Cell Reports, Volume *22*

Supplemental Information

Loss of an Androgen-Inactivating and Isoform-Specific *HSD17B4* Splice Form Enables Emergence of Castration-Resistant Prostate Cancer

Hyun-Kyung Ko, Michael Berk, Yoon-Mi Chung, Belinda Willard, Rohan Bareja, Mark Rubin, Andrea Sboner, and Nima Sharifi

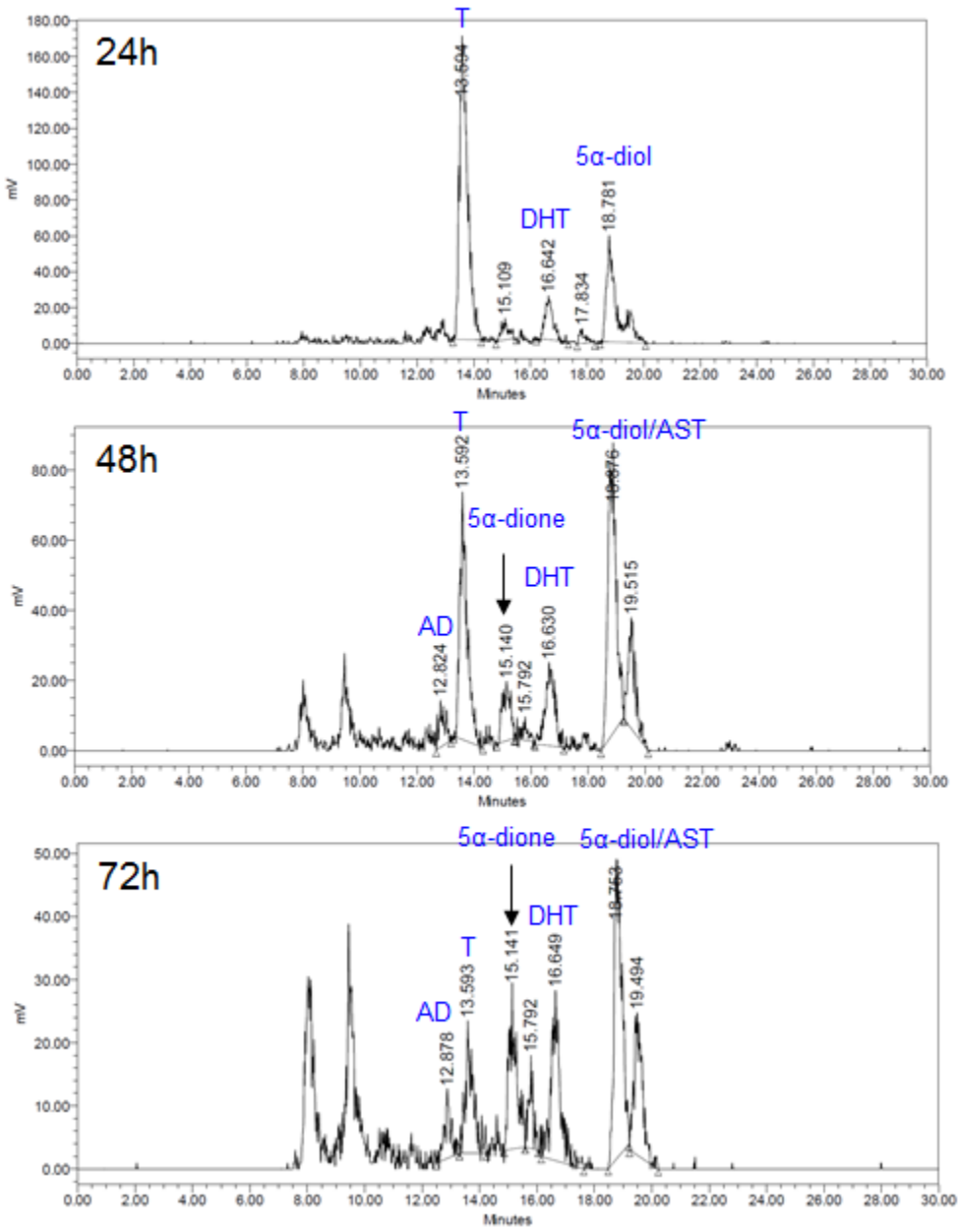
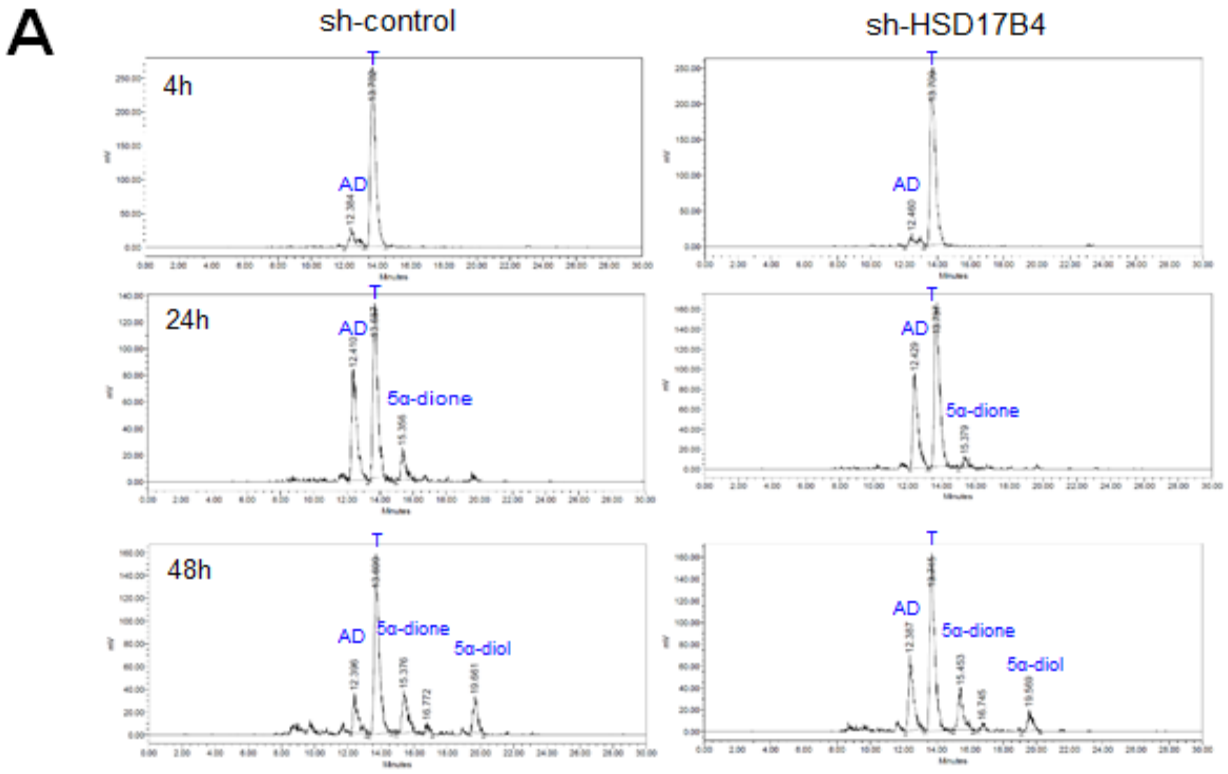


