network in BPDCN.**
A) BPDCN cell lines were infected with either Ctrl or TCF4 shRNA. TCF4 immuno-histo-chemistry (IHC) was performed on paraffin embedded cell pellets after selection and induction of shRNA expression for 1 day. **B)** Cal-1 cells were infected with either Ctrl or TCF4 shRNA. TCF4 expression was analyzed by intracellular flow after induction of shRNA expression for 1 day. **C)** The top 10 enriched motives predicted by Clover are shown for the indicated TCF4 binding regions (Frith et al., 2004). **D)** The gene expression changes induced by TCF4 knock-down in BPDCN cell lines are shown, for selected TCF4 target genes, as a Log2 fold change (FC) versus a Ctrl shRNA. **E)** A representative stain for surface CD56 and CD123 is shown for Cal-1 and Gen2.2 cells infected with either Ctrl or TCF4 shRNAs. **F)** Heat-map comparing the RNA-Seq expression of TCF4 activated genes and TCF4 repressed genes in BPDCN cells with that of normal pDC isolated from 6 healthy donors.

Table S4, related to Figure 4. Provided as a separate Excel file.

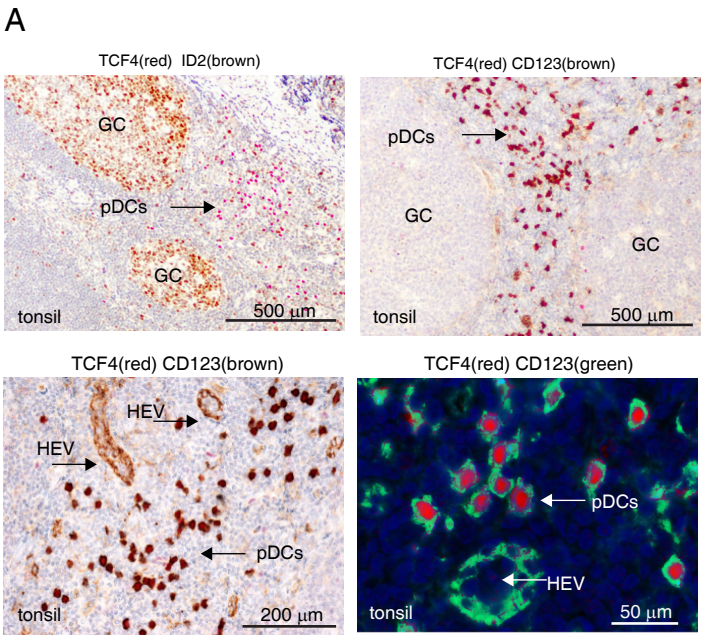

Figure S4, related to Figure 4: **TCF4 expression facilitates BPDCN diagnosis**.
A) A normal, reactive human tonsil was used to perform double TCF4/ID2 IHC (top left panel) or TCF4/CD123 IHC (top right panel, and both bottom panels). The ID2 antibody was used to identify germinal centers, as well as a negative control for pDCs. Germinal Center: GC; High Endothelial Venules: HEV; cluster of plasmacytoid dendritic cells are indicated as pDCs.

Table S5, related to Figure 5. Provided as a separate Excel file.

