Clinical trials of high-dose androgen therapy for prostate cancer have shown promising efficacy but are limited by lack of criteria to identify likely responders. To elucidate factors that govern the growth-repressive effects of high-dose androgens we applied an unbiased integrative approach utilizing genetic screens and transcriptional profiling of prostate cancer cells with or without demonstrated phenotypic sensitivity to androgen-mediated growth repression. Through this comprehensive analysis, we identified genetic events and related signaling networks that determine the response to both high-dose androgen and androgen withdrawal. We applied these findings to develop a gene signature that may serve as an early indicator of treatment response and identify men with tumors amenable to high dose androgen therapy.
Molecular determinants of response to high dose androgen therapy in prostate cancer

Authors: Michael D Nyquist¹, Alexandra Corella¹, Osama Mohamad², Ilsa Coleman¹, Arja Kaipainen¹, Daniel A Kuppers¹, Jared M Lucas¹, Patrick J Paddison¹, Stephen R Plymaste³,⁴, Peter S Nelson¹,³, Elahe A Mostaghel¹,³,⁴*

Affiliations:
¹Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
²Faculty of Medicine, Benha University, Benha 13518, Egypt.
³Department of Medicine, University of Washington, Seattle, Washington
⁴Geriatric Research, Education and Clinical Center, VA Puget Sound Health Care System, Seattle, Washington

*To whom correspondence should be addressed:
Elahe A. Mostaghel MD PHD
Geriatric Research, Education and Clinical Center S-182
VA Puget Sound Health Care System
1660 South Columbian Way
Seattle, WA 98108-1597
emostagh@fredhutch.org

The authors have declared that no conflict of interest exists.
Abstract: Clinical trials of high-dose androgen therapy for prostate cancer have shown promising efficacy but are limited by lack of criteria to identify likely responders. To elucidate factors that govern the growth-repressive effects of high-dose androgens we applied an unbiased integrative approach utilizing genetic screens and transcriptional profiling of prostate cancer cells with or without demonstrated phenotypic sensitivity to androgen-mediated growth repression. Through this comprehensive analysis, we identified genetic events and related signaling networks that determine the response to both high-dose androgen and androgen withdrawal. We applied these findings to develop a gene signature that may serve as an early indicator of treatment response and identify men with tumors amenable to high dose androgen therapy.
INTRODUCTION

Prostate cancer (PC) is driven by androgen receptor (AR) signaling. Androgen deprivation therapy (ADT) is the frontline treatment for metastatic prostate cancer. While initially effective at suppressing growth of PC, ADT inevitably fails after 1-2 years giving rise to castration-resistant prostate cancer (CRPC), in addition to being associated with significant quality of life complications(1). Targeting the AR program with next-generation AR-signaling inhibitors like enzalutamide (ENZ) or abiraterone can be effective but responses are not durable and most men will progress within months(2, 3). CRPC continues to be driven primarily by AR signaling through diverse mechanisms that include AR amplification, AR mutations, intra-tumoral synthesis of androgens, and the production of ligand independent AR splice variants(4).

AR signaling can be dichotomous, promoting growth at lower androgen levels while suppressing PC growth when hyper-stimulated by high doses of androgen(5, 6). The observation that PC cells can adjust to too little or too much AR-signaling over time by up- or down-tuning AR-activity(7) has led to the development of an approach termed bipolar androgen therapy (BAT) in which periods of high-dose testosterone (T) injections, given every 28 days, are cycled with periods of ADT, theoretically preventing cells from adapting to a static androgen environment while improving quality of life(5, 8). This approach has achieved 40-60% response rates in clinical trials of CRPC patients(9, 10). A recent study of BAT for ADT-naive PC patients met its primary endpoints for efficacy and indicated improved quality of life(10). A phase II study of BAT in CRPC patients previously treated with ENZ reported BAT as safe, effective, and able to re-sensitize 52% of PC to ENZ(11). These trials and anecdotal reports of exceptional responses to BAT(12) highlight the potential of this therapy.

A diverse array of mechanisms have been proposed to explain the growth-repressive effects of high-dose androgen (HDA): 1) Transcription of cell cycle inhibitory genes like p16 and p21(13-15); 2) Transcriptional suppression of MYC, CDKI, and BCL2(16, 17); 3) Induction of DNA-damage via transcription stress induced by TOP2B-mediated double strand DNA breaks(6, 18); 4) Induction of DNA-
damage and senescence by the production of reactive oxygen species (ROS) via acute SRC-AKT activation\(^{19, 20}\); 5) Induction of ROS through AR-induced SMAD signaling\(^{(21)}\); 6) Induction of ROS and p21 in a MTOR dependent manner\(^{(22)}\); 7) Prevention of replication relicensing through the stabilization of AR on DNA\(^{(6, 23)}\); 8) Androgen induced expression of endoplasmic reticulum stress response genes such as \(NDRGI\)\(^{(24, 25)}\); and 9) RB1-dependent suppression of cell cycle-related genes like \(MCM2\), \(MCM7\), and \(FANCI\)\(^{(26)}\). Despite these mechanistic insights, the genetic factors that mediate sensitivity or resistance to HDA are poorly understood and criteria for prospectively identifying responders or non-responders to HDA-based therapies have not been established.

To identify factors that govern response to this novel treatment modality we applied an integrative approach utilizing genetic screens and RNA-seq analysis to identify and functionally annotate genetic determinants mediating sensitivity and resistance. By characterizing the transcriptomic changes associated with low, proliferative, and high AR activities, we identified novel modalities of the AR-signaling program that governed growth and survival. In a clinical context, coordinate use of these novel gene expression signatures may identify castration resistance prostate cancers that are likely to respond to AR-directed therapies.

