Data Article

Data of relative mRNA and protein abundances of androgen receptor splice variants in castration-resistant prostate cancer

Tianfang Ma\textsuperscript{a}, Nathan Ungerleider\textsuperscript{b}, Derek Y. Zhang\textsuperscript{c}, Eva Corey\textsuperscript{d}, Erik K. Flemington\textsuperscript{b,\ast}, Yan Dong\textsuperscript{a,\ast}

\textsuperscript{a} Department of Structural and Cellular Biology, Tulane University School of Medicine, Tulane Cancer Center, New Orleans, LA, USA
\textsuperscript{b} Department of Pathology, Tulane University School of Medicine, Tulane Cancer Center, New Orleans, LA, USA
\textsuperscript{c} Lusher Charter School, New Orleans, LA 70115, USA
\textsuperscript{d} Department of Urology, University of Washington, Seattle, WA, USA

\textbf{A R T I C L E   I N F O}

Article history:
Received 5 January 2021
Revised 14 January 2021
Accepted 15 January 2021
Available online 18 January 2021

Keywords:
Androgen receptor
Splice variants
Castration-resistant prostate cancer
Androgen deprivation therapy

\textbf{A B S T R A C T}

These data include secondary analysis of publicly available RNA-seq data from castration-resistant prostate cancer (CRPC) patients as well as RT-qPCR and Western blotting analyses of patient-derived xenograft models and a CRPC cell line. We applied Spearman correlation analysis to assess the relationship between canonical androgen receptor (AR) splicing and alternative AR splicing. We also assessed the ratio of AR splice variants (AR-Vs) to the full-length AR (AR-FL) at the RNA and protein levels by absolute RT-qPCR and Western blotting, respectively. These data are critical for studying the mechanisms underlying upregulated expression of AR-Vs after AR-directed therapies and the importance of AR-Vs to castration-resistant progression of prostate cancer. Data presented here are related to the research article by Ma et al., “Increased transcription and high translation efficiency lead
to accumulation of androgen receptor splice variant after androgen deprivation therapy”, Cancer Lett. In Press [1].

© 2021 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 Table

| Subject | Biological sciences |
|---------|---------------------|
| Specific subject area | Prostate cancer |
| Type of data | Table |
| How data were acquired | The RNA-seq data of metastatic CRPC samples were downloaded from dbGaP with accession numbers phs000909.v1.p1, phs001141.v1.p1, and phs000915.v2.p2. The RT-qPCR and Western blotting data were generated using RNA and protein samples isolated from the LNCaP95 CRPC cell line and the hormone sensitive and castration-resistant LuCaP 35 and 96 patient-derived xenografts (PDXs). |
| Data format | Raw Analyzed Filtered |
| Parameters for data collection | The parameters considered were the quality and abundance of the RNA-seq reads, the melting curve pattern and the RT-qPCR amplicon for each primer pair to guarantee the specificity of RT-qPCR, as well as the number and size of the bands to ensure the specificity of Western blot analysis. |
| Description of data collection | The RNA-seq reads were aligned to the hg38 genome using STAR. The number of reads spanning each AR splice junction in each sample was quantified, normalized to total splice junction reads in that sample, and subjected to Spearman correlation analysis. Serial dilutions of the pLVX-AR-FL and pLVX-AR-V7 plasmids were used as templates in RT-qPCR analysis to generate standard curves for quantifying absolute copy numbers of AR-FL and AR-V7 transcripts. Protein samples were subjected to Western blotting using a pan-AR antibody, and the levels of AR were normalized to that of α-tubulin. |
| Data source location | Institution: Tulane University School of Medicine |
| Country: USA | |
| Latitude and longitude for collected samples/data: 29.956888 and -90.08018 |
| Primary data sources: dbGaP accession numbers phs000909.v1.p1 (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000909.v1.p1), phs001141.v1.p1 (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001141.v1.p1), and phs000915.v2.p2 (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000915.v2.p2). |
| Data accessibility | Repository name: Mendeley Data |
| Data identification number: DOI:10.17632/7hyzhp8wrf.1 |
| Related research article | T. Ma, S. Bai, Y. Qi, Y. Zhan, N. Ungerleider, D. Y. Zhang, T. Neklesa, E. Corey, S. M. Dehm, K. Zhang, E. K. Flemington, and Y. Dong. Increased transcription and high translation efficiency lead to accumulation of androgen receptor splice variant after androgen deprivation therapy, Cancer Lett. In Press [1]. |

Value of the Data

- Recognizing the relative abundance of AR-Vs in castration-resistant prostate cancer is critical to understanding the role of AR-Vs in mediating disease progression after AR-directed therapies.
- These data are useful to all researchers who would like to understand the role of AR-Vs in mediating castration-resistant progression of prostate cancer as well as those who are interested in studying mechanisms of AR-V generation after AR-directed therapies.
• These data might be a valuable reference and a starting point for identifying factors involved in translational or posttranslational regulation of AR-FL and AR-Vs.

