DNA methylation profiling of hepatosplenic T-cell lymphoma

Anke K. Bergmann,[1,2,3] Virginie Fataccioli,[4] Giancarlo Castellano,[5] Nadine Martin-Garcia,[4] Laura Pelletier,[4] Ole Ammerpohl,[1,7] Juri Bergmann,[8] Jaydeep Bhat,[9] Enrique Carrillo-de Santa Pau,[10] José I. Martín-Subero,[5,6] Andrea B. Moffitt,[11] Alfonso Valencia,[12,13] Hans-Heinrich Oberg,[9] Daniela Wesch,[9] Sandrine Jayne,[14] Martin J.S. Dyer,[14] Dieter Kabelitz,[9] Philippe Gaulard[4] and Reiner Siebert[1,7]

[1]Institute of Human Genetics, Christian-Albrechts-University Kiel & University Hospital Schleswig-Holstein, Campus Kiel, Germany; [2]Department of Pediatrics, Christian-Albrechts-University Kiel & University Hospital Schleswig-Holstein, Campus Kiel, Germany; [3]Institute of Human Genetics, Medical School Hannover (MHH), Germany, [4]Department of Pathology, AP-HP, Groupe hospitalier Henri Mondor; Université Paris-Est, Faculté de Médecine; Inserm U955, Créteil, France; [5]Departamento de Anatomía Patológica, Farmacología y Microbiología, Universitat de Barcelona, Institut d'Investigacions Biomèdiques August Pi I Sunyer (IDIBAPS), Spain; [6]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Institució Catalana de Recerca i Estudis Avançats (ICREA), and Departament de Fonaments Clinics, Universitat de Barcelona, Spain; [7]Institute of Human Genetics, University of Ulm & University Medical Center Ulm, Germany; [8]Anatomical Institute, Christian-Albrechts-University Kiel, Germany; [9]Institute for Immunology, Christian-Albrechts-University Kiel & University Hospital Schleswig-Holstein, Campus Kiel, Germany; [10]Computational Biology Group. Precision Nutrition and Cancer Research Program. IMDEA Food Institute. Madrid, Spain; [11]Duke Center for Genomics and Computational Biology, Duke University, Durham, NC, U S A; [12]Barcelona Supercomputing Centre (BSC), Barcelona Spain; [13]ICREA, Pg. Lluis Companys 23, Barcelona, Spain and [14]Ernest and Helen Scott Haematological Research Institute, University of Leicester, UK

Correspondence: ANKE K. BERGMANN
bergmann.anke@mh-hannover.de
doi:10.3324/haematol.2018.196196

Supplementary Methods

Patient Cohort

All patients had spleen, liver and bone marrow involvement without lymphadenopathies. Nine patients had been included in previous reports [1, 2]. HSTL was diagnosed according to the criteria of the 2008 WHO classification [3].The tumor cells were small-medium-sized lymphocytes with a CD3+, CD5-, CD4-/CD8-, TIA1+, GrB-, CD56+ phenotype. T-cell receptor (TCR) lineage was determined by flow cytometry or/and immunohistochemistry of TCRβ and TCRδ chain expression. Isochromosome 7 was identified in 5 out of 7 analyzed patients. The other cases are not known or there was no material left for FISH. Mutations in the *STAT5B* or *STAT3* gene were identified in 6, respectively 2 patients out of 11 patients. Mutations in the *SETD2* gene were identified in 5 out of 11 patients.

Control samples

Benign cells were sorted according to manufacturer's instructions using following kits: Anti-TCR γδMicroBead Kit (No.:130-050-701); CD4+ T Cell Isolation Kit II (No. 130-091-155); CD8 MicroBeads kit (No. 130-045-201) (Milteny, BergischGladbach, Germany).

DNA extraction

Whole genomic DNA was extracted from frozen tumor cells containing samples and the benign controls using the Gentra purification kit (Qiagen, Hilden, Germany). DNA quality was verified by measuring the DNA with the Qubit® Fluorometer using the Qubit dsDNA BR Assay Kit (Life Technologies, Darmstadt, Germany). For the following array-based DNA methylation analysis 1µg of whole genomic DNA was applied according to manufacturer´s instructions.