Figure S1, related to Figure 1.
DU145 cells have limited T conversion to AD, even at longer incubation times of T up to 72hr. The formation of metabolites in media from [^{3}H]-T (100 nM) was assessed by HPLC at the indicated time points after incubation with DU145 cells.



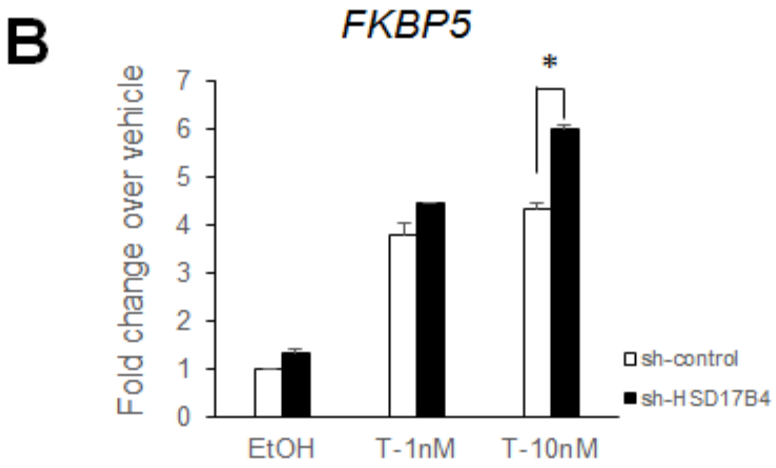


Figure S2, related to Figure 2.
(A) 17βHSD4 knockdown does not alter 17β-OH-oxidation of testosterone in VCaP. The formation of metabolites in media from [^{3}H]-T (100 nM) was assessed by HPLC at the indicated time points after incubation with VCaP cells stably expressing sh-Control or sh-HSD17B4. **(B)** Knockdown of 17βHSD4 in 22Rv1 significantly increases *FKBP5* expression in the presence of 10nM T. 22Rv1 cells expressing sh-control and sh-HSD17B4 were treated with 1 or 10nM T for 48h and gene expression was assessed as described in Figure 2E.

