Blastic plasmacytoid dendritic cell neoplasm: genomics mark epigenetic dysregulation as a primary therapeutic target

Maria Rosaria Sapienza,[1*] Francesco Abate,[2,3*] Federica Melle,[4] Stefania Orecchioni,[5] Fabio Fuligni,[6] Maryam Etebari,[1] Valentina Tabanelli,[4] Maria Antonella Laginestra,[1] Alessandro Pileri,[7,8] Giovanna Motta,[4] Maura Rossi,[1] Claudio Agostinelli,[1] Elena Sabattini,[1] Nicola Pimpinelli,[8] Mauro Truni,[9] Brunangelo Falini,[10] Lorenzo Cerroni,[11] Giovanna Talarico,[5] Rossana Piccioni,[12] Stefano Amente,[13] Valentina Indio,[14] Giuseppe Tarantino,[14] Francesco Brundu,[2] Marco Paulli,[15] Emilio Berti,[16] Fabio Facchetti,[17] Gaetano Ivan Dellino,[12,18] Francesco Bertolini,[5] Claudio Tripodo,[19*] Raul Rabadan[2,3*] and Stefano A. Pileri[4†*]

[1]Hematopathology Unit, Department of Experimental, Diagnostic, and Specialty Medicine, S. Orsola-Malpighi Hospital, University of Bologna, Italy; [2]Department of Systems Biology, Columbia University College of Physicians and Surgeons, New York, NY, USA; [3]Department of Biomedical Informatics, Columbia University College of Physicians and Surgeons, New York, NY, USA; [4]Division of Haematopathology, European Institute of Oncology, Milan, Italy; [5]Laboratory of Hematology-Oncology, European Institute of Oncology, Milan, Italy; [6]Department of Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON, Canada: [7]Dermatology Unit, Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Italy; [8]Division of Dermatology, Department of Surgery and Translational Medicine, University of Florence, Italy; [9]Pathological Anatomy Histology & Cytogenetics, Niguarda Cancer Center, Niguarda-Ca' Granda Hospital, Milan, Italy; [10]Institute of Hematology and Center for Hemato-Oncology Research (CREO), University and Hospital of Perugia, Italy; [11]Universitätsklinik für Dermatologie und Venerologie, LKH-Universitatsklinikum Graz, Austria; [12]Department of Experimental Oncology, European Institute of Oncology, Milan, Italy; [13]Department of Molecular Medicine and Medical Biotechnologies, University of Naples 'Federico II', Italy; [14]"Giorgio Prodi" Cancer Research Center, University of Bologna, Italy; [15]Unit of Anatomic Pathology, Department of Molecular Medicine, University of Pavia and Fondazione IRCCS San Matteo Policlinic, Pavia, Italy; [16]Department of Dermatology, Fondazione IRCCS Ca' Granda - Ospedale Maggiore Policlinic and Milan University, Milan, Italy; [17]Pathology Section, Department of Molecular and Translational Medicine, University of Brescia, Italy; [18]Department of Oncology and Hemato-Oncology, University of Milan, Italy and [19]Tumor Immunology Unit, Department of Health Science, Human Pathology Section, University of Palermo School of Medicine, Italy

**MRS, FA, CT, RR and SAP contributed equally to this work. †Alma Mater Professor, Bologna University*

 doi:10.3324/haematol.2018.202093

Received: July 16, 2018.

Accepted: October 30, 2018.

Pre-published: October 31, 2018.

Correspondence: *MARIA ROSARIA SAPIENZA*

mariarosaria.sapienza@gmail.com

Supplementary Appendix

Immunohistochemistry and cytogenetics

Fourteen BPDCN cases were investigated and their clinical characteristics are reported in Table S1. All the skin biopsies at diagnosis were reviewed by a panel of at least three expert hematopathologists (CA, EB, FF, LC, MP, ES, CT, MT, and SAP) according to the WHO Classification criteria. [1]

Immunohistochemistry (Dako Denmark) used the following antibodies: CD4,CD56, TdT (Novacastra) CD123 (BD Biosciences PharMingen), CD303/BDCA2 (Dendritics), TCL1 (Cell Marque), CD68PGM1, MPO, CD34, CD3, CD30 (Dako Denmark). If necessary, additional antibodies were evaluated accordingly to the specific requests of the single case (not reported). For the cytogenetic investigation of the BPDCN patients we used the whole exome sequencing data in order to map at highest resolution the chromosome 9 at cytoband p21 and evaluate the presence of the CDKN2A gene deletion in homo/heterozigosity; the most relevant alteration recognized by the WHO in the BPDCN context.

Immunohistochemistry and cytogentic characteristics are summarized in Table S2.

DNA sample extraction

We used MagCore Genomic DNA Tissue Kit (RBC Bioscience Corp, Taiwan) for DNA extraction from cryopreserved tumoral skin biopsies, Oragene DNA kit (DNA Genotek Inc., ON, Canada) for DNA from saliva samples and MagCore Cultured Cells DNAKit (RBC Bioscience Corp, Taiwan) for DNA extraction from the CAL-1 cell line. All the samples were then loaded on the semi-automatic MagCore nucleic acid extractor (RBC Bioscience Corp, Taiwan). DNA quantity was evaluated by the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen Life technologies, UK) and all the samples passed the quality check.

Informed consent was obtained from each patient in accordance with the guidelines of the Institutional Review Board of the Department of Experimental, Diagnostic, and Specialty Medicine of the University of Bologna and the Declaration of Helsinki.

Sanger Sequencing

We validated by conventional Sanger sequencing two candidate nonsense somatic mutations of SUZ12 and ASXL1 occurring in the patients BPDCN_38 and BPDCN_39, respectively, as reported by WES analysis (Figure S1).We sequenced both tumor and normal DNA. The following PCR primers were custom-designed using Primer3 on line software (http://bioinfo.ut.ee/primer3-0.4.0/primer3/):

ASXL1-Forward-GGACTCACAGATGGGCTAGG;

ASXL1-Reverse-AGAATGGGACCATTGTCTGC;

SUZ12-Forward TCATGCCTGTATGCTGTTTG;

SUZ12-Reverse- GAAGCAGATTCCCCCTTTTC.

The PCR products were sequenced both forward and reverse with ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit (Version 3) and loaded on ABI PRISM 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), according to manufacturer's instructions. Tumor sequences were compared with the corresponding germline sequences using FinchTV version 1.4 software (Geospiza Inc., Seattle City, WA, USA).

Libraries Preparation

Whole Exome Sequencing (WES) libraries

We performed paired-end sequencing of matched tumor/normal DNA samples (9 cases), tumor only DNA samples (5 cases), and the CAL-1 cell line (Table S3).Whole exome sequencing libraries were prepared by using the TruSeq Exome Kit (Illumina, San Diego, CA, USA) and Nextera Rapid Capture Exome kit (Illumina, San Diego, CA, USA). According to the kit instructions the genomic DNA of each patient was fragmented to provide DNA fragments with a base pair peak at 350 bp, ligated at the both ends with specific adapters and then purified by Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA). The DNA was then amplified by ligation-mediated PCR, purified, and hybridized. Hybridized fragments bounded to strepavidin beads whereas non-hybridized fragments were washed out. Captured DNA library amplification products were assessed for quality by Agilent 7500 DNA assay (Agilent, Santa Clara, CA, USA) and quantified by Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen Life technologies, UK), according to the manufacturer's protocol. Each captured library was subjected to cluster generation on cBot instrument (Illumina, Inc., San Diego, CA, USA) and finally paired-end sequencing was performed on the Illumina HiScan SQ platform (Illumina, San Diego, CA, USA) to generate 100 bp paired-end reads (2 × 100PE). All the libraries passed the quality check.

