Base-pair resolution DNA methyleome of the EBV-positive Endemic Burkitt lymphoma cell line DAUDI determined by SOLiD bisulfite-sequencing

Leukemia (2013) 27, 1751–1753; doi:10.1038/leu.2013.4

The Burkitt translocation t(8;14), first identified in the 1970s in biopsies and cell lines from Burkitt lymphoma (BL),1,2 and its variants juxtapose the MYC oncogene to one of the immunoglobulin (IG) loci.3 Nowadays, it is assumed that (nearly) all BL carry an IG-MYC translocation, rendering this somatic mutation a diagnostic marker for all three subtypes of BL (endemic, sporadic and immunodeficiency-related BL). In contrast to many other lymphomas, BL show a quite simple karyotype, that is, with few if any secondary chromosomal changes in addition to the IG-MYC translocation.4 Although there is evidence for some few recurrent secondary genetic changes the number of epigenetic alterations in BL as compared with normal B-cell subsets seems to outnumber the genetic changes by far.5–7 Indeed, along with others, we have identified several hundred genes showing de novo DNA methylation in aggressive B-cell lymphoma, including BL as compared with normal B-cell subsets.8–10 Nevertheless, the mentioned DNA changes the number of epigenetic alterations in BL as compared with normal B-cell subsets seems to outnumber the genetic changes by far.5–7 Indeed, along with others, we have identified several hundred genes showing de novo DNA methylation in aggressive B-cell lymphoma, including BL as compared with normal B-cell subsets.8–10 Nevertheless, the mentioned DNA methylation studies focused on a maximum of probably 10% (by HELP assays) of the CpGs of the genome, and were biased toward promoter regions and CpG islands and did not systematically analyze non-CpG methylation.8–11 Therefore, we here aimed at generating a complete DNA methyleome of a BL, allowing for unbiased analyses of all cytosines in the genome.

To this end, we chose the archetypal DAUDI cell line, established from an endemic BL (eBL) that was derived from a 16-year-old African male patient in 1967.12,13 We selected this cell line as it has been pivotal for the identification of t(8;14), still carries a simple karyotype despite being many years in culture and because it shows the prototypic features of eBL. Moreover, considering the strong association of eBL with Epstein–Barr virus (EBV) infection, the EBV-positive DAUDI cell line offers the opportunity for a direct comparison of its lymphoma and EBV methyleomes.

To obtain a base-pair resolution DNA methyleome of a prototypic eBL, we subjected DNA of the DAUDI cell line to full bisulfite-sequencing (BS-seq) using the SOLiD two-base encoding (color-space) approach (for details, see Supplementary Information). Two bisulfite-converted SOLiD fragment libraries were constructed. Briefly, 15 μg of genomic DNA were sheared to approximately 125 bp. After end-repair of the DNA fragments, methyl-P1 and -P2 adaptors were ligated. The DNA was then size selected and nick translated with a modified deoxyribonucleoside triphosphate (dNTP) mix containing methyl-deoxyctydinetriphosphate (dCTP) instead of regular dCTPs. Bisulfite conversion was carried out in solution and recovered DNA fragments were PCR amplified using eight cycles. The bisulfite-converted fragment library was cloned on SOLiD P1 beads using emulsion PCR. Templated (P2 positive) beads were then enriched and deposited on a slide for sequencing. Technical details on the genomic characterization of the cell line including karyotyping, single-nucleotide polymorphism (SNP) array analysis and exome sequencing as well as on control DNA methylation analyses using Illumina 450K BeadArray analysis (Illumina, San Diego, CA, USA), limited BS-seq using Illumina HiSeq2000 technology as well as Luminometric methylation assay (LUMA) are provided in the Supplementary Information.

By karyotyping and SNP-array analysis, we confirmed that the cells under study show the typical features of BL, including the t(8;14) plus a few secondary chromosomal changes. Exome sequencing followed by filtering for known SNPs revealed a total of 2313 non-synonymous mutations (Supplementary Table 1a). Owing to the lack of a germline control from the patient from which the DAUDI cell line has been established, it is not possible to reliably differentiate somatic (lymphoma-associated) mutations from germline variants. Despite this limitation, exome sequencing in line with recent reports,6,7,12,17 identified inactivating mutations in the I3D gene, which have been shown to cooperate with MYC activation in the pathogenesis of BL.5–7,14 In addition, Sanger sequencing confirmed sequence variants in the genes B2M, TET2 and KIT (Supplementary Table 1b).

