Genomics Data 3 (2015) 33–35

Contents lists available at ScienceDirect

Genomics Data

journal homepage: http://www.journals.elsevier.com/genomics-data/

Data in Brief

Genome-wide redistribution of BRD4 binding sites in transformation resistant cells

Han Si [a,1], Paola Scaffidi [b,2], Anand Merchant [a], Maggie Cam [c], Eric Stahlberg [a], Tom Misteli [b,*], Patricia Fernandez [b,*]

[a] CCRIFX, National Cancer Institute, NIH, Bethesda, MD 20892, United States
[b] Cell Biology of Genomes Group, LRBGE, National Cancer Institute, NIH, Bethesda, MD, 20892, United States
[c] OSTR, National Cancer Institute, NIH, Bethesda, MD, 20892, United States

ARTICLE INFO

Article history:
Received 31 October 2014
Received in revised form 11 November 2014
Accepted 11 November 2014
Available online 18 November 2014

Keywords:
Progeria
BRD4
ChIP-sequencing
Tumor protection

ABSTRACT

Hutchinson–Gilford progeria syndrome (HGPS) patients do not develop cancer despite a significant accumulation of DNA damage in their cells. We have recently reported that HGPS cells are refractory to experimental oncogenic transformation and we identified the bromodomain-containing 4 protein (BRD4) as a mediator of the transformation resistance. ChIP-sequencing experiments revealed distinct genome-wide binding patterns for BRD4 in HGPS cells when compared to control wild type cells. Here we provide a detailed description of the ChIP-seq dataset (NCBI GEO accession number GSE61325), the specific and common BRD4 binding sites between HGPS and control cells, and the data analysis procedure associated with the publication by Fernandez et al., 2014 in Cell Reports 9, 248-260 [1].

Specifications	
Organism/cell line/tissue	Homo sapiens/dermal fibroblast cell lines expressing transforming factors from HGPS and age-matched control individuals
Sex	Male and female cell lines
Sequencer or array type	Illumina HiSeq 2000
Data format	Fastq
Experimental factors	HGPS vs. wild type
Experimental features	ChIP-seq, transcription factor binding sites, peak diffing
Consent	NA
Sample source location	Original cell lines were obtained from Coriell Institute

Direct link to deposited data

Deposited data can be found at: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE61325.

Experimental design, materials and methods

The overall aim was to identify specific and common genomic binding sites for BRD4 in experimentally transformed dermal fibroblasts derived from HGPS patients and control individuals through the analysis of BRD4 ChIP-sequencing data.

Cell culture

Transformed cell lines were obtained by retroviral introduction of TERT (T), V12-HRAS (R) and SV40 large and small T antigens (S) of primary dermal fibroblasts from HGPS patients (TRS-HGPS) and age-matched control wild-type individuals (TRS-WT) as previously described [1–3]. Cells were grown in MEM containing 15% fetal bovine serum (FBS), 2 mM L-glutamine, 100 U ml^{-1} penicillin and 100 μgml^{-1} streptomycin, at 37 °C in 5% CO_2. Two independent cell lines from each group were selected for ChIP-sequencing experiments.

ChIP-sequencing

ChIP was performed as previously described [4] with modifications. Briefly, 10×10^6 cells were cross-linked for 10 min with formalin 1% at room temperature, followed by 5 minute quenching with 125 mM glycine. After three washes with chilled PBS, cell pellets were snap-frozen and stored at −80 °C. Cells were resuspended in lysis buffer containing 1 mM EDTA, 0.8% SDS, 20 mM Tris–HCl pH 8 and protease inhibitor

* Corresponding authors.
E-mail addresses: mistelit@mail.nih.gov (T. Misteli), fernandezferrp@mail.nih.gov (P. Fernandez).
[1] Present address: Laboratory of Immunology, National Eye Institute, NIH, Bethesda 20892, United States.
[2] Present address: Cancer Epigenetics Laboratory, Cancer Research UK London Research Institute, Lincoln's Inn Fields Laboratories, London WC2A 3PX, UK and UCL Cancer Institute, University College London, London WC1E 6DD, UK.

