



Data in Brief

Genome-wide chromatin accessibility, DNA methylation and gene expression analysis of histone deacetylase inhibition in triple-negative breast cancer



Matias A. Bustos^{a,1}, Matthew P. Salomon^{a,1}, Nellie Nelson^b, Sandy C. Hsu^b, Maggie L. DiNome^c, Dave S.B. Hoon^{a,b}, Diego M. Marzese^{a,*}

^a Department of Translational Molecular Medicine, John Wayne Cancer Institute at Providence Saint John's Health Center, Santa Monica, CA 90404, USA

^b Sequencing Center, John Wayne Cancer Institute at Providence Saint John's Health Center, Santa Monica, CA 90404, USA

^c Department of Surgery, David Geffen School of Medicine, University of California Los Angeles (UCLA), CA, USA

ARTICLE INFO

Article history:

Received 3 January 2017

Accepted 25 January 2017

Available online 14 February 2017

ABSTRACT

Triple-negative breast cancer (TNBC), especially the subset with a basal phenotype, represents the most aggressive subtype of breast cancer. Unlike other solid tumors, TNBCs harbor a low number of driver mutations. Conversely, we and others have demonstrated a significant impact of epigenetic alterations, including DNA methylation and histone post-translational modifications, affecting TNBCs. Due to the promising results in pre-clinical studies, histone deacetylase inhibitors (HDACi) are currently being tested in several clinical trials for breast cancer and other solid tumors. However, the genome-wide epigenetic and transcriptomic implications of HDAC inhibition are still poorly understood. Here, we provide detailed information about the design of a multi-platform dataset that describes the epigenomic and transcriptomic effects of HDACi. This dataset includes genome-wide chromatin accessibility (assessed by ATAC-Sequencing), DNA methylation (assessed by Illumina HM450K BeadChip) and gene expression (assessed by RNA-Sequencing) analyses before and after HDACi treatment of HCC1806 and MDA-MB-231, two human TNBC cell lines with basal-like phenotype.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Specifications	
Organism/cell line/tissue	<i>Homo sapiens</i> /cancer cell lines (HCC1806 and MDA-MB-231)
Sex	Female
Sequencer or array type	- Illumina HiSeq 2500 - Illumina HumanMethylation450 BeadChip
Data format	Raw (FASTQ) and analyzed (BigWig) for ATAC sequencing data Raw (FASTQ) and analyzed (TXT) for RNA sequencing data Raw (IDAT) and analyzed (TXT) for DNA methylation arrays data
Experimental factors	Human triple-negative cell lines treated with HDAC inhibitor (LBH589)
Experimental features	HCC1806 and MDA-MB-231 treated with 10 nM of LBH589 for 28 days
Consent	N/A
Sample source location	N/A

* Corresponding author at: Department of Translational Molecular Medicine, John Wayne Cancer Institute, 2200 Santa Monica Blvd, Santa Monica, CA 90404, USA.

E-mail address: marzese@jwci.org (D.M. Marzese).

¹ These authors (MAB and MPS) contributed equally to this work.

1. Direct link to deposited data

Study SuperSeries (GSE92911) Access: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=mlodeuqanxslvmz&acc=GSE92911>
 ATAC-Sequencing SubSeries (GSE92898) Access: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=qjudwycgtfsfnuv&acc=GSE92898>
 RNA-Sequencing SubSeries (GSE92910) Access: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=svsnwcvzudffif&acc=GSE92910>
 DNA methylation SubSeries (GSE92909) Access: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=grgxqoouphyjhmz&acc=GSE92909>

2. Experimental design and cell culture conditions

HCC1806 (cat#CRL-2335; ATCC, Manassas, VA) and MDA-MB-231 (cat#HTB-26; ATCC, Manassas, VA) cells were cultured in RPMI 1640 supplemented with 10 mM HEPES, 10% heat-inactivated fetal bovine serum (FBS) and 1% penicillin-streptomycin (complete medium)

containing 10 nM of LBH589 (Panobinostat; cat#S1030; Selleck Chemicals, Houston, TX) during 120 h followed by 48 h of incubation in RPMI-1640 medium alone. To maintain the drug concentration, the medium was refreshed every 24 h. This cycle was repeated 4 times to complete 28 days of treatment. In parallel, HCC1806 and MDA-MB-231 cells were cultured in complete medium containing drug vehicle (DMSO) alone for the same period of time (28 days). RNA-Sequencing (RNA-Seq), Assay for Transposase-Accessible Chromatin Sequencing (ATAC-Seq), and DNA methylation arrays were performed after the 4 cycles. Finally, to evaluate the stability of HDACi-induced epigenetic reprogramming, HCC1806 and MDA-MB-231 treated during 28 days with 10 nM of LBH589 were cultured in complete medium alone for 14 additional days, and ATAC-Seq was performed. Gene expression and DNA methylation profiling were also performed after the initial 96 h of incubation to evaluate early response to 10 nM of LBH589. A description of the cell lines analyzed, the regimens used and the data analyses performed are summarized in Table 1.