We aligned 79.9 Gb of BS-seq of the SOLiD platform.15 These were compared with 7.8 Gb aligned BS-seq of the HiSeq 2000 platform and the results to the DNA methylation levels determined by HumanMethylation450 BeadChip analysis (Supplementary Figure S1). We observed high correlation of the SOLiD data with both the sequence-based HiSeq 2000 (Pearson r = 0.86; Supplementary Figure S2) and the array-based (Pearson r = 0.96; Supplementary Figure S3) methylation levels. This led us to focus our further analyses on the most extensive data set derived from SOLID BS-seq.

In total, 91.1% of all CpG sites and 90.2% of all non-CpG sites of the genome were covered by at least five SOLID reads (Supplementary Figure S4). On the genome-wide level, 68.99% cytosines in CpG dinucleotides were methylated, which is in line with previous pyrosequencing-based determinations using LUMA. In contrast, the 450K BeadArray shows a mean methylation level of 59.24%, which is mostly due to the selection bias of the array loci, which are predominantly located within regions upstream of genes. We observed a mean CpG methylation level of 65.79% in LINEs and 78.84% in SINEs (Supplementary Figure S5). BS-seq shows the DNA methylation patterns on the forward and reverse strand to be comparably established (Pearson r = 0.90).

Considering the recent description of non-CpG methylation in embryonic stem cells (ESCs), and the fact that the MYC oncogene deregulated in BL is also one of the four factors used to induce a stem cell-like phenotype in differentiated cells,16 we analyzed the level of non-CpG methylation. The genome-wide fraction of methylated cytosines in a non-CpG context does not exceed the respective threshold of 0.003 given by the unmethylated lambda control DNA that was tested in parallel.17 Moreover, we confirmed absence of non-CpG methylation at hallmark sites described in ESC11 by bisulfite pyrosequencing (Supplementary Figure S6). Despite this overall low frequency of non-CpG methylation, we could identify a remarkable 6.7-fold enrichment of methylated non-CpG sites within genes (P < 2.2 × 10−16; Figure 1b). Such non-CpG methylation might be linked to transcriptional activity (Supplementary Figure S7).

Next, we determined the sequence-based methylation status of 969 genes recently shown by us to exhibit de novo promoter hypermethylation in mature aggressive B-cell lymphoma (including BL) as compared with normal B cells.8
We could confirm that in DAUDI cells 91.21% of these genes have a DNA methylation level ≥ 60% in their promoter region and lack transcription. As compared with all other RefSeq genes, the mean CpG methylation level within promoter regions of the 969 genes was significantly higher (84% vs 41%; Supplementary Tables S2).

Gene expression analyses confirmed that DAUDI cells show the typical signature of molecular BL. Correlating methylation and expression patterns in our data revealed that significant presence of transcripts is associated with absence of DNA methylation particularly at and closely around the transcription start site (TSS). In contrast, DNA methylation exactly at the TSS correlates with lack of transcription (Figure 2). Although the group of non-expressed genes showed an overall high mean DNA methylation level across the whole gene with highest methylation levels in exons, genomic regions comprising expressed genes were characterized by particular high methylation levels in the first intron. Moreover, the patterns of both expressed and non-expressed genes were characterized by sharp transitions of methylation levels at exon–intron borders (Figure 2). Overall, when compared with transcriptional activity the patterns of DNA methylation across different parts of genes were similar to those recently determined in non-neoplastic tissues including blood.18

Finally, we studied the DNA methylation of the mitochondrial and EBV genomes of DAUDI cells.19–21 We estimated 80 EBV and 370 mitochondrial copies per DAUDI cell based on coverage analyses, which is in accordance with previous studies.22 Although mitochondrial DNA is mostly unmethylated (mean methylation 6.43%; Figure 1a), CpG methylation in the human and EBV genome is comparably distributed, although the EBV genome exhibits hardly any fully methylated sites (Figure 1a). Overall, EBV shows a high level of DNA methylation (mean methylation 80.18%), as it was previously shown for BL cell lines.23 Nevertheless, DNA methylation within the EBV genome correlates with expression only at high transcript levels (fragments per kilobase of exon model per million mapped (FPKM) ≥ 15) (Supplementary Figure S8).