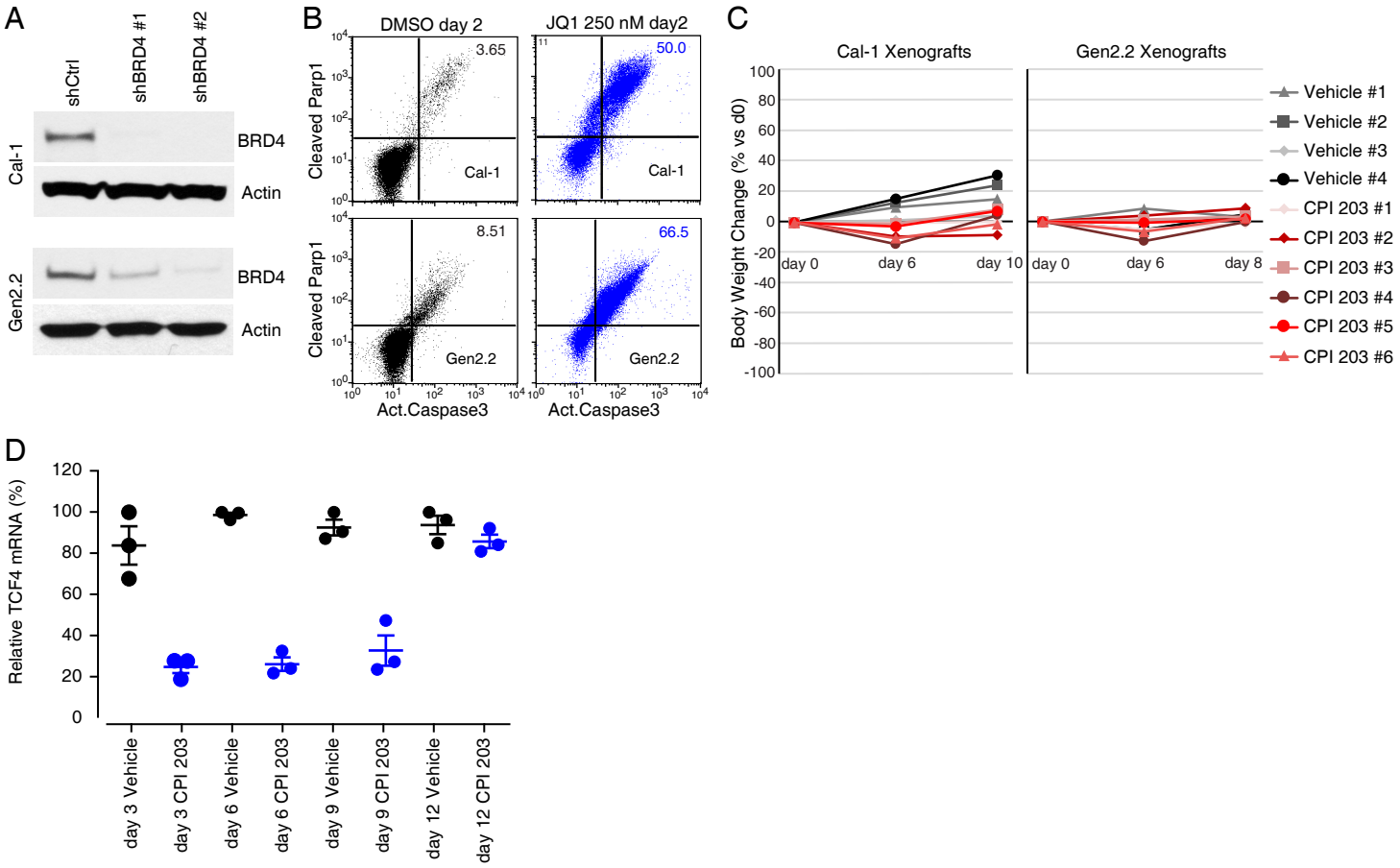

Figure S5, related to Figure 5: **The bromodomain and extra-terminal domain BRD4 is required for BPDCN survival.** **A)** Western-blot analysis of BRD4 protein levels 24 hr after the inducible expression of either Ctrl or BRD4 shRNAs in BPDCN lines. Actin was used as loading control. **B)** A representative flow cytometry stain for active Caspase-3 and cleaved Parp1 is shown for DMSO and JQ1 treated BPDCN cell lines at day 2 post-treatment. **C)** Cal-1 BPDCN cells were established as a subcutaneous tumor in NOD/SCID mice and treated by intraperitoneal injection with vehicle or BET Inhibitor CPI 203 (5 mg/kg, BID IP). Body weight was measured at the indicated time points and percent change was calculated vs day 0. **D)** Relative TCF4 mRNA levels in Cal-1 xenografts from mice treated with either vehicle or CPI 203 for the indicated time points. 3 mice for each treatment/time point combination were used. B2M was used as the housekeeping control gene. For each time point, data were scaled to the highest vehicle value, and scatter dot plots were generated accordingly. Lines refer to the mean value +- SEM.

Table S6, related to Figure 6. Provided as a separate Excel file.

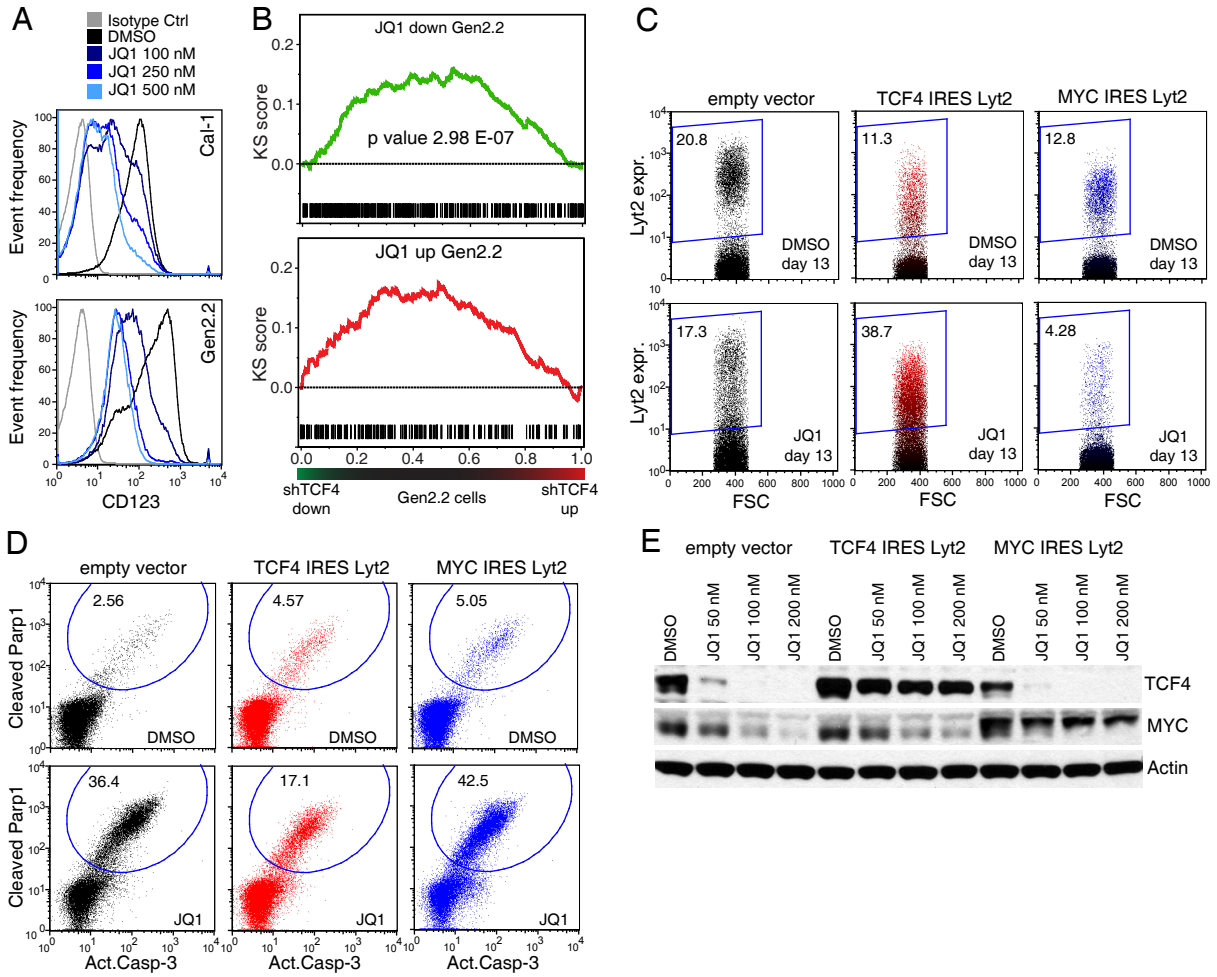

Figure S6, related to Figure 6: **A TCF4 and BRD4 dependent transcriptional network in BPDCN.**
A) BPDCN cell lines were treated with either DMSO or the indicated JQ1 amounts and expression of CD123 was measured by flow cytometry 24 hr after treatment. Shown is a representative stain, compared to the isotype control. **B)** Gene set enrichment analysis (GSEA) comparing JQ1 and shRNA TCF4 gene expression datasets in the Gen2.2 BPDCN cell line. Genes were ranked based on the gene expression changes induced by TCF4 shRNAs and the distribution of JQ1 dependent genes was analyzed accordingly.
C) Cal-1 cells were infected with the indicated rescue vectors and the fraction of Lyt2$^+$ cell was monitored over a 2-weeks treatment with either DMSO or the indicated JQ1 amounts. A representative Lyt2 stain is shown. **D)** Cal-1 cells were infected with the indicated rescue vectors. Apoptosis induction was monitored by flow cytometry for active Caspase-3 and cleaved Parp1. A representative stain is shown for either DMSO or JQ1 treated cells (JQ1 150 nM, 2 day). **E)** Cal-1 cells were infected with the indicated rescue vectors and Lyt2$^+$ cells were positively selected with CD8a Magnetic Beads. Lyt2$^+$ pools were treated with the indicated JQ1 amounts for 24 hr and TCF4 and MYC protein expression was determined by western blot. Actin was used as loading control.