3. Assay for Transposase-accessible chromatin sequencing (ATAC-Seq)

ATAC-Seq was performed as previously described by Buenrostro et al. [1]. Briefly, 50,000 cells were washed in 50 μ l cold PBS and resuspended in 50 μ l cold lysis buffer [1]. The nuclei were centrifuged at 500 \times g for 10 min at 4 $^{\circ}$ C. The pellet was then incubated with transposition reaction mix for 30 min at 37 $^{\circ}$ C and then purified using the MinElute PCR Purification kit (Qiagen, Hilden, Germany). The eluted transposed DNA was barcoded and amplified for 5 cycles, and then 5 μ l of the product was used for a qPCR side-reaction to prevent amplification saturation. The optimal amount of additional PCR cycles was determined based on the number of cycles that represented one-third of maximum fluorescence intensity in the qPCR reaction after 20 cycles. The final transposed DNA was amplified for a total of 9–11 cycles. The final amplified library was purified using MinElute PCR Purification kit and eluted in 20 μ l elution buffer. The ATAC libraries were sequenced on the Illumina HiSeq 2500 in Rapid Mode using 50 bp paired-end at the John Wayne Cancer Institute (JWCI) Sequencing Center.

4. ATAC-Seq data processing

Raw genomic sequence reads were mapped to the 1000 Genomes (b37) build of the human genome reference using BWA-MEM (version 0.7.5a) with default settings [2]. Alignments were further processed using GATK (version 2.8–1) [3] for INDEL realignment and PicardTools for duplicate marking (version 1.103) (<http://broadinstitute.github.io/picard/>). ATAC-Seq peaks were identified using the callpeaks function in MACS2 [4] with a threshold set to $-q = 0.01$. The resulting peak calls were filtered for sequences that mapped to the mitochondria using shell scripts, and BigWig files were generated using bedGraphToBigWig [5].

Table 1

Summary of the data analyses included in the dataset.

Name	Treatment	ATAC-Seq	HM450K	RNA-Seq
HCC1806-DMSO-96hs	DMSO for 96 h	x	x	✓
HCC1806-LB-96hs	LBH589 (10 nM) for 96 h	x	✓	✓
HCC1806-DMSO-4w	DMSO for 28 days	✓	✓	✓
HCC1806-LB-4w	LBH589 (10 nM) for 28 days	✓	✓	✓
HCC1806-LB-4w + complete medium-2w	LBH589 (10 nM) for 28 days + complete medium for 14 days	✓	x	x
MDA-MB-231-DMSO-96hs	DMSO for 96 h	x	x	✓
MDA-MB-231-LB-96hs	LBH589 (10 nM) for 96 h	x	✓	✓
MDA-MB-231-DMSO-4w	DMSO for 28 days	✓	✓	✓
MDA-MB-231-LB-4w	LBH589 (10 nM) for 28 days	✓	✓	✓
MDA-MB-231-LB-4w + complete medium-2w	LBH589 (10 nM) for 28 days + complete medium for 14 days	✓	x	x

5. RNA sequencing (RNA-Seq)

Total RNA from HCC1806 and MDA-MB-231 cells was extracted using ZR-Duet DNA/RNA MiniPrep Plus (cat# D7003; Zymo Research, Irvine, CA). RNA samples with high quality (RIN \geq 8.0) and high purity (OD 260/280 = 1.8–2.0) scores were used to generate libraries using the Illumina TruSeq RNA Sample Preparation Kit v2 (Illumina Inc., San Diego, CA). The mRNA libraries were sequenced on the Illumina HiSeq 2500 in Rapid Mode using 101 bp paired-end reads at the JWCI Sequencing Center [6].