**RESULTS**

**Androgen-mediated suppression of prostate cancer growth**

To investigate the molecular determinants of response and resistance to HDA we first characterized the responses of PC models to dose-ranges of androgen. Under normal growth conditions (10% FBS), increasing androgen concentrations potently repressed the growth of LNCaP (Figure 1A), VCaP (Figure 1B), 22PC (Figure 1C), and LAPC4 (Figure 1D) starting around 10 pM R1881, whereas R1-AD1 (Figure 1E) and 22Rv1 (Figure 1F) were essentially resistant to the growth-modulating effects of androgen. Charcoal-stripped serum (CSS), in which steroidal compounds and cytokines such as IGF are depleted\(^{(27)}\), is routinely used to mimic low androgen conditions and investigate the effects of high-dose androgens \textit{in vitro}\(^{(6, 18)}\). LNCaP (Figure 1G) and VCaP cells (Figure 1H) displayed a biphasic growth
response when treated with R1881 in CSS media, with maximal proliferation induced at 100 pM R1881 and less proliferation induced by doses over 1nM. R1AD1 and 22Rv1 cells treated with androgen in CSS displayed a weak biphasic response (Supplementary Figure 1A,B). However, in CSS media no dose of androgen suppressed viability below baseline, suggesting experiments performed in CSS inadequately model the growth suppressive effects of androgen. Furthermore, FBS-containing media represents a more physiological environment, with androgen levels similar to those observed in CRPC patients(28), and thus may be more appropriate for molecular studies evaluating the phenotype and mechanism of androgen-mediated growth suppression.

Independent regulation of mitotic and biosynthetic AR functions

To characterize the transcriptional changes caused by HDA we performed RNA-seq analysis comparing three phenotypically-defined androgen-signaling induced growth states (all in 10% FBS containing media at 48 hours): 1) growth suppression in response to ADT (10uM ENZ); 2) proliferation under normal growth conditions (vehicle control); and 3) growth suppression in response to HDA (10nM R1881) (Supplementary Figure 2A). We observed substantial overlap in genes regulated by HDA vs vehicle across the four cell lines (FDR < 0.05) with 1626 genes upregulated (Supplementary Figure 2B) and 1667 genes downregulated (Supplementary Figure 2C) between VCAP, 22PC, LNCaP, and LAPC4. Geneset Enrichment Analysis (GSEA)(29) revealed positive enrichments of Hallmark genesets associated with canonical AR functions and luminal differentiation (e.g. Androgen_response, Adipogenesis, and Cholesterol_homeostasis), and suppression of genesets related to proliferation and cell cycle (e.g. E2F_targets, DNA_repair, G2/M_checkpoint, Mitotic_spindle, and MYC_targets_V1/2 (Figure 2A)). Fewer genes were significantly regulated by ENZ across all four cell lines, with 384 genes upregulated (Supplementary Figure 2D) and 679 genes downregulated (Supplementary Figure 2E). As expected, inhibition of AR activity with ENZ strongly downregulated genesets related to canonical AR signaling, as well as genesets associated with cell-cycle and proliferation (Figure 2B).
To define the molecular signaling underlying ADT- and HDA-mediated growth repression we identified genes that were concordantly repressed (termed biphasic genes, Figure 2C, blue) or concordantly induced (termed inverse-biphasic, Figure 2C, red) in response to both ENZ-mediated ADT and R1881-mediated HDA. We also identified genes that were oppositely regulated by ENZ and R1881, including AR-induced genes (Figure 2C, green) with decreased expression at low AR activity and increased expression at high AR activity, and AR-repressed genes (Figure 2C, yellow) with increased expression at low AR activity and decreased expression at high AR activity. Genes with biphasic expression comprised the largest subset, with 288 genes (66%) similarly regulated across all four cell lines (LNCaP, 22PC-EP, VCaP, LAPC4) (Figure 2D, Supplementary Figure 2F, Supplementary Table 1). AR-induced genes comprised the next largest category with 66 genes (15%). When the stringencies were relaxed to include genes that were commonly regulated in three out of four cell lines, the number of genes in the biphasic subset increased to 608 and in the AR-induced subset increased to 436 (Supplementary Figure 2G, Supplementary Table 1).

Pathway analysis using PANTHER Gene Ontology(30) (GO) overrepresentation test revealed that each of these genesets mediated distinct biologic functions. Consistent with the GSEA findings above (Figure 2A,B), biphasic genes were associated with proliferation and DNA replication with pathway enrichments for cell-cycle, mitosis, chromosome segregation, DNA-metabolizing pathways including de novo pyrimidine deoxyribonucleotide synthesis, and DNA replication (Figure 2E). The inverse-biphasic category of genes did not include any statistically overrepresented pathways across all four cell lines, but in three of four cell lines the data showed pathways relating to PDGF signaling and transcription factor activity were enriched (Figure 2F). This category also included notable PC-related transcription factors NCOA3 and FOXO3 (Supplementary Table 1). AR-induced genes consisted of metabolic and biosynthetic programs such as lipid and cholesterol biosynthesis, as well as glycolysis and gluconeogenesis (Figure 2G). Of note, AR activity drives lipid biosynthetic programs largely through the transcription factor SREBF1(31, 32), and SREBF targets were overrepresented in the AR-induced
genes. AR-repressed genes were involved in the regulation of protein translation and ribosome biosynthesis (Figure 2H). Notably, GATA2, a known effector of castration resistance, was represented in the AR-repressed subset (33, 34). These data suggest that not all AR responsive genes are linearly AR-induced or AR-repressed, and that each subset of AR regulated genes associates with distinct biologic functions. For example, metabolic and biosynthetic functions may be regulated independently from cell-cycle and proliferative functions.