1. Data Description

In the present study, we reported data of correlation between all alternative AR splicing and canonical AR splicing in RNA-seq data from CRPC patient samples as well as relative mRNA and protein abundances of AR-Vs in CRPC PDX models and a CRPC cell line. The patient RNA-seq data included polyA-selected RNA-seq data from 34 CRPC adenocarcinoma samples in the Beltran cohort [2] and 74 CRPC samples in the Prostate Cancer Medically Optimized Genome-Enhanced Therapy (PROMOTE) cohort [3] as well as exome-captured RNA-seq data from 181 CRPC samples in the Stand Up To Cancer (SU2C) cohort [4]. The normalized numbers of reads spanning each alternative or canonical AR splice junction in individual samples were subjected to Spearman correlation analysis. The correlations in the polyA-selected RNA-seq data and the exome-captured RNA-seq data are shown in Figs. 1 and 2, respectively. Canonical AR splicing is indicated as black, and alternative AR splicing is indicated as blue. The splice junctions are segregated by donor or acceptor. The correlation coefficient scores are encoded by different colors of the squares ranging from dark blue to dark red. The darker the red color, the more positive the correlation. The darker the blue color, the more negative the correlation. No squares bear blue color.

We then quantified the ratios of AR-V7 to AR-FL mRNA in LNCaP95 cells cultured in androgen-deprived or -replenished condition as well as in hormone-sensitive and castration-resistant LuCaP 35 and 96 PDXs. As presented in Fig. 3, the P value comparing androgen-deprived and androgen-replenished conditions for LNCaP95 cells was 0.15, and the mean ratios of AR-V7 to AR-FL mRNA were between 0.02 and 0.03 in both conditions. The P value comparing the hormone-sensitive and castration-resistant LuCaP PDXs was even higher, and the mean ratios of AR-V7 to AR-FL mRNA in these PDXs were between 0.05 and 0.06 (Fig. 3). We also assessed the ratios of AR-V7 to AR-FL protein in these same cells and PDXs, and the data are shown in Fig. 4. Despite the lack of statistically significant change in the ratios of the transcripts (Fig. 3), androgen depletion and castration resistance led to a much higher ratio of AR-V7 to AR-FL protein (Fig. 4). The sequences of qPCR primers and probes are listed in Table 1. The raw data are deposited to Mendeley Data with DOI: 10.17632/7hzyhp8wrf.1.
Fig. 2. Spearman correlation analysis of correlations between AR canonical (black) and alternative (blue) splicing in 181 CRPC samples in the SU2C exome-captured RNA-seq dataset. The red-blue gradient bar indicates the scale of correlation coefficient scores.

Fig. 3. AR-V7/AR-FL mRNA ratio in androgen-deprived or –replenished condition. A. Standard curves generated using pLVX-AR-FL and pLVX-AR-V7 plasmids for RT-qPCR absolute quantification of AR-FL and AR-V7 transcript copy numbers. B & C. AR-V7/AR-FL mRNA ratios in LNCaP95 cells cultured in androgen-deprived or –replenished condition (B) or in hormone sensitive (HS) and castration-resistant (CR) LuCaP 35 and 96 PDXs (C) were calculated using the equations in Panel A. R1881: 1 nM.
2. Experimental Design, Materials and Methods

2.1. Cell culture

Phenol-red free RPMI-1640 medium (Gibco, catalog No.11835-030) with 10% charcoal-stripped Fetal Bovine Serum (FBS) (Atlanta, catalog No. S11650) was used to culture LNCaP95 cells. For androgen treatment, cells were incubated with vehicle or 1 nM R1881, a synthetic androgen, for 24 h. Cells were evaluated for mycoplasma contamination monthly.

