Supplemental Methods

Single Molecule Sequencing and AR transcriptional signature

RNA was extracted from cell lines using the RNeasy Micro Kit (QIAGEN), and then reverse transcribed using the Superscript III First-Strand Synthesis System (Invitrogen). Purified RNA was subjected to Digital Gene Expression (DGE) sample prepping and analyzed on the HeliScope™ Single Molecule Sequencer from Helicos BioSciences (Cambridge, MA) (1). To compute digital gene expression (DGE) from Helicos sequence data, we used the DGE pipeline of the helisphere software (open.helicosbio.com) using the Human.Txome (ftp://ftp.helicosbio.com/pub/distribution) reference file. We ignored DGE output corresponding to mitochondrial and ribosomal RNA and spike-ins. Comparisons of Helicos digital gene expression (DGE) profiles were performed using the DEGexp function of version 1.4.3 of the Bioconductor (www.bioconductor.org) DEGseq package (**Supplemental Figure S1a**) (2). The "method" argument was set to "MARS" for the MA-plot-based method with Random Sampling. To generate the AR transcriptional signature, DGE profiles were compared for LNCaP cells cultured in media containing 10% charcoal-stripped serum for 3 days followed by treatment with 1 nM R1881, 10 μM bicalutamide, or 0.01% DMSO for 24 hours. The genes with the most significant p-values that were more highly expressed with R1881 (12 genes) or bicalutamide treatment (12 genes) were identified as "AR-on" and "AR-off," respectively (**Supplemental Figure S1b**). Because CTC samples are often contaminated with white blood cells, we excluded from the AR signature transcripts that are highly expressed in white blood cells (above 20 per 100,000) based on Helicos DGE profiles of white blood cells from a healthy donor. Also excluded from consideration for the signature were transcripts that did not correspond to RefSeq genes.

HBCTC-Chip Fabrication

Microfluidic HBCTC-Chip devices were made of PDMS bonded to glass substrates using soft lithography techniques, and functionalized with epithelial cell adhesion molecule antibody (EpCAM, R&D Systems) using avidin-biotin chemistry, as previously described (3).

HBCTC-Chip Blood Processing

All blood specimens were collected into Vacutainer (Becton-Dickson) tubes containing the anticoagulant EDTA and were processed through the HBCTC-Chip within 6 hours of blood draw. Samples were run on the previously described microfluidic processing machine (3). Briefly, a 5 mL aliquot of blood was placed in an airtight conical tube on a rocker assembly, and 2~4 mL of blood were pneumatically driven through the chip at a flow rate of 1-1.5 mL/hour. Following processing, the HBCTC-Chip was flushed with 2.5 mL of PBS at 2.5 mL/hour to remove nonspecifically bound cells. Isolated CTCs were then subjected to immunofluorescence staining, as described below.

Automated Fluorescence Microscopy

An automated upright fluorescence microscopy scanning system (BioView) fitted with a precision motorized stage and xenon arc lamp (Lumen 2000, Prior Scientific) was used to comprehensively image each CTC-chip under 10X magnification in seven z-planes. Due to inherent autofluorescence of biological samples that can interfere with the specificity of stains in the green spectra, the negative control marker, CD45, was paired with the Alexa Fluor 488 secondary antibody. This choice reduced the sensitivity of the staining assay (potentially more false negatives), but lessened the risk of false positives through natural autofluorescence. Conversely, PSA, our highest affinity antibody (and subsequently

strongest fluorescent signal) was placed in the Cy5 channel, where the quantum efficiency of traditional monochrome CCD sensors is reduced. To successfully achieve distinct, non-overlapping fluorescent signals in four colors while maximizing the fluorescent intensity output, modified Magnetron sputter-coated filter sets were selected for the Cy3 and Cy5 spectra (Chroma). Additional filter sets for the DAPI and FITC channel were used and exposure times were optimized. All samples were subsequently imaged at predetermined exposure times. Potential CTC targets were automatically classified using a previously described algorithm based on predetermined fluorescence intensity and cell morphology criteria (4), followed by manual validation by a blinded human reviewer. CTC counts were tabulated based on the total number of cells that were positive for PSA and/or PSMA, and negative for CD45. Normalized counts (CTC/mL) were calculated by dividing the total CTC count by the total volume of blood processed. Based on analysis of blood from healthy donors, a signal intensity threshold of detection was determined to be 4 CTC/mL (**Supplemental Figure S5**), and normalized counts which fell below this threshold were considered to be false positive events and were not included in the final analysis.

References for Supplement

1. Lipson D, Raz T, Kieu A, Jones DR, Giladi E, Thayer E, et al. Quantification of the yeast transcriptome by single-molecule sequencing. Nat Biotechnol 2009; 27: 652-8.
2. Wang L, Feng Z, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 2010; 26: 136-8.
3. Stott SL, Hsu CH, Tsukrov DI, Yu M, Miyamoto DT, Waltman BA, et al. Isolation of circulating tumor cells using a microvortex-generating herringbone-chip. Proc Natl Acad Sci U S A 2010; 107: 18392-7.
4. Stott SL, Lee RJ, Nagrath S, Yu M, Miyamoto DT, Ulkus L, et al. Isolation and characterization of circulating tumor cells from patients with localized and metastatic prostate cancer. Sci Transl Med 2010; 2: 25ra3.