*Enrichment of cell cycle drivers in overexpression screen for HDA resistance*

While RNA-seq studies provide a global picture of the effects of AR signaling on cellular pathways involved in regulating the growth-repressed phenotypes, they do not provide functional information on specific genetic mechanisms. We therefore pursued an unbiased functional genetics approach employing whole genome overexpression and knockdown screens to identify molecular pathways functionally relevant to the growth-repressed phenotype, and determined the overlap of these findings with the signaling pathways identified in the RNAseq analysis. We performed a whole-genome open reading frame (ORF) library screen on LNCaP cells under conditions of HDA. Briefly, LNCaPs were transduced with the ORF library and cultured in normal growth media with vehicle or 10nM R1881 for 25 days at which time DNA was harvested and sequenced (primers in Supplementary Table 2). Barcode read-counts at the end of the 25 days were compared to day 0 to determine depletion or enrichment of the associated ORF. Overexpression of genes in several signaling programs mediated resistance to HDA, including i) cell-cycle factors that govern the G1/S transition such as D-type cyclins and their binding partners CDK4 and CDK6; ii) modulators of AR-signaling such as PIAS1 and FGFR1; and iii) stem cell-related and reprogramming factors such as SOX2, KLF4, LIN28A, and WT1 (Figure 3A).

Given that amplification and overexpression of factors that induce cell cycle are common events in CRPC (35), and that these factors regulate the same cellular processes as the biphasic genes associated with the growth repressive phenotype, CDK4/6 and D-type cyclins were chosen for further validation in competitive enrichment assays. As ADT and HDA both arrest cells in the G1-phase of the cell cycle (23),
we hypothesized that cyclin D and CDK4/6 overexpression should mediate resistance to both ENZ and R1881. A competition assay with the two empty vector controls established that the fluorescent proteins themselves did not impart a selective advantage (Supplementary Figure 3A,B). Overexpression of Cyclin D1 (CCND1), Cyclin D2 (CCND2) and Cyclin D3 (CCND3) were strongly enriched in LNCaPs in response to R1881 and ENZ and modestly enriched in VCaPs (Figure 3C). Overexpression of CDK4 and CKD6 similarly promoted growth in LNCaPs in response to R1881 and ENZ, with more modest enrichment in VCaPs. These results demonstrate that increased expression of cell-cycle drivers can override AR-mediated control of proliferation and implicate processes associated with biphasic genes as primary mediators of response to ADT and HDA.

**Enrichment of cell cycle inhibitors in whole genome CRISPR/CAS9 knockout screens for HDA sensitivity**

To complement the ORF screen, we performed whole genome CRISPR/CAS9 knockout screens to identify genes required for response to AR-mediated growth repression. LNCaP cells were transduced with a pooled whole-genome sgRNA library encoding 10-sgRNAs per gene(36), cultured in normal growth media with vehicle, 10uM ENZ, or 10nM R1881 for 25 days and sequenced to identify the sgRNAs enriched or depleted in the ENZ or HDA-treated cell populations. Genes associated with enriched sgRNAs (where knockdown promoted cell growth) or depleted sgRNAs (where knockdown promoted cell loss) were determined by MAGECK analysis(37, 38). We performed four biological replicates and determined overlap between the top-2000 scoring genes for enrichment and depletion across all four replicates.

In agreement with the pathways identified by the RNAseq studies and the ORF screen, sgRNAs targeting critical drivers of cell cycle progression and mitosis like CCNA2, CCND1, CDK2, CDK4, CDC20/27/45, and MYC decreased growth and were depleted regardless of treatment condition (Figure 4A). Of note, SOX4, which positively regulates plasticity factor EZH2 (39) and oncogenic AKT signaling(40) was depleted in all four replicates of the HDA treatment but not the vehicle or ADT conditions. sgRNAs that
target tumor suppressors that inhibit cell-cycle progression in response to stress such as CDKN-family genes, GADD45G, and CHEK2 were enriched under HDA and ADT conditions (Figure 4A). These targets were validated in competitive enrichment assays using separate, design-optimized sgRNAs (Supplementary Figure 3C, Supplementary Table 2). Knockouts of CDKN1A, CDKN1B, CDKN1C, CDKN2B, GADD45G, and CHEK2 in LNCaP cells were selected in multiple conditions but most strongly in the presence of R1881 (Figure 4B). In VCaP cells, loss of CDKN1B, CDKN2A and CHEK2 were modestly enriched in multiple conditions, while loss of CDKN1C, CDKN2B and GADD45G were not strongly selected for (Figure 4B).

Loss of TP53 was enriched in all treatment conditions in both cell models (Figure 4B). Interestingly, RB1 knockout did not influence response to R1881 in LNCaP, while loss of RB1 was enriched in VCaP cells (Figure 4B). Of note, VCaP cells are heterozygous for mutant TP53(41), and TP53 loss can synergize with RB1 loss to promote tumorigenesis(42). To determine whether TP53 loss is necessary for RB1 loss to mediate resistance to HDA in LNCaP, we performed a dual competitive enrichment assay wherein LNCaP cells were transduced with sgRNAs targeting one, both, or neither of TP53 and RB1 and treated with 10nM R1881. Dual loss of TP53 and RB1 was the most strongly enriched (Figure 4C, from 19% to 45%), suggesting therapy resistance may be a result of multiple synergistic events. Overall, these data demonstrate that loss of cell cycle inhibitors and tumor suppressors uncouple both high and low dose AR signaling from control of proliferation, suggesting a mechanism by which resistance to HDA and ADT may be co-acquired during cancer progression.

Functional annotation of AR-regulated genes through targeted CRISPR/CAS9 screens

While sensitivity to ADT and HDA may be influenced by loss or gain of cell-cycle related tumor suppressors and oncogenes, respectively, it is unclear which specific AR-regulated genes are mediators of growth suppression by HDA. To identify these genes, we performed a focused CRISPR/CAS9 screen consisting of AR-responsive genes in LNCaP cells under 10nM R1881 conditions. Targeted screens allow specific components of the transcriptome to be examined in isolation at greater sgRNA-per-gene depth.
and read coverage. Moreover, by removing genes essential for survival that do not change with androgens, AR-responsive genes that are critical for growth and survival can be identified with higher confidence and less noise (Supplementary Table 3).