Table S7, related to Figure 7. Provided as a separate Excel file.

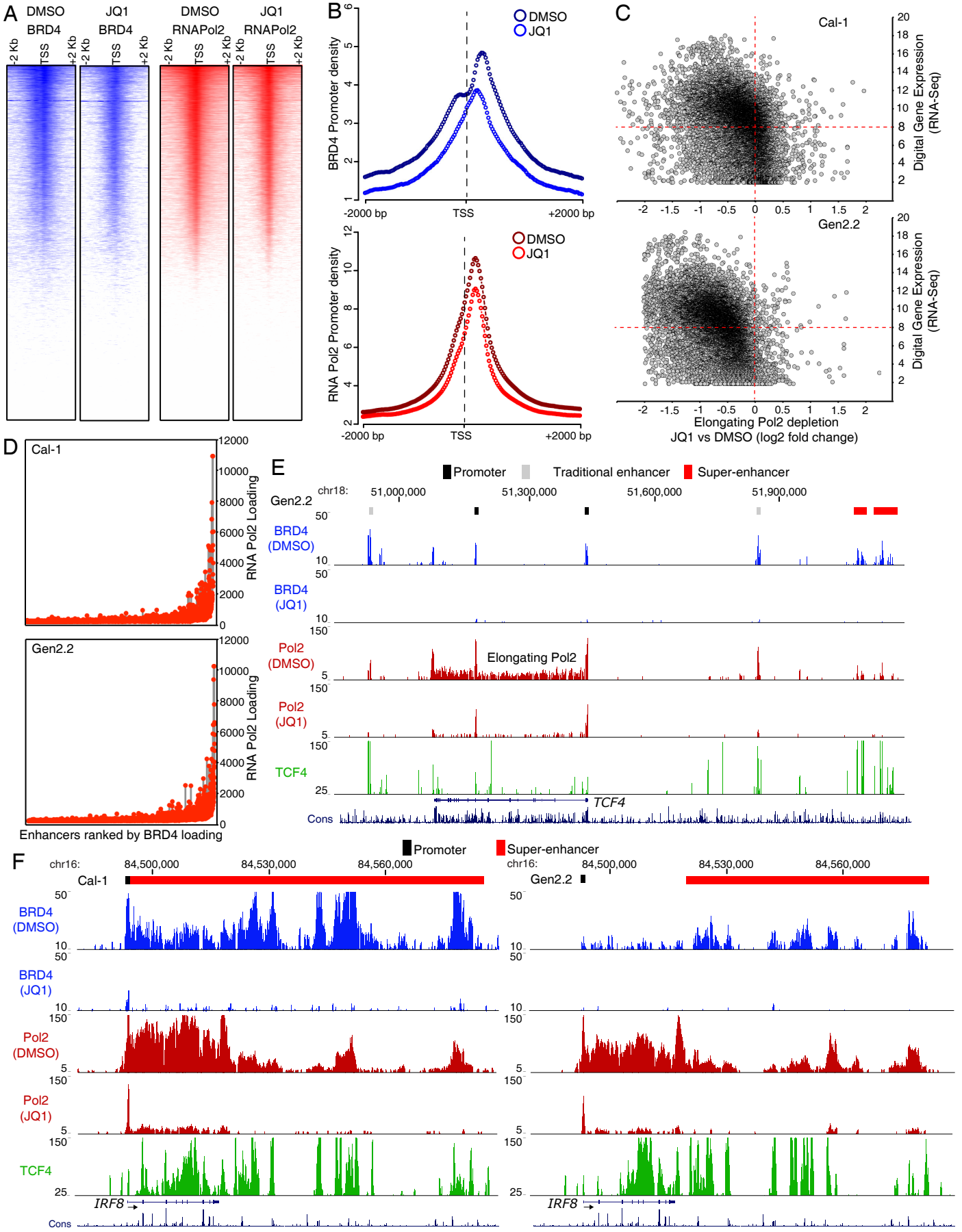

Figure S7, related to Figure 7: **Mapping the BRD4 dependent super-enhancers in BPDCN.**
A) Heat-maps of promoter RNA Pol2 (blue) and BRD4 (red) density in Gen2.2 cells, after 12 hr treatment with either DMSO or 250 nM JQ1. Genes were ranked by RNA Pol2 density in DMSO treated cells and the other heat-maps were displayed accordingly. **B)** Meta-promoter profiles of BRD4 (blues) and RNA Pol2 (reds) ChIP-Seq data in Gen2.2 cells. **C)** For each BPDCN cell line, the depletion of elongating RNA Pol2 (JQ1 vs DMSO, Log2 fold change) was plotted against the digital gene expression value of the corresponding RefSeq, derived from an RNA-Seq experiment in the same cell line. **D)** For each BPDCN cell line, enhancers were ranked based on increasing BRD4 loading. The corresponding signal from RNA Pol2 ChIP-Seq was displayed accordingly. **E)** The *TCF4* locus ChIP-Seq tracks for BRD4 (blue), RNA Pol2 (red) and TCF4 (green) are shown for Gen2.2 cells. **F)** The *IRF8* locus ChIP-Seq tracks for BRD4 (blue), RNA Pol2 (red) and TCF4 (green) are shown for both Cal-1 and Gen2.2 cell lines.

Table S8, related to Figure 8. Reported as a separate Excel file.

