Defining oncogenic determinants in prostate cancer initiation and progression

Huiyun Sun

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

John K. Lee, Chair

David Macpherson

Li Xin

Program Authorized to Offer Degree:

Molecular Engineering and Sciences

University of Washington

Abstract

Defining oncogenic determinants in prostate cancer initiation and progression

Huiyun Sun

Chair of the Supervisory Committee:

John K. Lee

School of Medicine

Prostate tumors harbor substantial biological heterogeneity with highly various treatment response. In recent years, the next-generation sequencing technology has enabled comprehensive genomic analysis of primary and metastatic prostate cancer and uncovered previously uncharacterized genetic alterations in cancer genome. Functionally association the genetic abnormalities with the prostate cancer initiation and progression is necessary for a better understanding of the prostate cancer biology. In this work, we describe a combinatorial genetic strategy applied to an organoid transformation assay to rapidly generate diverse, clinically relevant prostate cancer models. Importantly, coupling with single-cell or spatially resolved next-generation sequencing we are able to resolve the clonal architecture of the resultant tumors to

uncover polygenic drivers of cancer phenotypes. We successfully recapitulated prostate tumor initiation with a focal giant cell histology and identified potential genetic alteration KMT2C associated with this phenotype. Lineage plasticity is recognized as a common mechanism for treatment resistance in cancer. The shift from an androgen receptor (AR)-positive lineage to an AR-negative lineage has increased over the past decade due to the use of 2nd generation AR inhibitors. However, the genetic determinants driving the progression of prostate cancer to an AR-null state and the acquisition of neuroendocrine differentiation remain largely unknown. We have delineated critical roles of the pioneer factors ASCL1 and NEUROD1 in neuroendocrine transdifferentitation and uncovered their abilities to silence AR expression and signaling by remodeling chromatin at the somatically acquired AR enhancer and global AR binding sites with enhancer activity. We have demonstrated a tuft cell lineage driven transcription factor POU2F3 is associated with downregulation of AR signaling pathways under low androgen conditions. In summary, our work has contributed to a deeper understanding of the genetic determinants involved in the initiation and progression of prostate cancer. By employing innovative genetic strategies and leveraging next-generation sequencing technologies, we have made significant strides in unraveling the complexities of prostate cancer biology.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am genuinely thankful to be a member in the Lee Lab. I vividly recall the challenging time I was not yet settled down in a lab and desperately needed to find a lab to do rotation at the last quarter of my first year due to the scarcity of research funding that year. It was John who took me in and I am grateful that I can join the lab after completing my rotation.

I would like to thank my mentor Dr. John K. Lee, for all his support and mentorship for the past five years. John is one of the smartest people I have ever met. His guidance has been invaluable in helping me think critically, explore unconventional approaches, and ultimately find solutions to research problems. I am truly appreciative of his constant challenge to broaden my perspective and see things comprehensively. I am sincerely grateful for his dedication and commitment to my academic and professional growth.

I would also like to express my gratitude to the members of my Supervisory Committee: Dr. David MacPherson, Dr. Li Xin, and Dr. Georg Seelig. Their expertise, advice, and support have been instrumental throughout my scientific training.

I would like to express my gratitude to all the members in the lab who have played a significant role in my scientific journey. Their willingness to share their knowledge, teach me various techniques, and provide valuable advice and perspectives on my projects has been crucial in my

growth as a researcher. Their support and collaboration have created a stimulating and enriching environment that has enhanced my learning experience. Special thanks to Shan and Ailin for being my “teacher”, and their support and companionship to help me go through all the frustrations in the lab.

Words cannot adequately express the depth of my gratitude for my family. Thanks to my parents, grandparents, aunt and uncle, for their constant encouragement, understanding and sacrifices, which have made it possible for me to pursue my passion and achieve the milestone of completing my PhD.

Thanks to my amazing friends here, Rene, Gloria, Hao, Jiajie and Wenwen and many others for being my academic buddies, party buddies, game buddies and Kpop buddies. I am grateful for their presence in my life.

Thanks to my significant other, Shin-Ya (Emerson), for taking good care of me these years. He is always there for me, patiently supporting me through all the ups and downs. His belief in me and constant encouragement have given me the strength and confidence to pursue my dreams.

Lastly, I would like to thank some of my emotional support: my cats, Douhua and Doubanjiang, my idol, Ohno Satoshi, for being my motivation, and inspiration.

DEDICATION

This dissertation is dedicated to my family: my parents, Yuanwei Sun and Shuanghong Guo, my grandparents, Jingshun Guo and Qifen Long

Chapter 1. INTRODUCTION

Prostate cancer (PCa) is the second most common cancer in men, with 288,300 new cases and 34,700 deaths estimated in the United States in 2023 [1]. Because of early detection of localized prostate tumors by rapid uptake of prostate-specific antigen (PSA) testing, the incidence rate for prostate cancer dropped significantly in the early 2010s. Most men diagnosed with early and local prostate cancer can be cured with surgery and/or radiation therapy. However, in many cases, patients present with a more aggressive and metastatic disease at the time of diagnosis or may develop metastatic recurrence years after their initial diagnosis despite the local therapy. The pioneering treatment for advanced PC has been to suppress the androgen receptor (AR) signaling axis by androgen deprivation therapy (ADT) by reducing circulating androgen via chemical or surgical castration [2]. The persistent recurrence of cancer after ADT therapy results in the development of castration-resistant prostate cancer (CRPC), among which metastatic castration-resistant prostate cancer (mCRPC) is the lethal form of prostate cancer, which is mostly found to remain reliance on AR activity [3]. These findings provided the rationale for a drug discovery of next-generation AR pathway inhibitors that antagonize steroidogenesis (abiraterone) or effectively block AR signaling (enzalutamide) [4]. These agents have demonstrated benefit in mCRPC and are now used in combination with ADT in first-line treatment of advanced PC with significant extension of survival [5, 6]. Nevertheless, prostate cancer at this advanced stage, will develop resistance to these treatments and is not considered curable.

In order to develop new therapeutics for lethal CRPC, a more comprehensive understanding of the genetic components involved in the initiation and progression of prostate cancer. The advent of

high-throughput next-generation sequencing technology has facilitated a more systematic exploration of the genomic landscape of primary and metastatic prostate cancer, thereby uncovering both previously known and unknown genetic correlates of prostate cancer [7-9]. The ongoing challenge is to functionally characterize and validate these large number of genetic abnormalities, as their individual and contextual contributions to PCa initiation and progression remain largely unknown. Understanding the biology of prostate cancer by recapitulate prostate cancer genesis and progression with relevant model systems will advance precision oncology and the innovation of novel therapeutics.

1.1 GENETIC DETERMINANTS OF PROSTATE CANCER

With assessment of genomic landscape of primary and metastatic prostate tumor samples, people have identified recurrent somatic mutations, copy number alterations, and oncogenic structural DNA rearrangements involved in multiple pathways, including androgen signaling, phosphoinositide 3-kinase (PI3K)-AKT signaling, MAPK signaling, WNT signaling, DNA repair, and cell cycle.

In 2005, chromosomal rearrangements were identified in prostate cancer that the androgen-responsive promoter elements of TMPRSS2 was fused with two family members of erythroblast transformation-specific (ETS) transcription factors, ERG and ETV1 [10], leading to over-expression of ERG and ETV1. Subsequent studies confirmed ETS gene fusions in approximately 50% of prostate cancers from prostate-specific antigen (PSA) screened surgical cohorts [11]. As the essential survival and proliferation axis of prostate, androgen signaling pathway was demonstrated to be altered in both primary and advanced prostate cancer [12]. Deletion of the

androgen-regulated homeobox gene NKX3.1 plays an important role in prostate cancer initiation and progression [13, 14]. Function as a tumor suppressor by targeting AR for degradation, mutations of SPOP was detected in over 13% of prostate cancer [15, 16]. Research has identified recurring mutations in the AR collaborating factor FOXA1, which are present in both untreated localized prostate cancer and CRPC and demonstrated that mutated FOXA1 inhibits androgen signaling, resulting in increased tumor growth [15, 17]. AR, on the other hand, is the most frequently aberrant gene in androgen independent CRPC (62.7%) [7], including point mutations in AR [17], AR amplification [18], AR variants emergence [19].

In many cases, mutations in the kinase itself, or an upstream regulator, render these kinases constitutively active, driving tumorigenesis. The PI3K pathways play an essential role in the tumorigenesis across many cancer types [20]. In prostate cancer, PI3K-signaling pathway is found to be altered in over 40% of primary and 70% of metastatic tumors [7, 8]. Loss of PTEN by homologous loss and truncating mutations is identified as a main mechanism of PI3K pathway activation [21], along with activating mutations of subunits of PI3Ks, such PIK3CA [22] and PIK3R2 [23]. Sometimes in concert with PI3K pathway activation, the MAPK pathway is also dysregulated in prostate cancer, from activating mutations of MAP kinases, such as KRAS, BRAF and ERK1 [24, 25]. The WNT signaling is involved in cell proliferation, differentiation, and epithelial-mesenchymal interactions [26, 27], deregulation of which has been implicated in cancer development. More than 20% of advanced prostate cancers harbor WNT-activating mutations, which are associated with accelerating antiandrogen resistance of prostate cancer [28, 29].