In agreement with the whole genome CRISPR screen, sgRNAs mediating loss of AR-regulated stress-response genes such $TP53$, $CDKN1A$, and $GADD45G$ countered the growth repressive effect of HDA and were strongly enriched in response to R1881 (Figure 5A). sgRNAs targeting tumor suppressors $PPPIR15A$ and $NDRG4$ were also enriched (Figure 5A). In contrast, sgRNAs targeting AR-regulated genes involved in promoting PC growth and progression were strongly depleted, including $GATA2$, $SOX9$, $BRCA1$, $EZH2$, $BARD1$, $BCL2L1$, $FOXA1$, $WEE1$, and $MYC$ (Figure 5B).

The RNAseq studies described earlier demonstrated that tumor suppressor genes, which were enriched by knockdown in the CRISPR screens, were largely upregulated at the transcript level by HDA (Figure 5C, top). In contrast, with the exceptions of $SOX9$ and $BCL2L1$, genes that supported growth and survival, which were depleted by knockdown in the CRISPR screen, were largely downregulated by HDA (Figure 5C, bottom).

Underlining their functional importance, biphasically regulated genes were strongly clustered among the most depleted of all AR-responsive genes (Figure 5D). Compared to the median rank of 1865, the biphasic genes were more highly represented amongst the top ranked/most depleted genes with a median rank of 422 and 316 in each replicate, respectively, supporting the functional relevance of this geneset for the growth repressive effects of HDA and ADT (Figure 5D). These findings suggest that AR exerts control over the proliferation of PC though multiple factors that act in parallel, including both upregulation of tumor suppressors and suppression of oncogenes and biphasic genes.

**Suppression of biphasic gene expression is an innate function of AR**

A number of the molecular mechanisms proposed for AR mediated growth suppression rely on the premise that CRPC cells have uncontrolled growth and are thus more susceptible to the
transcriptional(18), mitochondrial oxidative(19-21), or mitotic stresses(6, 23) induced by acute doses of androgen. However, AR signaling has been previously reported to arrest the growth of various untransformed cell types including prostate epithelial cells(43). To determine whether AR signaling suppresses biphasic gene expression in untransformed cells in a manner similar to that observed in PC cells, we utilized the hTERT-immortalized prostate epithelial cell (PREC) line 957E/TERT-AR cells, which ectopically express AR(43). These cells were treated with 1nM R1881 or vehicle control for 24hrs and analyzed by RNA-seq. Androgen strongly repressed the biphasic geneset signature and induced the inverse-biphasic gene signature (Figure 6A, top). Among the top 250 genes repressed by androgen were 89 biphasic genes like FOXM1, RAD54L, E2F1, and CCNA2 (Figure 6A, Supplementary Table 4). Inverse-biphasic genes such as FOXO3, MXD4, and SOS2 were among the top 250 upregulated genes, as were AR-targets MAF and FKBP5 (Figure 6A, Supplementary Table 5). As in the cancer cells, biphasic genes partially overlap with E2F1 and MYC targets (Figure 6A). These data suggest that AR-mediated growth suppression is achieved through transcriptional repression of cell-cycle gene expression in both untransformed and transformed cells. 7621010 - 64630.

Previous studies in PCa cells have found E2F1 and AR to coordinately upregulate cell-cycle(44) and DNA-damage response (DDR) genes(45) whereas AR and RB1 coordinately repressed cell-cycle genes(26). To explore the possibility that AR coordinates with RB1 and E2F1 to regulate biphasic genes, we determined the percentage of genes in each of the four AR subsets that associated with peaks bound by either AR, RB1 or E2F1, using previously published ChIP-seq datasets for AR(7, 46), E2F1(44), and RB1(26). Not surprisingly, binding of AR to genes in the AR-induced gene set was enriched. RB1 and E2F1 were also associated with these genes, but at lower levels. (Figure 6B). Binding of RB1 and E2F1 to biphasic genes was enriched by 17% and 33%, respectively, while AR associated with biphasic genes at a lower percentage in these studies. However, these data do not support or exclude the possibility that AR coordinates with E2F1 and RB1 to regulate biphasic targets. AR typically regulates gene expression via binding to intra- or intergenic enhancers, whereas E2F1/RB1 more directly associates with promoters,
making it difficult to unambiguously determine whether an AR-bound enhancer regulates a given biphasic gene. Notably, previous studies have identified AR binding sites associated with the regulation of DDR and cell-cycle genes: AR-induced DDR genes (45) (AR_DDR_UP), AR/RB1-repressed cell cycle genes (26) (AR_CC_DN), and AR/E2F1-activated cell-cycle genes (44) (AR_CC_UP), and there is a high-concordance of these genelists with the biphasic gene category (Figure 6C). The percentage of biphasic genes in these three data sets ranged from 53% to 67% whereas biphasic genes made up only 4% of all genes. Interestingly, biphasic genes MCM7 and FANCI were found to be activated by AR in the AR_DDR_UP study and repressed by AR in the AR_CC_DN study, suggesting that the conventional binary classification of AR-induced or repressed does not fully capture the complex regulatory modalities of AR signaling. Together these data indicate that suppression of biphasically-regulated genes may be a fundamental function of AR that is disrupted upon oncogenic transformation but may be re-established by HDA.