Mutation analysis of the *STAT3, STAT5B, SETD2* genes

The exons known to be affected by mutations in HSTL (*STAT3*: exon 21 *STAT5B*:exons 15, 16, 17) were amplified by polymerase chain reaction from genomic DNA using primers and conditions described in Supplementary Table 1. Amplicons were purified (MinElute 96 UF PCR Purification Kit, Qiagen, Hilden, Germany) and cycle-sequenced using fluorescent dye-termination (Big Dye Terminator V1.1 Cycle Sequencing Kit, Applied Biosystems, Darmstadt, Germany) and an ABI 3100 or ABI 310 automatic capillary genetic analyzer. Sequencing results were compared to the sequence of the wild-type human *STAT3* (HGNC: 11364, ENSG00000168610) and *STAT5B* (HGNC:11367, ENSG00000173757). Identified single nucleotide polymorphisms (SNP) were compared to the known COSMIC variations (http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/) and to dbSNP (http://ncbi.nlm.nih.gov/SNP).

An AmpliSeq DNA custom panel (ThermoFisher Scientific, Waltham, MA, USA) was used as previously described [4] to sequence SETD2 in two patients (ID 1 and 2) . This panel, designed to study mutations in T-cell lymphomas, covers the entire coding sequence of SETD2 and includes 8 other genes (hotspot regions or full gene) [4].

Array-based DNA Methylation analysis

The HumanMethylation450 BeadChip allows the interrogation of 485,577 assays (482,421 CpG sites, 3091 non-CpG sites and 65 random SNPs) in parallel at a single-nucleotide resolution per sample [5]. Arrays were scanned using the Illumina iScan. Raw hybridization signals were processed using the GenomeStudio software (version2011.1; Methylation Analysis Module version 1.9.0, Illumina) applying the default settings and internal controls for normalization. All samples met the internal quality criteria referring to the technical results of

hybridization of a gene call rate above >99% per sample. CpG loci with detection p-value below 0.01, as well as loci covering SNPs (rs-tags) and loci located on the X- or Y chromosome were excluded. Tags including polymorphisms with an allele frequency >5% in the population which are located within a distance of 0-3 bp next to the cytosine analyzed on the array were also excluded. We included 2,902 CH loci of the array representing non CpG methylation targets. All 23 hybridizations were normalized together and met the described quality criteria and entered with 450,959 CpG loci further analysis. Principal component analysis (PCA) and hierarchical cluster analysis were performed using the QlucoreOmics Explorer 2.3 (Version 2.3(45), 64-bit, Qlucore, Lund, Sweden).

Unsupervised analysis of all cases and controls revealed 5,376 differentially methylated CpG loci ($\sigma/\sigma_{max}= 0.5$). Supervised analysis of cases and controls revealed 4,113 CpG loci to be significantly differentially methylated (see supplemental table: supervised analysis 1; t-test using FDR= $1*10^{-5}$). Further described supervised analyses include: I) αβ HSTL and benign αβ T-cells ($CD4^+$ and $CD8^+$ T-cells) (see supplemental table: supervised analysis 2, t-test using FDR= $1*10^{-5}$, revealed 2,165 differentially methylated loci); and II) γδ HSTL and benign γδ T-cells (see supplemental table: supervised analysis 3, t-test using FDR= 0.05, revealed 5,896 differentially methylated loci). Neither unsupervised nor supervised analyses did show clustering by array, thus there was no evidence for major batch effects.

Gene expression data

For Supplementary Table 9 gene-expression profiling data from HSTL patients were included from Traver*et al* [1]. E-MTAB-638 accession number.

Immunohistochemical analysis of selected candidate genes

Immunohistochemistry was performed on deparaffinized tissue sections for five hyperplastic spleens and ten HSTL patients of the current series (six on full sections, four on tissue micro-array sections). CD5 (Novocastra, Newcastle, UK) was detected according to standard procedures [6]. CXCR6 was detected using a standard indirect avidin-biotin immunoperoxidase method revealed in brown with 3,3'-diaminobenzidine (D.A.B Vector Laboratories, Burlingame, CA, USA). After a pH8 antigen retrieval pretreatment, sections were incubated with goat polyclonal anti-CXCR6 (Abcam, Cambridge, England) followed by biotinyled anti-goat and avidin-peroxydase steps. Peroxidase activity was revealed with Diaminobenzidine.

Gene ontology analysis

The package topGO (Alexa, A. and Rahnenführer, J. topGO: Enrichment analysis for Gene Ontology. R package version 2.12.0) was used for testing GO terms from Gene Ontology biological processes.

Genomic annotation, transcriptional and epigenomic characterization of CpG and non CpG sites

DNA methylation array data was annotated using data from hg19 version of the UCSC Genome Browser database. CpG and non-CpG sites were annotated for their location with respect to genes, CpG islands and chromatin states.