DNA damage response gene alterations are observed in up to 30% of mCRPCs, including mutations in homologous recombination repair (HRR) genes BRCA1/2 and CDK12, and the DNA damage checkpoint activator ATM [30, 31], which opens probability of treatment for advanced prostate cancer with PARP inhibitors. Tumor suppressors, TP53 and RB1, are observed to be inactivated in prostate cancer but are more frequent in advanced disease, resulting in dysregulation of cell cycle pathways and metastatic progression of prostate cancer [17, 32]. MYC plays a central role in the oncogenic process. It has been found that deregulation of the MYC family oncogene occurs in over 50% of human cancers, and this deregulation is often linked to a poor prognosis and reduced patient survival. MYC plays a prominent role in early-stage of prostate cancer, and confers castration-resistance to advanced aggressive prostate cancer [33, 34].

Despite the incidence of genetic alterations with high frequency in primary and metastatic prostate cancer, there is a long-tail distribution, with many genes mutated in less than 3% of cases and significant heterogeneity in prostate tumors [23]. Biological systems that effectively model these newly identified genetic abnormalities in the genomic landscape studies in prostate cancer must be developed to understand their functional roles in prostate cancer initiation and progression, and to develop new therapeutic approaches.

1.2 MECHANISMS OF CASTRATION RESISTANCE IN PROSTATE CANCER

Androgen deprivation therapy has been the most widely used first-line treatment for advanced prostate cancer that recurs after surgery or radiotherapy. Although 80%-90% patients respond to androgen ablation, the response is usually short-lived and the disease recurs as castration-resistant prostate cancer [35]. People have been investigated the mechanisms responsible for the emergence

of CRPC over the past several decades. Various findings have supported the continuous activation of AR signaling in castration-resistant prostate cancer, reasoning the importance of AR signaling for prostate cancer progression [36]. As a result of the growing clinical use of second-generation AR-targeted therapies, abiraterone and enzalutamide, it is increasingly recognized that some relapse samples are present with clinically aggressive variants of prostate cancer exhibiting reduced or absent AR expression [37]. Therefore, mechanisms involved in castration resistance are generally divided into two categories, AR signaling dependence or independence.

Despite the considerable decline in circulating testosterone levels after ADT, there are remaining serum androgen, either converted from adrenal hormones or other *de novo* androgen synthesis, that can activate AR signaling [38]. Point mutations of AR are rarely found in the early-stage of prostate cancer, but become more frequent in CRPC. These point mutations, mostly occurring in the ligand binding domain of AR, can activate AR by loss of specificity for the ligand, thereby allowing binding to other hormones. T877A mutation was the first AR mutation identified in prostate cancer [39]. Other mutations, including V715M and H874Y, have been identified to enhance the transcriptional sensitivity of AR to nonandrogenic steroids [11, 40]. AR gene amplification has been reported as a common mechanism of castration resistance, by either elevation of AR copy number or dysregulation of AR co-regulators and collaborating factors [8, 41, 42]. Androgen receptor splice variants (ARVs) with varying lengths of AR ligand binding domain have been observed greatly increased in CRPC samples. Although the functional role of ARVs in resistance remains controversial, their capability to induce ligand-independent activation of AR signaling, suggests resistance to castration [43]. Recent studies have identified tandem

duplications on AR enhancer regions specifically enriched in castration-resistant prostate cancer, suggesting role in promoting castration resistance and prostate cancer progression [43, 44].

Independent of AR, other steroid hormone nuclear receptor family members, which share homology in ligand binding domain with AR, can function to regulate AR target genes in prostate cancer [45]. The induction of glucocorticoid receptor (GR) expression was identified as a common feature to enzalutamide resistance [46]. There are other oncogenic pathways associated with AR independence, such as elevated tyrosine kinase pathways activity such as FGFR and MAPK in mCRPC [37]. Recent new therapies targeting AR signaling pathway have increased prevalence of CRPC with devoid of AR expression [37]. Some of the tumors have histological features of small-cell carcinoma (SCC), which can be classified as a rare AR-negative neuroendocrine carcinoma. The transition of lineage to AR-null prostate cancers, expressing markers of neuroendocrine differentiation or not, represents another possible mechanism of escaping androgen ablation [42]. Further studies to define underlying mechanisms for prostate cancer lineage plasticity and pathogenesis of these AR-null CRPC are crucial for improvements in the clinical management of this lethal phase of prostate cancer.

1.3 PATHOGENESIS OF NEUROENDOCRINE PROSTATE CANCER

De novo neuroendocrine prostate cancer (NEPC) accounts for an extremely rare proportion (less than 2%) in prostate cancer patients [46]. Incidence of NEPC increased significantly (10-17%) in prostate cancer patients that initially diagnosed with adenocarcinoma and recurs after treatment of ADT or AR pathway inhibitors (ARPIs), which is described as treatment-induced NEPC [47]. In spite of having similar genetic and transcriptional background, it is unclear whether *de novo* NEPC

and treatment-emergent NEPC originate from same progenitor cells and represent the same disease [48]. Given its rarity and the limited number of cases available for study [49], current treatment for NEPC is mostly platinum-based chemotherapy based on chemotherapeutic experience with these agents in small cell lung cancer (SCLC), because of similar histology and disease progression of NEPC, especially the small-cell NEPC and SCLC [50-52]. Although NEPC patients generally respond to the initial treatment of chemotherapy, the duration of response is short, and the overall survival is approximately one year due to acquired drug resistance [53]. There is an urgent need for the development of novel and highly effective therapeutic strategies for NEPC.

Clinically, NEPC consists of a wide range of histology, including prostate adenocarcinoma with NE differentiation, prostate adenocarcinoma with Paneth-cell NE differentiation, carcinoid tumor, small-cell carcinoma, large-cell neuroendocrine carcinoma and mixed neuroendocrine carcinoma and acinar adenocarcinoma [52]. Immunohistochemistry can also be performed to characterize NEPC by the positive expression of NE differentiation markers such as synaptophysin (SYP), neural cell adhesion molecule 1 (CD56), chromogranins (CHGA and CHGB) and neuron specific enolase (NSE), and the absence of AR and PSA secretion [54]. It is worth noticing that recent studies shown co-expression of AR signaling and neuroendocrine markers in aggressive prostate tumors, lacking NE morphology features, which is defined as amphicrine prostate carcinoma (AMPC) [55, 56], highlighting the need for IHC evaluation for AR and neuroendocrine markers to distinguish AMPC from NEPC.

The origins of NEPC remain uncertain. One hypothesis is that NEPC originates from pre-existing neuroendocrine cells in benign prostate epithelium, which can secrete neuro-proteins and growth

factors to regulate growth and homeostasis of the surrounding epithelial cells [57]. There are similarities shared between NE and NEPC cell, such as expression of NE markers. A genetically engineered mouse model CR2-TAg where simian virus 40 T antigen (SV40 TAg) was targeted expressed in the NE cells of prostate lobes in mice, demonstrated that NEPC can be originated from transformation of normal prostatic NE cells [58]. Although evidence suggests that NEPC could arise from normal NE prostate cells, more studies are in favor of the transdifferentiation of NEPC directly from adenocarcinoma cells.

Linear plasticity is considered to drive the process of the adenocarcinoma to lose AR signaling dependence and acquire neuroendocrine lineage program during long-term of ADT [59]. Studies showed that TMPRSS2-ERG occurs in NEPC at around 45%, a frequency similar to that observed in prostate adenocarcinoma [60, 61]. Hansel et al identified shared TP53 mutation between adenocarcinoma and adjacent small cell carcinoma in the adenocarcinoma and NEPC mixed tumors [62]. A temporal study revealed that divergent clonal evolution of NEPC from adenocarcinoma by acquisition of new genomic and epigenomic drivers to adapt the ARPI treatment [63]. A single-cell RNA sequencing study of clinical CRPC samples showed that small-cell NEPC expressed high luminal phenotype indicating neuroendocrine transdifferentiation from luminal-like cells [64]. There are also preclinical models supporting transdifferentiation hypothesis. *In vitro* studies demonstrated that neuroendocrine transdifferentiation of the adenocarcinoma cell line LNCaP can be induced by cyclic AMP treatment [65] or androgen deprivation [66]. Lineage tracing of the *in vivo* mouse model of the double knock out of Tp53 and Pten revealed NEPC cells originated from the luminal adenocarcinoma cells [67]. The xenograft model LTL331 and its corresponding castrate resistant LTL331R tumors, which showed NE

markers expression and AR decrease, share similar genomic profiles, suggesting neuroendocrine lineage transition in response of ADT in LTL331 model [68]. Based on these studies, it is widely believed that NEPC originated from lineage reprogramming of prostate adenocarcinoma. However, with the possibility of clonal evolution of NE cells, the detailed role of each hypothesis in NEPC progression remains unclear.