Expression profiling of de novo resistance to high-dose androgen therapy

We next sought to determine whether cells with de novo resistance to HDA would maintain expression of biphasic genes during HDA. We compared the transcriptional responses of the R1AD1 (47) cell line, that was minimally growth-repressed by HDA (Figure 1), to the four cell lines which were strongly growth repressed by HDA. Compared to the HDA-sensitive cell lines, GSEA analysis of HDA-treated R1AD1 cells revealed a positive enrichment of pathways associated with the biphasic geneset, including cell cycle-related E2F_targets and G2M_checkpoint genesets, as well as plasticity-associated epithelial_mesenchymal_transition (Figure 7A). This model was negatively enriched for genesets associated with AR-induced genes and luminal prostate differentiation such as androgen_response, oxidative phosphorylation, adipogenesis, cholesterol_homeostasis, protein_secretion, and fatty_acid_metabolism (Figure 7A).

Accordingly, numerous oncogenes and drivers of cell cycle, including 29 biphasic genes, were within the top 250 genes expressed most highly expressed in R1881-treated R1AD1 cells (Figure 7B,
**Supplementary Table 6.** Biphasic genes, including several E2F targets, were only weakly repressed by ENZ and R1881. In addition, oncogenes that drive proliferation such as MYCL, CDK6, and CCND1 were highly expressed (Figure 7B). This attenuated repression of biphasic genes by R1881 can be seen in FOXM1 and CCNA2 (Figure 7C,D).

Consistent with the GSEA data showing negative enrichment of the AR-induced geneset in R1AD1 cells, the 250 most downregulated genes with R1881 between R1AD1 and HDA-sensitive lines included 82 AR-induced genes such as TMPRSS2, PEMPA1, and AMACR (Figure 7E, Supplementary Table 7). Several important AR-coregulators were also expressed at very low levels in R1AD1 cells including HOXB13, GATA2, SOX9, and FOXA1(Figure 7E). Interestingly, MYC was significantly suppressed under R1881 conditions in R1AD1 cells (Figure 7F), suggesting HDA resistance in R1AD1 cells was not dependent on maintaining MYC expression. Taken together, these data indicate that induction of cellular differentiation and inhibition of mitogenic pathways are partially uncoupled from the control of AR in R1AD1 cells, resulting in resistance to AR-directed therapies.

**Predicting therapy response using Biphasic and AR gene signatures**

Geneset Variation Analysis (GSVA)(48) was then used to compare the four geneset signatures identified in this study to previously published gene sets defining AR-regulation (ARG.10), cell-cycle/proliferation (CCP.31), epithelial to mesenchymal transformation (EMT.12), and the neuroendocrine phenotype (NE.10)(49) (Supplementary Table 8). HDA-sensitive cell models showed strong upregulation of AR-induced and ARG.10 genes and strong repression of the AR-repressed, biphasic, and CCP.31 genesets in response to HDA, whereas none of these gene sets were strongly affected by HDA in R1AD1(Figure 8A). R1AD1 also had constitutively low AR-repressed and inverse-biphasic gene signatures. Neither the epithelial to mesenchymal transition (EMT.12) or the Neuroendocrine signatures (NE.10) distinguished R1AD1 from the other cell lines (Figure 8A). However, the HDA-resistant model R1AD1 could be clearly distinguished from the HDA-sensitive models by having a high biphasic/proliferative but a low AR-induced gene signature under all treatment conditions, including vehicle (Figure 8A).
Analogous tumor samples with a high biphasic/proliferative but a low AR-induced gene signature can be identified in the RNA-seq profiles of 4212 CRPC metastases in the PCF/Stand-Up-To-Cancer cohort (Figure 8B). Similar to R1AD1, tumors that have high biphasic/proliferation signatures but a low AR-induced signatures (quadrant 1, Q1) have uncoupled AR-signaling from the control of proliferation and would likely be poor candidates for HDA-based therapies (Figure 8B). Tumors that have high biphasic/proliferation and high AR-induced signatures (Q2) are more likely to have achieved castration resistance through re-activation of AR signaling and may be good candidates for HDA therapy (Figure 8B). Interestingly, several of the candidate responders in Q2 also had high neuroendocrine signature scores (NE.10), analogous to VCaP cells, raising the possibility that some NE+/AR+ or amphicrine PC subtypes may also benefit from HDA-based therapies.

To determine whether AR signatures can be used as an early indicator of therapy response we applied them to RNA-seq data from a recent study of the LuCaP CRPC PDX models treated with HDA(50). RNA from responder tumor lines (LuCaP96CR and LuCaP35CR) and one non-responder line (LuCaP77CR) were harvested at an early timepoint, five days into treatment. While all lines had significant reductions (p < 0.05) in biphasic gene signature scores in response to testosterone, expression of biphasic genes remained high in the non-responder line (Figure 8C). In contrast, all PDX lines had similar increases in expression of the inverse-biphasic (Figure 8D) and AR-induced genesets (Figure 8E), and decreases in the AR-Repressed gene signature (Figure 8F).

**DISCUSSION**

PC patients display a range of radiographic and PSA responses to HDA-based therapies(11) but the factors that mediate response or resistance to AR-mediated growth suppression are not well understood. Studies of AR signaling commonly use CSS-media with or without high-doses of androgen. As a result, genes are either classified as AR activated or suppressed. However, by tracking how individual genes changed through a phenotypically-defined AR activity spectrum consisting of AR-inhibited growth repression, normal proliferation, and AR-hyperactive growth repression, AR signaling modalities can be
subset into four categories: i) AR-induced genes – representing biosynthetic processes; ii) AR-repressed genes – involved in regulation of protein translation; iii) Inverse-biphasic genes - upregulated in cell-cycle suppressed states and including important stress-response genes in the FOXO family; and iv) Biphasic genes – representing critical drivers of proliferation.

While numerous mechanisms have been proposed to explain the growth suppressive effect of HDA, they primarily conform to one of two models. In one model, AR activity suppresses an oncogene or upregulates a tumor suppressor, either directly or in conjunction with a metabolic stress response to ROS or ER stress(13-17, 26, 51). In another model, acute activation of AR by HDA, when preceded by several days of androgen withdrawal, induces DNA damage leading to a stress response and cell cycle suppression (6, 18-20). In the former model, no single factor regulated by AR has been identified that can completely account for the growth suppressed phenotype. In the latter model, the connection between the acute and transient DNA damage signaling foci observed after acute AR stimulation and the subsequent phenotype of growth suppression after culturing cells with androgens continuously for several days has not been established (9).