Chromatin State Segmentations:

ChIP-seq data for H3K4me3;H3K4me1;H3K36me3;H3K27Ac;H3K27me3 and H3K9me3 from BLUEPRINT consortium samples C002TWH2 (CD4 naïve) and C003UQH1 (CD8 naïve) were used (segmentation data are available as supplemental material).

Chromatin segmentations were carried out with ChromHMM package (v1.10) [7] using the BLUEPRINT 12 states chromatin model generated by BLUEPRINT consortium for 20150128 release (for more details visit ftp://ftp.ebi.ac.uk/pub/databases/blueprint/releases/20150128/homo_sapiens/secondary_analysis/Segmentation_of

_ChIP-Seq_data/) (Figure 1). Further, segmentations for each sample from the 11 states model were collapsed into 5 chromatin states as described previously by Carrillo-de-Santa-Pau *et al* [8] (Figure 1, color code).

State	Description	Label
State1	Repressed Polycomb High signal H3K27me3 with H3K4me1/3	Rpro
State2	Repressed Polycomb Low signal H3K27me3	Rhet
state3	Low signal	Rhet
State4	Heterochromatin High Signal H3K9me3	Rhet
State5	Transcription High signal H3K36me3	TranR
State6	Transcription Low signal H3K36me3	TranR
State7	Genic Enhancer High Signal H3K4me1 & H3K36me3	TranR
State8	Enhancer High Signal H3K4me1	RegE
State9	Active Enhancer High Signal H3K4me1 & H3K27Ac	RegE
State10	Distal Active Promoter (2Kb) High Signal H3K4me3 & H3K27Ac & H3K4me1	Apro
State11	Active TSS High Signal H3K4me3 & H3K4me1	Apro
state12	Active TSS High Signal H3K4me3 & H3K27Ac	Apro

Figure 1: Chromatin states. Chromatin segmentations according to the BLUEPRINT 12 states chromatin model generated by BLUEPRINT consortium. The color code is depicting the 5 chromatin states as described previously by Carrillo-de-Santa-Pau *et al* [8].

Enrichment analysis

Transcription factor binding site (TFBS) information was obtained by ChIP-seq data in the ENCODE project and available at the UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgTrackUi?db=hg19&g=wgEncodeHaibTfbs). A total of 79 TFBS were used or the analysis. The relative enrichment of each TFBS was calculated with respect to the background.

References:

1. Travert M, Huang Y, de Leval L, *etal.*Molecular features of hepatosplenic T-cell lymphoma unravels potential novel therapeutic targets. *Blood* 2012; 119: 5795-5806.
2. McKinney M, Moffitt AB, Gaulard P, *et al.* The Genetic Basis of Hepatosplenic T-cell Lymphoma. Cancer Discov. 2017 Apr;7(4):369-379.
3. Swerdlow SH, Campo E, Harris NL, *et al.* eds. World Health Organization Classification of Tumours of Haematopoietic and Lymphoid Tissues.IARC: Lyon, France; 2008.
4. Lemonnier F, Cairns RA, Inoue S,*etal.*The IDH2 R172K mutation associated with angioimmunoblastic T-cell lymphoma produces 2HG in T cells and impacts lymphoid development. ProcNatlAcadSci U S A. 2016 ;113(52):15084-15089.
5. Bibikova M, Barnes B, Tsan C, *et al.*.High density DNA methylation array with single CpG site resolution. *Genomics* 2011;98:288-295.
6. Moroch J, Copie-Bergman C, de Leval L, *etal.*Follicular peripheral T-cell lymphoma expands the spectrum of classical Hodgkin lymphoma mimics. *Am J SurgPathol* 2012; 36: 1636-1646.
7. Ernst J., Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 2012; 9:215–216
8. Carrillo-de-Santa-Pau E, Juan D, Pancaldi V, *et al.*Automatic identification of informative regions with epigenomic changes associated to hematopoiesis.NucleicAcids Res. 2017;45(16):9244-9259.