The molecular characterization of NEPC comparing adenocarcinoma identified multiple genetic factors that might be involved in the progression of this disease. Zhou et al. found loss of RB1 and TP53 drives aggressive metastatic prostate cancer progression with neuroendocrine differentiation [69]. Beltran et al. identified that MYCN and AURKA amplification were detected in approximately 40% of NEPCs, but only up to 5% in prostate adenocarcinomas [54, 63]. Lee et al. demonstrated MYCN overexpression with AKT1 activation drives the transformation of prostate basal epithelial cells to adenocarcinoma and NEPC [70]. The combination of loss of TP53 and RB1, together with overexpression BCL2, MYC and AKT1 activation, are capable of reprogramming normal human prostate epithelial tissues to a lethal small cell prostate cancer lineage [71]. Studies have shown epigenetic regulation is also crucial for driving NEPC progression. Upregulation of multiple members of the polycomb group (PcG) family, such as EZH2, are preferentially enriched in NEPC [72, 73]. Critical dysregulation of lineage determinant transcription factors (such as BRN2 [75], SOX2 [76], ONECUT2 [77], ASCL1) also drives plasticity and maintains the NE state.

Well-defined experimental models that reproducibly capture the lineage transition from prostate adenocarcinomas to NEPC are lacking. One human model of spontaneous NE transdifferentiation

of PC is the LTL331 patient-derived xenograft (PDX) tumor that starts as ARPC, partially regresses upon castration of the host, and relapses after 6-8 months (as LTL331R) as NEPC [68]. The other model is in which LNCaP cells were induced to undergo NE transifferentiation by ADT treatment [66]. However due to the potential slow process, it is difficult to interrogate the molecular mechanisms and temporal trajectories underlying NE transdifferentiation. A model that can recapitulate the abrupt transition of ARPC to NEPC is critical to understand the pathogenesis of treatment-induced NEPC and how NEPC can be prevented and managed by novel targeted therapies.

1.4 THE MOLECULAR LANDSCAPE OF BLADDER CANCER

Bladder cancer is one of the most frequently diagnosed cancers in western countries. Bladder cancers are mostly classified as either non-muscle-invasive bladder cancer (NMIBC) or muscle-invasive bladder cancer (MIBC). Up to 70% of the patients present NMIBC at the time of diagnosis, with nearly 90% of five-year survival [78]. MIBCs, which emerge *de novo* at an approximate frequency of 25%, commonly exhibit unfavorable five-year survival rates and have a tendency to rapidly advance into metastatic stages [79]. TERT promoter were found to be mutated in 70-80% of patients with urothelial bladder cancer, in contrast to the low incidence observed in normal urothelium [80]. FGFR3 mutations have been identified as early events in urothelial malignancy [81]. Alterations of FGFR3 were usually associated with a luminal papillary subtype [82]. Homozygous deletion of CDKN2A, resulting from deletion of chromosome 9, were found to be associated with initiation of bladder cancer [83]. There are evidence showing PIK3CA mutations present in low grade bladder cancer [84]. Mutations in DNA repair and DNA damage

response genes, including ERCC2, ATM, and FANCA, can contribute to extensive chromosomal alterations observed in MIBCs [85]. MIBC has defects in cell cycle checkpoint, such as inactivation of TP53 and RB1 [86], whose cooccurrence demonstrated a NE-like phenotype [87]. Bladder cancer exhibit significant epigenetic dysregulation, with hypermethylation in MIBCs [88]. Many genes involved in chromatin modification are shown to be mutated in MIBCs, such as KDM6A, KMT2A, KMT2C, KMT2D, ARID1A, EP300 and CREBBP [89]. The study to associate molecular and histopathologic in bladder cancer might provide an improvement for bladder cancer classification and produce new insights into the therapeutic development of this tumor type.

Chapter 2. COMBINATORIAL GENETIC STRATEGY ACCELERATES THE DISCOVERY OF CANCER GENOTYPE-PHENOTYPE ASSOCIATIONS

This work has been submitted for publication and is available on bioRxiv.

2.1 MAIN

Most cancers are not driven by a single oncogenic driver but are rather the sum of multiple genetic perturbations that occur during tumor evolution. However, the functional impact of most genomic abnormalities found in cancers remains largely unknown. Genetically-engineered mouse models are a standard approach to functionally define genetic drivers in cancer but they are costly, slow, and do not allow facile manipulation of more than two or three genes. To address this issue, we developed a methodology incorporating barcoded lentiviral (LV) libraries encoding cancer-associated genetic events into primary epithelial cells at a high multiplicity-of-infection (MOI) that are engrafted in mice for tumorigenic selection. This system is amenable to LV barcode sequencing of tumors to identify cooperative oncogenic drivers of malignant transformation and specific cancer phenotypes.

A major barrier to higher-order genetic studies in primary epithelial cells has been inefficient transgenesis using available LV transduction protocols. We hypothesized that enforced cell-virus contact in a constrained volume of gel matrix could increase LV transduction. Primary mouse bladder urothelial (mBU) and prostate epithelial (mPE) cells were isolated by fluorescence activated cell sorting (FACS) based on a $Lin^{-}(CD45^{-}CD31^{-}Ter119^{-})$ $EpCAM^{+}CD49f^{high}$

immunophenotype (Extended Data Fig. 1a), as these populations self-renew at high frequencies [90] and readily establish organoids in culture (Extended Data Fig. 1b). Cells were mixed into cold Matrigel containing concentrated LV expressing green fluorescent protein (GFP) prior to seeding and polymerization of organoid droplets [91]. mBU and mPE cells could be completely transduced with delivery of up to 10-20 copies per cell (Figs. 1a and 1b). We next characterized the distribution of unique proviral copies per cell. LV constructs were barcoded with matching 10 nucleotide sequences distal to the 5' long terminal repeat (LTR) and proximal to the 3' LTR and produced as a pool. In parallel, a custom single-cell amplicon panel was designed on the Mission Bio Tapestri platform to enable sensitive enumeration of multiple uniquely barcoded LVs per cell (Extended Data Fig. 1c) and was validated using a defined population of NIH 3T3 cells engineered with LVs to harbor up to 4 unique LV barcodes per cell (Extended Data Fig. 1d). mPE were transduced with a diverse barcoded LV pool at varying MOIs and single-cell amplicon sequencing showed relatively normal distributions of proviral copies per cell (Extended Data Fig. 1e and Fig. 1c).

To determine the utility of this approach in understanding the initiation/progression of bladder and prostate cancer, we selected commonly mutated genes from cancer genome sequencing studies [6, 89, 92] (Extended Data Fig. 2a) and cloned these as open reading frames (ORF) or short hairpin RNAs (shRNA) into barcoded LV constructs to mimic gain- or loss-of-function events (Extended Data Fig. 2b). At least three shRNA from The RNAi Consortium (TRC) targeting each gene was tested for knockdown in 3T3 cells by quantitative polymerase chain reaction (Extended Data Fig. 2c). shRNA demonstrating the most potent knockdown of target gene expression was incorporated into the LV libraries. A bladder urothelial LV pool (BU-LVp) of 33 genes and a prostate epithelial

LV pool (PE-LVp) of 24 genes were produced in arrayed format to avoid LV barcode re-combination and concentrated by ultracentrifugation (Extended Data Fig. 3a). Infectivity (representation) of each LV was evaluated by transducing either mBU or mPE cells with the respective LV pool and performing bulk amplicon sequencing of LV barcodes (Extended Data Fig. 3b). Initial LV pools demonstrated over 10-fold overrepresentation of shRNA vectors relative to ORF vectors (Extended Data Fig. 3c), presumably due to more efficient viral packaging because of the reduced length between LTRs of the transfer plasmid [93]. This data was applied to adjust the cell surface area of producer cells for subsequent arrayed LV library production, leading to near normalization of the representation of shRNA and ORF vectors (Extended Data Fig. 3d).

We adopted an approach in which primary mBU and mPE cells infected with BU-LVp or PE-LVp at high MOI in organoids were recombined with inductive mouse E16.5 embryonic bladder mesenchyme (EBLM) [94] or urogenital sinus mesenchyme (UGSM) [95] and subsequently grafted subcutaneously in NOD scid gamma (NSG) mice to enable biological selection for tumorigenic clones (Fig. 1d). A representative tumor derived from primary mBU cells transduced with BU-LVp exhibited three morphologically distinct regions consistent with papillary urothelial carcinoma with in-verted growth pattern, urothelial carcinoma with squamous differentiation, and sarcomatoid carcinoma which were also supported by Gata3, p63, and pan-cytokeratin immunostaining patterns (Fig. 1e). Single-cell DNA amplicon sequencing was performed to enumerate the LV barcodes for the determination of clonal architecture and deconvolution of LV-delivered genetic events putatively involved in tumorigenesis. Three major clones were identified (Fig. 1f) but spatial resolution was lost due to single-cell dissociation. To associate histology with clonality, we performed laser capture microdissection (LCM) of the three regions on stained tissue

sections and performed bulk DNA amplicon sequencing (Fig. 1g). The papillary urothelial carcinoma was uniquely associated with Fgfr3 S243C, shAtm, and Zfp703, in addition to the common Ywhaz, Pik3ca E545K, Pparg, and Pvrl4 observed in all three dominant clones. Cancer genomics studies have shown that activating mutations in FGFR3 are highly enriched in luminal papillary urothelial carcinomas [89, 96]. We validated these findings in the mouse urothelial transformation assay using a de-fined LV pool of Fgfr3 S243C, Ywhaz, Pik3ca E545K, Pparg, and Pvrl4 (Extended Data Fig. 4a) which generated a tumor with papillary urothelial carcinoma with inverted growth pattern (Extended Data Fig. 4b).