Targeted sequencing libraries

We performed targeted sequencing of the 14 BPDCN tumor samples, 7 normal matched saliva samples, and the CAL-1 cell line (Supplemental Material, Table S4).

To validate the WES results we used the MiSeq TruSeq Custom Amplicon (Illumina, Inc., San Diego, CA, USA) a highly multiplexed targeted sequencing assay planned with DesignStudio, an online software, available at Illumina website. We developed a custom amplicon panel to specifically interrogate 9 genes (Table S5). We used the TruSeq Custom Amplicon Kit with 250 ng of DNA per sample and the amplicon libraries were loaded on MiSeq instrument (Illumina, Inc., San Diego, CA, USA) to generate 2 × 151-bp paired reads, according to the manufacturer's instructions. All libraries passed the quality check and the MiSeq targeted sequencing approach allowed us to reach a median coverage depth of 647X, a mean coverage depth of 1600X (ranging from 98X to 4056X).

RNA sequencing libraries

We performed RNA sequencing of 9 BPDCN samples – 5 cases belonging to the discovery set and 4 cases belonging to the extended set (Table S6). Total RNA from the nine biopsies was extracted using TRIzol reagent (Invitrogen). Paired-end libraries (2x75 base pair) were prepared according to the TruSeq RNA sample preparation v2 protocol (Illumina, San Diego, USA). Briefly, 2 µg of Poly(A)+ RNA were purified from total RNA using poly-T oligo attached magnetic beads and then used for fragmentation into 130–290 bp fragments. First strand of cDNA synthesis was performed using reverse transcriptase enzyme (SuperScript II, Invitrogen, Life Technologies, USA) and random hexamer primer, followed by generation of double-stranded cDNA. AmpureXP beads (Beckman Coulter, Brea CA) were used to purify the ds cDNA and an End Repair step was

performed to convert the overhangs, resulting from fragmentation, into blunt ends by 3' to 5' exonuclease activity. A single "A" nucleotide was added to the 3'ends of the blunt fragments to prevent them from ligating to one another during the adapter ligation reaction. This approach was adopted to ensure a low rate of chimera (concatenated template) formation. Subsequently, sequencing adapters were added to the ends of the ds cDNA fragment and a PCR reaction was used to selectively enrich those ds cDNA fragments that had adapter molecules on both ends, amplifying the amount of ds cDNA in the final libraries. Lastly, PCR library products were purified by AmpureXP beads and quality control analysis was assessed using a DNA-1000 (Agilent, USA). The quantification was performed by the Quant-it PicoGreen dsDNA Assay Kit per manufacturer's protocol (Invitrogen, Life Technologies, USA). The resulting libraries were sequenced on an Illumina HiScan SQ (Illumina, San Diego, USA) following the manufacturer's instructions.

Pathology tissue-chromatin immunoprecipitation (PAT-ChIP) sequencing libraries

PAT-ChIP experiments were performed as in Fanelli *et al*[2] with the following modification: sonication for chromatin extraction was performed in 400 ml. Antibodies used were: anti-Histone H3acetylK27 (ab472) and anti-trimethyl-Histone H3K27 (07-449). Immunoprecipitated DNA was purified with QIAGEN columns and, after library preparation, sequenced with a HiSeq2000 in multiplexed run to obtain 50 bp single-end reads following manufacturer protocols.

Bioinformatic Analysis of Sequencing Data

Whole exome sequencing analysis

Illumina HiScanSQ analysis produced an average of 70 million paired-end reads per sample. Average coverage breadth, defined as the percentage of the captured coding sequence of a haploid reference covered by reads, was 98% (92% at 20x, 72% at 50x). We computed breadth of coverage using cnvkit (version 0.9.0)[3], pysam (version 0.12.0.1) and samtools (version 1.6).[4] Paired-end reads were mapped to the hg19 reference genome using the Burrows-Wheeler Aligner (BWA version 0.5.9)[5] alignment tool. We detected variants as sites that differed from the reference in each sample independently. To assess statistical significance of variant calling we used the SAVI algorithm (Statistical Algorithm for Variant Identification) developed at Columbia University.[6] Briefly, SAVI constructs empirical priors for the distribution of variant frequencies for each sample. High-credibility intervals (posterior probability ≥ 1–10−5) are constructed for the corresponding change in frequency between tumor and normal samples. A discrete set of frequencies will be the base for constructing prior and posterior distribution and posterior probability is connected to the prior by a modified binomial likelihood. Variant with total depth in tumor and normal lower than 10x were filtered. Then, we first selected somatic variant with frequency greater than 10% in tumor samples and less than 3% in normal samples. In samples without a matched normal control we selected variants with depth greater than 20x and frequency greater than 25%. To remove systematic errors, we created an internal

database with all the variants present in normal samples, and excluded all variants that were found to be present in any of the normal samples. WES results were validate by MiSeq (Table S7).

Functional Enrichment Analysis

To perform the functional mutation enrichment analysis, we selected only genes affected by deleterious SNVs (nonsense or frameshift) and/ or recurrently mutated in ≥ 3 samples, for a total of 54 genes. We analyzed the selected gene list by WebGestalt toolkit[7], using the Overrepresentation Enrichment Analysis (ORA) method and the Gene Ontology/Biological Process functional database. In Table S8 the top 10 most significant biological processes emerged from the analysis.

WES analysis recognized several deleterious mutations of *ASXL1* and *TET2* genes (Table S9) . Overall, 25 epigenetic modifier genes were found mutated (Table S10).

Copy Number Analysis

We obtained copy number variation calls using cnvkit (version 0.9.0)[3]. Where possible, default parameters were used. We used GRCh37.75 human genome as reference. Starting from this reference, we built 10 kbases anti-target regions (as required by cnvkit). Then, we computed copy ratio, splitting the normal samples with respect to their exome sequencing preparation kit. The two normal reference copy number profiles were then used to estimate the copy ratio for the tumor samples (Figure S2 and S3). Finally, the copy ratios for each sample were discretized into absolute copy number calls using the "call" command of the cnvkit suite. Tumor purity (100% for the cell line, and 90% for the other samples) was the single

parameter of the "call" step. For each sample and computation, we used the appropriate bedfile provided by each exome sequencing preparation kit: Nextera rapid capture exome kit (v1.2) and TruSeq Exome Kit (2012).

Targeted sequencing analysis

Reads were aligned to the UCSC hg19 reference genome using BWA-MEM.[5] Aligned reads were analyzed using the SAVI algorithm[6], and variants were selected based on coverage depth and frequency. Specifically, SAVI constructed empirical priors for the distribution of variant frequencies in each sample, from which we obtained a corresponding high-credibility interval for the frequency of a particular allele. To obtain estimates for alleles with frequencies as low as 0.5%, we chose logarithmically spaced precision for the priors and posteriors. Furthermore, we considered variants detected in the normal samples and absent in the tumor as false positive calls and determined that alleles with lower bound interval of posterior probability less than 0.5 produced a false discovery rate < 3%.

