Supplementary Online Content


eMethods
eReferences
eFigure 1. Study design
eFigure 2. cfDNA concentration after extraction from whole blood
eFigure 3. Overview of cfDNA AR sequencing assay
eFigure 4. Positive controls for the AR exon 2-8 sequencing assay
eFigure 5. Copy number changes detected in cfDNA from enzalutamide patients
eFigure 6. Comparison of AR mutant allele fraction between baseline and progression cfDNA samples
eFigure 7. Regression of W742L/C positive clone after switch to enzalutamide treatment
eFigure 8. Inter-platform validation of AR mutation status in cfDNA from patients at progression on enzalutamide
eFigure 9. Germline DNA repair defects and somatic loss-of-heterozygosity from cfDNA profiling

This supplementary material has been provided by the authors to give readers additional information about their work.
eMethods

Deep AR gene sequencing from cfDNA

To generate deep sequencing data from the AR gene we designed a PCR-based assay to amplify exons 2-8 of the AR gene (eFigure 3). From each sample, a total of fifteen ~150bp fragments coding for exons 2-8 of the AR were amplified in three pools of PCR reactions, using Q5 Hot Start High-Fidelity DNA polymerase and buffer (NEB), 1ng of cfDNA template, and 35 PCR cycles. Each pool of amplicons was run on an ethidium bromide gel to confirm success, and clean-up was performed using Agencourt AMPure XP beads (Beckman Coulter). In the second step, universal adapters specific for Illumina sequencing were added using PCR with ultramer primers (Integrated DNA Technologies) complimentary for each of the fifteen amplicons and tagged with the universal sequence. This PCR used the Q5 Hot Start High-Fidelity DNA polymerase and buffer, 2ul of template, and 6 PCR cycles. Gel visualization and clean-up was performed as described. Finally, amplicons were pooled by sample and a sample-specific barcode added using PCR as described above, and appropriately tagged ultramer primer oligonucleotides. After clean-up, libraries were normalized by measuring concentration with a KAPA library quantification kit (KAPA biosciences) and diluting each library to 4nM. Sequencing was performed using an Illumina MiSeq, in multiplexed batches of 12-48 libraries, and with a 1 million (M), 15M or 25M read kit, according to the manufacturer’s instructions. Each sequencing run included positive and negative controls in the form of prostate cancer cell line DNA with known AR mutation status titrated down to 1% or 5% dilution. Reads were mapped to the human genome (hg19) with bowtie2. At each base position the number of reads with all four possible bases was determined using bam-readcount v.0.7.4. Only bases with phred score higher than 25 and mapping score of greater than 18 were counted. Raw base counts were converted to percentage values relative to total high quality (phred > 25, mapQ > 18) coverage at a given location. Candidate mutations were identified as non-reference base calls with coverage greater than 100 reads, non-reference percentage >1%, and being present in <75% of samples. Mutation calls were manually curated using the Integrative Genomics Viewer (IGV) to remove false positives (e.g. due to amplicon bias). To calibrate the assay we tested the detection limits of the AR H875Y and T878A mutations present in prostate cancer cell lines 22RV1 and LNCaP respectively, when diluted with ‘wildtype’ AR from the LAPC4 cell line to varying concentrations. All patient cfDNA sequencing reads are available at the European Nucleotide Archive (ENA), accession number PRJEB11659.

For previously validated mutations we set a detection threshold of 1% in cfDNA. These included four recurrent mutations in the AR that confer broadened ligand specificity and therefore therapy resistance during progression to CRPC (L702H, W742L/C, H875Y and T878A)\(^1\), as well as the F877L mutation detected in a few long-term treated patients progressing on enzalutamide or ARN-509, shown to convert antagonist to agonist\(^2\). To minimize the risk of false positives we required ‘novel’ mutations to be detected at >2%.

Targeted sequencing of prostate cancer genes

We designed an Ion Ampliseq Custom DNA panel (Life Technologies) to capture the exonic regions of the following 19 genes: AR, ASXL1, BRCA1, BRCA2, CHD1, CHEK2, CTNNB1, FOXA1, HSDB31, KDM6A, MED12, KMT2A, MYC, OR5A1, PIK3CA, PTEN, SCN11A, SPOP and TP53. A minimum of 10ng of cfDNA or leukocyte DNA was then used to generate sequencing libraries with the Ion AmpliSeq™ Library Kit 2.0 (Life Technologies). Libraries were sequenced in batches of 8 to 12 on an Ion Proton (Life Technologies) with an Ion PI sequencing reagents kit 200 v3 and Ion PI chip kit v3. Reads were trimmed to a maximum length of 200 bp, then aligned with bowtie2 against GRCh37\(^1\). Somatic mutations were called by comparing allele read counts between cfDNA and paired WBC samples. We required a somatic mutation in a cfDNA sample to have an allele fraction of at least 1.0%, to have a 20 times higher allele fraction than the paired WBC, to have a 20 times higher allele fraction than the median of all WBC samples, and to have a statistically significant difference in the mutant allele read count relative to paired WBC (p < 1e-20, chi-square test). Finally, all somatic mutation candidates were inspected visually for
strand bias, amplicon bias, and positional bias within reads. Germline variants were called by searching for positions where at least one WBC sample displayed an alternate allele fraction of 40% or more, and no WBC sample displayed an alternate allele fraction between 10% and 35% (to remove false positives due to bases with high error rates). Both somatic mutations and germline variants were annotated for biological significance using ANNOVAR version 170715. To call copy number changes, aligned reads were first counted at each amplicon using bedtools-2.25.0. Sample-specific differences in overall read coverage were multiplicatively corrected using medians-of-ratios calculated across amplicons located in chromosomes 1, 3, 11, 20 and 22. These chromosomes were selected because they have historically displayed low rates of copy number alterations in prostate cancer. After normalization, amplicon coverage logratios were calculated between cfDNA and paired WBC samples. Gene coverage logratios were calculated as the median coverage logratio across all amplicons inside the gene. The logratio thresholds for copy number calls in sequencing data (and in aCGH data) were as follows: single copy gain > 0.5 (0.1); single copy loss < -0.5 (-0.1); amplification > 1 (1.2); homozygous loss < -1 (-1.2).