Several tumors in the Fred Hutch Bladder Tumor (FHBT) series have been generated using this methodology including those with pure urothelial carcinoma and others with mixtures of variant histologies (Fig. 2a). The urothelial origin of these tumors is supported by GFP staining (Fig. 2b-d), even with low/absent pan-cytokeratin staining as seen in regions of sarcomatoid carcinoma (Fig. 2b). We conducted molecular profiling of these tumors and their regional tumor histologies by LCM and RNA-seq analysis. Principal component analysis (PCA) of the gene expression data showed that squamous and sarcomatoid histologies clustered together separately from urothelial and papillary urothelial histologies (Fig. 2e). The BASE47 subtype predictor [97] generally called the tumors with papillary and papillary squamous histologies as luminal and the squamous and sarcomatoid histologies as basal, consistent with an established relationship between sarcomatoid histology and the basal subtype [98] (Fig. 2f). The Consensus Molecular Classifier [87] revealed that the non-papillary urothelial histologies showed neuroendocrine (NE)-like gene expression with low/absent luminal and basal gene signatures (Fig. 2g and Extended Data Fig. 5a). Gene set enrichment analysis (GSEA) from comparing these tumor histologies in a pairwise manner

revealed enrichment of genes associated with epithelial-to-mesenchymal transition in sarcomatoid carcinoma, as expected from prior molecular analysis of human tumors [98] (Fig. 2h). We further con-firmed the relevance of our FHBT models by projecting them onto PCA plots of human metastatic bladder cancer samples obtained at rapid autopsy [89] (Fig. 2i) and established N-butyl-N-(4-hydroxybutyl)-nitrosamine (BBN)-induced mouse bladder cancer models [99] (Extended Fig. 5b) to show that they occupy overlapping space.

mPE cells transduced with PE-LVp and engrafted in mice (Extended Data Fig. 6a) also gave rise to mixed cancer morphologies. One tumor showed high-grade prostate adenocarcinoma with interspersed pleomorphic giant cell carcinoma (Extended Data Fig. 6b), a rare variant associated with a poor prognosis [100] that may contribute to therapeutic resistance and lethality [101]. Immunostaining revealed Hoxb13 and Ar expression in both histologies with pronounced nuclear Trp53 expression in the pleomorphic giant cells (Extended Data Fig. 6b). We isolated large (pleomorphic giant cell carcinoma) and small (adenocarcinoma) cells from dissociated tumors using a flow cytometry-based strategy, propagated these cells briefly in organoid cultures, then dissociated the cells and stained with the nuclear dye Hoechst 33342 to further isolate cells based on DNA content for downstream single-cell LV barcode enumeration (Extended Data Fig 6c). This single-cell clonality analysis revealed striking enrichment of shKmt2c in the putative pleomorphic giant cell clones (Extended Data Fig. 6d). Recent studies have established that KMT2C mediates DNA damage response in cancer [102, 103] and DNA damage repair alterations are common in human prostate adenocarcinoma with pleomorphic giant cell features [104].

In sum, we describe a set of technologies that form a functional cancer genomics assay with efficient delivery of compound genetic perturbations from barcoded LV libraries and single-cell sequencing to rapidly investigate genotype-phenotype relationships in cancer initiation/progression from primary epithelial cells. These initial studies provide proof-of-principle of this approach but additional optimization is ongoing to enable the high-throughput study of tumorigenesis in the native tissue microenvironment in immune-competent hosts and more facile spatial association of genotype-phenotype.

2.2 MATERIALS AND METHODS

Lentiviral constructs and lentiviral library production

Double-barcoded lentiviral vectors were generated from FU-CGW by sequentially inserting matched 10-nucleotide barcodes into the PacI site distal to the HIV FLAP using the Quick Ligation Kit (New England Biolabs) and PCR amplification of the WPRE sequence and barcode with insertion into the ClaI sites proximal to the 3' LTR by HiFi DNA Assembly (New England Biolabs). ORFs were cloned into the EcoRI site of the double-barcoded lentiviral vectors by HiFi DNA Assembly. To generate shRNA lentiviral vectors, the Ubiquitin C promoter sequence was excised from the double-barcoded plasmid by digesting with PspXI and EcoRI. U6 promoter and shRNA cassettes were isolated by digesting pLKO.1 TRC shRNA clones with PspXI and EcoRI and were inserted into the digested double-barcoded plasmid using the Quick Ligation Kit. Individual lentiviruses were generated in arrayed format in 293T cells by co-transfection of each double-barcoded lentiviral ORF or shRNA plasmid with the helper plasmids pVSV-G, pMDL, and pRev using FuGENE HD Transfection Reagent (Promega). Lentiviral supernatants were collected 36 hours after transfection, pooled, and concentrated by ultracentrifugation in V-bottom

polypropylene centrifuge tubes on a SW 32 Ti in an Optima XE 90 (Beckman Coulter) at a speed of 22,000 rpm at 4°C for 2 hours. Supernatants were aspirated and lentiviral pellets were resuspended in residual media and cryopreserved.

shRNA screening

The top three to five shRNA sequences identified from The RNAi Consortium for each target gene were identified from the Genetic Perturbation Platform Web Portal at the Broad Institute. shRNA sequences were cloned into pLKO.1. pLKO.1-TRC control and pLKO.1-shRNA lentiviruses were generated and used to transduce 3T3 cells. 72 hours after lentiviral transduction, 3T3 cells were collected and RNA harvested using a RNeasy Mini Kit (Qiagen). Reverse transcription of RNA was performed using SuperScript IV Reverse Transcriptase (Invitrogen) per manufacturer's instructions. qPCR was performed on a QuantStudio 6 using SYBR Green qPCR Master Mix (ThermoFisher Scientific) and primers specific to each target gene and *Ubc* as a control. All primers used for these studies are listed in Extended Data Table 1. Relative expression was calculated using ddCT analysis.

Embryonic bladder mesenchyme and urogenital sinus mesenchyme preparation

Urogenital sinus mesenchyme was isolated and propagated as previously described [95]. E16 fetal bladders were also collected at the same time as the urogenital sinus and subjected to similar steps for preparation of embryonic bladder mesenchyme. Urogenital sinus mesenchyme and urogenital sinus mesenchyme were passaged less than five times prior to use in engraftment studies.

Mouse bladder and prostate dissociation and organoid culture

Bladder and prostates from eight- to twelve-week-old male C57BL/6 mice (The Jackson Laboratory) were dissected and enzymatically dissociated in Collagenase/Dispase 1 mg/ml solution in DMEM media for 1 hour at 37ºC with continuous rocking. Cells were pelleted and supernatant aspirated. Trypsin-EDTA 0.25% solution was added to the cell pellet and incubated for 5 minutes at 37ºC. Trypsin was neutralized and cells were washed in DMEM + 10% FBS before filtering through a 40 μm cell strainer. Cells were stained with antibodies for fluorescence-activated cell sorting on a Sony SH800 Cell Sorter with collection of $Lin^{-}CD49f^{high}EpCAM^{high}$ cells. 1-2 x 10^4 bladder urothelial and prostate epithelial cells were resuspended in a total of 15 μl of growth factor-reduced Matrigel (Corning) +/- concentrated lentivirus and seeded as droplets in each 48-well tissue culture plate well. Cells were cultured as previously described [91]. Mouse bladder organoid culture media consisted of Advanced DMEM-F12, 10 mM HEPES, 2 mM GlutaMAX, B27 supplement, 1.25 mM N-acetylcysteine, 50 ng/ml hEGF, 100 ng/ml hNoggin and 500 ng/ml hR-spondin, 200 nM A83-01, and 10 μM Y-27632. Mouse prostate organoid culture media consisted of mouse bladder organoid culture media with the addition of 1 nM dihydrotestosterone.

Organoid transformation assay

After 5-7 days of culture, transduced mouse bladder urothelial or prostate epithelial organoids were liberated by dissociating Matrigel matrix with Dispase 5 U/ml (STEMCELL Technologies). Organoids were washed with PBS and resuspended in ice cold Matrigel with either 10^5 EBLM or UGSM and subcutaneously injected into the flanks of six- to eight-week-old male NSG (NOD-SCID-IL2Ry-null) mice (The Jackson Laboratory). For prostate epithelial transformation studies, mice were supplemented with testosterone through the subcutaneous implantation of 90-day

release testosterone pellets (Innovative Research of America). Tumors were harvested when they reached 1 cm in maximal diameter. All animal care and studies were performed in accordance with an approved Fred Hutchinson Cancer Center Institutional Animal Care and Use Committee protocol and Comparative Medicine regulations.

Copy number assay

DNA was extracted from organoids using a GeneJET Genomic DNA Purification Kit (ThermoFisher Scientific). Copy number analysis was performed by TaqMan Real-Time PCR Assay (ThermoFisher Scientific) using the TaqMan Copy Number Reference Assay, mouse, Tfrc (4458366) and EGFP TaqMan Copy Number Assay (Mr00660654_cn) on a QuantStudio 6. Genomic DNA extracted from the tails of transgenic C57BL/6 mice with one or two copies of GFP was used as a calibrator sample. GFP copy number was determined using ddCT analysis where sample copy number = calibrator copy number x 2-ddCT.