RNA sequencing analysis

We mapped 9 BPDCN samples – 5 cases belonging to the discovery set and 4 cases belonging to the extended set- by means of STAR aligner (version 2.4.0)[8] on human reference genome hg19 and obtained an average of 70 million paired-end mapped reads per sample. Differential expression analysis and mRNA quantification was performed by means of DeSeq.[9]

Gene set enrichment analysis was separately performed on the discovery, extended set and CAL-1 cell line by means of GSEA software and Molecular Signature Database (MSigDB)[10] on the previously ranked gene list based on regularized base-2 DeSeq logarithm transformation (Figure S4 and S5).

Integration of RNA and PAT-ChIP sequencing

We recognized 86 genes up-regulated and also marked by H3K27acetylation. The Gene ontology Analysis was conducted on the 86 genes by WebGestalt toolkit,[7] using the Overrepresentation Enrichment Analysis (ORA) method and the Gene Ontology/Biological Process functional database. In Table S11 the top 10 most enriched biological processes emerged from the analysis.

Pathology tissue-chromatin immunoprecipitation (PAT-ChIP) sequencing analysis

PAT-Chip FASTQ files were quality checked and filtered with NGS QC Toolkit (versione 2.3.3) using default parameters. Alignments were performed with Burrows-Wheeler Aligner (version 0.7.10)[5] to hg18 using default parameters. SAMtools (version 1.2) and BEDtools (version 2.24) were used for filtering steps and file formats conversion. Duplicate reads were discarded and peaks were identified from uniquely mapping reads, using MACS (version 2.1.0)[11] callpeak with default parameters and -broad, -SPMR, -shiftsize 73 options. The q-value cutoff used to call significant regions was 0.05. UCSC tools and genome browser was used for data visualization.

ChIP-seq signals of peaks called by MACS were subjected to unbiased clustering, using the seqMINER 1.3.2 platform.[12] Linear Kmeans was used for clustering, with the following parameters: left and right extension = 5 kb, internal bins (with respect to the peaks) = 160, number of cluster = 25. seqMINER was also used to generate the heatmaps and the average profiles of read density for the different clusters.

Mouse Model

Experiments were carried out on nonobese diabetic severe combined immunodeficient NOD/SCID interleukin-2 receptor g (IL-2Rg)–null (NSG) mice, 6 to 8 weeks old. NSG mice were bred and housed under pathogen-free conditions in the animal facilities at the European Institute of Oncology–Italian Foundation for Cancer Research (FIRC) Institute of Molecular Oncology (IEO-IFOM, Milan, Italy) as previously reported.[13] All animal experiments were carried out in accordance with the applicable Italian laws (D.L.vo 26/14 and following amendments) and the institutional guidelines. All *in vivo* studies were ratified by the Italian Ministry of Health. For induction of BPDCN in mice, 5.000 CAL-1 cells were injected intravenously (i.v.) through the lateral tail vein in non-irradiated mice. Human engraftment was defined by means of percentage of human cells in peripheral blood from tail vein of the recipient animals.

Human cell engraftment in the peripheral blood, bone marrow of the femur and of the spine, spleen and liver was investigated by flow cytometry and immunohistochemistry from 39 days after transplant onward in a representative mouse vehicle-treated. The phenotype of human cells in NSG mice was evaluated by flow cytometry using the following anti-human antibodies: anti-CD38-APC (clone LS198-4-3), -CD45-APC-Cy7 (clone J33), -CD56-PE (clone N901) from Beckman-Coulter and anti-mouse CD45-FITC (clone 30-F11) from Becton Dickinson (BD) to exclude murine cell contamination from the analysis. After red cell lysis, cell suspensions were evaluated by a FACSCalibur (BD) using analysis gates designed to

exclude dead cells, platelets and debris. Percentages of stained cells were determined and compared to appropriate negative controls. Seven-aminoactinomycin D (7AAD; Sigma-Aldrich) was used to enumerate viable, apoptotic and dead cells. Hematoxylin and eosin (H&E) staining was performed on the bone marrow and spleen of transplanted mice vehicle-treated after 39 days. Immunohistochemistry was performed on samples obtained from the spine of transplanted mice vehicle-treated after 39 days by using an anti-BDCA2/CD303 antibody (Dendritics, Lyon, France; Clone 124B3; dilution: 1:20).

In vivo treatments

Bortezomib, 5'-Azacytidine, Decitabine provided by Sigma-Aldrich (Sigma-Aldrich Corporation, St. Louis, MO, USA) and Romidepsin provided by Santa Cruz (Santa Cruz Biotechnology, Santa Cruz, CA), were dissolved in saline (0.9% w/v NaCl) and injected intraperitoneally into the mice: Bortezomib was administrated at 0.5 mg/kg two times weekly for 4 weeks, 5'-Azacytidine 5 mg/kg 5 doses (2-day intervals), Decitabine 2.5 mg/kg 3 doses (2-day intervals) and Romidepsin 0.5 mg/kg every day for 4 weeks. Drug dosages were previously defined as non-toxic in mice not injected with CAL-1. Administration started one day after CAL-1 cells injection. Mice were monitored for survival daily until reaching humane end-points. The log-rank test was used to compare survival between different groups. All experiments were carried out in duplicate, a total of 110 animals having been treated.

Table S1. Patients clinical characteristics

Sample	Sex	Age (Y)	Tissue	Other sites	F. Up	(mo)	Main Therapy	Transplant
BPDCN_40	F	9	Skin	BM	DOD	71.5	AIEOP AML 2002/HR	Auto-SCT
BPDCN_39	M	89	Skin	NA	LOST	NA	NA	NA
BPDCN_23	M	19	Skin	//	DOD	76	CHOP + MTX	Allo-SCT
BPDCN_43	M	67	Skin	BM, LN	DOD	6.3	GIFOX	NO
BPDCN_25	M	62	Skin	BM, PB, LN	ADF	21	Hyper-CVAD	NO
BPDCN_45	F	66	Skin	LN	DOD	36	Hyper-CVAD	Auto-SCT
BPDCN_46	M	29	Skin	PH	ADF	57	Hyper-CVAD	Auto-SCT
BPDCN_41	M	60	Skin	//	DOD	41.3	ICE	Auto + Allo-SCT
BPDCN_47	M	78	Skin	//	ADF	74	Local RT	NO
BPDCN_49	F	49	Skin	//	DOD	28	Local RT	NO
BPDCN_42	F	73	Skin	BM, PB, LN, PL	DOD	7.6	MICE	NO
BPDCN_37	M	75	Skin	NA	DOD	9	NA	NA
BPDCN_38	M	69	Skin	BM	DOD	6.4	NILG AML 02/06	Allo-SCT
BPDCN_50	M	37	Skin	LN	LOST	NA	NA	NA

Abbreviations: BPDCN, blastic plasmacytoid dendritic cell neoplasm; y, years; mo, months; M, male; F, female; BM, bone marrow; LN, lymph node, PB, peripheral blood; PH, pharynx; PL, pleura; DOD, died of disease; AWD, alive with disease; ADF, alive disease free; LOST, lost at follow-up; NA, not available. AIEOP AML 2002/HR, Associazione Italiana di Ematologia e Oncologia Pediatrica acute myeloid leukemia high-risk children 2002/01 trial; CHOP, cyclophosphamide, doxorubicin, vincristine, prednisone; CHOP+MTX, cyclical chemotherapy with high-dose methotrexate and CHOP; GIFOX, gemcitabine, ifosfamide, and oxaliplatin; Hyper-CVAD, alternate cycles of hyper-fractionated cyclophosphamide, vincristine, doxorubicin, dexamethasone, and methotrexate and cytarabine; ICE, idarubicin, cytarabine, etoposide; RT, radiotherapy; MICE, mitoxantrone, cytarabine, etoposide; NILG AML 02/06, Northern Italy Leukemia Group acute myeloid leukemia 02/06 trial; Auto-SCT, autologous stem cell transplant; Allo-SCT, allogenic stem cell transplant.