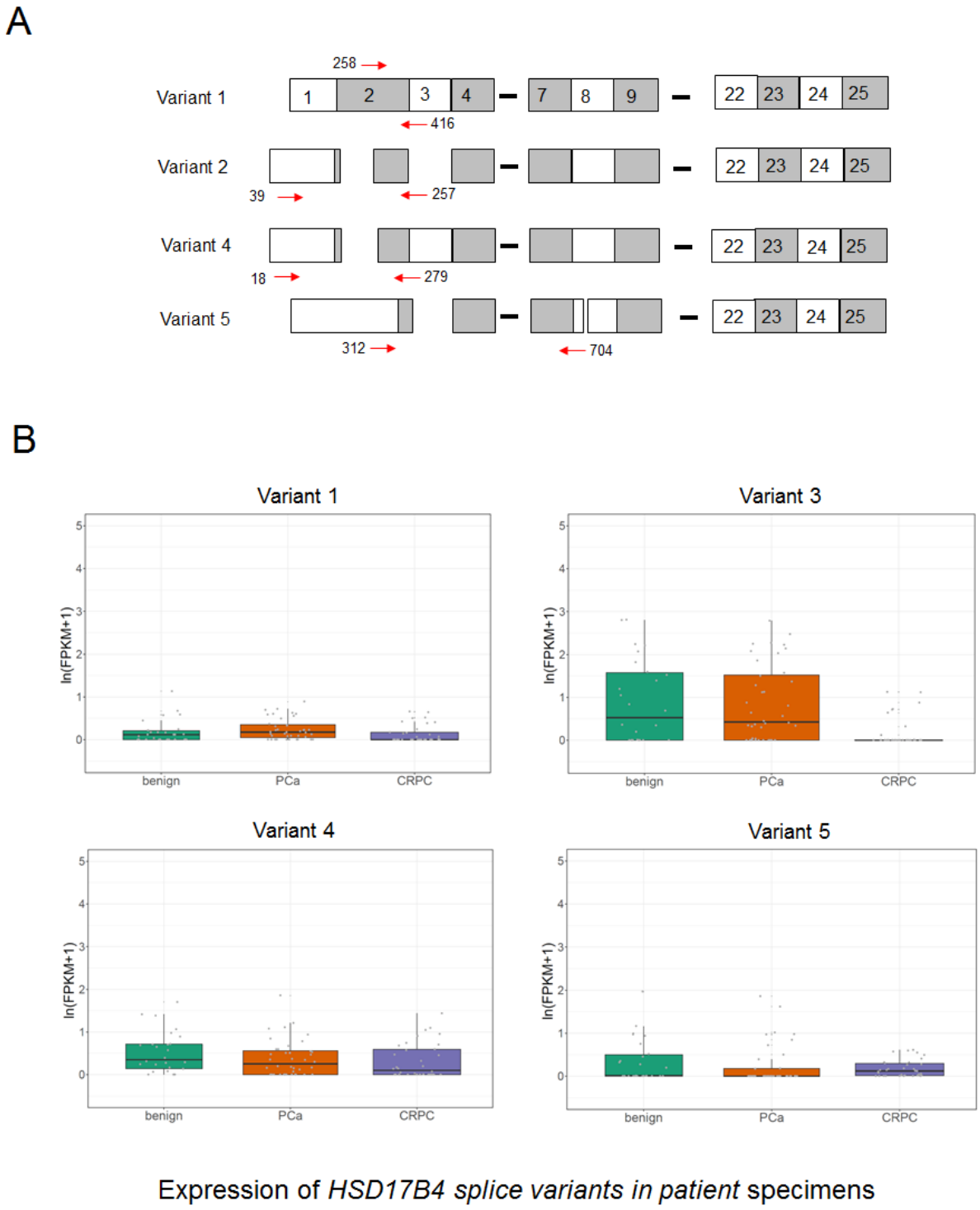


Figure S3, related to Figure 3.
(A) Schematic of *HSD17B4* splice variants and primers designed and used for detection**.** Primer pairs for detection of each variant are indicated by arrows and the 5' nucleotide position (with respect to the transcript) is indicated.
(B) RNA-seq xpression of *HSD17B4* splice variants in benign, PCa, and CRPC clinical tissues. The values are expressed as log2(FPKM+1). FPKM, fragments per kilobase of transcript per million mapped reads.