Single-cell DNA amplicon sequencing library preparation and sequencing

A custom panel was designed for the Mission Bio Tapestri to amplify segments of ten mouse genes at two exons each, the 5' and 3' lentiviral barcodes, and lentiviral GFP. Libraries were generated either from cryopreserved or freshly dissociated tumor cells using the Mission Bio Tapestri Single-cell DNA Custom Kit according to the manufacturer's recommendations. Single cells (3,000 to 4,000 cells per μl) were resuspended in Tapestri cell buffer and encapsulated using a Tapestri microfluidics cartridge, lysed, and barcoded. Barcoded samples were subjected to targted PCR amplification and PCR products were removed from individual droplets, purified with KAPA Pure Beads (Roche Molecular Systems), and used as a template for PCR to incorporate Illumina P7

indices. PCR products were purified by KAPA Pure Beads, and quantified by Qubit dsDNA High Sensitivity Assay (ThermoFisher Scientific). Sample quality was assessed by Agilent TapeStation analysis. Libraries were pooled and sequenced on an Illumina MiSeq or HiSeq 2500 with 150 bp paired-end reads in the Fred Hutchinson Cancer Center Genomics Shared Resource.

Laser capture microdissection and DNA/RNA isolation for high-throughput sequencing

10 μm thick sections were cut from FFPE tumor tissues blocks and mounted onto PEN membrane frame slides (ThermoFisher Scientific). Sections were fixed with 95% ethanol for 1 minute, stained with 3% cresyl violet, and dehydrated through graded alcohols and xylene. Histology review and annotation was performed by a pathologist. Laser capture microdissection was performed on an Arcturus XT Laser Capture Microdissection System (ThermoFisher Scientific). Microdissected specimens were collected for DNA and RNA extraction. DNA was extracted using a GeneRead DNA FFPE Kit (Qiagen) and RNA was extracted using a RNeasy FFPE Kit (Qiagen) according to manufacturer's protocols. Two-step PCR for amplification and sequencing library adaptor ligation was performed. The first PCR reaction consisted of 2x KAPA HiFi HotStart ReadyMix, 100 nM of 1º FWD primer (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAAAATTTTCGGGTTTATTACAGG-3'), 100 nM of 1º REV primer (5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCCGCTCGAGGACTATTAAG-3'), and 80 ng of genomic DNA. Thermal cycling conditions were 95ºC for 3 minutes; (95ºC for 30 seconds, 64ºC x 30 seconds, 72ºC x 30 seconds) x 25 cycles; 72ºC for 5 minutes, and hold at 4ºC. PCR cleanup was conducted using the Wizard SV Gel and PCR Clean-Up System (Promega) with elution in 30 μl of double distilled water. The second PCR reaction consisted of 2x KAPA HiFi

HotStart ReadyMix, 140 nM of 2º i7 primer, 140 nM of 2º i5 primer, and 5 µl of elution from the PCR cleanup of the 1º PCR. Thermal cycling conditions were 95ºC for 3 minutes, (95ºC for 30 seconds, 61ºC x 30 seconds, 72ºC x 30 seconds) x 8 cycles; 72ºC for 5 minutes, and hold at 4ºC. The sequences of 2º primers used to incorporate dual-indexed Illumina sequencing adapters are displayed in Extended Data Table 2. PCR cleanup was conducted using the Wizard SV Gel and PCR Clean-Up System with elution in 30 µl of double distilled water. Sample quality was assessed by Agilent TapeStation analysis. Sequencing was performed on an Illumina MiSeq or HiSeq 2500 instrument using 150 bp single-end reads. PhiX sequences were excluded from the sequencing reads by Bowtie 2 [105]. Cutadapt [106] was used to trim the reads to the barcode region. Then the trimmed reads were aligned to custom DNA references containing all barcodes using Bowtie 2. Samtools [107] was used to extract read counts for each barcode. The RNA-seq libraries were prepared using a SMARTer Stranded Total RNA-Seq Kit v3 - Pico Input Mammalian (Takara Bio) and sequenced on an Illumina NovaSeq 6000 using a NovaSeq S4 flow cell with 100 bp paired end reads by MedGenome, Inc. Sequencing reads were mapped to mouse genome reference GRCm39 and gene expression was quantified and normalized using the UC Santa Cruz Computational Genomics Lab Toil RNA-seq pipeline (v4.1.2) [108].

Transcriptional subtype analysis and PCA projections

All computational analyses were carried out in RStudio Version 4.1.0. Mouse Ensembl genes were converted to MGI gene symbols using the biomaRt package. MGI gene symbols were then converted to their human orthologs by referencing the mouse-human ortholog database available from The Jackson Laboratory. The human ortholog matrix was used for downstream analysis in transcriptional subtype analysis. FHBT samples were classified using the BASE47 subtype

predictor gene list [97] and the ConsensusMIBC package v1.1. Z-score means of genes and signature scores were calculated for each sample. Heatmaps of both the BASE47 and ConsensusMIBC results were generated using the pheatmap package v1.0.12. For PCA analysis, the FPKM human ortholog matrix was normalized by $\log_2 +1$ transformation before performing mean-centered PCA using the prcomp package v3.6.2. Visualization of the PCA plot was performed using the factoextra package v1.0.7 and ggpubr package v0.6.0.

For PCA projections, RNA-seq count data from the FHBT, GSE220999, and UW TAN BLCA datasets were transformed to CPM and normalized to compare across each dataset using the DGEobj.utils package v1.0.6. PCA projection of the FHBT data onto the UW TAN BLCA space was done by first generating a PCA of the UW TAN BLCA samples from the common genes between the FHBT and HSE220999 data. A PCA for both the FHBT and GSE220999 samples was then scaled by the eigenvalues of the UW TAN BLCA using the base package v3.6.2 (https://www.rdocumentation.org/packages/base/versions/3.6.2/topics/scale). A plot was constructed overlaying the reference UW TAN BLCA samples with either FHBT or GSE220999 tumor projections using ggplot2 v3.4.1. UW TAN BLCA samples were colored by their Consensus Molecular Classifier subtype with 95% confidence ellipses drawn over each group.

Differential gene expression analysis was performed pairwise between FHBT histologies using the DESeq2 package v1.38.3[109]. P-values were generated by Wald test and P-adj by Benjimini-Hochberg correction. Pre-ranked Gene Set Enrichment Analysis (Broad Institute) was conducted by inputting a ranked list of differentially expressed genes based on $\log_{10}$ transformed p-values from the DESeq2 analysis for each pairwise comparison. Dot plots were generated by plotting the

Normalized Enrichment score and log transformed FDR for each pre-ranked GSEA output using ggplot2.

Single-cell lentiviral barcode enumeration and clonality analysis

Raw sequencing reads were trimmed to the amplicon regions using the awk command. Barcode sequences in the reads were filtered and extracted using UMI-tools[110]. Processed reads were aligned to custom references containing all amplicon sequences using bwa-mem[111]. Samtools was used to extract amplicon counts for each barcode. Mouse cells with no GFP amplicon counts were removed. Counts per cell were normalized to total counts for each barcode. A minimum threshold normalized count of 1% of total counts was used to define the presence of a barcode in a cell. The clonal architecture of cells was determined by enumerating all cells containing each distinct combination of barcodes.

Immunohistochemistry

Tumor samples were formalin-fixed and paraffin-embedded, sectioned to 5 μm thickness, and placed on positively charged glass slides. For each tumor, slides were stained with a standard H&E protocol. For immunohistochemical studies, tissue sections were deparaffinized in xylene and rehydrated in graded alcohols. Antigen retrieval was performed in pH 6.0 citrate buffer boiled for 30 minutes in a pressure cooker heated by microwave. Slides were washed with PBS, blocked with 2.5% horse serum for 30 minutes at room temperature, and then incubated with primary antibodies overnight at 4°C. Vector Laboratories ImmPRESS-HRP anti-rabbit or anti-mouse IgG (peroxidase) polymer detection kits were used to detect proteins. Liquid DAB+ Kit (Agilent

Technologies) was used for chromogenic visualization of immunostaining. All sections were counterstained with hematoxylin, dehydrated, and mounted with a glass coverslip.

Antibodies

Antibodies used for FACS: Human/mouse/bovine integrin alpha 6/CD49f PE-conjugated antibody (FAB13501P, R&D Systems); PE/Cyanine 7 anti-mouse CD325 (Ep-CAM) antibody (118216, BioLegend); CD31 (PECAM-1) monoclonal antibody (390), FITC (11-0311-82, eBioscience); CD45 monoclonal anitbody (30-F11), FITC (11-0451-85, eBioscience); TER-119 monoclonal antibody (TER-119), FITC (11-5921-82, eBioscience). Antibodies used for immunohistochemistry: Rabbit polyclonal panCK (ab9377, Abcam, 1:100); rabbit monoclonal GFP antibody (clone D5.1, Cell Signaling, 1:100); rabbit polyclonal p63 antibody (12143-1-AP, Proteintech, 1:200); mouse monoclonal p53 antibody (clone 1C12, Cell Signaling, 1:500); rabbit monoclonal HOXB13 antibody (clone D7N8O, Cell Signaling, 1:50); rabbit polyclonal AR antibody (06-680, Millipore, 1:2,000); rabbit monoclonal GATA3 antibody (clone D13C9, Cell Signaling, 1:200); rabbit monoclonal CD44 antibody (clone E7K2Y, Cell Signaling, 1:100).