eReferences

eFigure 1: Study design

Consort diagram showing high success rate of cfDNA extraction and genomic profiling at baseline, on-treatment (12-weeks post-commencement) and at progression. Note that 40/65 of the baseline enzalutamide samples were ‘end-of-treatment’ samples from our previous study¹¹, where aCGH was performed. Therefore aCGH was not required to be re-performed for these baseline samples and they were subjected to deep AR sequencing only. CRPC = castration-resistant prostate cancer; ENZ = enzalutamide; Hx and Ex = history and physical examination; WBC = white blood cells.
Plasma samples collected while patients were on-treatment yielded significantly less cfDNA than those from patients at progression (median 1.00 vs. 3.11 ng/ul; p=0.025 (t-test)). This likely reflects a reduction in the tumor contribution to total cfDNA, commensurate with a lower tumor burden in patients responding to therapy. Note that total elution volume varied from 20-60µl. enza = enzalutamide.
**eFigure 3**: Overview of cfDNA AR sequencing assay.

A) Step 1 of the assay involves three sets of multiplexed PCR reactions, each from 1ng of cfDNA, to generate amplicons spanning exons 2-8 of the AR gene. Small amplicon sizes are required, given the degraded nature of cfDNA. B) After cleanup, universal sequence tags are added by PCR to each amplicon pool. C) Finally, sample-specific barcodes are added by PCR, and libraries are sequenced (after quantification and normalization) using an Illumina MiSeq machine.
eFigure 4: Positive controls for the AR exon 2-8 sequencing assay.

A) Two cell lines (22rv1 and 49C [an LNCaP derivative]) with different AR mutations were sequenced at varying admixture with wildtype AR cell line LAPC4. 1ng DNA from each dilution was sequenced in triplicate. Bar plots show the frequency of detected mutations +/- standard error. Below 2% mutant input, true positive detection was more variable, as sampling probability begins to affect the detection of a single mutant molecule among the 150 diploid genomes in 1ng DNA. B) Sequence coverage for each amplicon, demonstrating that depth across each mutation (in exon 8; AR8-1) is >50,000X.
eFigure 5: Copy number changes detected in cfDNA from enzalutamide patients.

A) Overview of autosomal profiles showing changes consistent with prostate cancer. In this Figure, blue indicates gain; red loss (due to output of Nexus Biodiscovery software. Frequency of select aberration at baseline are shown at the top. Prog = progression samples. B) Chromosome X profiles showing focal amplification or gain of AR in the cohort at baseline on enzalutamide. C) Focal PIK3CB amplification detected in VC-085 at progression.
**eFigure 6:** Comparison of AR mutant allele fraction between baseline and progression cfDNA samples.
eFigure 7: Regression of W742L/C positive clone after switch to enzalutamide treatment.

Shows the clinical history of an independent case (outside of the current ‘enzalutamide cohort’) where treatment was changed in response to detection of W742L/C AR mutation in a cfDNA sample. The top panel shows serum PSA levels over time together with treatment regimens. Annotated below the chart are the three samples obtained during progression: a diagnostic biopsy, and two minimally-invasive plasma cfDNA samples. Two AR mutations were identified in the cfDNA sample obtained in September 2014, at progression on bicalutamide. Both mutations are known to convert bicalutamide from AR antagonist to agonist. Neither mutation was detectable in the diagnostic biopsy, nor in the second cfDNA sample obtained after the patient responded to a switch to enzalutamide and abiraterone.
**eFigure 8:** Inter-platform validation of AR mutation status in cfDNA from patients at progression on enzalutamide.

A) Absolute detection of AR mutation was consistent between Illumina and Ion Proton sequencing workflows in all except one sample. B) The allelic-frequency of AR mutation was also consistent.
**eFigure 9:** Germline DNA repair defects and somatic loss-of-heterozygosity from cfDNA profiling.

a) Defects to DNA damage repair genes detected in 30 patients at progression on enzalutamide. b) Sanger sequencing traces confirming selected events in patient germline DNA. c) Protein schematics showing mutation location. d) Allelic-frequency of selected germline mutations showing evidence of loss-of-heterozygosity in patient cfDNA at progression on enzalutamide, and demonstrating the potential of cfDNA to act as a biomarker for bi-allelic BRCA2 loss in patient tumors.