Genomics Data 2 (2014) 170–172

Contents lists available at ScienceDirect

Genomics Data

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

Data in Brief

Genome-wide binding of transcription factors in inv(16) acute myeloid leukemia

A. Mandoli, K. Prange, J.H.A. Martens *

Department of Molecular Biology, Faculty of Science, Nijmegen Centre for Molecular Life Sciences, Radboud University, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 22 May 2014
Accepted 11 June 2014
Available online 20 June 2014

Keywords:
ChIP-seq
RNA-seq
inv(16) AML
Transcription factor

ABSTRACT

The inv(16) translocation is associated with 5% of AML cases and gives rise to expression of the oncofusion protein CBFβ-MYH11. Although different molecular mechanisms for the oncogenic activity of this fusion protein have been proposed these were mostly based on *in vitro* experiments or single loci analysis. Recently, we investigated the genome-wide action of this fusion protein in the context of other hematopoietic transcription factors (Mandoli et al., 2014). Here, we describe in detail the ChIP-seq and RNA-seq methods used to generate the data associated with this study. Our analysis of CBFβ-MYH11 as well as multiple other hematopoietic transcription factors using ChIP-seq data revealed RUNX1 dependent binding of CBFβ-MYH11 as well as interaction of the RUNX1/CBFβ-MYH11 complex with other hematopoietic regulators. Further RNA-seq based analysis suggested that CBFβ-MYH11 can act both as activator and repressor.

Specifications	
Organism/cell line/tissue	*Homo sapiens*/ME-1, inv(16) patients, U937
Sex	Male
Sequencer or array type	Illumina
Data format	Analyzed
Experimental factors	inv(16) Acute myeloid leukemia cell line, inv(16) AML model cell lines and inv(16) AML primary patient cells
Experimental features	ChIP seq and RNA seq
Consent	All patients gave their consent
Sample source location	Groningen, The Netherlands

Direct link to deposited data

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE46044

Experimental design, materials and methods

Cell culture and patient samples

In this study we used ChIP-seq to generate genome-wide binding profiles of proteins in 3 cell types: ME-1 cells harboring the inv(16) translocation and expressing the CBFβ-MYH11 oncofusion protein [5]; an inducible U937 cell line (U937CM) that on induction expresses CBFβ-MYH11[1]; and a mononuclear $CD34^+$ inv(16) AML blast population isolated from peripheral blood of a *de novo* AML patient which was studied after informed consent was obtained in accordance with the Declaration of Helsinki. In addition, an inducible knockdown cell line was generated by cloning of CBFβ-MYH11 specific shRNA in a FH1tUTG lentiviral construct as described previously [2]. For this, lentiviral particles were produced in Cos-7 cells after which filtered and concentrated viral supernatants were used to infect ME-1 cells. GFP positive cells were sorted by FACS and shRNA expression was induced by addition of doxycycline.

The ME-1 and U937-Tet-off CBFβ-MYH11 cells were cultured in RPMI 1640 supplemented with 10% FCS and 1% penicillin/streptomycin at 37 °C in 5% CO_2. For conditional expression of CBFβ-MYH11, U937 cells were washed 5 times in 50 ml phosphate-buffered saline (PBS) and seeded at a density of 2×10^5 cells/ml in the absence of tetracycline. The increase of CBFβ-MYH11 expression was detected by RT-PCR and Western blot. The different cell types were used for ChIP-seq experiments using different antibodies as outlined in Supplementary Table 1. In addition, for the cell lines RNA-seq data was generated to determine the correlation between transcription factor binding and gene expression.

Chromatin immunoprecipitation (ChIP)

For ChIP 25 million cells were crosslinked with 1% formaldehyde for 20 min at room temperature with trembling. The reaction was stopped by adding 0.125 M glycine and cells were centrifuged and washed once with PBS. Cell pellets were suspended in 50 ml of buffer A (0.25% Triton

* Corresponding author. Tel.: +31 24 3610525.
E-mail address: j.martens@ncmls.ru.nl (J.H.A. Martens).

http://dx.doi.org/10.1016/j.gdata.2014.06.014

X 100, 10 mM EDTA, 0.5 mM EGTA, 20 mM HEPES pH 7.6) and incubated at 4 °C for 10 min with rotation. The suspension was centrifuged and the pellet was resuspended in 50 ml of buffer B (150 mM NaCl, 10 mM EDTA, 0.5 mM EGTA, 20 mM HEPES pH 7.6). After 10 min rotation at 4 °C the suspension was centrifuged and resuspended in incubation buffer (0.15% SDS, 1% Triton X 100, 150 mM NaCl, 10 mM EDTA, 0.5 mM EGTA, 20 mM HEPES pH 7.6 and protease inhibitor cocktail). Chromatin was sonicated using the bioruptor (Diagenode) for 20 min at high power, 30 s ON, 30 s OFF. Sonication is an important step of the ChIP procedure and it varies for different cell types and sonicators. To find the proper number of cycles for sonication, *i.e.* to obtain a fragment size range between 150 and 500 base pairs, we first performed sonication on a fraction of the chromatin suspension. In our case, 20 cycles were optimal for ME-1 cells, and 16 cycles were optimal for U937 cells.