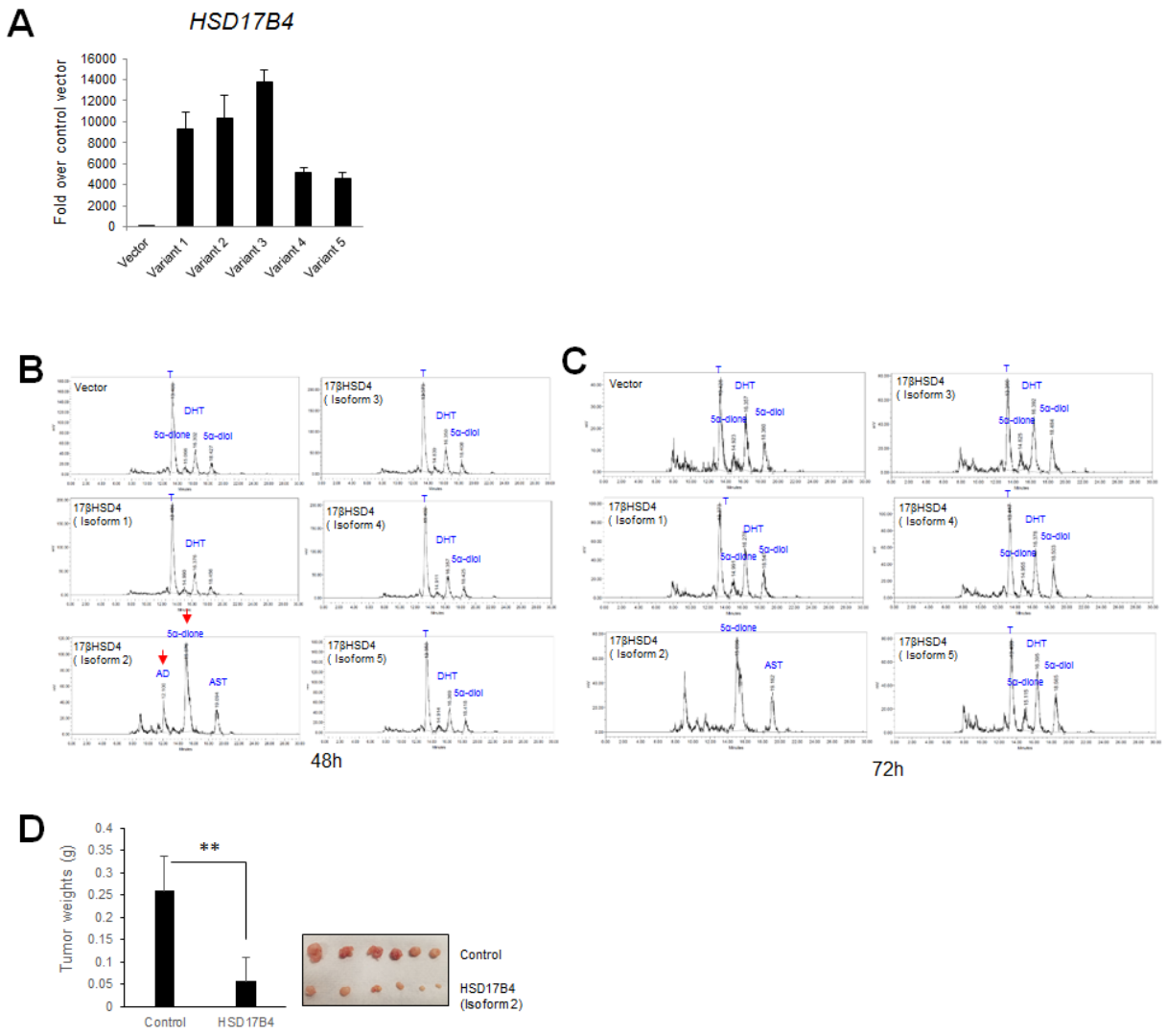


Figure S4, related to Figure 4.

(A) Transcription levels with ectopic expression of *HSD17B4* splice variants 1-5 in LAPC4 cells. Cells were transiently transfected with plasmid constructs expressing the designated *HSD17B4* variant, or empty vector, and expression was assessed by qPCR, using *RPLP0* as an internal control. **(B and C)** Expression of 17βHSD4 isoform 2 results in complete loss of T by 48 hrs. However, no change in T metabolism is detectable with expression of 17βHSD4 isoforms 1 and 3-5 compared with vector even at longer incubation times of T up to 72hr. LAPC4 cells expressing isoform 1-5 or vector were incubated for up to 72h in the presence of [^{3}H]-T (100nM). Steroids were separated and detected by HPLC. Data