Table S2 . Patients immunohistochemical and cytogenetic characteristics

Case No	CD3	CD4	CD30	CD34	CD56	CD68PGM1	CD123	CD303/BDCA2	TCL1	MPO	TdT	DEL_CDKN2A Y/N
BPDCN_23	(-)	(+)	(-)	(-)	(+)	(+ -)	(+)	(+)	(+)	(-)	(-)	Y homozigosity
BPDCN_25	na	(+)	na	(-)	(+)	(+ -)	(+)	(+)	(+)	(-)	(-)	Y heterozigosity
BPDCN_37	(-)	(+)	na	na	(+)	(- +)	(+)	na	na	(-)	(-)	N
BPDCN_38	(-)	(+)	na	(-)	(+ -)	(- +)	(+)	(+)	(+)	(-)	na	Y heterozigosity
BPDCN_39	(-)	(+ -)	(-)	(-)	(+)	(- +)	(+)	(+)	na	na	(-)	Y heterozigosity
BPDCN_40	na	(- +)	na	(-)	(+)	na	(+)	(+)	(- +)	na	(+ -)	N
BPDCN_41	(-)	(+)	(-)	(- +)	(+ -)	(- +)	(+)	(+)	(+ -)	(-)	(-)	N
BPDCN_42	na	(+)	na	na	(+)	na	(-)	(+)	na	na	na	Y homozigosity
BPDCN_43	(- +)	(-)	(-)	(-)	(+)	na	(+)	(+)	na	(-)	(-)	N
BPDCN_45	(-)	(- +)	na	(-)	(+ -)	(- +)	(+ -)	(+)	(+)	(-)	(-)	N
BPDCN_46	(-)	(+)	(-)	(- +)	(+)	(- +)	(+)	na	(-)	na	na	Y heterozigosity
BPDCN_47	(-)	(+ -)	(-)	(-)	(+)	(+ -)	(+)	(+)	(+ -)	(-)	(-)	N
BPDCN_49	(-)	(+)	(-)	(-)	(+)	(-/+)	(+)	na	na	(-)	(-)	Y heterozigosity
BPDCN_50	(-)	(+)	na	(-)	(+)	(-/+)	(+)	(-)	na	(-)	(-)	Y homozigosity

Abbreviations: BPDCN, blastic plasmacytoid dendritic cell neoplasm; (+), positive > 75% cells ; (+ -), positive 50-75% cells; (- +), positive 25-50%; (-), negative, no cell; DEL, deletion.

Table S3. BPDCN samples sequenced by Whole-Exome Sequencing.

Sample	Wes	Tumor cells %
BPDCN_37	Matched	≥90%
BPDCN_38	Matched	≥90%
BPDCN_39	Matched	≥90%
BPDCN_40	Matched	≥90%
BPDCN_43	Matched	≥90%
BPDCN_45	Matched	≥90%
BPDCN_46	Matched	≥90%
BPDCN_47	Matched	≥90%
BPDCN_49	Matched	≥90%
BPDCN_23	Unmatched	≥90%
BPDCN_25	Unmatched	≥90%
BPDCN_41	Unmatched	≥90%
BPDCN_42	Unmatched	≥90%
BPDCN_50	Unmatched	≥90%
CAL-1	Unmatched	100%

Abbreviations: BPDCN, blastic plasmacytoid dendritic cell neoplasm; CAL-1, blastic plasmacytoid dendritic cell line

Table S4. BPDCN samples sequenced by MiSeq targeted sequencing.

Sample	MiSeq	Tumor cells %
BPDCN_37	matched	≥90%
BPDCN_39	matched	≥90%
BPDCN_40	matched	≥90%
BPDCN_43	matched	≥90%
BPDCN_45	matched	≥90%
BPDCN_46	matched	≥90%
BPDCN_49	matched	≥90%
BPDCN_23	unmatched	≥90%
BPDCN_25	unmatched	≥90%
BPDCN_38	unmatched	≥90%
BPDCN_41	unmatched	≥90%
BPDCN_42	unmatched	≥90%
BPDCN_47	unmatched	≥90%
BPDCN_50	unmatched	≥90%
CAL-1	unmatched	100%

Abbreviations: BPDCN, blastic plasmacytoid dendritic cell neoplasm; CAL-1, blastic plasmacytoid dendritic cell line

Table S5. The genes investigated by MiSeq targeted sequencing.

Genes	Regions selected for Miseq panel
IDH2	Hot spot region (Ex 4)
KRAS	Hot spot region (Ex 2, 3)
BRAF	Hot spot region (Ex 11, 15)
ZRSR2	All coding Exons
TET2	All coding Exons
TNFRSF13B	All coding Exons
ASXL1	All coding Exons
SUZ12	All coding Exons
NRAS	Only Exon 2

Table S6. Samples analyzed by RNA sequencing.

Sample	WES	Set	Tumor cells %
BPDCN_21	N	Extension Set	≥90%
BPDCN_20	N	Extension Set	≥90%
BPDCN_22	N	Extension Set	≥90%
BPDCN_24	N	Extension Set	≥90%
BPDCN_23	Y	Discovery Set	≥90%
BPDCN_25	Y	Discovery Set	≥90%
BPDCN_37	Y	Discovery Set	≥90%
BPDCN_42	Y	Discovery Set	≥90%
BPDCN_43	Y	Discovery Set	≥90%
pDC_1	N	Both	//
pDC_2	N	Both	//
pDC_3	N	Both	//
pDC_4	N	Both	//

Abbreviations: BPDCN, blastic plasmacytoid dendritic cell neoplasm; pDC, plasmacytoid dendritic cells; samples already sequenced by WES Y, yes; N, no.

Table S7. MiSeq targeted sequencing validation results.

Cases	Frequency %	ref/var	Gene	CCDS	AA	Validation MiSeq Y/N
BPDCN_25	28	-/G	*ASXL1*	CCDS13201.1	G642+	Y
BPDCN_38	55	G/T	*ASXL1*	CCDS13201.1	G710*	Y
BPDCN_39	36	T/A	*ASXL1*	CCDS13201.1	L775*	Y
BPDCN_49	10	C/T	*ASXL1*	CCDS13201.1	R965*	Y
BPDCN_47	49	A/G	*ASXL1*	CCDS13201.1	N986S	Y
BPDCN_41	41	-/A	*TET2*	CCDS47120.1,CCDS3666.1	S657+	Y
BPDCN_42	27	C/T	*TET2*	CCDS47120.1,CCDS3666.1	Q770*	Y
BPDCN_38	33	G/A	*TET2*	CCDS47120.1,CCDS3666.1	W1003*	Y
BPDCN_43	51	-/A	*TET2*	CCDS47120.1,CCDS3666.1	Q1084+	Y
BPDCN_42	41	C/T	*TET2*	CCDS47120.1	Q1466*	Y
BPDCN_37	80	C/T	*TET2*	CCDS47120.1	R1516*	Y
BPDCN_42	33	C/G	*KRAS*	CCDS8702.1,CCDS8703.1	L19F	Y
BPDCN_CAL-1	99	C/G	*KRAS*	CCDS8702.1,CCDS8703.1	G12A	Y
BPDCN_43	34	C/G	*NRAS*	CCDS877.1	G12A	Y
BPDCN_41	32	C/T	*NRAS*	CCDS877.1	G12S	Y
BPDCN_23	38	C/G	*BRAF*	CCDS5863.1	G469A	Y
BPDCN_46	25	C/A	*BRAF*	CCDS5863.1	G464V	N
BPDCN_39	42	C/T	*SUZ12*	CCDS11270.1	R654*	Y
BPDCN_50	42	C/T	*TNFRSF13B*	CCDS11181.1	R122Q	Y
BPDCN_25	90	C/G	*ZRSR2*	CCDS14172.1	Y373*	Y
BPDCN_25	40	C/T	*IDH2*	CCDS10359.1	R140Q	Y