Sonicated chromatin was centrifuged at 13,000 rpm for 10 min at 4 °C and supernatant was collected for ChIP. For every ChIP 100 μl of chromatin (corresponding to 1.25×10^6 cells) was incubated overnight at 4 °C with rotation in incubation buffer supplemented with 0.1% BSA, protease inhibitor cocktail, 30 μl of protein A/G-sepharose beads (Santa Cruz), and 4 μg of antibody (see Supplementary table). Beads were washed sequentially with four different wash buffers at 4 °C: two times with buffer 1 (0.1% SDS, 0.1% DOC, 1% Triton, 150 mM NaCl, 10 mM Tris pH 8, 0.1 mM EDTA, and 0.5 mM EGTA), one time with buffer 2 (0.1% SDS, 0.1% DOC, 1% Triton, 500 mM NaCl, 10 mM Tris pH 8, 0.1 mM EDTA, and 0.5 mM EGTA), one time with buffer 3 (0.25 M LiCl, 0.5% DOC, 0.5% NP-40, 10 mM Tris pH 8, 0.1 mM EDTA, and 0.5 mM EGTA), and two times with buffer 4 (10 mM Tris pH 8, 0.1 mM EDTA, and 0.5 mM EGTA). The immunoprecipitated chromatin was eluted from the beads by adding 200 μl of elution buffer (1% SDS, 0.1 M $NaHCO_3$, 10 μg/ml RNase and 250 μg/ml proteinase K) and rotated for 20 min at RT. Chromatin was extracted by centrifugation at maximum speed for 2 min and protein-DNA crosslinks were reversed at 65 °C for 4 h in the presence of 200 mM NaCl, after which DNA was purified with the QIAquick PCR purification kit (Qiagen). For each ChIP-seq, 4–5 ChIPs were pooled using a single column, DNA was eluted in 35 μl of elution buffer and measured with the Qubit dsDNA HS Assay kit (Invitrogen).

The quality of ChIPs was checked using an aliquot of the eluted DNA in quantitative PCR (qPCR). Relative occupancy at putative CBFβ-MYH11 binding regions was calculated as fold over background, for which the promoter of the H2B gene was used (Fig. 1 and Supplementary Table 1 for primer sequences).

Strand specific RNA sequencing

For RNA-seq total RNA was extracted from ME-1 cells using the RNeasy kit and on-column DNase treatment (Qiagen). The concentration of RNA was measured with the Qubit fluorometer (Invitrogen). 250 ng of total RNA was used in the Ribo-Zero rRNA Removal Kit (Epicentre) to remove ribosomal RNAs according to manufacturer instructions. 16 μl of purified RNA was fragmented by addition of 4 μl of 5× fragmentation buffer (200 mM Tris acetate pH 8.2, 500 mM potassium acetate, and 150 mM magnesium acetate), incubated at 94 °C for exactly 90 s and placed on ice immediately afterwards. Fragmented RNA was precipitated by adding 2 μl NaAC (3 M), 2 μl glycogen (10 mg/ml) and 60 μl of absolute ethanol in 1.5 ml tubes. After mixing by inverting, tubes were incubated at −80 °C for at least 1 h. Next, tubes were spun at 4 °C for 30 min and supernatant was discarded. RNA pellet was washed once with 70% ethanol, air dried for 5 min and dissolved in 10 μl of nuclease free H_2O.

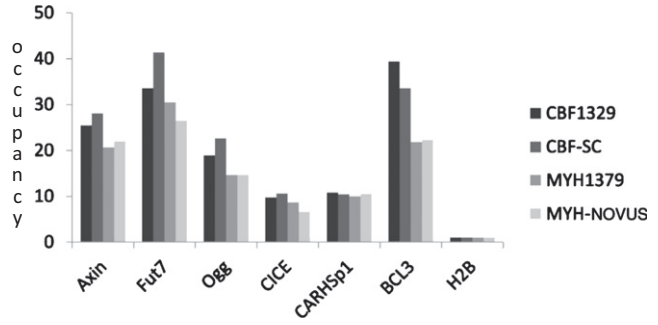

Fig. 1. ChIP-qPCR using CBFβ and MYH11 antibodies in ME-1 cells. Five putative binding sites of CBFβ and CBFβ-MYH11 were selected and occupancies at these sites were determined by qPCR. Data was normalized using H2B as a negative control.