2.2. Splice junction and correlation analysis

RNA-seq data from 3 cohorts of metastatic CRPC samples downloaded from dbGaP were used in the correlation analyses. These include poly-A selected RNA-seq data from 74 samples from the PROMOTE cohort (dbGaP accession phs001141.v1.p1) [3] and 34 adenocarcinoma samples from the Beltran study (dbGaP accession phs000909.v1.p1) [2] along with exome-capture RNA-seq data from 181 samples in the SU2C cohort (dbGaP accession phs000915.v2.p2) [4]. The cDNA libraries for the exome-capture RNA-seq dataset were generated using Agilent SureSelect Human All Exon V4 platform, excluding the polyA selection step. Of note, the PROMOTE cohort contains 81 samples [3]. However, 7 samples, SRR4043802, SRR4043827, SRR4043849, SRR4043853, SRR4043855, SRR4043985, and SRR4044025, have a low mapping rate and were removed from our analysis. STAR was used to align the raw RNA-seq reads to the hg38 genome. The number of reads spanning every AR splice junction in each sample was quantified and normalized to total splice junction reads in that sample. Spearman correlation analysis was performed on the normalized read counts for each canonical and alternative AR splice junction, after filtering out the ones spanned by very few reads using R 3.6.0. The R code files are deposited to Mendeley Data with DOI:10.17632/7hyzhp8wrfl.1.

2.3. Absolute quantification of transcript copy numbers

Absolute quantification of AR-FL and AR-V7 transcript copy numbers was carried out as in a previous study [5]. Serial dilutions of the pLVX-AR-FL and pLVX-AR-V7 plasmids were used.
as templates in RT-qPCR analysis to generate standard curves for quantifying absolute numbers of AR-FL and AR-V7 transcripts, respectively. Total RNA was isolated from at least three independent experiments using the TRIzol reagent (Thermo Fisher, catalog No. 15596026). The iScript cDNA synthesis kit (Bio-rad, catalog No. 1708891) was used to reverse transcribe 1 μg RNA following the manufacturer’s instructions. qPCR analyses were conducted in duplicates with iQ SYBR Green Supermix and iQ Supermix (Bio-rad, Catalog No. 1708880 & 1708862). Primer-probe sets were purchased from IDT. Each amplification reaction contained 3 μL of the respective cDNA diluted 1:3, 0.5 μL of the forward primer, 0.5 μL of the reverse primer, 10 μL of the 2X iQ SYBR Green Supermix or iQ Supermix, and 6 μL of DNase/ RNase free H2O, resulting in a 20 μL reaction. The cycle threshold (Ct) values were used to generate standard curves of transcript copy numbers versus Ct values. The amplicon of every sample and the melting curve for each primer pair were assessed to assure qPCR specificity.

2.4. Western blotting

Cells were lysed in lysis buffer (Cell Singaling, Catalog No.9803S) with additional protease inhibitor cocktail (Sigma, Catalog No. P8340-5ML). Thirty μg of protein per sample was separated on a 10% acrylamide gel and transferred to a polyvinylidine difluoride membrane. The membrane was blocked in 5% Blotting-Grade blocker (Bio-rad, Catalog No. 1706404) in PBST and then incubated with a primary antibody diluted 1:1000 in PBST at 4 °C overnight. The membrane was subsequently washed with PBST three times for 5 min each time and incubated with the IRDye® 800CW Goat anti-Mouse IgG Secondary Antibody (Licor, Catalog No. 926-32210) at room temperature for 1 h. The membrane was washed again with PBST three times for 5 min each time and imaged using LI-COR Odyssey CLx. The primary antibodies that were used were anti-AR (441, Santa Cruz Biotechnology, catalog No. sc-7305) and anti-α-tubulin (Cell Signaling Technology, catalog No.3873S). The experiments were conducted at least three times, and AR levels were normalized to α-tubulin levels.

2.5. Statistical analyses

The mean differences between two groups were determined by the Student two-tailed t-test. P < 0.05 was considered significant. Data are presented as mean ± SEM.

Ethics Statement

This work does not involve human subject. The animal experiments comply with the ARRIVE guidelines and were carried out in accordance with the National Institutes of Health guide for the care and use of Laboratory animals (NIH Publications No. 8023, revised 1978).

CRediT Author Statement

Tianfang Ma: Methodology, Investigation, Data curation, Writing- Original Draft, Visualization, Validation; Nathan Ungerleider: Software, Formal analysis; Derek Y. Zhang: Software, Formal analysis; Eva Corey: Resources; Erik K. Flemington: Conceptualization, Writing- Reviewing and Editing; Yan Dong: Conceptualization, Writing- Reviewing and Editing.