Table S8. Top 10 biological processes emerged from gene functional analysis of WES data

Term	Description	Count	P-value	FDR	Fold Enrich.
EPIGENETIC PROCESS					
GO:0016569	covalent chromatin modification	8	0.000101184	0.333669805	5.362358277
GO:0006325	chromatin organization	9	0.000117091	0.333669805	4.586177715
GO:0048096	chromatin-mediated maintenance of transcription	2	0.000332385	0.473593505	72.25777778
GO:0045815	positive regulation of gene expression, epigenetic	3	0.000470088	0.519119857	19.70666667
HEMATOPOIETIC STEM CELL HOMEOSTASIS					
GO:0061484	hematopoietic stem cell homeostasis	2	0.000111579	0.333669805	120.4296296
GO:0042592	homeostatic process	13	0.000284612	0.473593505	2.928151843
RAC SIGNALING					
GO:0035020	regulation of Rac protein signal transduction	2	0.000485783	0.519119857	60.21481481
GABA SECRETION					
GO:0014051	gamma-aminobutyric acid secretion	2	0.000266377	0.473593505	80.28641975
GO:0015812	gamma-aminobutyric acid transport	2	0.000573097	0.54437804	55.58290598

Table S9. SNVs of *ASXL1* and *TET2* genes.

Gene	SNV	DNA	Sample	AA	SNV	Position	COSMIC
TET2	Somatic	Unmatched	BPDCN_41	S657+	frameshift	before interaction with DNA domain	//
	Somatic	Unmatched	BPDCN_42	Q770*	nonsense	before interaction with DNA domain	//
	Somatic	Matched	BPDCN_43	Q1084+	frameshift	before interaction with DNA domain	//
	Somatic	Unmatched	BPDCN_42	Q1466*	nonsense	before interaction with DNA domain	//
ASXL1	Somatic	Unmatched	BPDCN_25	G642+	frameshift	inside interaction with NCOA1 domain	COSM G642*
	Somatic	Matched	BPDCN_38	G710*	nonsense	before domain for interaction with RARA	COSM1283534
	Somatic	Matched	BPDCN_39	L775*	nonsense	before domain for interaction with RARA	COSM52930
	Somatic	Matched	BPDCN_49	R965*	nonsense	before domain for interaction with RARA	COSM132978

Table S10. Twenty-five epigenetic modifier genes mutated in BPDCN.

Case	chr:pos	ref/var	Gene	AA	Cosmic_v66_ Gene	MutComFocal Mut Score	Polyphen2 HDIV_score	Polyphen2 HVAR_score	SIFT_score
BPDCN_45	chr1:27106539	G/A	*ARID1A*	W1833*	790	1.66E-05	.|.	.|.	.|1
BPDCN_23	chr1:155448058	G/A	*ASH1L*	R1535C	273	5.00E-07	1.0;1.0	0.99;0.996	0
BPDCN_25	chr20:31022441	-/G	*ASXL1*	G642+	849	5.43E-04			
BPDCN_38	chr20:31022643	G/T	*ASXL1*	G710*	849	5.43E-04	.	.	0.15
BPDCN_39	chr20:31022839	T/A	*ASXL1*	L775*	849	5.43E-04	.	.	1
BPDCN_49	chr20:31023408	C/T	*ASXL1*	R965*	849	5.43E-04	.	.	0.67
BPDCN_47	chr20:31023472	A/G	*ASXL1*	N986S	849	5.43E-04	0.009;0.009	0.005;0.005	0.45
BPDCN_46	chr18:31323569	C/T	*ASXL3*	P1253S	131	1.15E-06	0.208	0.029	0.13
BPDCN_50	chr14:21854308	G/A	*CHD8*	R2125W	205	5.87E-06	0.999	0.988	0
BPDCN_23	chr14:21861393	G/A	*CHD8*	L1835F	205	5.87E-06	0.004	0.003	0.71
BPDCN_CAL-1	chr14:21873903	C/T	*CHD8*	G731R	205	5.87E-06	0.999;0.999|.	0.997;0.995|.	0.02|.
BPDCN_CAL-1	chr4:144457820	T/G	*SMARCA5*	L495*	72	3.68E-07	.	.	1
BPDCN_CAL-1	chr4:95173830	A/C	*SMARCAD1*	N318T	92	5.64E-07	0.028;0.021	0.004;0.009	0.43
BPDCN_25	chr12:50492757	C/T	*SMARCD1*	Q508*	35	1.60E-06	.	.	0.2
BPDCN_39	chr17:30325762	C/T	*SUZ12*	R654*	54	5.58E-05	.	.	1
BPDCN_CAL-1	chr3:4345101	A/G	*SETMAR*	E16G	30	1.34E-06	0.286;0.165;0.009	0.055;0.02;0.003	0.21
BPDCN_42	chr16:30732644	C/T	*SRCAP*	P1130S	275	6.08E-07	0.988;0.979	0.794;0.628	0.09
BPDCN_41	chr4:106157069	-/A	*TET2*	S657+	929	2.90E-04			
BPDCN_42	chr4:106157407	C/T	*TET2*	Q770*	929	2.90E-04	.	.	0.34
BPDCN_38	chr4:106158108	G/A	*TET2*	W1003*	929	2.90E-04	.	.	1
BPDCN_43	chr4:106158349	-/A	*TET2*	Q1084+	929	2.90E-04			
BPDCN_42	chr4:106193934	C/T	*TET2*	Q1466*	929	2.90E-04	.	.	0.25
BPDCN_37	chr4:106196213	C/T	*TET2*	R1516*	929	2.90E-04	.	.	1
BPDCN_25	chr15:90631934	C/T	*IDH2*	R140Q	607	3.04E-06	1	0.998	0
BPDCN_47	chr12:49448165	C/G	*MLL2*	W145C	600	1.51E-05	1	1	0
BPDCN_41	chr7:151843784	A/C	*MLL3*	V4644G	774	1.51E-06	.|0.993;0.804;0.902	.|0.907;0.342;0.415	0.02|0.01

