Prostate cancer is the most frequently diagnosed male cancer. For prostate cancer that has progressed to an advanced or metastatic stage, androgen deprivation therapy (ADT) is the standard of care. ADT inhibits activity of the androgen receptor (AR), a major regulator transcription factor in normal and cancerous prostate cells. The major limitation of ADT is the development of castration-resistant prostate cancer (CRPC), which is almost universally drug-resistant. One mechanism of AR transcriptional re-activation is in expression of AR-V7, a truncated, constitutively active AR variant (AR-V) arising from alternative AR pre-mRNA splicing. Noteworthy, AR-V7 is being developed as a predictive biomarker of primary resistance to androgen receptor (AR)-targeted therapies in CRPC. Multiple additional AR isoforms are expressed in clinical CRPC, but the extent to which these may be co-expressed with AR-V7 or predict resistance is not known.

Here, we utilized long-read sequencing to identify and quantify AR isoforms expressed in CRPC. To unambiguously characterize all AR isoforms, we prepared iso-seq libraries via 3′ rapid amplification of cDNA ends (RACE) with RNA isolated from prostate cancer cell lines and xenograft tissues using a forward primer anchored in AR exon 1. 3′RACE reactions were subjected to Single Molecule, Real-Time (SMRT) long-read sequencing with a PacBio RS II System.

Our work identified AR-V9 as a truncated isoform that is frequently co-expressed with AR-V7 in CRPC. Mechanistically, our work re-annotated AR-V7 and AR-V9 mRNAs, showing these two species shared a common 3′ terminal exon containing separate splice acceptor sites. Using this new information, novel AR isoforms and antibodies that could distinguish between AR-V7 and AR-V9 were designed, validated and used to measure the relative expression of these new AR isoforms in CRPC cells with a view to determining the potential of AR-V9 as a predictive biomarker of primary resistance to AR-targeted therapies.

### Method Overview

**SMRTbell 3’ RACE (Rapid Amplification of cDNA Ends)**

3′ or 5′ RACE works best for targeted sequencing of cDNAs when the gene of interest includes alternative transcriptional termination or start sites. Full-length first strand cDNA is generated using an edge primer annealing to an arbitrary primer sequence appended to the end. The gene of interest is specifically primed using a custom primer The and Iso-seq template by ligation of hairpin adapters to cDNA and sequencing on a PacBio Single Molecule, Real-Time (SMRT) Sequencing System.

Long sequencing reads from PacBio enable sequencing near full-length cDNA amplicons by using a gene-specific primer near the 5′ and 3′ of the gene.

**Deep Sequencing of AR RNA in Castration-Resistant Prostate Cancer.** AR-V7 RNA-Seq coverage across disease subtypes of the AR gene is shown. RNA-Seq analysis was performed through AR-V7 sequencing data. It is difficult to detect the contiguous full-length mRNA isoforms. This makes it a challenging understanding expression of multiple variants.

**Single-Molecule Real-Time (SMRT) Sequencing Identifies Novel Isoforms in Castration-Resistant Prostate Cancer.**

(A) Schematic of AR gene structure in the 22Rv1 cell line. (B) Genomic visualization of short (454/CE4) and long (PacBio 22Rv1 exons) SMRT Sequencing data. (C) Exon composition and quantification of the 20 most abundant AR transcripts from 22Rv1 cell line. (D) Enlarged view of the alternative 3′ splice acceptor site (SAS) flanked by two canonical exons. (E) Boxplot of the log fold change in expression of AR-V7 relative to AR-WT isoforms in cell line.

### Summary and Resources

- 3′ RACE combined with PacBio SMRT Sequencing enables targeted full-length isoform identification
- SMRT Sequencing transcriptome analysis of full-length mRNA information
- Alternatively spliced AR transcripts that incorporate cryptic exons result in proteins that are likely to be truncated
- AR-V7 is widely and frequently co-expressed with AR-V7 and both share a common 3′ terminal exon
- The newly-annotated features of the human AR gene are important for the design and interpretation of targeted assays for developing AR-V7 as biomarkers in prostate cancer

More information on full-length transcript sequencing (iso-seq Application) can be found on the PacBio website: [http://pacific.biomarx.org](http://pacific.biomarx.org)