Data in Brief

Androgen receptor DNA binding and chromatin accessibility profiling in prostate cancer

Ekaterina Nevedomskaya a,b,1, Suzan Stelloo a,1, Henk G. van der Poel c, Jeroen de Jong d, Lodewyk F.A. Wessels c, Andries M. Bergman c,⁎, Wilbert Zwart a,⁎

a Division of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands
b Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands
c Division of Urology, The Netherlands Cancer Institute, Amsterdam, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands
d Division of Pathology, The Netherlands Cancer Institute, Amsterdam, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands
e Division of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

ABSTRACT

Prostate cancer (PCa) is the second most common cancer in men. The Androgen Receptor (AR) is the major driver of PCa and the main target of therapy in the advanced setting. AR is a nuclear receptor that binds the chromatin and regulates transcription of genes involved in cancer cell proliferation and survival. In a study by Stelloo et al. (1) we explored prostate cancer on the level of transcriptional regulation by means of Formaldehyde-Assisted Isolation of Regulatory Elements and Chromatin Immunoprecipitation coupled with massive parallel sequencing (FAIRE-seq and ChIP-seq, respectively). We employed these data for the assessment of differences in transcriptional regulation at distinct stages of PCa progression and to construct a prognostic gene expression classifier. Genomics data includes FAIRE-seq data from normal prostate tissue as well as primary, hormone therapy resistant and metastatic PCa. Furthermore, ChIP-seq data from primary and resistant PCa were generated, along with multiple input controls. The data are publicly available through NCBI GEO database with accession number GSE65478. Here we describe the genomics and clinical data in detail and provide comparative analysis of FAIRE-seq and ChIP-seq data.

© 2015 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/).

Keywords:
Prostate cancer
androgen receptor
ChIP-seq
FAIRE-seq

1. Experimental design, materials and methods

1.1. Clinical samples and experimental design

Fresh frozen tissue samples were obtained through postoperative needle biopsies targeting both tumor and normal areas of prostatectomy specimens at The Netherlands Cancer Institute (Amsterdam, The Netherlands). Tissue samples from androgen deprivation resistant tumors (from transurethral resection of the prostate (TURP)) and lymph node metastases were obtained from the Erasmus University Medical Center (Rotterdam, The Netherlands). Slides stained with hematoxylin and eosin (H&E) of the cases were reviewed by our pathologists. Clinical and pathological parameters of the selected patients are provided in Table 1. Leftover anonymized tissue, which cannot be traced back to the patient and does not interfere with care and/or prognosis, and would have been discarded otherwise, has been used in accordance with the Code of Conduct of the Federation of Medical Scientific Societies in The Netherlands. NKI and Erasmus MC institutional medical ethics committees have approved the study.
and three resistant tumors (Fig. 1). Androgen Receptor ChIP-seq was carried out on four primary, three therapy resistant tumors and three lymph node metastases.

1.2. Formaldehyde-assisted isolation of regulatory elements (FAIRE)

FAIRE-seq was performed on four normal samples, four primary, three therapy resistant tumors and three lymph node metastases (Fig. 1). Androgen Receptor ChIP-seq was carried out on four primary and three resistant tumors (Fig. 1).

1.3. Chromatin immunoprecipitation (ChIP)

Chromatin immunoprecipitation was carried out as described before [3,4]. 10 μg of AR-N20 (sc-618; Santa Cruz) antibody was used for immunoprecipitation, with 100 μl of Protein A magnetic beads (Invitrogen).

1.4. DNA sequencing

Libraries were prepared according to Illumina DNA Sample Kit instructions. Sequencing was performed on the Illumina Hiseq 2000 Genome Analyzer using 51-bp reads. Reads were aligned to the Human Reference Genome (assembly hg19, February 2009) using bwa 0.5.9.

1.5. Data analysis

Reads that map uniquely to the genome, with MAPQ quality score above 20, were used for the analysis. FAIRE-seq and ChIP-seq peaks were called with two algorithms, MACS 1.4 [5] and DFilter 1.0 [6], against mixed input controls corresponding to each group. MACS was run with default parameters, except for p = 10^{-7} for ChIP-seq data. DFilter was run with bs = 100, ks = 50 for FAIRE-seq data and bs = 50, ks = 30, refine, nonzero for ChIP-seq data. Peaks detected by both algorithms were used for further analysis. Sequencing read depths and number of called peaks can be found in Table 2.