BPDCN_CAL-1	chr19:36216708	C/T	*MLL4*	R1292W	223	1.07E-06	1	0.981	0
BPDCN_50	chr11:94731791	C/T	*KDM4D*	R419W	54	2.46E-06	0.002	0.001	0.01
BPDCN_46	chr9:96429387	C/T	*PHF2*	S738L	105	1.64E-06	0.064;0.002	0.005;0.001	0.21
BPDCN_CAL-1	chr12:9085218	C/T	*PHC1*	Q389*	65	5.63E-07	.	.	0.15
BPDCN_CAL-1	chr1:33836638	C/A	*PHC2*	A131S	72	7.53E-07	0.009;0.009;0.049	0.014;0.006;0.012	1
BPDCN_41	chr22:41573648	G/C	*EP300*	G1978A	326	9.51E-07	0.734	0.196	0.46
BPDCN_CAL-1	chr12:132445627	C/T	*EP400*	P155S	346	4.73E-07	0.196;0.121;0.196;0.245;0.036	0.044;0.044;0.044;0.094;0.028	0.16
BPDCN_CAL-1	chr12:132471135	C/A	*EP400*	S668Y	346	4.73E-07	0.631;0.631;0.631;0.98;0.753	0.329;0.329;0.329;0.851;0.329	0.05
BPDCN_46	chr20:45618703	T/A	*EYA2*	D18E	71	1.12E-05	0.0;0.0;0.0;0.0	0.001;0.001;0.001;0.001	0.8
BPDCN_42	chr20:45702876	C/T	*EYA2*	P188L	71	1.12E-05	0.996;0.846;0.875;0.875	0.797;0.131;0.173;0.173	0.02
BPDCN_41	chr8:41789947	T/C	*MYST3*	M1931V	193	4.16E-06	0.969	0.914	0.15
BPDCN_23	chr10:76784746	C/T	*MYST4*	R1135C	209	2.84E-06	1.0;1.0;1.0	0.997;0.998;0.988	0

Table S11. Top 10 significant biological process emerged by Gene Ontology analysis of 86 up-regulated genes marked by H3K27-promoter acetylation

Term	Description	Count	P-Value	Fold Enrichment	FDR
GO:0007067	mitotic nuclear division	15	3.69034691694026e-10	8.027320606	3.15487757929223e-06
GO:0007059	chromosome segregation	13	1.26807986333688e-09	9.177333912	3.71276365696005e-06
GO:0051783	regulation of nuclear division	10	1.63092095384343e-09	14.42590949	3.71276365696005e-06
GO:0000819	sister chromatid segregation	11	1.8299290971413e-09	11.88292359	3.71276365696005e-06
GO:0098813	nuclear chromosome segregation	12	2.55158560946711e-09	9.883282675	3.71276365696005e-06
GO:0000278	mitotic cell cycle	20	2.60575294674936e-09	4.778953557	3.71276365696005e-06
GO:0000070	mitotic sister chromatid segregation	9	5.32755506377214e-09	15.71664876	6.50646689145543e-06
GO:0007088	regulation of mitotic nuclear division	9	6.9189063367503e-09	15.25776851	7.39371628410979e-06
GO:0051301	cell division	15	1.00188266571877e-08	6.299922501	9.0508192280403e-06
GO:0022402	cell cycle process	22	1.05869917277346e-08	3.967125111	9.0508192280403e-06

Supplemental Figures

Figure S1. Sanger Sequencing validation experiments. Representative chromatograms of matched tumor DNA samples and germline DNA of saliva samples showing somatic mutations in exon 13 and exon 16 of *ASXL1* and *SUZ12* of two patients. respectively. Mutations were detected in both strands of tumor DNA and absent from germline DNA.

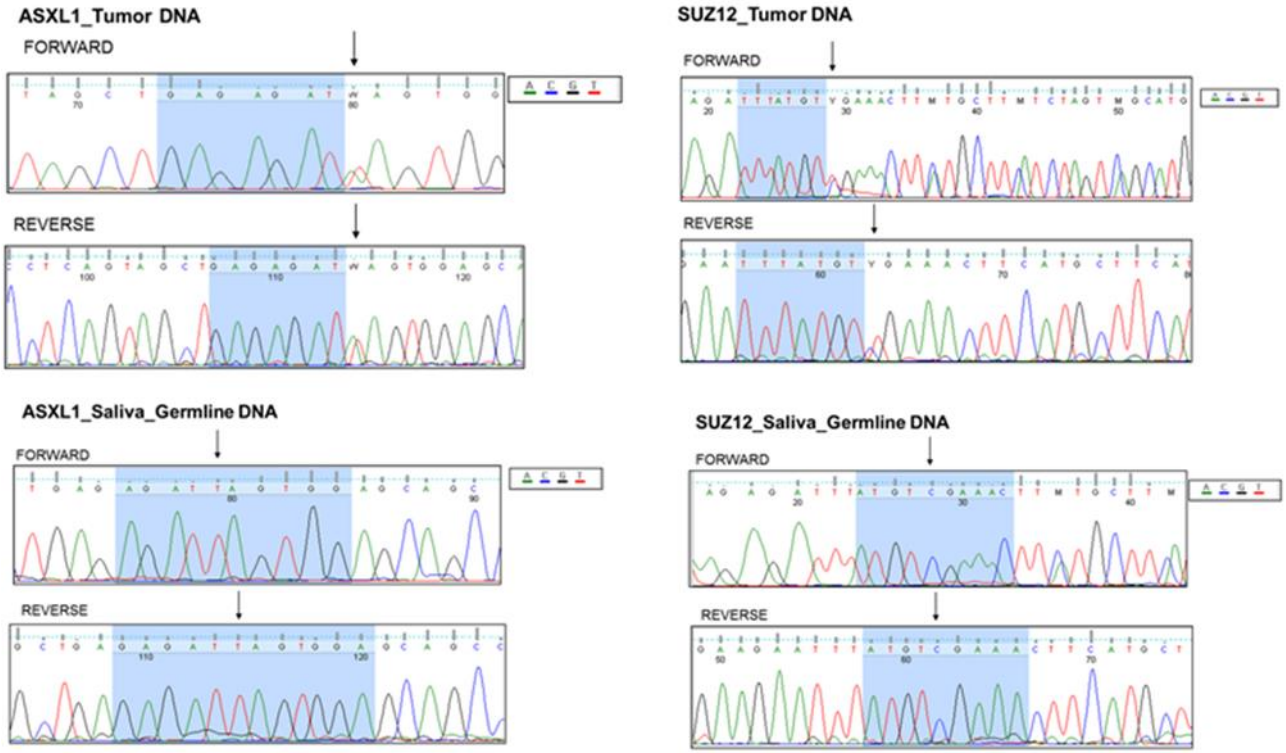


Figure S2 - Chromosome 9 copy ratio heatmap. Blu color represents chromosome segment loss (homozygous deletion is darker than single-copy loss). Red color represents chromosome segment gain. Values of this heatmap are the ratios between each tumor sample segment with respect to the corresponding normal reference segment. . (BPDCN_01_U represent the CAL-1 tumor cell line) . Therefore, a lower ratio (negative number) means that the tumor sample segment is less represented with respect to the corresponding normal reference segment.