5 μg random hexamers were added to the RNA followed by incubation at 70 °C for 10 min and chilling on ice. Next, first-strand cDNA was synthesized using a RNA primer mix and adding 4 μl of5× first-strand buffer, 1 μl of RNase inhibitor (Promega # N2618), 2 μl of 100 mM DTT, 1 μl of 10 mM dNTPs, 132 ng of actinomycin D, and 200 U SuperScript III (Invitrogen), followed by 2 hours incubation at 48 °C. First strand cDNA was purified using Qiagen mini elute columns to remove dNTPs and eluted two times with 17 μl of elution buffer. Second-strand cDNA was synthesized by adding 91.8 μl of a mix containing 5 μg random hexamers, 4 μl of 5× first-strand buffer, 2 μl of 100 mM DTT, 4 μl of 10 mM dNTPs with dTTP replaced by dUTP, 30 μl of 5× second-strand buffer, 40 U of *Escherichia coli* DNA polymerase (NEB # M02095S), 10 U of *E. coli* DNA ligase (NEB # M0205L) and 2 U of *E. coli* RNase H (Ambion # AM2293), and incubated at 16 °C for 2 h. This was followed by addition of 10 U T4 polymerase (NEB # M0203) and incubation at 16 °C for 10 min. Double stranded cDNA was purified using Qiagen mini elute columns and an adenosine base was added to the ends as described below for DNA obtained from ChIP experiments to facilitate ligation. All the subsequent steps are the same for the RNA-seq and ChIP-seq sample prep protocol (see below) other than the addition of 1 U USER enzyme (NEB # M5505L) to RNA-seq samples before PCR and incubation at 37 °C for 15 min, followed by 5 min at 95 °C.

Illumina high throughput sequencing

For library preparation, 2–10 ng of ChIP DNA was used. End repair was performed using 40 μl of ChIP DNA, 5 μl T4 DNA ligase buffer with 10 mM ATP, 2 μl dNTP mix (10 mM each), 1 μl of T4 DNA polymerase (NEB # M0203L), 1 μl diluted Klenow DNA polymerase (1:5, NEB # M0210L), and T4 PNK (NEB # M0201L) in a total volume of 50 μl and incubation at 20 °C for 30 min, followed by purification using the QIAquick PCR purification kit and elution in 34 μl of EB.

To prepare the DNA fragments for adaptor ligation an adenosine base was added to the 3′ ends of the repaired DNA by addition of 5 μl Klenow buffer, 10 μl dATP and 1 ul Klenow exo– (NEB # M0212L). The reaction was incubated at 37 °C for 30 min followed by purification using the Qiagen mini elute reaction clean up kit. DNA was eluted in 10 μl of EB.

Adaptors were ligated to 10 μl of eluted DNA by addition of 15 μl of 2× T4 DNA ligase buffer, 1 μl of Nextflex adaptor (see the Bio Scientific ChIP-Seq barcodes protocol for adaptor dilution; # 514120), and 4 μl T4 ligase (Promega #M180B). The reaction was incubated at room temperature for 15 min. DNA was purified using the Qiagen mini elute reaction clean up kit, eluted in 10 μl of EB, and followed by PCR. The PCR reaction was assembled in a total volume of 50 μl by adding 10 μl DNA, 2 μl Nextflex primer, 25 μl Kapa 2× master mix (Kapa #KK2612), and 13 μl of H_2O. PCR was performed for 4 cycles using amplification conditions as mentioned in the Kapa protocol followed by purification using the Qiagen mini elute reaction clean up kit. DNA was eluted in 10 μl of EB and size selected using an E-gel system (Invitrogen # G661002) according to the manufacturer's protocol. 23 μl of DNA was collected in the size range of 290–320 bp (ChIP fragment + adapters). Size selected DNA was PCR amplified using 2 μl Nextflex primer and 25 μl Kapa 2× master mix in a total volume of 50 μl. The PCR reaction was purified by AMPure

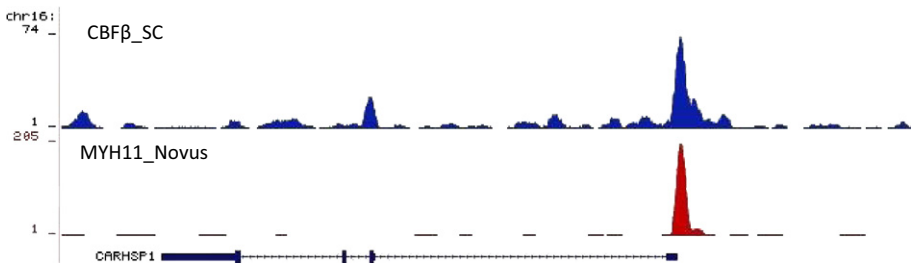

Fig. 2. ChIP-seq screen shot of a CBFβ-MYH11 binding site using two antibodies, one against CBFβ and one recognizing MYH11. Overview of the *CARHSP1* CBFβ-MYH11 binding site in ME-1 cells. In red the MYH11 ChIP-seq data using a Novus antibody is plotted and in blue the CBFβ data using a Santa Cruz antibody.