In summary, we have characterized the nuclear DNA methylome of an eBL along with its mitochondrial and EBV methylome using colorspace BS-seq. We unravel significant differences between the different sub-methylomes and moreover show that gene transcription is associated with complex patterns of methylation, extending beyond simple promoter and CpG

Figure 1. Distribution of DNA methylation in DAUDI cells. (a) The graphs show genome-wide distributions of CpG methylation of the human nuclear, EBV and mitochondrial genomes. The y axis indicates DNA methylation levels assessed by SOLiD BS-seq. Green: human nuclear CpG methylation; red: EBV CpG methylation; blue: mitochondrial CpG methylation. (b) Significantly methylated non-CpG sites of DAUDI within RefSeq genes (comprising 424,969,306 non-CpGs) are 6.7-fold enriched compared with those outside of RefSeq genes (comprising 689,750,660 non-CpGs). Red: fraction of significantly methylated non-CpGs of DAUDI within RefSeq genes; blue: fraction of significantly methylated non-CpGs of DAUDI outside of RefSeq genes. OR: odds ratio.

Figure 2. Correlation of DNA methylation levels and transcriptional states. CpG methylation levels were averaged for annotated RefSeq gene regions and transcripts are clustered by their expression level in present (n = 7,662 transcripts) and absent (n = 5,429 transcripts) calls. A strong dependency of the location of CpGs related to their distance to the TSS and the transcript expression level can be observed. Green: average methylation pattern for present transcripts; red: average methylation pattern for absent transcripts.
methylome. As the DAUDI cell line has been used over decades in many laboratories in the world, the obtained methylome data might serve as a ‘reference epigenome’ for future studies.

Data availability: methylome data are available at sftp://134.245.63.215/export/home/daudi (login: daudi; password: daudismethylome2012).

CONFLICT OF INTEREST
OA and RS received research support from Illumina on BeadChip Analyses. CABC is and VLB used to be employed by Life Technologies. The remaining authors declare no conflict of interest.

ACKNOWLEDGEMENTS
This study was financed through a grant from the ‘Medizinueschuss Schleswig-Holstein’. The project received support through the BMBF IGC/C-ESGI project (FKZ 01KU1002A and 01KU1002E) and the EU Seventh Framework Programme (FP7/ 2007-2013, grant number 262055, ESGI), and infrastructure through the DFG Cluster of Excellence ‘Inflammation at Interfaces’. Bisulfite pyrosequencing was supported by KinderkrebsInitiative Buchholz/Holm-Seppensen and gene expression data were obtained in the framework of the Deutsche Krebshilfe-Network ‘Molecular Mechanisms in Malignant Lymphomas (MMML)’. This article as dedicated to Professor Lore Zech for her groundbreaking research on chromosomal aberrations in hematologic neoplasms in general and BL in particular.

AUTHOR CONTRIBUTIONS
BK, JR, OA, PR, AF and RS designed the research, interpreted data and wrote the paper. PR, AF, OA, SSc and RS provided infrastructure and obtained grant support. IV, EMMP, HGD and RAFM performed cell line characterization. MH obtained in the framework of the Deutsche Krebshilfe-Network ‘Molecular Mechanisms in Malignant Lymphomas (MMML)’. This article as dedicated to Professor Lore Zech for her groundbreaking research on chromosomal aberrations in hematologic neoplasms in general and BL in particular.