Our data suggests that sensitivity or resistance to HDA and ADT is multifactorial, and a continuum of phenotypes likely exist. Genetic screening established that numerous genes mediating co-resistance to ADT and HDA converged to regulate cell-cycle and biphasic gene expression. Integrating RNA-seq datasets with a custom, AR-focused, CRISPR/CAS9 screen revealed that HDA-mediated AR-activity suppressed growth through the upregulation of tumor suppressors like CDKN-family genes, PPP1R15A, and GADD45G as well as through the suppression of factors that drive PC growth like EZH2, MYC, and E2F1. These data also support the conclusion that de novo resistance to HDA therapy may be co-acquired with ADT resistance through the uncoupling of AR from control of the cell cycle, potentially explaining why not all CRPC patients are sensitive to HDA therapy. Another means by which cells could evade suppressive AR functions is through changes to cellular plasticity. In this regard, the ORF screen data suggests that overexpression of reprogramming factors LIN28A and SOX2(52) many enhance resistance.
Additionally, the HDA-resistant line R1AD1 overexpressed oncogenes not seen in the HDA-sensitive lines such as *MYCL*, *MET*, and *BCL2*.

Homologous recombination deficiency (HRD) has been suggested to indicate sensitivity to HDA (12) however, the genomic changes in HRD related genes in our cell models precluded a meaningful analysis. For example, while the responder cell line LNCaP has a heterozygous *CHEK2* mutation, the other responder cell lines, including our non-transformed epithelial cell line, are not known to have HRD mutations (53). In contrast, the non-responder R1AD1 is a subclone of the CWR-R1 tumor which may have a defect in the HR gene *MRE11* (53). Moreover, while a responder PDX model, LuCaP 96 CR, has *BRCA2* loss, the other responder, LuCaP 35 CR, is HRD proficient (50). Our data does not support or refute the hypothesis that sensitivity to HDA therapy is related to DNA-damage, but demonstrates that factors other than DNA-repair are involved in response to HDA.

AR is a master-regulator of cell homeostasis controlling the maintenance of prostatic cellular identity and metabolism. Accordingly, AR signaling is likely to be integrated into negative-feedback mechanisms that dynamically regulate cell physiology when hormone signaling is perturbed. Overall, these experiments support a model wherein AR signaling influences a network of tumor suppressors and oncogenes that regulate cell-cycle progression in normal and transformed prostatic tissues (*Figure 8G*). In this model, partial uncoupling of AR from control of cell cycle through gain of oncogenes such as *MYC* and *TMPRSS2-ERG*, or loss of tumor suppressors such as *PTEN*, leads to oncogenic AR-activities and PC. CRPC results from further dysregulation of growth through loss of tumor suppressors or activation of promiscuous and/or ligand independent AR-signaling though AR mutation or amplification. In the latter case, growth suppression and re-differentiation may be induced in CRPC through hyperstimulation of AR (*Figure 8H*).

CRPC samples with high-proliferation scores have variable AR-activity scores. This diversity may be a manifestation of the varied and complex paths taken to castration resistance or androgen independence.
Tumors able to uncouple cell-cycle from AR-control though gain of oncogenes, loss of tumor suppressors, or acquisition of a plastic phenotype are likely to be highly resistant to any AR-directed therapies and will be manifested by high biphasic and low AR-induced signature scores. Alternatively, cells that gain castration resistance mainly through aberrant activation of AR-signaling, achieved by AR mutation, AR amplification, or transcriptional induction of active AR variants, are likely to be sensitive to HDA. Cancers with persistent AR signaling in context of AR-mutations that render them resistant to antiandrogens may be excellent candidates for HDA therapy. Tumors with AR variants generated by structural gene rearrangements that eliminate the ligand binding domain(47) are unlikely to be sensitive to HDA. However, these represent a minority of tumors, and expression of the more commonly encountered AR variants generated by alternative splicing does not preclude sensitivity to HDA, as repletion of androgen is associated with suppression of AR variant expression(54).

Suppression of PC growth by HDA is a new therapeutic approach with the potential to prolong the efficacy of AR-directed therapies while improving quality of life. There is a critical need to identify cancers with de novo resistance either prospectively or soon after treatment has begun so that alternative therapies can be administered. Given the diversity of AR-signaling and the multitude of genetic factors that regulate the cell-cycle, gene signatures may have a superior ability to identify responders over monogenic biomarkers such as PSA or MYC. Evaluation of the LuCaP PDX models suggests that solid or liquid tumors biopsies taken early (within one week) after initiation of HDA that show repression of the biphasic gene signature derived from this investigation may identify men with tumors amenable to HDA-based therapies and serve as an early measure of therapy response. To implement these signatures into clinical practice, gene expression profiles of CRPC patient tumors or circulating tumor cells prior to and during HDA therapy should be correlated with outcomes to define the optimal criteria for their use.

METHODS

Dose-response assays
Cells lines were obtained from ATCC and used within 20 passages or provided by other laboratories (see acknowledgements). Dose response Assays were carried out in 96-well plates (n=4) in 100uL/well of 10% FBS or 10% CSS supplemented phenol-free RPMI-1640 (Thermo, #11835-030) for LNCaPs, R1AD1, 22Rv1, and 22PC-EP, or DMEM/F12 (Thermo, #11039021). Cells were cultured for 5 days then harvested with CellTiter-Glo 2.0 (Promega, #G9241) and measured on a Synergy plate reader (Biotek). 22PC-EP were a gift from the lab of Dr. Charles Sawyers.