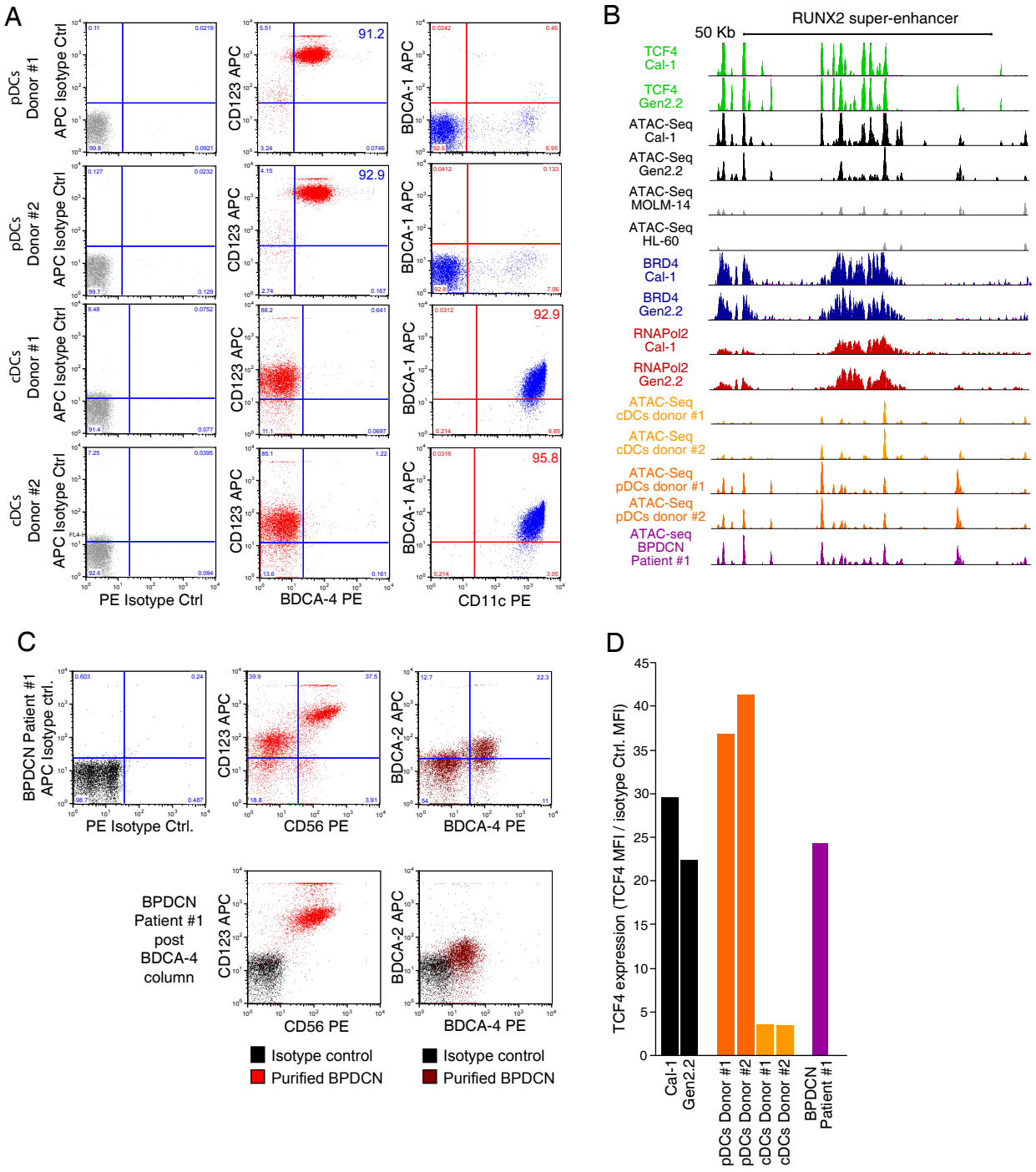

Figure S8, related to Figure 8: **ATAC-Seq mapping of the regulatory landscape of BPDCN.**
A) pDCs and conventional CD11c^{+} dendritic cells (cDCs, also referred as mDC1) were isolated from 2 healthy donors. The purity of the isolated population (>90%) was confirmed by flow for CD123 and BDCA-4 for pDCs, and BDCA-1 and CD11c for cDCs. **B)** ATAC-Seq was used to map the open chromatin regulatory landscape of the *RUNX2* super-enhancer. Samples are indicated on the left, and color-coded as reported in Figure 8A, B and C. **C)** PBMCs isolated from a patient with a diagnosis of BPDCN were analyzed by flow cytometry. Stains for CD123 vs CD56 and BDCA-2 vs BDCA-4 are shown before and after positive selection of BPDCN primary cells. **D)** TCF4 expression was measured by intracellular flow. The mean fluorescence intensity (MFI) ratios of TCF4 vs Isotype Ctrl are shown for the indicated cell lines and primary samples.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cell lines

Cal-1, CCRF-CEM, Jurkat, HL-60, MOLM-14 and SKM-1 cells were cultured in RPMI-1640 medium supplemented with penicillin/streptomycin and 10% fetal bovine serum (Tet tested, Atlanta Biologicals). Gen2.2 cells were cultured in RPMI 1640 medium supplemented with non-essential amino acids, sodium pyruvate, penicillin/streptomycin and 10% fetal bovine serum. The mouse MS-5 stromal feeder cell line was used to support growth of Gen2.2 cells. All cell lines were maintained in a humidified, 5% CO2 incubator at 37 °C. To perform inducible, shRNA toxicity experiments, all cell lines were engineered to express a feline endogenous virus (FEV) ecotropic retroviral receptor (blastacidin as selectable marker) and a bacterial tetracycline repressor (TETR, selectable marker neomycin), as previously described (Ngo et al., 2006).

shRNA library toxicity screening

The shRNA library toxicity screening was performed essentially as previously described (Ngo et al., 2006). Briefly, pools of 1,000 shRNAs were used to retrovirally transduce Cal-1 cells. After puromycin selection, shRNA expression was induced by doxycycline (50 ng/ml). Uninduced cultures were kept in parallel as a control. After 3 weeks, genomic DNA from both un-induced and induced cultures was harvested and depletion of specific shRNAs was quantitatively assessed by next generation sequencing.

shRNA toxicity screening

For individual shRNA toxicity screens, each shRNA was cloned into the pRSMX-PG vector that enables co-expression of the shRNA and of a GFP-Puromycin fusion, and retrovirally transduced into target cells. When required, stable pools of shRNA expressing cells were selected for 2-4 days with puromycin. Two days after retroviral transduction, shRNA expression was induced by doxycycline (50 ng/ml) and the fraction of GFP^+, shRNA-expressing cells was monitored over time by flow cytometry. Toxicity was assessed as depletion of the GFP^+ cell population over time.
For rescue experiments, Cal-1 cells were first infected with a pBMN-IRES-Lyt2 vector (empty, MYC or TCF4) that enables internal ribosome entry site (IRES) based surface expression of mouse CD8a (Lyt2) together with expression of the indicated cDNA. Following mouse CD8a Dynabeads selection (Invitrogen), rescue pools were superinfected with pRSMX-PG vectors co-expressing GFP and the shRNAs of interest. Viability of GPF^+ cells was monitored over time as described above.
The RNAi sequences of the shRNA vectors used in this study are:

shCtrl:	CTCTCAACCCTTTAAATCTGA
shMYC:	CGATTCCTTCTAACAGAAATG (3' UTR)
shTCF4 #1:	GGGACGACTTTCTTTAACATA (3' UTR)
shTCF4 #2:	CAAGTCTGAGTAGTTATGAAT (3' UTR)
shBRD4 #1:	GCACTGACTTTGCCTTGAACA (3' UTR)
shBRD4 #2:	GCAGAGCCTGTCTGCCTTACG (3' UTR)

RT-PCR

Total RNA was extracted from BPDCN cell lines with TRIzol (Life Technologies). For each sample, 1 ug of total RNA was reverse-transcribed with random primers and SuperScript II reverse transcriptase (Invitrogen) for 2h at 42 °C. A 1:25 dilution of the cDNA stock was used to set-up TaqMan Q-RT-PCR on an ABI7500 instrument (Applied Biosystems). The TaqMan assays used for this study are as follow: B2M (Hs00984230_m1), TCF4 (Hs00162613_m1), TLR9, (Hs00370913_s1), MYC (Hs00153408_m1) and BCL2 (Hs00608023_m1).