REFERENCES
1 Manolov G, Manolova Y. Marker band in one chromosome 14 from Burkitt lymphomas. Nature 1972; 237: 33–34.
2 Zech L, Haglund U, Nilsson K, Klein G. Characteristic chromosomal abnormalities in biopsies and lymphoid-cell lines from patients with Burkitt and non-Burkitt lymphomas. Int J Cancer 1976; 17: 47–56.
3 Salaverria I, Siebert R. The gray zone between Burkitt’s lymphoma and diffuse large B-cell lymphoma from a genetics perspective. J Clin Oncol 2011; 29: 1835–1848.
4 Boerma EG, Siebert R, Kluin PM, Baudis M. Translocations involving b2q4 in Burkitt lymphoma and other malignant lymphomas: a historical review of cytogenetics in the light of today’s knowledge. Leukemia 2009; 23: 225–234.
5 Richter J, Schlesner M, Hoffmann S, Kreuz M, Leich E, Burkhardt B et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. Nat Genet 2012; 44: 1316–1320.
6 Love C, Sun Z, Jima D, Li G, Zhang J, Miles R et al. The genetic landscape of mutations in Burkitt lymphoma. Nat Genet 2012; 44: 1321–1325.
7 Schmitz R, Young RM, Ceribelli M, Jhavar S, Xiao W, Zhang M et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. Nature 2012; 490: 116–20.
8 Martin-Subero JI, Kreuz M, Bibikova M, Bentink S, Ammerpohl O, Wickham-Garcia E et al. New insights into the biology and origin of mature aggressive B-cell lymphomas by combined epigenomic, genomic, and transcriptional profiling. Blood 2009; 113: 2488–2497.
9 Shaknovich R, Geng H, Johnson NA, Tsikitas L, Cerchietti L, Greally JM et al. DNA methylolation signatures define molecular subtypes of diffuse large B-cell lymphoma. Blood 2010; 116: e81–e89.
10 Shaknovich R, Cerchietti L, Tsikitas L, Kornaksson M, De S, Figueroa ME et al. DNA methylolationtransferase 1 and DNA methylolation patterning contribute to germinal center B-cell differentiation. Blood 2011; 118: 3559–3569.
11 Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 2009; 462: 315–322.
12 Klein E, Klein G, Nadkarni JS, Nadkarni JJ, Wiggell H, Clifford P. Surface IgM-kappa specificity on a Burkitt lymphoma cell in vivo and in derived culture lines. Cancer Res 1968; 28: 1300–1310.
13 Nadkarni JS, Nadkarni JJ, Clifford P, Manolv G, Fenyö EM, Klein E. Characteristics of new cell lines derived from Burkitt lymphomas. Cancer 1969; 23: 64–79.
14 Sander S, Calado DP, Sinivasan L, Köchert K, Zhang B, Rosolowski M et al. Synergy between PI3K signaling and MYC in Burkitt lymphomasigenes. Cancer Cell 2012; 22: 167–179.
15 Krueger F, Kreck B, Franke A, Andrews SR. DNA methylation analysis using short bisulfite sequencing data. Nat Methods 2012; 9: 145–151.
16 Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006; 126: 663–676.
17 Hummel M, Bentink S, Berger H, Klapper W, Weisendorf S, Barth TF et al. A biologic definition of Burkitt’s lymphoma from transcriptional and genomic profiling. N Engl J Med 2006; 354: 2419–2430.
18 Li Y, Zhu J, Tian G, Li N, Li Q, Ye M et al. The DNA methylolation of human peripheral blood mononuclear cells. PLoS Biol 2010; 8: e1000533.
19 Lin Z, Xu G, Deng N, Taylor C, Zhu D, Flemington EK. Quantitative and qualitative RNA-seq-based evaluation of Epstein-Barr virus transcription in type I latency Burkitt lymphomas. Curr Biol 2010; 20: 1350S–1358S.
20 Flower K, Thomas D, Heather J, Ramasubramanyan S, Jones S, Sinclair AJ. Epi-genetic control of viral life-cycle by a DNA-methylation dependent transcription factor. PloS ONE 2011; 6: e25992.
21 Desangs LR, Teuton JR, Feng H, Jacobs JM, Camp 2nd DG, Purvine SO et al. Coupled transcriptome and proteome analysis of human lymphotropic tumor viruses: insights on the detection and discovery of viral genes. BMC Genomics 2011; 12: 625.
22 Leenman EE, Panzer-Gru¨ mayer RE, Fischer S, Leitch HA, Horsman DE, Lion T et al. Rapid determination of Epstein-Barr virus latent or lytic infection in single human cells using in situ hybridization. Mod Pathol 2004; 17: 1564–1572.
23 Kim DN, Song YJ, Lee SK. The role of promoter methylation in Epstein-Barr virus (EBV) microRNA expression in EBV-infected B cell lines. Exp Mol Med 2011; 43: 401–410.

This work is licensed under a Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by/3.0/

Supplementary Information accompanies the paper on the Leukemia website (http://www.nature.com/leu)