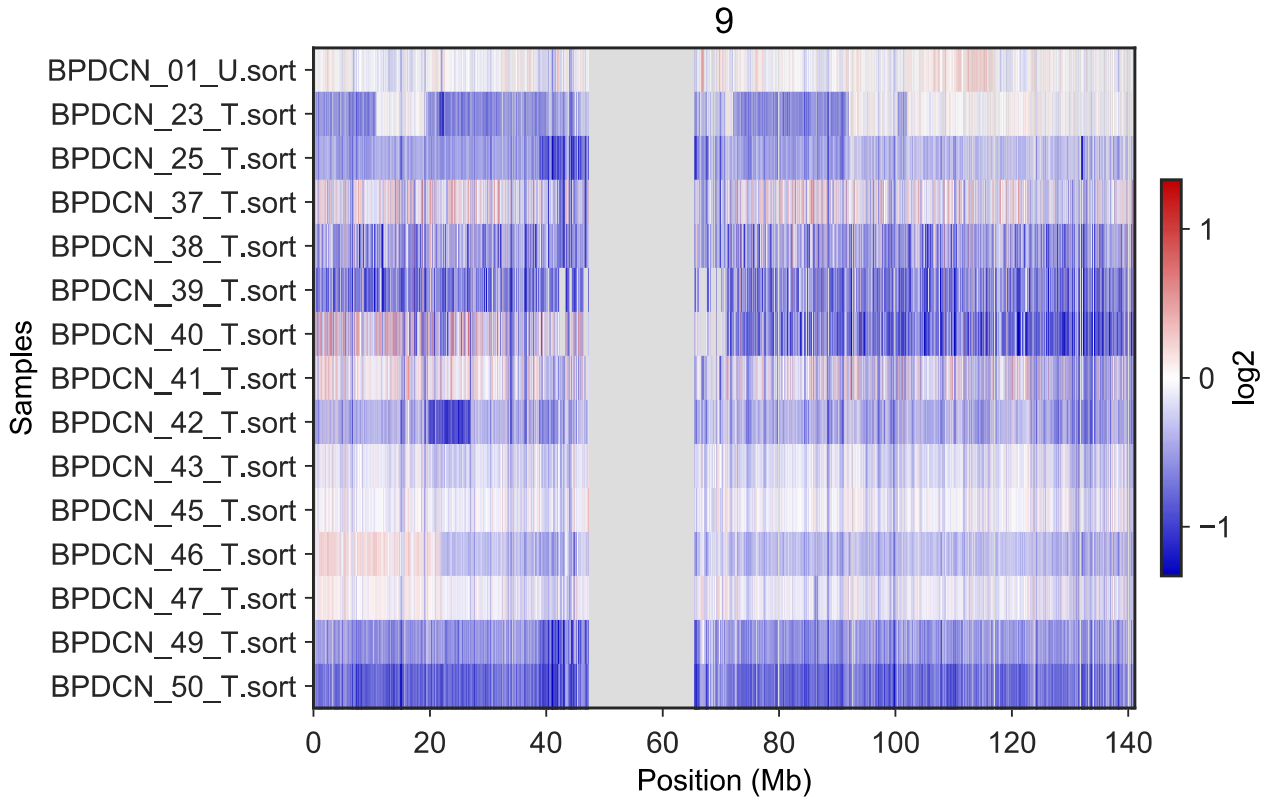


Figure S3. Copy number variants of 9 recurrently mutated genes (≥ 3 BPDCN samples). Each row represents a gene and each column a BPDCN sample.

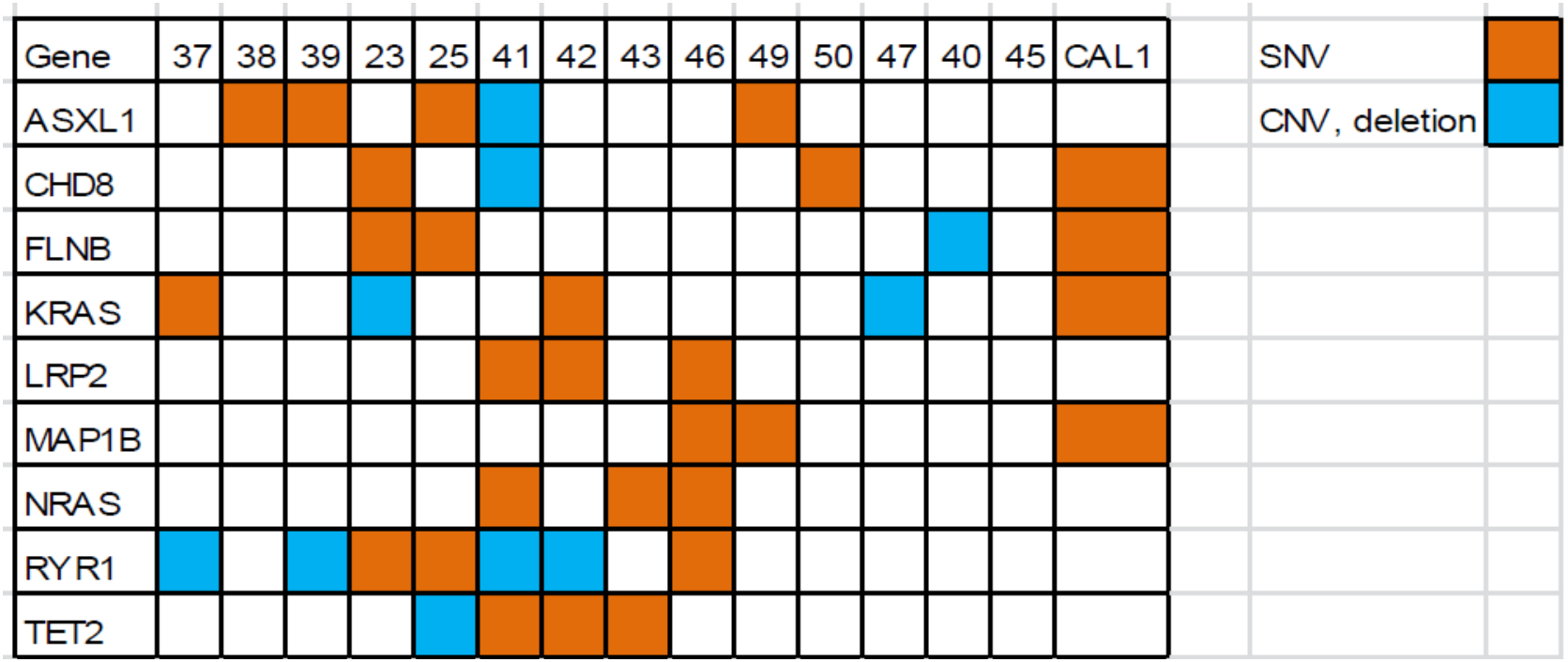


Figure S4 Gene Set Enrichment Analysis (GSEA) in BPDCNs extension set . Representative plots illustrate. in BPDCN patients the enrichment of the same gene signatures recognized also in BPDCN discovery set: the *KDM5B* and *PRMT5* gene signatures reported in literature[14,15] and also the enrichment of a set of genes. described by Missiaglia *et al*[16] as responsive to a hypomethylating treatment. namely the Decitabine. NES normalized enrichment score ≥ 1.8 ; FDR q-val. false discovery rate ≤ 0.02