Development of a rabbit anti-TCF4 monoclonal antibody

The TCF4 rabbit monoclonal antibody was developed in collaboration with Epitomics (now Abcam). Briefly, the protein coding sequence corresponding to Exon 14-17 of the TCF4 gene was recombinantly expressed as an HIS tag fusion (Immunogen sequence:
GTAVWSRNGGQASSSPNYEGPLHSLQSRIEDRLERLDDAIHVLRNHAVGPSTAMPGGHGDMHGII
GPSHNGAMGGLGSGYGTGLLSANRHSLMVGTHREDGVALRGSHSLLPNQVPVPQLPVQSATSPD
LNPPQDPYRGMPPGLQGQSVSSGSSEIKSDDEGDENLQDTKSSEDKKLDDDKKDIKSITRSRSS).

Following purification of the soluble fraction, the recombinant TCF4 Ex14-17 was used to immunize 2 rabbits in parallel (3 boosts). The best responding rabbit was selected, splenectomy was performed and multi-clone hybridomas were isolated. Hybridomas were screened by ELISA against the recombinant TCF4 Ex14-17. The corresponding region of the TCF12 gene was also used for counter-screening purposes (59.1% identity in 225 residues overlap including the HIS tag). The top 10 multiclones with a TCF4 specific signal were selected and further tested by dot-blot and western blot. Based on this testing, the top 3 multiclones were selected, 3 Immuno Globulin heavy chains and 2 light chains were cloned and tested by ELISA, western blots and ChIP in all the 6 possible pairwise combination. The actual TCF4 rabbit mAb was then selected based on the results of these final tests. Further details are available upon request.

Western Blot

BPDCN cells were pelleted by centrifugation, washed once with ice-cold PBS and lysed on ice for 30 min in RIPA buffer (50 mM Tris HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with protease inhibitor cocktail (Sigma) and 1 mM PMSF. Following centrifugation at 13,000 rpm at 4°C for 5 min, supernatants were recovered and protein concentration was determined by BCA protein assay (Pierce). 20-40 μg of whole cell lysates were separated on 4-12% Bis-Tris gels and transferred to 0.45 uM nitrocellulose membranes (Invitrogen). For BRD4 blots, lysates where separated on 3-8% Tris-Acetate gels. Proteins were detected using the following antibodies: Actin-HRP (Santa Cruz Biotechnology, sc-1615), Bcl2 (Cell Signaling, #2870), Caspase-3 (Cell Signaling, #9662), Caspase-3-cleaved (Cell Signaling, #9661), Brd4 (Bethyl, A301-985A), Myc (Epitomics/Abcam, ab32972), Parp1-cleaved (Cell Signaling, # 9541), Tcf4 (Rabbit monoclonal, developed by our lab in collaboration with Epitomics/Abcam), Tlr9 (Cell Signaling, # 5845).

Apoptosis analysis

To measure the induction of apoptosis following expression of specific shRNA or after treatment with small molecule inhibitors, BPDCN cells were seeded at a concentration of $2{,}5X10^5$ cell/ml and treated as indicated. For each dose/time point, 0.5-1 MLN cells were collected, washed once with PBS, and fixed for 10' at room temperature with 1 ml of 2% para-formaldehyde (in PBS). After fixation, cells were pelleted by centrifugation, 5' at 1500 rpm, and permeabilized with 1 ml of ice-cold methanol. Methanol treated samples were kept over-night at 4 °C and moved at -20 °C for longer storage. Before staining, cells were washed twice with FACS buffer (PBS + 5% fetal bovine serum, FBS) and then incubated with PE anti Active Caspase-3 (1:20 dilution in FACS buffer; 51-68655X, BD Pharmingen) and Alexa 647 anti Cleaved Parp1 (1:1000 dilution in FACS buffer; 558710, BD Pharmingen). Following 1h incubation at 4°C, cells were washed twice with FACS buffer and analyzed with a FACScalibur BD instrument.

Cell cycle analysis

To measure cell cycle distributions following inducible expression of specific shRNA or after treatment with small molecule inhibitors, BPDCN cells were seeded at a concentration of $2.5x10^5$ cell/ml and treated as indicated. For each dose/time point, 0.5-1 MLN cells were collected, washed once with PBS, and fixed in 1 ml of 70% ice-cold ethanol. Fixed samples were stored at -20 °C for up to 2 weeks. Before staining, cells were washed twice with FACS buffer (PBS + 5% FBS) and then stained in 1 ml of PBS plus PI (propidium iodide, 50 ug/ml final) and RNAse A (500ug/ml final). Following 1h incubation at 37°C in the dark, cells were washed once with FACS buffer and analyzed with a FACScalibur BD instrument.

Immuno-histo-chemistry (IHC)

For the immuno-histo-chemistry (IHC) analysis of primary BPDCNs, reactive tonsils and cell lines were used as controls. Formalin-fixed paraffin-embedded tissue sections were stained with the following antibodies: TCF4 (Louis Staudt Lab, rabbit-monoclonal anti-human TCF4, clone 6A developed by Epitomics), Id-2 (CalBioreagents, CA; Rabbit monoclonal anti-mouse/human Id2, Clone 9-2-8), CD123 (Biolegend, CA; mouse monoclonal, clone 6H6). The TCF4 and the Id-2 staining procedure were similar. Briefly, tissue sections were de-paraffinized in xylene and rehydrated in graded-alcohol. The antigens were retrieved in a low pH "Target retrieval solution" (Dako, Carpinteria, CA) in a microwave for 6 minutes. After cooling for 15 minutes, tissue sections were blocked with serum Tris-goat for 15 minutes and incubated with primary antibody at 1:500 dilution for 60 minutes. For CD123 staining, after de-paraffinizing and rehydration, antigen was retrieved in a Low pH Dako target retrieval solution and microwaved for 20 minutes. After cooling and blocking, the sections were incubated with CD123 primary

antibody at 1:200 dilution for 60 minutes. Ventana *ultra*View Universal DAB detection kit (Ventana Medical Systems, AZ) was used for detection of all stains on an automated system (Ventana BenchMark XT). Images were captured with an Olympus BX50 microscope, Digital camera Olympus DP7, using Olympus DP Controller Version 3.3 and imported to Adobe Photoshop CS4. The TCF4 staining was scored using both distribution (0-4; negative, 0-25%, 25-50%, 50-75%, 75-100%) and intensity (0-3; negative, weak, moderate, strong).