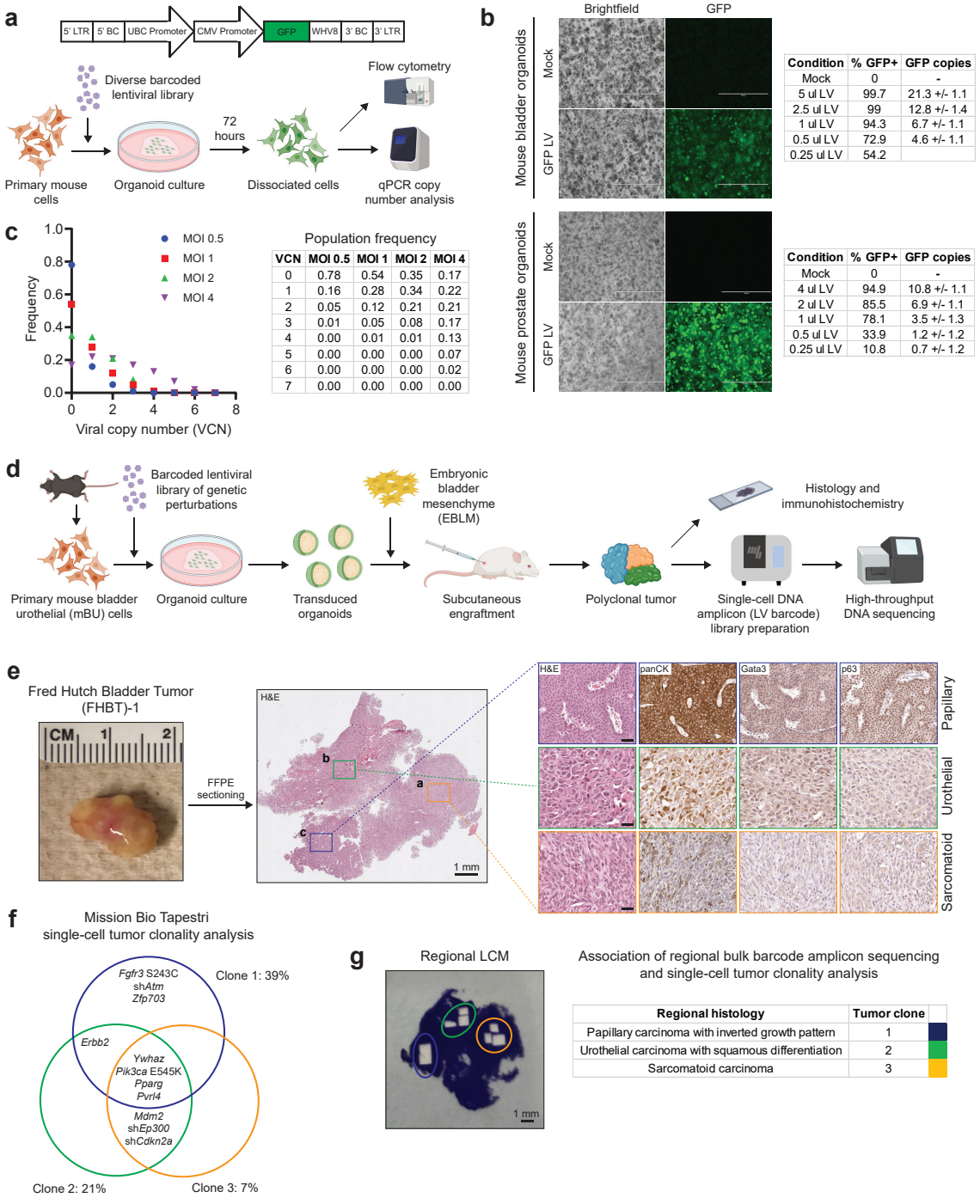
Statistical analyses

Data analysis was performed on GraphPad Prism 9 (GraphPad Software, Inc.). qPCR results were analyzed in Excel. Statistical significance was determined using the unpaired two-tailed Student t test. Results are depicted as mean + SD unless stated otherwise. For all statistical tests, p-values of <0.05 were considered significant.

.

2.3 FIGURES

Fig. 1



Condition	% GFP+	GFP copies
Mock	0	-
5 ul LV	99.7	21.3 +/- 1.1
2.5 ul LV	99	12.8 +/- 1.4
1 ul LV	94.3	6.7 +/- 1.1
0.5 ul LV	72.9	4.6 +/- 1.1
0.25 ul LV	54.2	

Condition	% GFP+	GFP copies
Mock	0	-
4 ul LV	94.9	10.8 +/- 1.1
2 ul LV	85.5	6.9 +/- 1.1
1 ul LV	78.1	3.5 +/- 1.3
0.5 ul LV	33.9	1.2 +/- 1.2
0.25 ul LV	10.8	0.7 +/- 1.2

Population frequency

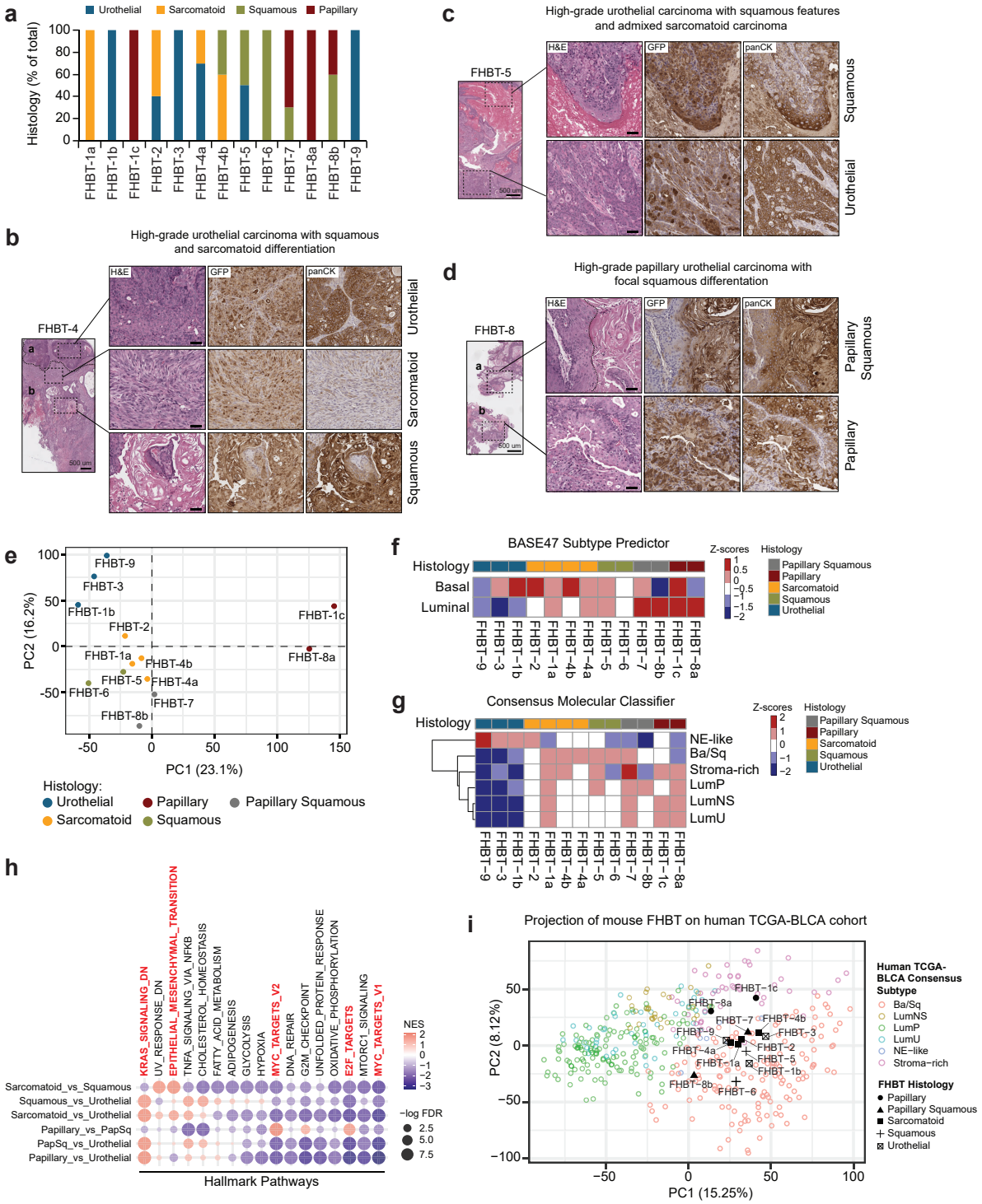
VCN	MOI 0.5	MOI 1	MOI 2	MOI 4
0	0.78	0.54	0.35	0.17
1	0.16	0.28	0.34	0.22
2	0.05	0.12	0.21	0.21
3	0.01	0.05	0.08	0.17
4	0.00	0.01	0.01	0.13
5	0.00	0.00	0.00	0.07
6	0.00	0.00	0.00	0.02
7	0.00	0.00	0.00	0.00