Legend to Supplemental Tables 1- 10

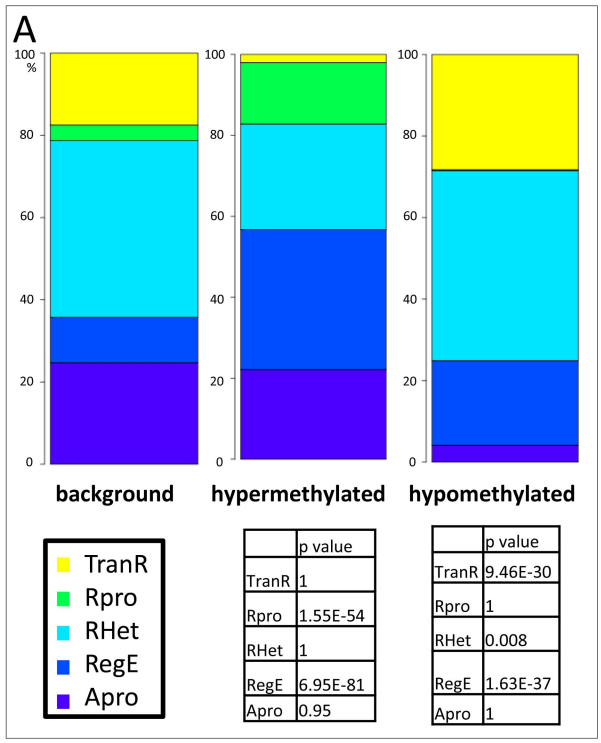
Suppl Tab 1: STAT5B and STAT3 primers and conditions
Suppl Tab 2: Supervised analysis of 11 HSTL compared to 12 benign T-cell subsets
Suppl Tab 3: Enriched GO terms of analysis from Supplementary Table 2
Suppl Tab 4: Hypermethylated CpGs in repressed promoters of genes
Suppl Tab 5: Supervised analysis of αβ HSTL compared to benign αβ T-cell subsets
Suppl Tab 6: Enriched GO terms of analysis from Supplementary Table 5
Suppl Tab 7: Supervised analysis of γδ HSTL compared to benign γδ T-cell subsets
Suppl Tab 8: Enriched GO terms of analysis from Supplementary Table 7
Suppl Tab 9: High confidence genes differentially methylated in HSTL
Suppl Tab 10: Function, methylation status and gene expression of high confidence genes

Figure legend Supplementary Figures

Figure S2: Annotation of HSTL methylation signature with CD4/CD8 chromatin states. Differentially methylated loci between all HSTL and all benign T-cells were annotated by thechromatin segmentation data of the CD4 (A) and CD8 (B) T-cells from the BLUEPRINT project (segmentations are depicted in different colors in %, color code explained in supplementary methods file). The background bar represents the loci of the array and the bars on the right side present the hyper- respectively hypomethylated loci in HSTL. The tables below the bars indicate the corresponding p values. TranR=Transcribed region, Rpro=repressed promoter, Rhet=repressed Heterochromatin, RegE=regulatory element, Apro=active promoter

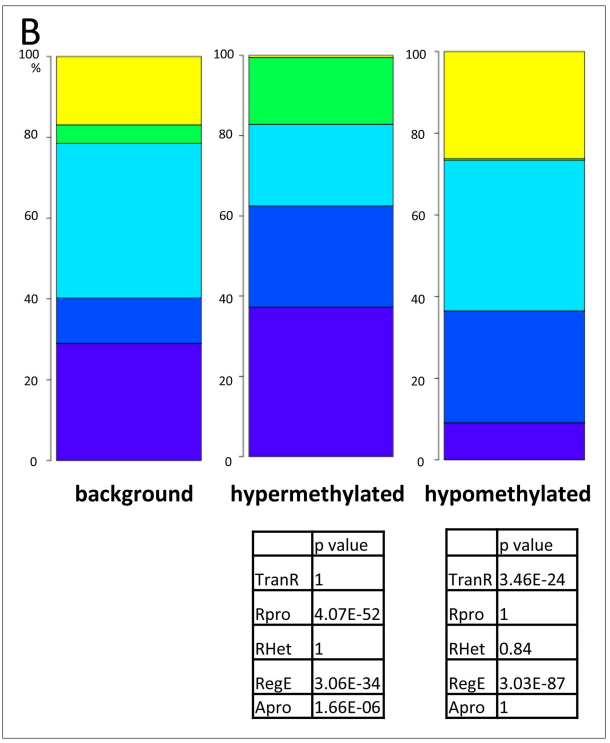
Figure S1: Genomewide methylation signature of αβ HSTL and γδHSTL

Heatmap and corresponding principal component analysis (PCA) of supervised cluster analysis of A) αβ HSTL and benign αβ T-cells ($CD4^+$ and $CD8^+$ T-cells), using FDR= $1*10^{-5}$, 2,165 of 450.959 differentially methylated CpG loci are visualized and B) γδHSTL and benign γδ T-cells , applying a FDR< 0.05, 5,896 of 450.959 differentially methylated CpG loci are visualized.