TCF4 intra-cellular flow (ICF)

For the detection of TCF4 expression by flow cytometry, cells were 0.5-1 MLN cells were collected, washed once with PBS, and fixed for 10' at room temperature with 1 ml of 1% para-formaldehyde (in PBS). After fixation, cells were pelleted by centrifugation, 5' at 1500 rpm, and permeabilized with 1 ml of ice-cold methanol. Methanol treated samples were kept over-night at 4 °C and moved at -20 °C for longer storage. Before staining, cells were washed twice with FACS buffer (PBS + 5% fetal bovine serum, FBS) and then incubated at 4 °C for 45' with a 1:500 dilution of the TCF4 rabbit mAb (in FACS buffer) or an equivalent amount of normal rabbit Ig-G as negative control (Santa Cruz Biotechnology). After incubation with the primary antibodies, cells were washed twice with FACS buffer and then incubated at 4 °C for 45' with a 1:500 dilution of Anti-rabbit IgG Fab2 Alexa Fluor® 647 (Cell Signaling Technology, 4414S, in FACS buffer). After incubation with the secondary antibody, cells were washed twice with FACS buffer and analyzed with a FACScalibur BD instrument.

Gene expression profiling

For shRNA gene expression profiling, Cal-1 and Gen2.2 cells were infected with either Ctrl or TCF4 shRNAs. Following puromycin selection, shRNA expression was induced for 24h and 48h in Cal-1 cells, and for 12h and 24h in Gen2.2 cells. These time-points were selected to maximize TCF4 mRNA knockdown in each cell line (not-shown). Total RNA was extracted with TRIzol (Invitrogen) and subsequently cleaned-up with RNeasy mini columns (Qiagen). Gene-expression profiling was performed using two-color human Agilent 4x44K gene-expression arrays, exactly as described by the manufacturer, comparing signal from shRNA control cells (Cy3) and shTCF4 cells (Cy5). Array elements were filtered for those meeting confidence thresholds for spot size, architecture, and level above local background. These criteria are a feature of the Agilent gene-expression software package. For each cell line, a gene universe was defined by selecting arrays elements with an average Cy3 or Cy5 value > 50. Array features were defined as down-regulated or up-regulated following inducible expression of TCF4 shRNA if the Log2 Fold Change (shTCF4 vs. shCtrl) was <-0.5 or >+0.5 for at least 2 time points, respectively. When multiple spots were present for the same gene, a representative spot was selected based on the smallest Log2 FC. For the signature enrichment analysis shown in Figure 3D, a second less stringent cut-off (Log2 FC <-0.35; >+0.35) was also applied in parallel. For JQ1 gene expression profiling, Cal-1 and Gen2.2 cells were treated with either DMSO or 100 nM JQ1 for 1h, 3h, 8h and 24h. Total RNA was extracted as described above. Gene-expression profiling was performed using the two-color human Agilent 8x60K gene-expression array, comparing signal from DMSO treated cells (Cy3) drug treated ones (Cy5). Array features were defined as down-regulated or up-regulated following Drug treatment if the Log2 Fold Change (Drug vs. DMSO) was <-0.5 or >+0.5 for at least 2 time points, respectively. For the signature enrichment analysis shown in Figure 5D, a second more stringent cut-off (Log2 FC <-0.75; >+0.75) was also applied in parallel. When multiple spots for the same gene were present, a representative spot was selected based on the smallest Log2 FC. The Gen2.2 24h JQ1 array was excluded from the final data analysis because of poor data quality. All the gene expression datasets (16 arrays) have been deposited in GEO, under the accession: GSE75650.

RNA-Seq

For RNA-Seq on the BPDCN cell lines Cal-1 and Gen2.2, total RNA from fresh samples was extracted with the AllPrep DNA/RNA kit (QIAGEN) and libraries were generated with the TruSeq RNA Sample Prep Kit-v2 (Illumina) according to the manufacturer's instruction. Pair-end sequencing was performed as previously described, on a HiSeq 2000 sequencer with v3 sequencing reagents (2x101bp reads) (Xiao et al., 2013). The RNA-Seq normal pDCs from 6 donors was publicly available (SRA study SRP071837).
For RNA-Seq on the FFPE samples, RNA was extracted with the AllPrep DNA/RNA FFPE Kit (QIAGEN). Smear analysis of the FFPE RNA was performed on the Agilent Bioanalyzer with the Agilent RNA 6000 Nano Kit. Only samples whose RNA was not excessively degraded (DV200 >30%) were used to generate

FFPE RNA-Seq libraries with the TruSeq RNA Access kit (Illumina) according to the manufacturer's instruction. Pair-end sequencing was performed on a NextSeq 500 sequencer with v2 sequencing reagents. To calculate Digital Gene Expression values, the pair-end reads were aligned to hg19 genome with STAR software. The total number of reads within each gene was counted using HTSeq software. Once the total number of reads within each gene was counted, we generated for each sample the 90% trimmed mean of the read counts (average of the read counts for the middle 90% of genes in that sample). We then normalized all counts in a given sample by dividing by the samples trimmed mean and multiplying by 500. Genes with normalized counts less than 1 were set equal to one. Finally we log2 transformed the normalized counts to generate the final gene expression values. All the RNA-Seq gene expression datasets have been deposited in SRA, under the accession: GSE84471.

Signature enrichment/over-representation analysis

For each signature, the number of signature genes overlapping with a given gene set was identified and a 2-by-2 contingency table with observed values for each association was generated. P-values for signature enrichments were calculated using a Fisher's exact test. 95% confidence intervals for enrichment were estimated by calculating the enrichment scores for all possible 2x2 tables with fixed marginals, and then assuming a conditional multinomial distribution for the cell contents with probabilities equal to the cell proportions observed in the true data.