FAIRE-seq, ChIP-seq data and clinical annotation of the samples that are deposited in NCBI GEO under accession number CSE65478.

For further analysis, a merged list of peaks present in all samples from each technique was generated. The number of peaks detected by FAIRE-seq was 25,797, while 20,703 peaks were detected by ChIP-seq data. DFilter was run with bs = 100, ks = 50 for FAIRE-seq data and bs = 50, ks = 30, refine, nonzero for ChIP-seq data. Peaks detected by both algorithms were used for further analysis.

2. Conclusions

In conclusion, we provide a unique dataset of genome-wide epigenetic profiling of prostate cancer tissue from different stages of the disease. The dataset consists of two parts: accessible chromatin profiling by FAIRE-seq and genome-wide androgen receptor binding to DNA by ChIP-seq. We previously used this dataset to identify changes in transcriptional regulation in prostate cancer upon acquisition of resistance to hormonal therapy, as well as to derive a prognostic gene expression signature for prostate cancer [1].

Acknowledgment

We thank Geert JLH van Leenders and Guido Jenster from Erasmus MC for providing tissue samples, clinical support and valuable discussions. We thank the NKI Genomics Core Facility for sequencing analyses and the Core Facility Molecular Pathology and Biobanking for help with tissue access.

This work was financially supported by Movember, grant number NKI01, and KWF/Alpe d’HuZes (Nkd 2014-6711).

Table 1

| Characteristic          | Number of patients |
|-------------------------|--------------------|
| Treatment type          |                  |
| Untreated               | 4                  |
| Bicalutamide/cyproteron acetate | 4                  |
| Bicalutamide/LHRH analogue | 4                  |
| Cyproteron acetate + LHRH analogue | 4                  |
| LHRH analogue Retry | 3                  |
| LHRH analogue, Cyproteron Acetate | 3                  |
| Gleason score           |                  |
| 6                       | 1                  |
| 7                       | 2                  |
| 8                       | 0                  |
| 9                       | 1                  |
| 10                      | 0                  |
| Initial PSA (ng/ml)     |                  |
| Mean                    | 8.7                |
| Range                   | 5.3–13.0           |

Fig. 1. FAIRE-seq and ChIP-seq analyses were performed on normal prostate tissue and prostate cancer samples from different stages of the disease.
References

[1] S. Stelloo, E. Nevedomskaya, H.G. van der Poel, J. de Jong, G.J. van Leenders, G. Jenster, et al., Androgen receptor profiling predicts prostate cancer outcome. EMBO Mol. Med. 7 (11) (2015) 1450–1464.

[2] P.G. Giresi, J.D. Lieb, Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements). Methods 48 (3) (Jul 2009) 233–239.

[3] W. Zwart, R. Koonstra, J. Wesseling, E. Rutgers, S. Linn, J.S. Carroll, A carrier-assisted ChIP-seq method for estrogen receptor–chromatin interactions from breast cancer core needle biopsy samples. BMC Genomics 14 (1) (2013) 232.

[4] M.P.H.M. Jansen, T. Krijnenburgh, E.A. Reijn, I. Simon, R. Kerkhoven, M. Droog, et al., Hallmarks of aromatase inhibitor drug resistance revealed by epigenetic profiling in breast cancer. Cancer Res. 73 (22) (Nov 15 2013) 6632–6641.

[5] Y. Zhang, T. Liu, C.A. Meyer, J. Eckelhouts, D.S. Johnson, B.E. Bernstein, et al., Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9 (9) (2008) R137.

[6] V. Kumar, M. Muratani, N.A. Rayan, P. Kraus, T. Lufkin, H.H. Ng, et al., Uniform, optimal signal processing of mapped deep-sequencing data. Nat. Biotechnol. 31 (7) (Jul 2013) 615–622.

[7] C.Y. McLean, D. Bristor, M. Hiller, S.L. Clarke, B.T. Schaar, C.B. Lowe, et al., GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. 28 (5) (May 2010) 495–501.