6. RNA-Seq data processing

Base calling and de-multiplexing were processed using CASAVA v1.8 (Illumina Inc., San Diego, CA), reads were mapped to the GENCODE release 19 reference using STAR version 2.4.2a [7], and read counts were generated using the `-quantMode GeneCounts` option in STAR. The Bioconductor package DESeq2 was used to detect fold change differences in expression between conditions.

7. DNA methylation profiling (HM450K)

Genomic DNA was extracted from HCC1806 and MDA-MB-231 cells using the Quick-gDNA MiniPrep kit (cat# D3025; Zymo Research, Irvine, CA). Following, 1 μ g of DNA was sodium bisulfite modified (SBM) using the EZ DNA Methylation-Direct kit (cat# D5021; Zymo Research, Irvine, CA). 200 ng of SBM-DNA was whole-genome amplified, enzymatically fragmented and hybridized on the HumanMethylation450 (HM450K) BeadChip (Illumina Inc., San Diego, CA). The chips were scanned with Illumina iScan (Illumina, Inc., San Diego, CA), as we previously described [8–11].

8. DNA methylation data processing

Data was extracted using the R package *methylumi*. The `'noob'` function in the R package *minfi* was used to process the data and then the `'dasen'` function in the R package *wateRmelon* for normalization and dye-bias correction. DNA methylation levels were reported as β -value ($\beta = \text{intensity of the methylated allele} / (\text{intensity of the unmethylated allele} + \text{intensity of the methylated allele})$) and calculated using the signal intensity value for each CpG site.

9. Discussion

In this article, we described a dataset involving transcriptomic and epigenomic analysis of TNBC cells that were exposed for a prolonged time to low doses of a clinically-approved HDACi. While the functional consequences of this treatment approach remains to be elucidated in the future, this multi-platform evaluation provides an opportunity to deeply understand the genome-wide effects of this powerful epigenetic drug and to identify novel transcriptomic and epigenomic clinically-relevant targets for TNBC tumors.

Funding

This work was supported by the Associates for Breast and Prostate Cancer Studies (ABCs) (Grant ID: 887377001-40000) award (M.B., D.H. and D.M.); the Fashion Footwear Association of New York (FFANY) (Grant ID: 88737700160000) foundation (M.D. and D.H.); the Margie and Robert E. Petersen Foundation (D.H.) (Grant ID: 88737401180001); and the AVON Foundation Breast Cancer Crusade (D.H. and D.M.) (Grant ID: 02-2015-061).

References

- [1] J.D. Buenrostro, P.G. Giresi, L.C. Zaba, H.Y. Chang, W.J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10 (2013) 1213–1218.
- [2] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25 (2009) 1754–1760.
- [3] A. McKenna, M. Hanna, E. Banks, et al., The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20 (2010) 1297–1303.
- [4] Y. Zhang, T. Liu, C.A. Meyer, et al., Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9 (2008) R137.
- [5] W.J. Kent, A.S. Zweig, G. Barber, A.S. Hinrichs, D. Karolchik, BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* 26 (2010) 2204–2207.
- [6] L. Lessard, M. Liu, D.M. Marzese, et al., The CASC15 long Intergenic noncoding RNA locus is involved in melanoma progression and phenotype switching. *J. Invest. Dermatol.* 135 (2015) 2464–2474.
- [7] A. Dobin, C.A. Davis, F. Schlesinger, et al., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29 (2013) 15–21.
- [8] D.M. Marzese, J.L. Huynh, N.P. Kwas, D.S. Hoon, Multi-platform genome-wide analysis of melanoma progression to brain metastasis. *Genom. Data* 2 (2014) 150–152.
- [9] D.M. Marzese, M. Liu, J.L. Huynh, et al., Brain metastasis is predetermined in early stages of cutaneous melanoma by CD44v6 expression through epigenetic regulation of the spliceosome. *Pigment Cell Melanoma Res.* 28 (2015) 82–93.
- [10] D.M. Marzese, R.A. Scolyer, M. Roque, et al., DNA methylation and gene deletion analysis of brain metastases in melanoma patients identifies mutually exclusive molecular alterations. *Neuro-Oncology* 16 (2014) 1499–1509.
- [11] D.M. Marzese, R.A. Scolyer, J.L. Huynh, et al., Epigenome-wide DNA methylation landscape of melanoma progression to brain metastasis reveals aberrations on homeobox D cluster associated with prognosis. *Hum. Mol. Genet.* 23 (2014) 226–238.