Regional histology	Tumor clone	
Papillary carcinoma with inverted growth pattern	1	
Urothelial carcinoma with squamous differentiation	2	
Sarcomatoid carcinoma	3	

2.3.1. Fig. 1. **Efficient lentiviral transduction of primary epithelial cells at high multiplicity-of-infection and transformation of urothelial cells to tumors with mixed cancer histologies.**

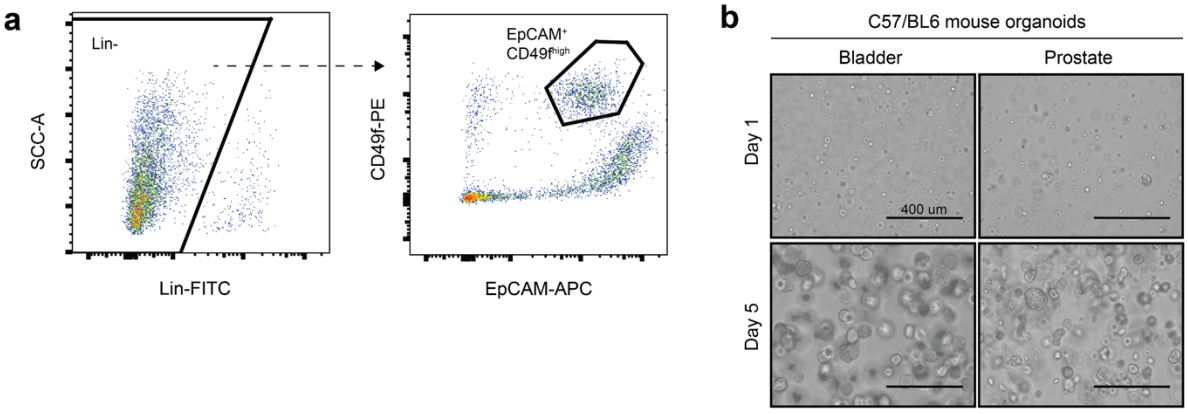
(**a**) *Top*, Schematic of a LV construct with matching barcodes (BC) at the 5' and 3' ends. *Bottom*, Overview of experiments with LV infection of primary mouse cells in organoid culture and quantification of transduction. (**b**) *Left*, Brightfield and GFP images of mouse bladder or prostate organoids 72 hours after mock or GFP LV transduction. Scale bar=400 μm. *Right*, Tables summarizing quantification of LV transduction by flow cytometry and LV copies of GFP (+/- SD) by qPCR. (**c**) *Left*, Plot of the distribution of LV copies per mPE cell at different MOIs 72 hours after transduction. *Right*, Table summarizing VCN population frequencies at varying MOIs. (**d**) Scheme of the mBU organoid transformation assay to uncover functional genotype-phenotype associations in bladder cancer. (**e**) *Left*, Gross image of a tumor arising from mBU transformed with BU-LVp. *Middle*, low-magnification image of the H&E-stained tumor section. *Right*, high-magnification images of H&E- and IHC-stained sections of regions with distinct histologies. Scale bars=50 μm. (**f**) Clonal architecture of the three dominant clones in the tumor as determined by Mission Bio Tapestri single-cell analysis of LV barcodes. (**g**) *Left*, Tumor tissue section after LCM of the histologically distinct regions. *Right*, Table showing the associations between regional tumor histologies and clones in **f** based on LCM and bulk DNA amplicon sequencing of LV barcodes.

Fig. 2



2.3.2. Fig. 2. **Rapid generation of a series of clinically relevant and phenotypically diverse bladder cancer models.**

(**a**) Bar graph showing the representation of cancer histologies present across a series of tumors (FHBT) generated using mBU transformed with BU-LVp. (**b-d**) Low- and high-magnification images of H&E-stained sections and high-magnification images of IHC-stained sections for GFP and pan-cytokeratin (panCK) expression depicting mixed histologies present within the same tumor. Scale bars=50 μm. (**e**) PCA plot showing FHBT series color-coded based on histology. Heatmaps showing the histologies of the FHBT series relative to (**f**) basal and luminal signature scores for the BASE47 subtype predictor and (**g**) signature scores for the Consensus Molecular Classifier. (**h**) Pre-ranked GSEA dotplot of Hallmark Pathways based on differentially expressed genes (FDR <0.001) in pairwise histology comparisons. (**i**) PCA projection plot of FHBT samples over the TCGA-BLCA samples color-coded by Consensus Molecular Classification.

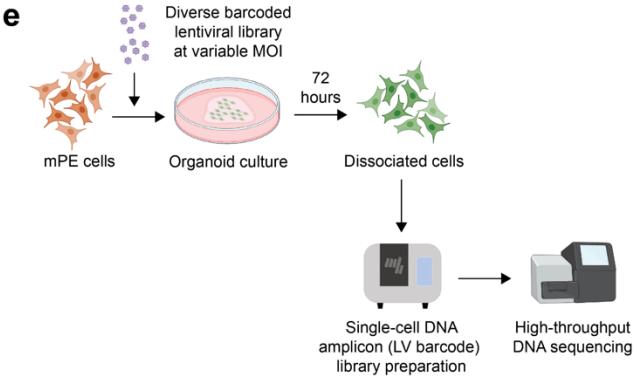


c Mission Bio Tapestri custom single-cell DNA amplicon sequencing panel design

Amplicons	Source	Chr	Start	End	Bases
5' BC	LV				10
3' BC	LV				10
GFP	LV				100
Trfc exon 5	Ms	chr16	32616748	32616899	148
Trfc exon 7	Ms	chr16	32618215	32618333	115
Actb exon 2	Ms	chr5	142905565	142905684	120
Actb exon 5	Ms	chr5	142904071	142904250	176
Gapdh exon 4	Ms	chr6	125162850	125163024	175
Gapdh exon 5	Ms	chr6	125162506	125162669	147
Hprt1 exon 3	Ms	chrX	53002100	53002255	156
Hprt1 exon 4	Ms	chrX	53008657	53008716	60
Pgk1 exon 3	Ms	chrX	106194360	106194504	145
Pgk1 exon 5	Ms	chrX	106196843	106196944	102
Sdha exon 9	Ms	chr13	74332115	74332261	147
Sdha exon 12	Ms	chr13	74327249	74327349	101
Ppia exon 3	Ms	chr11	6418243	6418284	42
Ppia exon 4	Ms	chr11	6419159	6419202	44
Rpl37 exon 2	Ms	chr15	5117297	5117410	114
Rpl37 exon 3	Ms	chr15	5117626	5117696	71
Eif3f exon 3	Ms	chr7	108938038	108938110	73
Eif3f exon 4	Ms	chr7	108938409	108938531	123
Eef2 exon 2	Ms	chr10	81177736	81177866	127
Eef2 exon 7	Ms	chr10	81179606	81179784	175

d

3T3 Cell Population	Expected Frequency (%)	Tapestri Clonal Frequency (%)
Unlabeled	28.13	28.59
BC01+BC02+BC04+BC06	12.46	9.98
BC02+BC04+BC06	1.41	6.12
BC06+BC08	56.26	54.8
BC05+BC07	0.14	0.51



2.3.3. Extended Data Fig. 1. **Isolation of mouse bladder urothelial and prostate epithelial cells for organoid culture and design/validation of a custom Mission Bio Tapestri single-cell DNA amplicon sequencing panel.**

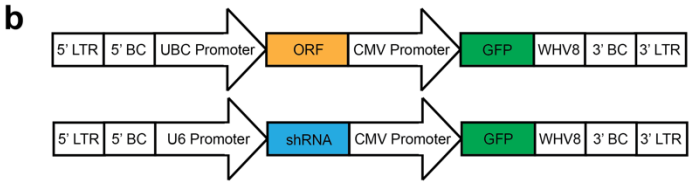
(**a**) Representative flow cytometry plot for the isolation of mouse bladder urothelial and prostate epithelial from dissociated tissues based on a $Lin^-(CD45^-CD31^-Ter119^-)$ $EpCAM^+CD49f^{high}$ immunophenotype. (**b**) Images of organoid cultures of mouse bladder urothelial and prostate epithelial cells on day 1 and day 5 after seeding. (**c**) Table showing the amplicons represented in a custom Mission Bio Tapestri single-cell DNA amplicon sequencing panel. (**d**) Table showing results of a validation study where a defined mixture of 3T3 cells with an unlabeled population and others labeled with combinations of lentiviruses encoding distinct barcodes were analyzed using the Mission Bio Tapestri single-cell DNA amplicon sequencing panel to determine clonality. ~2,000 cells were analyzed. (**e**) Overview of experiments with infection of mouse prostate epithelial (mPE) cells with a diverse barcoded lentiviral library in organoid culture across a range of multiplicity-of-infection (MOI) and quantification of viral copy number per cell across the population by single-cell amplicon sequencing.