XP beads (Beckman Coulter # A63881) and DNA was eluted in 20 μl of EB.

DNA integrity and size were checked on a Bio-Rad Experion machine and cluster generation and sequencing by synthesis was performed on the Illumina 1G or HiSeq genome analyzer according to Illumina instructions. The image files were processed using Illumina's Consensus Assessment of Sequence and Variation (CASAVA) for base calling and generation of FastQ file containing 35–45 bp tags. FastQ data were mapped to the human genome HG18 using the eland or Burrows–Wheeler Alignment Tool (BWA) program allowing 1 mismatch. For visualization purpose 35–45 bp sequence reads were directionally extended to 300 bp, corresponding to the length of the original fragments used for sequencing. For each base pair in the genome the number of overlapping sequence reads was determined and averaged over a 10 bp window and visualized as wiggle track in the UCSC genome browser (Fig. 2) (http://genome.ucsc.edu). Tag files as well as visualization tracks can be found in GSE46044. In addition, peak calls (see below) for transcription factor binding as well as expression values (RPKM, see below) for the RNA-seq are provided as supplementary information. Moreover, for visualization all tracks are available as trackhub: http://trackhub.science.ru.nl/hubs/CBFb/CBFb-MYH/hub.txt.

Identification of protein binding sites based on ChIP-seq results

Peaks (Supplementary table) were called by model-based analysis (MACS1.3.3) of ChIP-seq [6] at a p-value cut off of 10^{-6} and called regions are available in Supplementary tables.

Expression analysis

For expression analysis RPKM (reads per kilobase of gene length per million reads) [4] values for RefSeq genes were computed using tag counting scripts (Supplementary table) and used to analyze the expression level of genes in ME-1 and U937 cells.

Discussion

In this study we described the methods used to establish the genome wide binding profile of CBFβ-MYH11 and other hematopoietic transcription factors in an inv(16) cell line, in CBFβ-MYH11 inducible and knockdown cell lines, and in primary patient cells. The methodology was used to investigate the mechanisms by which CBFβ-MYH11 exerts its function in leukemogenesis. Our analysis, published in [3] revealed that CBFβ-MYH11 localizes to RUNX1 occupied promoters where it operates in the context of TBP associated factors (TAFs), the epigenetic enzymes EP300 and HDAC1, and several hematopoietic transcription factors. RNA-seq analysis in ME-1 knockdown cells revealed that CBFβ-MYH11 is involved in both transcriptional activation and repression, which was confirmed by ChIP-seq and RNA-seq analysis in an inducible U937 CBFβ-MYH11 cell system. Together these results reveal a role for CBFβ-MYH11 in regulating expression of genes important in inv(16) leukemogenesis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gdata.2014.06.014.

References

[1] D. Helbling, B.U. Mueller, N.A. Timchenko, J. Schardt, M. Eyer, D.R. Betts, M. Jotterand, S. Meyer-Monard, M.F. Fey, T. Pabst, CBFB-SMMHC is correlated with increased calreticulin expression and suppresses the granulocytic differentiation factor CEBPA in AML with inv(16). Blood 106 (4) (2005) 1369–1375.
[2] M.J. Herold, J. van den Brandt, J. Seibler, H.M. Reichardt, Inducible and reversible gene silencing by stable integration of an shRNA-encoding lentivirus in transgenic rats. Proc. Natl. Acad. Sci. U. S. A. 105 (47) (2008) 18507–18512.
[3] A. Mandoli, A.A. Singh, P.W. Jansen, A.T. Wierenga, H. Riahi, G. Franci, K. Prange, S. Saeed, E. Vellenga, M. Vermeulen, H.G. Stunnenberg, J.H. Martens, CBFB-MYH11/RUNX1 together with a compendium of hematopoietic regulators, chromatin modifiers and basal transcription factors occupies self-renewal genes in inv(16) acute myeloid leukemia. Leukemia 28 (4) (2014) 770–778.
[4] A. Mortazavi, B.A. Williams, K. McCue, L. Schaeffer, B. Wold, Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5 (7) (2008) 621–628.
[5] K. Yanagisawa, T. Horiuchi, S. Fujita, Establishment and characterization of a new human leukemia cell line derived from M4E0. Blood 78 (2) (1991) 451–457.
[6] Y. Zhang, T. Liu, C.A. Meyer, J. Eeckhoute, D.S. Johnson, B.E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, X.S. Liu, Model-based analysis of ChIP-seq (MACS). Genome Biol. 9 (9) (2008) R137.