Declaration of Competing Interest

This work was supported by the National Institutes of Health [grant numbers R01CA188609, R01AI106676, R01CA243793]. The Richard M Lucas Foundation supported the development of
the LuCaP models. The authors declare that they have no known competing financial interests or personal relationships that have or could be perceived to have influenced the work reported in this article.

**Acknowledgments**

We are grateful to Dr. Alan Meeker at Johns Hopkins University for providing LNCaP95 cells. We appreciate the support from the Tulane Cancer Next Generation Sequence Analysis core for utilization of resources and expertise for this work.

**References**

[1] T. Ma, S. Bai, Y. Qi, Y. Zhan, N. Ungerleider, D.Y. Zhang, T. Neklesa, E. Corey, S.M. Dehm, K. Zhang, E.K. Flemington, Y. Dong, Increased transcription and high translation efficiency lead to accumulation of androgen receptor splice variant after androgen deprivation therapy, Cancer Lett. (2021) In Press.

[2] H. Beltran, D. Prandi, J.M. Mosquera, M. Benelli, L. Puca, J. Cyrtz, C. Marotz, E. Giannopoulou, B.V. Chakravarthi, S. Varambally, S.A. Tomlins, D.M. Nanus, S.T. Tagawa, E.M. Van Allen, O. Elemento, A. Bioner, L.A. Garraway, M.A. Rubin, F. Demichelis, Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer, Nat. Med. 22 (2016) 298–305.

[3] L. Wang, S.M. Dehm, D.W. Hillman, H. Sicotte, W. Tan, M. Gormley, V. Bhargava, R. Jimenez, F. Xie, P. Yin, S. Qin, F. Quevedo, B.A. Costello, H.C. Pitot, T. Ho, A.H. Bryce, Z. Ye, Y. Li, P. Eiken, P.T. Vedell, P. Barman, B.P. McMenomy, T.D. Atwell, R.E. Carlson, M. Ellingson, B.W. Eckloff, R. Qin, F. Ou, S.N. Hart, H. Huang, J. Jen, E.D. Wieben, K.R. Kalari, R.M. Weinschilboum, L. Wang, M. Kohli, A prospective genome-wide study of prostate cancer metastases reveals association of wt pathway activation and increased cell cycle proliferation with primary resistance to abiraterone acetate-prednisone, Ann. Oncol. 29 (2018) 352–360.

[4] D. Robinson, E.M. Van Allen, Y.M. Wu, N. Schultz, R.J. Lonigro, J.M. Mosquera, B. Montgomery, M.E. Taplin, C.C. Pritchard, G. Attard, H. Beltran, W. Abida, R.K. Bradley, J. Vison, X. Cao, P. Vats, L.P. Kunju, M. Hussain, F.Y. Feng, S.A. Tomlins, K.A. Cooney, D.C. Smith, C. Brennan, J. Siddiqui, R. Mehra, Y. Chen, D.E. Rathkopf, M.J. Morris, S.B. Solomon, J.C. Durack, V.E. Reuter, A. Gopalan, J. Gao, M. Loda, R.T. Lis, M. Bowden, S.P. Balk, G. Gaviola, C. Sougne, M. Gupta, E.Y. Yu, E.A. Mostaghel, H.H. Cheng, H. Mulcahy, L.D. True, S.R. Plymate, H. Dvinge, R. Ferraldeschi, P. Plohr, S. Miranda, Z. Zafeiriou, N. Tunariu, J. Mateo, R. Perez-Lopez, F. Demichelis, B.D. Robinson, M. Schiffman, D.M. Nanus, S.T. Tagawa, A. Sigaras, K.W. Eng, O. Elemento, A. Bioner, E.L. Heath, H.I. Scher, K.J. Pienta, P. Kantoff, J.S. de Bono, M.A. Rubin, P.S. Nelson, L.A. Garraway, C.L. Sawyer, A.M. Chinnaian, Integrative clinical genomics of advanced prostate cancer, Cell 161 (2015) 1215–1228.

[5] L.L. Liu, N. Xie, S. Sun, S. Plymate, E. Mostaghel, X. Dong, Mechanisms of the androgen receptor splicing in prostate cancer cells, Oncogene 33 (2014) 3140–3150.