shown are representative of 3 independent experiments. **(D)** 17βHSD4 isoforms 2 suppresses CRPC tumor progression in the LAPC4 xenograft model, significantly reducing tumor weight. Mice were sacrificed at 56 days, the xenograft tumors were collected and weighed. ** P = 0.001.

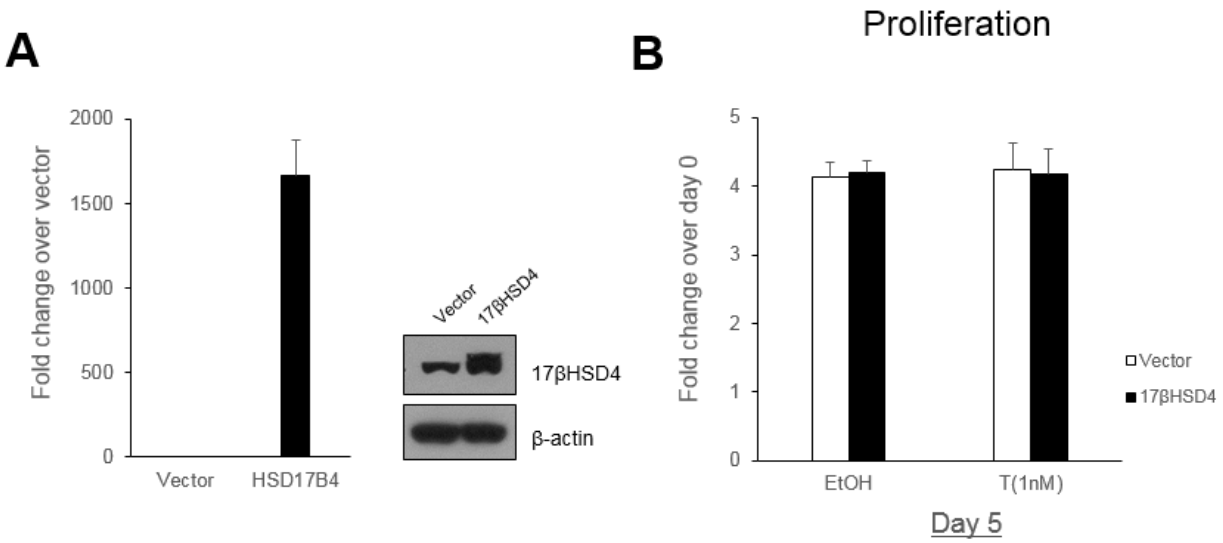


Figure S5, related to Figure 4.
(A) Ectopic expression of 17βHSD4 isoform 2 in DU145 cells was confirmed by qPCR (left) and immunoblot (right). **(B)** Isoform 2 expression does not change cell proliferation in the AR negative DU145 cell line model. 5 x10^4 cells were cultured in phenol red-free, 10% CSS medium and treated with or without T (1 nM) every other day, and were quantitated using the CellTiter Proliferation Assay at the indicated time points.