**Competitive enrichment assays**

To create the ORF vectors for the competition assay, ORFs were amplified from an arrayed ORF library(55), TOPO-TA cloned in to PCR8/GW (Thermo, #K250020), then gateway cloned into the FU-CGW lentiviral vector (a gift from the lab of Dr. Owen Witte) using LR clonase (Thermo, #11791100). Transduced cells were mixed in a 1:1 ratio with the FU-CRW control vector and cultured in normal growth conditions with DMSO-vehicle, 10nM R1881, or 10uM enzalutamide (Selleckchem, #S1250). A day 0 timepoint was analyzed by flow cytometry. The remaining cells were cultured for 25 (LNCaP) or 30 (VCAP) days then analyzed by flow cytometry. Dual TP53 and RB1 competition assays were cultured in R1881 for 42 days.

CRISPR/CAS9 competition assays were performed similarly using the pZHB-Z:U6_EFS-GFP (or mCherry)-Puro sgRNA vector (a gift from the lab of Dr. Patrick Paddison) which co-expressed either RFP or GFP with the non-targeting control and test sgRNAs respectively. Test sgRNAs were designed using Broad design tool (portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design) and cloned using ESP3I sites and overlapping annealed oligos (www.idtdna.com).

**RNA-seq analysis**

1.5*10^6 cells were plated in 6 well plates in normal growth media, 24 hours later 10uM ENZ (SelleckChem, #S1250) or 10nM R1881 (Sigma, #R0908) added and cultured for 48 hours at which point the cells were harvested using Qiagen RNAeasy kit (Qiagen, #74104). PREC lines were treated with 1ug/mL Doxycycline (Sigma, #D9891) and 1nM R1881 and harvested after 24 hours. Libraries were
prepared, sequenced, and aligned as previously described(49). Differentially regulated genes were analyzed using EdgeR(56). Geneset enrichment analysis performed as described(28) using Pre-ranked genelists with classic weighting. Geneset Variation Analysis performed using the GSVA R-package(48). PANTHER Gene Ontology overrepresentation tests used the online tool (www.pantherdb.org) as previously described(30).

**CRISPR/ORF screens**

Human GeCKOv2 CRISPR knockout pooled library was obtained from Feng Zhang (Addgene #1000000048, #1000000049). LNCaP cells were transduced with the pooled sgRNA libraries, cultured for 25 days with 10uM ENZ, 10nM R1881, or DMSO vehicle control. 4*10^7 cells were used per group for a 100-200 fold coverage of the library. DNA was harvested Blood & Cell Culture DNA Midi Kit (Qiagen, #13343). Barcode amplification was performed with two rounds of amplification using Phusion-HF polymerase (Thermo, #F530L) each 14 cycles using primers in Supplementary Table 2 and sequenced as previously described(36). Hits were identified using MAGECK analysis(37, 38).

The custom AR-regulated CRISPR screen was designed using optimized sgRNAs assembled from several studies(57-59) (Table S3). Oligos were ordered from custom array (www.customarrayinc.com), PCR amplified, and cloned into a custom vector pZHB-Z:U6_EFS-GFP-Puro, a gift from the lab of Patrick Paddison (FHCRC. Seattle, WA) using Esp3I enzyme sites (Thermo, #ER0451) and T4 ligase (NEB, #M0202S). The screen was performed and sequenced similarly to the whole genome except 2*10^7 cells were used to achieve 500x library coverage.

The TRC ORF screen, a gift from the lab of Patrick Paddison (FHCRC. Seattle, WA), was performed similarly to the custom CRISPR screen with two rounds of barcode amplification using primers in Supplementary Table 2.

**ChIP-seq Peak annotation**
ChIP-seq datasets were downloaded from the GEO database (www.ncbi.nlm.nih.gov/geo) as .bed files and converted to HG38 using the liftover tool (https://genome.ucsc.edu/cgi-bin/hgLiftOver). Peaks were annotated using the PAVIS tool (https://manticore.niehs.nih.gov/pavis2/) using default settings. The following datasets were used: AR in VCAP - GSM801011; AR in LNCAP - GSM980662; Rb VCAP – GSM1974981, Rb LNCAP – GSM1974982, Rb VCaP – GSM1974981, E2F1 LNCaP GSM2492421.

Statistics Overview

Significance was defined by a False Discovery Rate (FDR) value of less than 0.05 as determined by: EdgeR for differentially expressed RNA-seq values; PANTHER Gene Ontology and Gene Set Enrichment Analysis for pathway and gene set analysis; MACS2 for ChIP-seq peak calling; and PAVIS for gene annotation.

Study Approval

No human or animal subjects were used in this study.

AUTHOR CONTRIBUTIONS

MDN, AC, OM, and JL carried out in vitro experiments and RNA-seq analysis. PJP and DAK advised on the design and analysis of sgRNA screen. MDN, SP, PSN, EAM designed study and edited manuscript. MDN and EAM wrote the manuscript.

ACKNOWLEDGEMENTS

We would like to thank John Isaacs for providing PREC/AR cells and Charles Sawyers for providing 22PC-EP cells. This study was supported by the following grants: W81XWH-15-1-0319, W81XWH-16-1-0206, P01 CA163227, P50 CA097186, P30 CA015704, R21 CA230138.
REFERENCES


FIGURE LEGENDS
**Figure 1: Suppression of prostate cancer viability by androgens.** Relative viability (Celltiter-glo) of PC cell lines was measured in response to a dose-range of the androgen R1881 (n=4) in normal growth media (10% FBS) for A) LNCaP, B) VCaP, X) 22PC-EP, D) LAPC4, E) R1AD1, and F) 22Rv1. Dose-responses to R1881 in 10% CSS media after five days in culture for G) LNCaP and H) VCaP.
Figure 2: The androgen receptor signaling program is comprised of multiple modalities. A)
Normalized-enrichment scores (NES) of GSEA Hallmark genesets for all four cell lines. Significant
genesets comparing 10nM R1881 vs vehicle (FDR<0.05) are marked as circles. Androgen Receptor-
related genesets are in labeled red, cell cycle-related genesets in blue. (n=2) B) Same as panel “A” for
10uM ENZ-related gene expression changes. C) AR signaling modalities are diagramed: biphasic (blue),
inverse-biphasic (red), AR-repressed (yellow), AR-induced (green). D) Percentage of AR-responsive
genes in each category. Gene ontology (GO) pathways that are overrepresented in the E) biphasic gene
subset, F) inverse-biphasic subset, G) AR-induced subset, and H) AR-repressed subset.
Figure 3: ORF screen identifies cell cycle-regulator upregulation as a mechanism of dual resistance.