http://dx.doi.org/10.1016/j.gdata.2014.11.005

cocktail (Calbiochem) and sonicated in a Diagenode Bioruptor 300 for 12–15 cycles (30 s on, 60 s off). Chromatin was diluted in ChIP dilution buffer (1% Triton X-100, 1 mM EDTA, 20 mM Tris–HCl pH 8, 150 mM NaCl and protease inhibitor cocktail) and immunoprecipitated overnight with pre-coated anti-IgG magnetic beads (Dynabeads, Invitrogen) previously incubated with anti-BRD4 antibody (7 μg, Bethyl labs, lot A301-985) for 6 h at 4 °C. Beads were washed sequentially for 5 min each in low-salt (20 mM Tris–HCl pH 8, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), high-salt (20 mM Tris–HCl pH 8, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), LiCl (10 mM Tris pH 8.0, 1 mM EDTA, 250 mM LiCl, 1% NP-40, 1% Na-deoxycholate) and TE (10 mM Tris–HCl pH 8, 1 mM EDTA) buffer at 4 °C followed by a room temperature wash in TE buffer. Beads were eluted in 1% SDS and 100 mM $NaHCO_3$ buffer for 25 min at 65 °C and cross-linking was reversed for 8 h after addition of NaCl (final concentration 200 mM). After RNase A (1 mg/ml, Qiagen) and proteinase K (2 mg/ml, New England Biolabs) incubations, DNA was column-purified following the manufacturer's instructions (Qiaquick PCR purification kit, Qiagen).

Library construction (Illumina TruSeq 2.0 protocol) and sequencing were performed at the NCI-Sequencing Facility (SAIC-Frederick, MD) using Illumina HiSeq 2000. Paired-end sequencing with read lengths of 100 bp was performed. 4 input samples and 4 BRD4-ChIP samples (from TRS-WT and TRS-HGPS cells, two independent cell lines in each group) were multiplexed in two lanes. Raw reads files in fastq format were deposited to the SRA database at NCBI (Table 1).

ChIP-seq data analysis

The data analysis procedure is described in Fig. 1. The raw reads data quality was tested using FastQC [5] (V0.10.1). All samples showed all base quality ≥25 and a small number of TruSeq adapter sequences was detected in the WT1 BRD4 reads. Raw reads of each sample in fastq format were then aligned to reference human genome (hg19) using NovoalignMPI (V3.0.2, Novocraft) and the adapter sequence contamination in WT1 BRD4 sample was removed during mapping using "-a" option in command line. In addition, the program was instructed to only report the uniquely aligned reads by skipping the "-r" option. The mapping results in BAM format were preprocessed using Picard tools [6] (V1.119) to reorder, sort reads and the duplicated reads were removed as well. The clean bam files were in turn converted to bed format using "bamToBed" script in bedtools [7] (V2.19) in order to be analyzed using SICER [8] (V1.1). Given the reported binding nature of BRD4 [9], 200 bp window size and 400 bp gap size were chosen in the peak calling step using SICER. Statistically significant peaks were detected using 1% FDR as the cutoff. Peak numbers from each sample are shown in Table 2.

Peak signals were visualized using Integrative Genomics Viewer (IGV) [10] for basic quality check; in addition, DiffBind [11] was applied to the peak data to examine the correlation of WT and HGPS samples using all default settings. In order to investigate potential differences in BRD4 binding patterns in WT and HGPS samples, a pipeline was developed to detect sample-specific BRD4 bindings. First, the union of the peaks in each biological group was calculated using the "find_union_islands.py" script in SICER, and then the intersection of peaks in each group was achieved using the "intersect" command in bedtools, and finally, the sample-specific peaks were derived by comparing the union peak sets with the intersected peak sets from different samples using "-v" option in the "intersect" command. In this way, 3078 WT-specific and 5574 HGPS-specific BRD4 binding sites were detected and the annotation of the peaks was done using "annotatePeaks.pl" with default options in Homer [12]. Important genomic features e.g., promoter, 5′/3′ UTR, exon, intron, non-coding were used to annotate the peaks and differential enrichment in specific features were compared between wild type and HGPS samples, as interpreted in [1]. Common BRD4 binding sites were detected by "multiinter" command in bedtools and annotated using Homer. All annotated specific and common genomic binding sites between WT and HGPS cell lines are provided in Tables 3–5.