Chromatin Immuno-Precipitation (ChIP)

For the TCF4 ChIP-Seq, $5x10^7$ exponentially growing Cal-1 or Gen2.2 cells were collected by centrifugation, resuspended in room temperature RPMI without FBS ($2x10^6$ cells/ml) and cross-linked with 1% formaldehyde for 5 min at RT. For "Super-enhancer" ChIP-Seq, $1x10^8$ exponentially growing Cal-1 or Gen2.2 cells were treated for 12h with either DMSO or 250 nM JQ1, collected by centrifugation, resuspended in room temperature RPMI without FBS ($2x10^6$ cells/ml) and cross-linked with 1% formaldehyde for 5 min at RT. Cross-linking was quenched by addition of 125 mM Glycine for 5 min at RT. Cross-linked cells were first washed with ice-cold PBS and then resuspended in ice-cold RIPA buffer (10mM Tris-HCl pH8, 140 mM NaCl, 1mM EDTA pH 8, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS and 0.1% Sodium Deoxycholate) to a final concentration of $5x10^6$ cell/ml. DNA was sheared with a Misonix XL sonicator, by performing 12 x 45'' sonication cycles at power setting of 5.
For each immune precipitation reaction, $2x10^7$ chromatin cell equivalents were incubated overnight with 10 µg of Negative Ctrl antibody (normal rabbit Ig-G, Santa Cruz Biotechnology), BRD4 antibody (Bethyl, Cat No. A301-985A), total RNA Pol2 antibody (Covance, 8WG16) or with a rabbit monoclonal TCF4 antibody developed by Epitomics. The following day, chromatin/antibody complexes were incubated with 50 µl of Protein G/Protein A magnetic beads mix (Invitrogen, G to A ratio 3:1) for 4 h at 4°C. Protein G/A bound complexes were washed 4 times with RIPA Buffer, once with LiCl Buffer (10 mM Tris-HCl pH8, 250 mM LiCl, 0.5% NP40, 0.5% Sodium Deoxycholate, 1 mM EDTA), once with TE pH 8.0 and finally resuspended in 100 µl TE pH8 containing RNase A (0.2 µg/µl). Total chromatin samples were processed in parallel as input reference. Reverse cross-link was performed overnight at 65°C, followed by treatment with 20 µg Proteinase K (Invitrogen) for 2 h at 50°C. Final DNA purification was performed with QIAquick PCR Purification columns (QIAGEN). ChIP DNA was used to generate ChIP-Seq libraries with the NEXTflex™ Illumina ChIP-Seq Library Prep Kit (Bioo Scientific), according to manufacturer's instructions. The sequencing of the TCF4 ChIP-Seqs was performed on a GA2x (Illumina) sequencer, with single reads (SR) of 36 bp length. The sequencing of all the other ChIP-Seqs was performed on the High-Output flow-cell of a NextSeq 500 sequencer (Illumina), with single reads (SR) of 76 bp length. All the ChIP-Seq data have been deposited in GEO under the accession: GSE76147.

ChIP-Seq Analysis

For TCF4 ChIP-Seq analysis, 36 bp sequence tags were aligned to human genome build 36 (hg18) with Bowtie software. Redundant reads were removed and reads uniquely mapping to reference genome were used for further analysis. A maximum of two mismatches was allowed for each read. To perform "peak calling", the genome was divided into "bins" of 25 bp. Each sequencing tag was associated with the bin of its start site and an additional extension of 7 bins (for a total of 200 bp) along the direction of its read to match the average library size. For a given experiment, the number of "hits" in a bin was defined as the number of extended tags associated with that bin. Bins with fewer than 3 hits were removed. Peaks were defined by merging consecutive significant bins. The height of the peak was defined as the largest number

of hits in any of the bins forming the peak. The "apex" of the peak was defined as the bin that had this largest number of hits. In the case of ties, the earliest such peak was chosen as the apex. A threshold of apex >30 tags/bin was used to define high-stringency TCF4 binding peaks. A normal rabbit Ig-G immunoprecipitation was included as specificity control for each cell line. Experimental peaks overlapping with a control bin with a tag count of 5 or greater were excluded from further analysis. For BRD4 and RNA Pol2 ChIP-Seq analyses, 76 bp sequence tags were aligned to human genome build 36 (hg18) with Bowtie software. Redundant reads were removed and reads uniquely mapping to reference genome were used for further analysis. For each gene, a representative RefSeq was defined based on the RNA Pol2 ChIP-Seq data. In the case of multiple RefSeq deriving from alternative promoter usage, the RefSeq with highest RNA Pol2 promoter density (-+ 500 bp from TSS) was selected. In the case of multiple RefSeq sharing the same promoter but differing for exon composition, the RefSeq with the overall highest RNA Pol2 density was selected. Promoter peaks were defined as peaks whose apex was located within a -+ 2 Kb window from a representative TSS. Upstream binding peaks were defined as peaks whose apex was located within a -15 Kb to -2 Kb window from the TSS. Gene body binding peaks were defined as peaks whose apex was located in a window starting at + 2Kb from the TSS and ending with the end of the RefSeq. When multiple peaks were associated with the same representative RefSeq, the peak with the highest apex was used for further analyses. To compare the results of two experiments, ChIP-Seq data were normalized by total tags count. To exclude the background signal from the normalization factor, only tags mapping within 25 bp bins with a density of at least 5 tags/bin were used for normalization.

Super-enhancer analysis

For each cell line, the BRD4 ChIP-Seq of DMSO treated cells was used to define super-enhancers. The mapping of super-enhancer was performed essentially as previously described (Loven et al., 2013; Whyte et al., 2013). Briefly, we first defined BRD4 peaks and selected only high stringency peaks (apex >18 tags/bin) to perform further analyses. Next, we merged all the BRD4 peaks found within 15 Kb from each other and then ranked all the BRD4 bound regions by increasing total BRD4 occupancy. The data were finally scaled by normalizing to the largest value, so that both the x and y axis were ranging from 0 to 1. These plot revealed an obvious point at which the BRD4 loading began to increase rapidly, and enabled us to define super-enhancers in both BPDCN cell lines. Geometrically, this point was defined as the BRD4 loading point (x-axis) for which a tangent to the curve with slope 1 could be found. All the defined genomic regions above this point were defined as super-enhancers (SE). All the genomic regions below that point that did not map in the vicinity of an annotated promoter (−2/+2 kb from RefSeq TSSs) were conversely defined as traditional enhancers (TE).

To associate a representative RefSeq to a SE, both mapping and functional criteria were applied. In the case of a SE region directly overlapping with an annotated protein-coding gene RefSeq, the corresponding RefSeq was automatically associated to that SE region. For intergenic SE, either the upstream or downstream representative RefSeq was selected based on the highest depletion of elongating RNA Pol2 observed after JQ1 treatment.

ATAC-Seq

ATAC-Seq was performed essentially as previously described (Buernostro et al.) with few minor modifications. Briefly, cells were harvested and 50,000 cells were pelleted by centrifugation for 10' at 500g. Media was removed and pelleted cells were resuspended directly in room-temperature lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2 and 0.5% IGEPAL CA-630). Following a second centrifugation (10', 500g), isolated nuclei were tagmented in 50 ul transposase reaction mix (25 µl 2× TD buffer, 10 µl transposase (NextEra DNA Library Prep Kit, Illumina) and 15 µl nuclease-free water). The transposition reaction incubated for 30 min in a 37 °C water bath, and then purified with the Qiagen MinElute kit. Following purification, the tagmented DNA was amplified as described in the original protocol. Depending on the sample, 8-10 total PCR cycles were necessary. After PCR amplification, the ATAC-Seq libraries were purified with the Qiagen MinElute kit and further cleaned-up with Agencourt AMPure XP beads (Beckman Coulter) at a sample to beads ratio of 1 :1.75. ATAC-Seq. ATAC-Seq libraries were sequenced on the High-Output Next-Seq flow cell (Illumina, 150 cycles, paired end 2 x 75 bp). All the ATAC-Seq data have been deposited in GEO, under the accession: GSE84623.