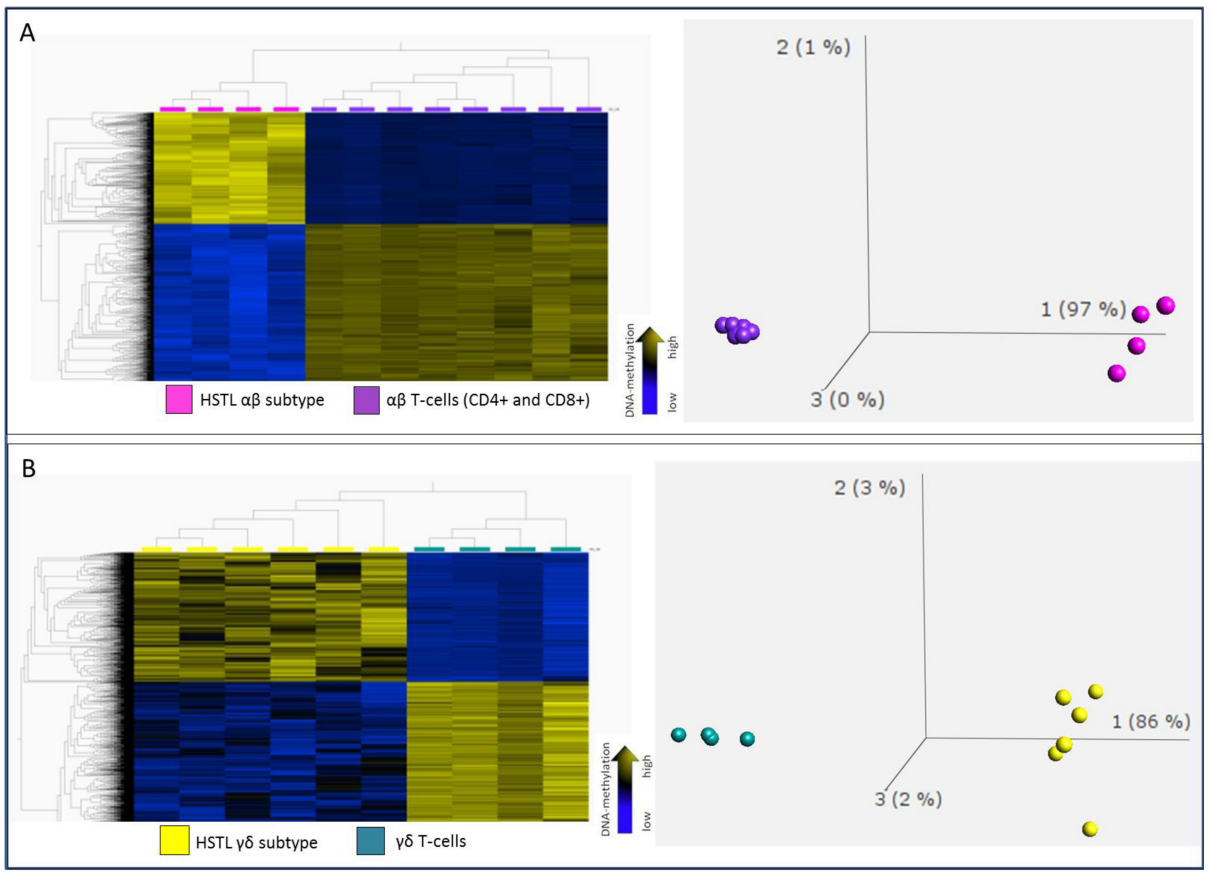
	p value
TranR	1
Rpro	1.55E-54
RHet	1
RegE	6.95E-81
Apro	0.95

	p value
TranR	9.46E-30
Rpro	1
RHet	0.008
RegE	1.63E-37
Apro	1



	p value
TranR	1
Rpro	4.07E-52
RHet	1
RegE	3.06E-34
Apro	1.66E-06

	p value
TranR	3.46E-24
Rpro	1
RHet	0.84
RegE	3.03E-87
Apro	1


A
HSTL αβ subtype
αβ T-cells (CD4+ and CD8+)
DNA-methylation
high
low
2 (1 %)
1 (97 %)
3 (0 %)
B
HSTL γδ subtype
γδ T-cells
DNA-methylation
high
low
2 (3 %)
1 (86 %)
3 (2 %)