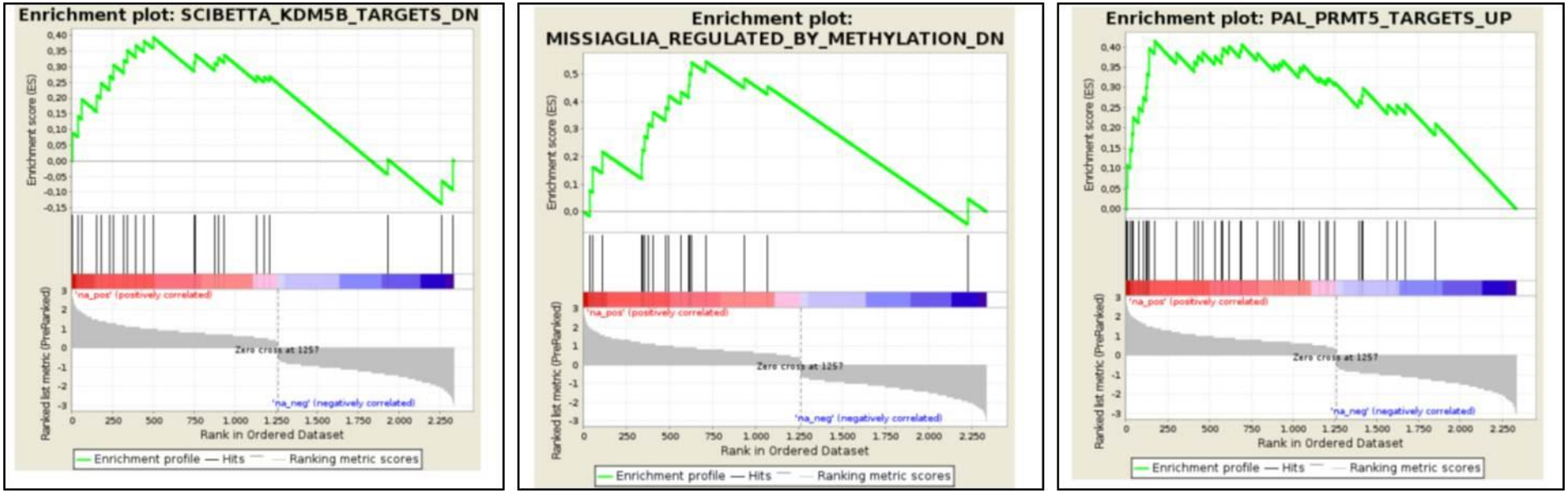
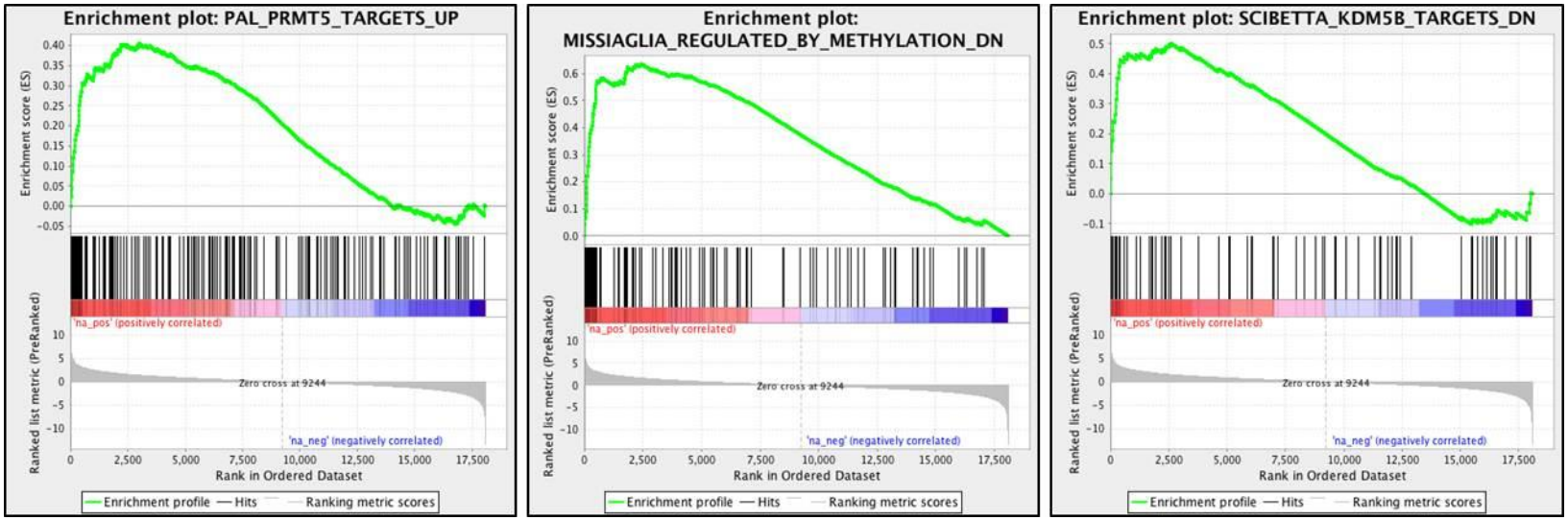


Figure S5. Gene Set Enrichment Analysis (GSEA) in CAL-1 cell line. Representative plots illustrate. in CAL-1 the enrichment of the same gene signatures recognized also in BPDCN samples: the *KDM5B* and *PRMT5* gene signatures reported in literature[14,15] and also the enrichment of a set of genes. described by Missiaglia[16] et al as responsive to a hypomethylating treatment. namely the Decitabine. NES normalized enrichment score ≥ 1.8 ; FDR q-val. false discovery rate ≤ 0.06



Supplemental References

1. Facchetti F, Petrella T, Pileri SA: Blastic plasmacytoid dendritic cell neoplasm. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Revised 4th Edition. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Editors. IARC Press; Lyon, France: 2008. pp. 145–147.28.

2. Fanelli M, Amatori S, Barozzi I, et al. Pathology tissue-chromatin immunoprecipitation, coupled with high-throughput sequencing, allows the epigenetic profiling of patient samples. Proceedings of the National Academy of Sciences. 2010;107(50):21535-21540

3. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS Comput Biol. 2016; 21;12(4).

4. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25(16):2078–2079.

5. Li H. Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-1760.

6. Trifonov V. Pasqualucci L. Tiacci E. Falini B. Rabadan R. SAVI: a statistical algorithm for variant frequency identification. BMC Systems Biology. 2013;7(Suppl 2):S2.

7. Wang J, Vasaikar S, Shi Z, Greer M, Zhang B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. Nucleic Acids Research. 2017;45(W1):W130-W137.

8. Dobin A. Davis C. Schlesinger F et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.

9. Anders S. Huber W: Differential expression analysis for sequence count data. Genome Biol. 2010; 11:R106.

10. Subramanian A. Tamayo P. Mootha V et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences. 2005;102(43):15545-15550.

11. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9(9):R137.

12. Ye T. Krebs A. Choukrallah M et al. seqMINER: an integrated ChIP-seq data interpretation platform.Nucleic Acids Research. 2011;39(6):e35-e35.

13. Agliano A, Martin-Padura I, Marighetti P, et al. Therapeutic effect of lenalidomide in a novel xenograft mouse model of human blastic NK cell lymphoma/blastic plasmacytoid dendritic cell neoplasm. Clinical Cancer Research. 2011;17(19):6163-6173.

14. Scibetta AG, Santangelo S, Coleman J, Hall D, Chaplin T, Copier J, et al. Functional analysis of the transcription repressor PLU-1/JARID1B. Molecular and Cellular Biology. 2007;27(20):7220-7235

15. Pal S, Vishwanath SN, Erdjument-Bromage H, Tempst P, Sif S. Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes. Molecular and Cellular Biology. 2004;24(21):9630-9645.

16. Missiaglia E, Donadelli M, Palmieri M, Crnogorac-Jurcevic T, Scarpa A, Lemoine NR. Growth delay of human pancreatic cancer cells by methylase inhibitor 5-aza-2'-deoxycytidine treatment is associated with activation of the interferon signalling pathway. Oncogene. 2005;24(1):199-211