ATAC-Seq analysis

To generate ATAC-Seq genome browser tracks, the reads were aligned to the human genome hg18 by Bowtie with default parameters. An in-house peak calling script was used to scan the read peaks on the alignment file with following criteria: maximum read mismatch <=4, minimum read depth >=3, bin size 25 bp, read extension 200 bp. Statistically significant peaks were selected based on a p value <= 10^{-6} (Poisson distribution probability) and saved in the standard WIG format for visualization in the genome browser. To perform the ATAC-Seq cluster analysis presented in Figure 8E, reads were aligned to human genome hg18 by Bowtie software with mismatches up to 2 nucleotides, and extended 100 bp in each direction from the 5' start. The frequencies of aligned reads in every 100-bp-bins of chromosome regions were further summarized by an in-house script. Only the bins with the depth greater than 15 reads were used for further analysis. For the actual clustering of ATAC-Seq samples, the height of bin i in experiment j, was defined as the total number of hits in bin i plus the number of tags for the bins on either side (that is the total number of tags in bins i -1, i, and i +1). This was done so as to avoid dividing a narrow peak that may have fallen on the borderline of two bins. Bins with heights less than 15 were set to 0. For each bin the sum of the heights of the 4 cell line experiments was calculated. Peaks were defined as regions consisting of one or more consecutive bins for which this total sum was greater than 200. This resulted in 82,873 peaks made up of 314,721 bins. To test whether a given peak K was higher in one pair of cell line experiments {A, B} versus a second pair {C, D} the following ratio was calculated:

$$R_K = \frac{\max_{i \in K}\left(\min\left(A_i, B_i\right)\right)}{\max\left(\max_{i \in K}\left(C_i\right), \max_{i \in K}\left(D_i\right), 14\right)}$$

where {Ai, Bi, Ci, Di} represent that heights of experiments {A, B, C, D}, in bin i where i runs through all the bins in peak K. Effectively this requires there to be a bin in the peak for which the heights in both A and B, was significantly higher that the height achieved by either of the experiments C and D at any point in the peak, or the height of 14, if neither C nor D achieved a height of at least 15.

Peaks with ratios greater than 15, were deemed to be differentially expressed. Of the 82,873, peaks, 497 met this threshold for being higher in both Cal-1 and Gen2.2 than in either MOLM-14 or HL-60, and 129 met this threshold in the reverse direction. For clustering purposes, we associated combination of experiment and differential peak, the greatest bin height achieved by that experiment over all bins in the peak (or 14 if no bin had a value of 15 or higher). These values were then log transformed, and the values of each peak centered at the mean of values of the 4 cell lines samples. They were then hierarchically clustered with average linkage based on an un-centered correlation metric.

High-throughput drug toxicity screen

The high-throughput single agent drug screen was performed essentially as previously described (Mathews Griner et al., 2014). Briefly, for each cell line tested, a total of 500 cells per well in 5 µl of media were dispensed in 1,536 wells tissue culture plates. For the generation of standard 11-point dose–response curves, the cells were plated, followed by the immediate pin-tool addition of 23 nl of control compound (bortezomib) and library compounds using a Kalypsys pin-tool. The plates were then covered with stainless steel cell culture Kalypsys lids and incubated at 37 °C with 5% CO_2 under 95% humidity. For cell proliferation assays, the cells were incubated for 48h and then 3 µl of CellTiter Glo luminescent cell viability assay reagent (Promega) were added to each well using a Bioraptor Flying Reagent Dispenser (Aurora Discovery-BD). The plates were then incubated for 15 min at room temperature. The signal was measured using a 10-s exposure with a ViewLux (Perkin-Elmer) with a luminescent filter. Relative luminescence units (RLU) for each well were normalized to the median RLUs from the DMSO control wells as 100% viability, and median RLUs from the bortezomib control wells as 0% viability. All the drug screening data have been deposited in PubChem with the following assay ID: Cal-1 cells: AID 1224825; Gen2.2 cells: AID 1224824.

Xenograft experiments

All animal experiments were carried out in accordance with the National Cancer Institute Animal Care and Use Committee (NCI ACUC) guidelines. NOD/SCID mice were inoculated by subcutaneous (s.c.) injection with 1x 10^6 cells in a suspension containing Matrigel (BD Matrigel Basement Membrane Matrix)/PBS at a ratio of 1:1. Mice were monitored daily for palpable tumor mass. When tumors reached an average 100 mm3 (approximately 7 days after tumor inoculation), mice were treated by intraperitoneal injection with

either vehicle control or CPI 203 (5 mpk BID IP). Tumor growth was monitored daily by measuring tumor size in two orthogonal dimensions by caliper measurement.

Primary cells and primary BPDCN sample

For the isolation of normal pDCs and conventional myeloid dendritic cells (cDCs) from healthy donors, an apheresis of $2x10^9$ PBMCs was collected from each donor. Following gradient centrifugation and removal of red blood cells (LymphoprepTM, StemCell Technologies), pan-dendritic cells were enriched by depletion of non-dendritic cells with the human Pan Dendritic Cells Enrichment Kit (Miltenyi Biotec). Following the removal of non-dendritic cells, conventional BDCA-1 expressing myeloid dendritic cells were positive selected with the human Blood Dendritic Cell Isolation Kit II (Miltenyi Biotec). The myeloid dendritic cell depleted flow-through was further processed to positively select BDCA-4 expressing pDCs with the human Plasmacytoid Dendritic Cell Isolation Kit II (Miltenyi Biotec).
For the processing of the viably frozen PBMCs from the BPDCN Patient #1, cells were thawed, stained with CD123, CD56, BDCA-2 and BDCA-4 and analyzed by FACS. BDCA-4 expressing cells were then positively selected with the human Plasmacytoid Dendritic Cell Isolation Kit II (Miltenyi Biotec).

This study was approved by the Institutional Review Board of the National Cancer Institute, Protocol 10-CN-074 C.

SUPPLEMENTAL REFERENCES

Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., Rahl, P. B., Lee, T. I., and Young, R. A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307-319.
Xiao, W., Tran, B., Staudt, L. M., and Schmitz, R. (2013). High-throughput RNA sequencing in B-cell lymphomas. Methods Mol Biol *971*, 295-312.