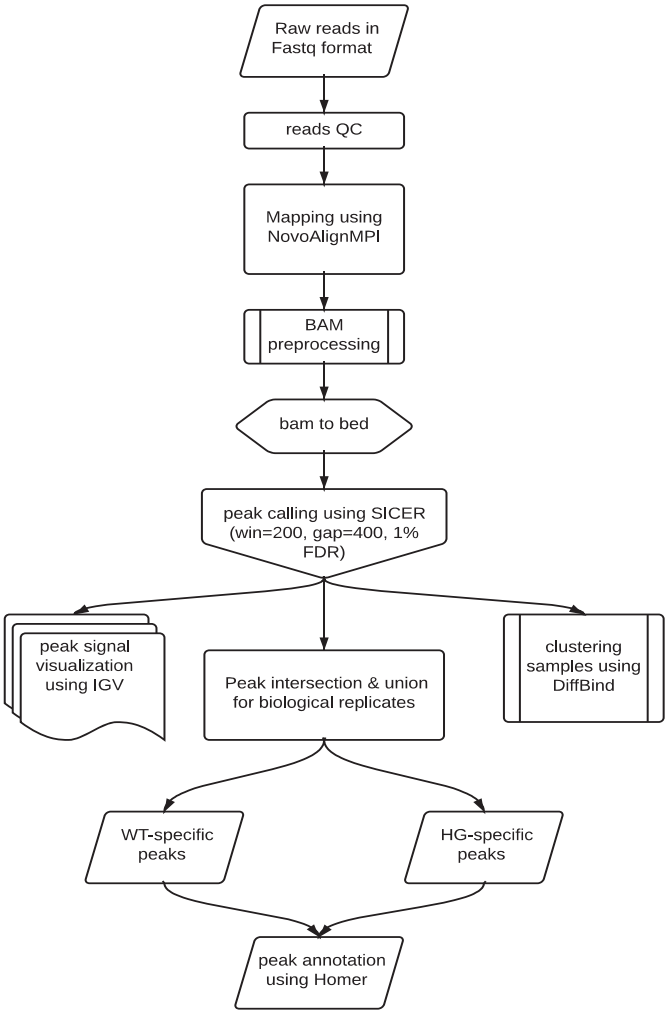

Fig. 1. Flow chart of data analysis.

Table 1
Sample names and their corresponding raw read file names in NCBI SRA archive database.

File names	Sample names	Group
SRR1574701	HG1 Input	TRS-HGPS cells
SRR1574702	HG2 Input	TRS-HGPS cells
SRR1574697	HG1 BRD4	TRS-HGPS cells
SRR1574698	HG2 BRD4	TRS-HGPS cells
SRR1574699	WT1 Input	TRS-WT cells
SRR1574700	WT2 Input	TRS-WT cells
SRR1574695	WT1 BRD4	TRS-WT cells
SRR1574696	WT2 BRD4	TRS-WT cells

Table 2
Number of peaks from each sample detected by SICER.

Samples	WT1	WT2	HG1	HG2
Peaks	29990	32239	35292	42784

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gdata.2014.11.005.

Acknowledgements

This work was supported by the Intramural Research Program of the National Institutes of Health (NIH), NCI, Center for Cancer Research.

References

[1] P. Fernandez, et al., Transformation resistance in a premature aging disorder identifies a tumor-protective function of BRD4. Cell Rep. 9 (1) (2014) 248–260.
[2] W.C. Hahn, et al., Creation of human tumour cells with defined genetic elements. Nature 400 (6743) (1999) 464–468.
[3] P. Scaffidi, T. Misteli, In vitro generation of human cells with cancer stem cell properties. Nat. Cell Biol. 13 (9) (2011) 1051–1061.
[4] R.F. Luco, et al., Regulation of alternative splicing by histone modifications. Science 327 (5968) (2010) 996–1000.
[5] FastQC, Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
[6] Picard, Available from. http://broadinstitute.github.io/picard/.
[7] A.R. Quinlan, BEDTools: the Swiss-army tool for genome feature analysis. Current Protocols in Bioinformatics, John Wiley & Sons, Inc., 2002
[8] C. Zang, et al., A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics 25 (15) (2009) 1952–1958.
[9] W. Zhang, et al., Bromodomain-containing protein 4 (BRD4) regulates RNA polymerase II serine 2 phosphorylation in human CD4+ T cells. J. Biol. Chem. 287 (51) (2012) 43137–43155.
[10] J.T. Robinson, et al., Integrative genomics viewer. Nat. Biotechnol. 29 (1) (2011) 24–26.
[11] C.S. Ross-Innes, et al., Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481 (7381) (2012) 389–393.
[12] S. Heinz, et al., Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38 (4) (2010) 576–589.