Supplemental Experimental Procedures

RNA isolation and qRT-PCR

Total RNA was isolated from cells using the RNeasy Kit (Sigma-Aldrich, St. Louis, Mo), and 1 ug RNA was used in a reverse-transcriptase reaction with the iScript cDNA synthesis kit (Bio-Rad, Hercules, CA). qRT-PCR was performed using the iTaq SYBR Green Supermix ROX kit (Bio-Rad) on an ABI-7500 Real-Time PCR system (Applied Biosystems, Foster City, CA). The qPCR analysis was carried out in triplicate using the following thermocycling program: 95°C for 2 minutes; indicated number of cycles of 95°C for 15 seconds, 60°C for 40 seconds, 72°C for 30 seconds; 72°C for 4 minutes, followed by cooling down to 4°C. Each mRNA transcript was quantitated by normalizing to the housekeeping gene RPLP0. Primer sequences for qPCR:

Primers	Sequences (5' to 3')
HSD17B4 forward	CGGGATCACGGATGACTCAG
HSD17B4 reverse	GTCCGTTTTCCACCAAAGCC
PSA forward	CATCAAATCTGAGGGTTGTCTGGA
PSA reverse	GCATGGGATGGGGATGAAGTAAG
FKBP5 forward	CCCCCTGGTGAACCATAATACA
FKBP5 reverse	AAAAGGCCACCTAGCTTTTTGC
RPLPO forward	CGAGGGCACCTGGAAAAC
RPLPO reverse	CACATTCCCCCGGATATGA

17βHSD4 overexpression

For 17βHSD4 isoform 2 expression, the human HSD17B4 cDNA (Dharmacon, CO, Clone ID:3502915, Catalog Number:MHS6278-202828086) was PCR-amplified using the Phusion High-Fidelity PCR Kit (New England Biology Labs, Ipswich, MA) and sub-cloned into the pcDNA3.1 vector by ligation into filled-in Kpn1 and Xba1 sites, resulting in the pcDNA3.1-HSD17B4 vector. The identity of the insert was further confirmed by DNA sequencing. Isoform 1 cDNA was cloned from the LNCaP cell line; total RNAs extracted from LNCaP cells were PCR amplified with isoform 1 specific primers and the PCR product was TA-cloned into the expression TOPO-TA vector (Thermo Fisher Scientific). cDNAs of isoforms 3, 4, and 5 were then generated by PCR amplification using isoform 1 and isoform 2 plasmid DNA, respectively, followed by sub-cloning into the same pCDNA3.1 vector. Plasmid DNA was transfected into cells in 6- or 12 well-plates using LipoD293 (SignaGen, Ijamsville, MD) reagent according to the manufacturer's instructions; cells were seeded in culture media 1 day before transfection. The next day, 1.0 and 0.75 ug of each DNA construct was added per well in 6- and 12–well plates, respectively. The transfected cells were harvested 18~24h after transfection for western blotting analysis or used for the subsequent experiments. Primer sequences for PCR:

17βHSD4	Primers	Sequences (5' to 3')
Isoform 1	Forward	ACGTGGTACC ATGGTTATTCTTGAGGCACCGCA
	Reverse	ACGTTCTAGA TCAGAGCTTG GCGTAGTCTTTAAG
Isoform 2	Forward	ACGTGGTACC ATGGGCTCAC CGCTG
	Reverse	ACGTTCTAGA TCAGAGCTTG GCGTAGTCTTTAAG
isoform 3	Forward	ATCGATCGATGTGAATGATT TGGGAGGGGA CTTC
	Reverse	ATCGATCGATTGCCCCCGCGCCGGTGACCAGTAC
Isoform 4	Forward	ACGTGGTACC ATGGAGAAGATCATTTCAC
	Reverse	ACGTTCTAGA TCAGAGCTTG GCGTAGTCTTTAAG
Isoform 5	Forward	ACGTGGTACC ATGAAGAAACAGAAGATTATTATGACTTCATCAGCTTCAGG
	Reverse	ACGTTCTAGA TCAGAGCTTG GCGTAGTCTTTAAG