[8] J. Yu, J. Yu, R.-S. Mani, Q. Cao, C.J. Brenner, X. Cao, et al., An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. Cancer Cell 17 (5) (May 18 2010) 443–454.

[9] K. Davies, J. Jacobs, M. Atkins, D. Potier, V. Christiaens, G. Halder, et al., in: P. McKinnon (Ed.), Discovery of Transcription Factors and Regulatory Regions Driving In Vivo Tumor Development by ATAC-seq and FAIRE-seq Open Chromatin Profiling, PLoS Genet., 11(2), Public Library of Science Feb 2015, p. e1004954.

Table 2

Sequencing and peak calling details.

| GEO accession     | Experiment  | Tissue | Total number of reads | Mapped reads | % mapped reads | No. peaks |
|-------------------|-------------|--------|-----------------------|--------------|---------------|----------|
| GSM1598204        | FAIRE-seq   | Normal | 19,147,127            | 17,986,187   | 93.94         | 50       |
| GSM1598205        | FAIRE-seq   | Normal | 21,599,945            | 19,883,501   | 92.05         | 472      |
| GSM1598206        | FAIRE-seq   | Normal | 26,080,719            | 25,043,481   | 96.02         | 61       |
| GSM1598207        | FAIRE-seq   | Normal | 23,167,347            | 22,177,458   | 95.73         | 2837     |
| GSM1598208        | FAIRE-seq   | Primary| 36,827,373            | 34,411,896   | 93.52         | 6450     |
| GSM1598209        | FAIRE-seq   | Primary| 18,306,926            | 17,002,416   | 92.87         | 1579     |
| GSM1598210        | FAIRE-seq   | Primary| 32,197,589            | 30,568,523   | 94.94         | 13,348   |
| GSM1598211        | FAIRE-seq   | Primary| 28,992,853            | 27,590,961   | 95.16         | 2243     |
| GSM1598212        | FAIRE-seq   | Resistant | 37,452,682           | 35,655,681   | 95.2          | 80       |
| GSM1598213        | FAIRE-seq   | Resistant | 28,372,546           | 26,836,918   | 94.59         | 3497     |
| GSM1598214        | FAIRE-seq   | Resistant | 27,545,618           | 26,061,843   | 94.61         | 5754     |
| GSM1598215        | FAIRE-seq   | Metastasis | 39,562,972           | 37,594,752   | 95.03         | 2043     |
| GSM1598216        | FAIRE-seq   | Metastasis | 29,130,845           | 27,291,106   | 93.68         | 281      |
| GSM1598217        | FAIRE-seq   | Metastasis | 27,253,810           | 25,789,354   | 94.63         | 1313     |
| GSM1598218        | AR ChIP-seq | Primary | 13,782,549            | 12,232,556   | 88.75         | 754      |
| GSM1598219        | AR ChIP-seq | Primary | 18,146,927            | 16,009,388   | 88.22         | 402      |
| GSM1598220        | AR ChIP-seq | Primary | 13,040,014            | 11,254,994   | 86.31         | 17,511   |
| GSM1598221        | AR ChIP-seq | Primary | 9,928,626             | 7,080,840    | 71.32         | 3278     |
| GSM1598222        | AR ChIP-seq | Primary | 12,243,485            | 11,160,623   | 91.16         | 7932     |
| GSM1598223        | AR ChIP-seq | Resistant | 16,518,987           | 14,727,645   | 89.16         | 739      |
| GSM1598224        | AR ChIP-seq | Resistant | 16,382,421           | 14,441,817   | 88.15         | 238      |
| GSM1598225        | AR ChIP-seq | Resistant | 15,621,538           | 13,067,477   | 89.41         | 1779     |
| GSM1598226        | Input       | Resistant | 28,171,838           | 26,825,849   | 95.22         |         |
| GSM1598227        | Input       | Metastasis | 24,117,145           | 22,902,755   | 94.96         |         |
| GSM1598228        | Input       | Primary   | 23,982,305            | 22,739,491   | 94.82         |         |
| GSM1598229        | Input       | Primary   | 27,642,177            | 26,387,234   | 95.46         |         |