A) Log2 fold change in frequency for two biological replicates of an ORF screen for genes that modulate responses to 10nM R1881 in LNCaP. B) Competition assay schematic to measure the selective advantage or disadvantage of gene. C) Heatmap of competition assays that validated screen hits.
**Figure 4:** Genetic loss of function screens identify factors that mediate response to AR-directed therapies. A) Summary of four whole-genome CRISPR/CAS9 screens. Values (1-4) indicate the number of times the gene appeared in the top 2000 most enriched(+) or depleted(-) sgRNAs by gene. B) Heatmap of validation study indicating the change in percentage of cells that harbor sgRNAs to the indicated gene. C) Competitive enrichment assay for cells harboring sgRNAs targeting either, both, or neither RB1 and TP53.
Figure 5: Focused CRISPR/CAS9 screens identify AR-responsive genes that mediate sensitivity to high-dose androgen. A) Gene enrichment ranks as measured by MAGECK analysis of two biological replicates of an AR-focused CRISPR/CAS9 screen. Known tumor suppressors enriched in both biological replicates are labeled. B) Gene depletion ranks of two biological replicates. Genes associated with prostate cancer growth and progression are labeled. C) RNA-seq gene expression heatmaps of mean-centered log2-CPM values representing genes identified in the whole-genome and focused CRISPR
screens. D) sgRNA depletion rank plot of biphasic genes. E) Median depletion ranks for biphasic genes (green) compared to all other genes (blue). Bars represent Q1 and Q3 quartile ranges.

**Figure 6: Suppression of biphasic gene expression is an innate function of AR.** A) Heatmaps of custom gene expression signature scores (top) and RNA-seq mean-centered log2CPM values (bottom) for the hTERT-immortalized prostate epithelial cell line 957E/hTERT that was engineered to overexpress AR. Cells were treated with either 1nM R1881 or EtOH. Examples were selected from the top 250 up- or downregulated genes. B) Percent of genes in each AR gene signature associated with at least one ChIP-seq peak for AR, RB1, or E2F1 subtracted by the percentage of genes bound by at least one peak in the whole-genome. ChIP-seq datasets for both LNCaP and VCAP were compared. C) Graph of DNA-damage response (DDR) and cell-cycle (CC) genes previously shown to be bound by AR and are regulated by AR activity. Blue squares
indicate the presence of a gene in an AR gene signature for 3/4 HDA-sensitive cell lines. Black bars represent the study that validated the gene as regulated (UP) or downregulated (DN) by androgens.

Figure 7: De novo resistance to high-dose androgen is associated with loss of differentiation and gain of a mitotic gene expression profile. A) GSEA plots of NES and FDR values for differentially expressed genes between R1AD1 and the four HDA-sensitive cell lines when treated with 10nM R1881 (FDR<0.05). B) Heatmaps RNA-seq Log2-CPM values, mean-centered across all cell lines, of a selection of the 250 most upregulated genes in R1AD1 cells treated compared to LNCaP, VCAP, LAPC4, and 22PC-EP when treated with 10nM R1881 (bottom). Black squares indicate positive status as a biphasic gene or Hallmark_E2F_target gene. Log2 fold change in gene expression from vehicle control when treated with 10nM R1881 for C) FOXM1 and D) CCNA2. E) Heatmaps of RNA-seq mean-centered Log2-CPM values of a subset of the 250 most downregulated genes in R1AD1 compared to LNCaP, VCAP, LAPC4, and 22PC-EP when treated with 10nM R1881 (bottom). Black squares indicate positive status as an AR-induced,
Hallmark_Androgen_Response, or AR-coregulator gene. F) Log2 fold change in MYC expression from vehicle control when treated with 10nM R1881.

Figure 8: AR signatures scores identify potential responders in clinical CRPC samples. A) Heatmaps of GSVA signature scores comparing previously published gene signatures to gene signatures derived in the current study. B) Correlation of biphasic and AR-induced signature scores in n=212 CRPC tumor specimens, colored by neuroendocrine (NE.10) signature score values. Quadrants Q1-4 are delineated by hatched lines. Signature scores are indicated for vehicle treated (veh) or high dose testosterone (HiT) treated responder (R) or non-responder (NR) PDX lines for C) biphasic, D) inverse-biphasic, E) AR-
induced, and F) AR-repressed genesets. G) Diagram of an interactive web of factors that mediate sensitivity (red) and resistance (green) to AR-directed therapies. H) The relationship between AR-signaling and prostate cancer throughout disease progression is diagramed.

SUPPLEMENTARY MATERIALS

List of supplementary materials

Fig S1: Suppression of prostate cancer viability by androgen in CSS conditions.
Fig S2: Suppression of biphasic gene expression is an innate function of AR.
Fig S3: Competition Assays
Supplementary TABLES
Table S1: AR genesets
Table S2: Primers
Table S3: Custom sgRNA library
Table S4: Genes upregulated in R1AD1
Table S5: Genes downregulated in R1AD1
Table S6: Genes downregulated in PREC/AR
Table S7: Genes upregulated in R1AD1
Table S8: Gene signatures ENTREZ ID