*Underlined indicates enzyme sites (KpnI, XbaI, and ClaI)

Detection of transcription variants

cDNAs were produced using total RNA extracted from each cell line that was PCR amplified with primer sets targeting specific variants. PCR products were sequenced. Primer sequences for qPCR:

HSD17B4	Primers	Sequences (5' to 3')
Variant 1	Forward	ATGGTTATTCTTGAGGCACCG
	Reverse	GATCTTCTCCATAGGGTTATTGC
Variant 2	Forward	TCGTCCCGCCCCCGCCAT
	Reverse	CCAAATCATTCACAACAACTAACGC
Variant 4	Forward	AAATCG GCAAGTCACTGACCCT
	Reverse	CTTCTCCATAGGGTTATTGCATAG
Variant 5	Forward	GGTCTCTCAAGCAGGATT
	Reverse	GTCATAATAATCTTCTGTTTCTT

Analysis of RNA sequencing data

We analyzed two publicly available datasets that our group reported in recent years (Beltran et al., 2016; Chakravarty et al., 2014). For this analysis, we selected 26 benign prostate tissues, 40 localized PCa, and 34 CRPC. RNA-sequencing and data processing was performed according to the protocol described in the respective papers (Beltran et al., 2016; Chakravarty et al., 2014). Briefly, RNA was extracted from frozen material for RNA-sequencing (RNA-seq) using Promega Maxwell 16 MDx instrument, (Maxwell 16 LEV simplyRNA Tissue Kit (cat. # AS1280)). Specimens were prepared for RNA sequencing using TruSeq RNA Library Preparation Kit v2 or riboZero as previously described (Beltran et al., 2016; Chakravarty et al., 2014). RNA integrity was verified using the Agilent Bioanalyzer 2100 (Agilent Technologies). cDNA was synthesized from total RNA using Superscript III (Invitrogen). Sequencing was then performed on GAII, HiSeq 2000, or HiSeq 2500, as paired ends. Similarly to the original analysis, all reads were independently aligned with STAR_2.4.0f1 (Dobin et al., 2013) for sequence alignment against the human genome sequence build hg19, downloaded via the UCSC genome browser [http://hgdownload.soe.ucsc.edu/goldenPath/hg19/bigZips/], and SAMTOOLS v0.1.19 (Li et al., 2009) for sorting and indexing reads. Cufflinks (2.0.2) (Trapnell et al., 2012) was used to estimate the expression values (FPKMS), and GENCODE v23 (Derrien et al., 2012) GTF file for annotation. Gene coordinates were mapped to hg19 using liftOver (Hinrichs et al., 2006). Since the sequenced samples from the published datasets were processed using different library preps, batch normalization was done using ComBat (Johnson et al., 2007) from sva bioconductor package (Leek et al., 2012). Rstudio (1.0.136) with R (v3.3.2) and ggplot2 (2.2.1) were used for the statistical analysis and the generation of figures depicting the expression levels for each isoforms in each of the classes: benign, PCa, and CRPC).

Supplemental References

Chakravarty, D., Sboner, A., Nair, S.S., Giannopoulou, E., Li, R., Hennig, S., Mosquera, J.M., Pauwels, J., Park, K., Kossai, M., *et al.* (2014). The oestrogen receptor alpha-regulated lncRNA NEAT1 is a critical modulator of prostate cancer. Nat Commun *5*, 5383.

Beltran, H., Prandi, D., Mosquera, J.M., Benelli, M., Puca, L., Cyrta, J., Marotz, C., Giannopoulou, E., Chakravarthi, B.V., Varambally, S., *et al.* (2016). Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. Nat Med *22*, 298-305.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562-578.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., *et al.* (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res *22*, 1775-1789.

Hinrichs, A.S., Karolchik, D., Baertsch, R., Barber, G.P., Bejerano, G., Clawson, H., Diekhans, M., Furey, T.S., Harte, R.A., Hsu, F., *et al.* (2006). The UCSC Genome Browser Database: update 2006. Nucleic Acids Res *34*, D590-598.

Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics *8*, 118-127.

Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E., and Storey, J.D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics *28*, 882-883.