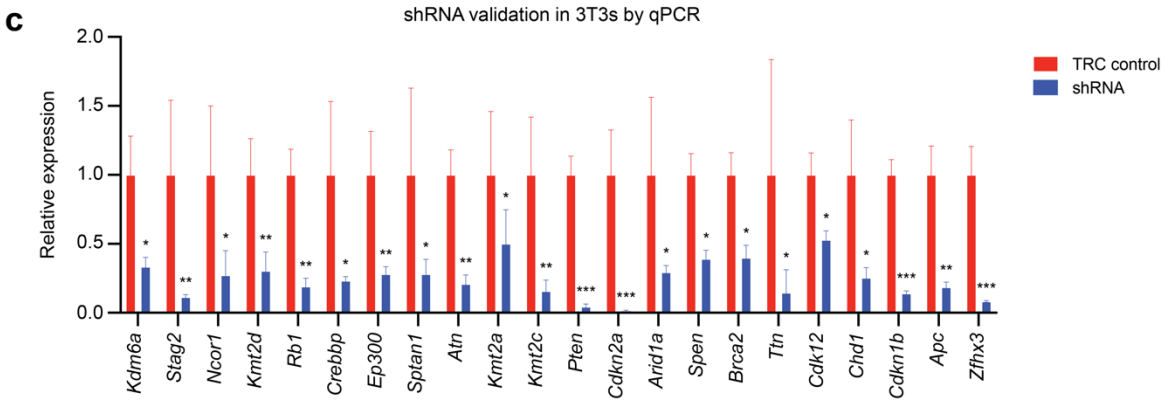
a Genetic alterations associated with bladder cancer

Gain-of-function	Loss-of-function
E2f3	Kmt2d
Pparg	Kdm6a
Mdm2	Arid1a
Ccnd1	Kmt2c
Egfr	Ep300
Pvrl4	Stag2
Ywhaz	Atm
Yap1	Fat1
Ccne1	Sptan1
Myc	Kmt2a
Zfp703	Cdkn2a
Trp53 R245Q	Rb1
Pik3ca E545K	Crebbp
Fgfr3 S243C	Ncor1
Erbb3 V104L	Pten
Erbb2 S311Y	

Genetic alterations associated with prostate cancer

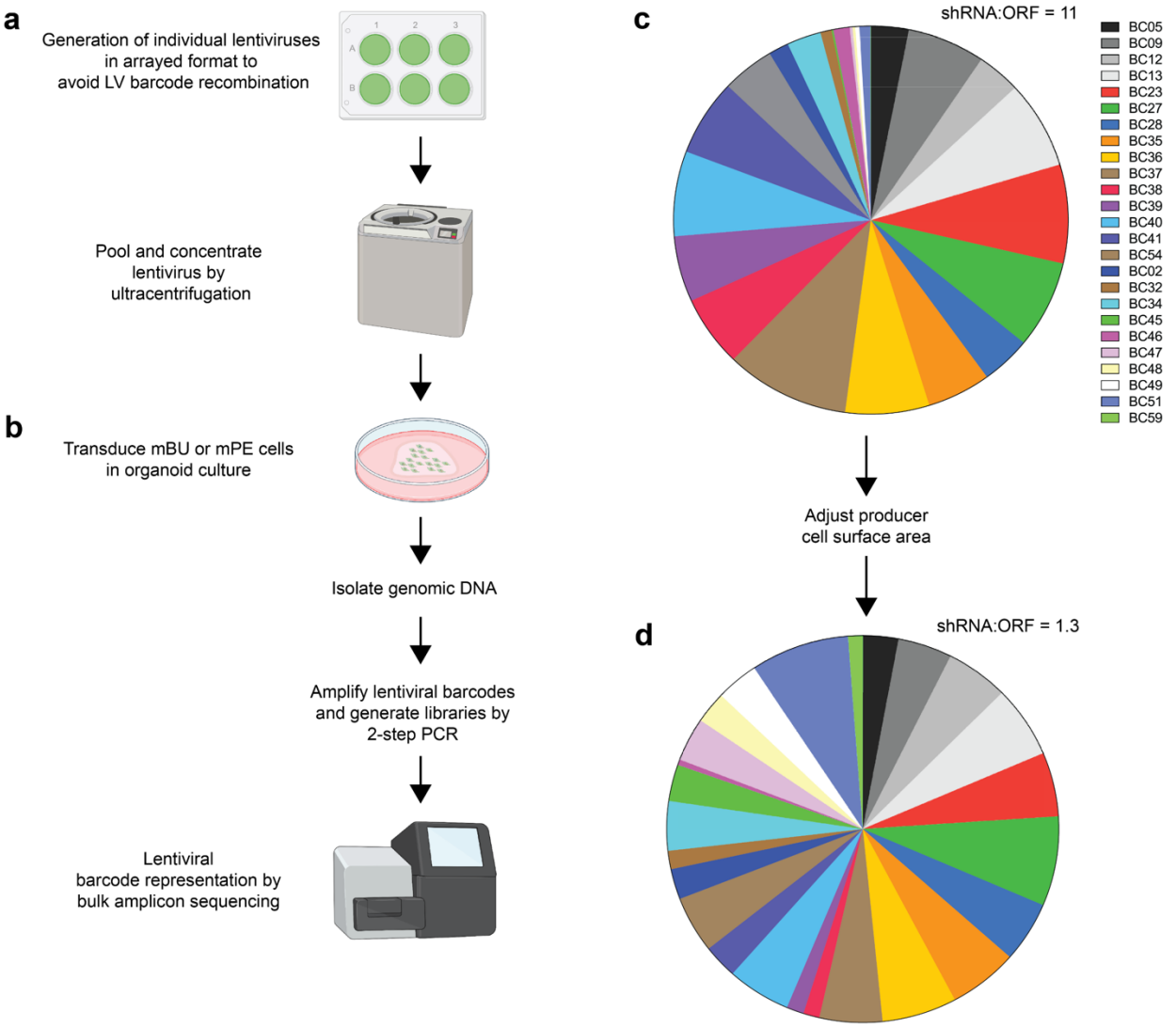
Gain-of-function	Loss-of-function
Ar	Pten
Erg	Ncor1
Etv1	Spen
Pik3ca E545K	Apc
Trp53 R245Q	Brca2
Spop F133C	Atm
Foxa1 R261C	Cdk12
Ctnnb1 D32A	Cdkn1b
Myc	Kmt2c
	Kmt2d
	Kdm6a
	Chd1
	Ttn
	Zfhx3





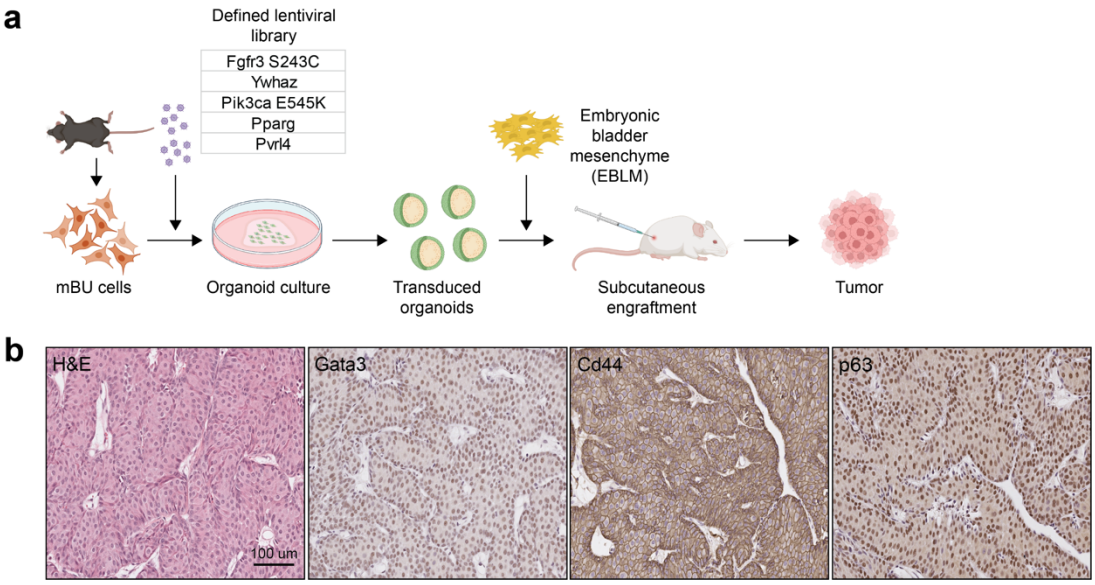
2.3.4. Extended Data Fig. 2. **Extended Data Fig. 2. Recurrent genetic alterations associated with bladder and prostate cancer encoded in barcoded lentiviral libraries.**

(**a**) Tables showing gain-of-function and loss-of-function genetic alterations associated with bladder and prostate cancer selected for representation in cancer-specific barcoded lentiviral libraries. (**b**) Schematics of barcoded lentiviral vectors expressing open reading frames (ORF) or short-hairpin RNA (shRNA). LTR=long terminal repeat; BC=barcode; UBC=Ubiquitin C; CMV=cytomegalovirus; GFP=green fluorescent protein; WHV8=Woodchuck hepatitis virus 8 post-transcriptional regulatory element. (**c**) Plot showing relative expression of target genes as determined by quantitative polymerase chain reaction (qPCR) in 3T3 cells 72 hours after lentiviral transduction with pLKO.1-TRC control or pLKO.1 expressing select shRNA previously screened and selected for inclusion in the barcoded lentiviral libraries based on the extent of gene knockdown. qPCR reactions were performed in quadruplicate. * $p<0.05$; ** $p<0.01$, *** $p<0.001$.



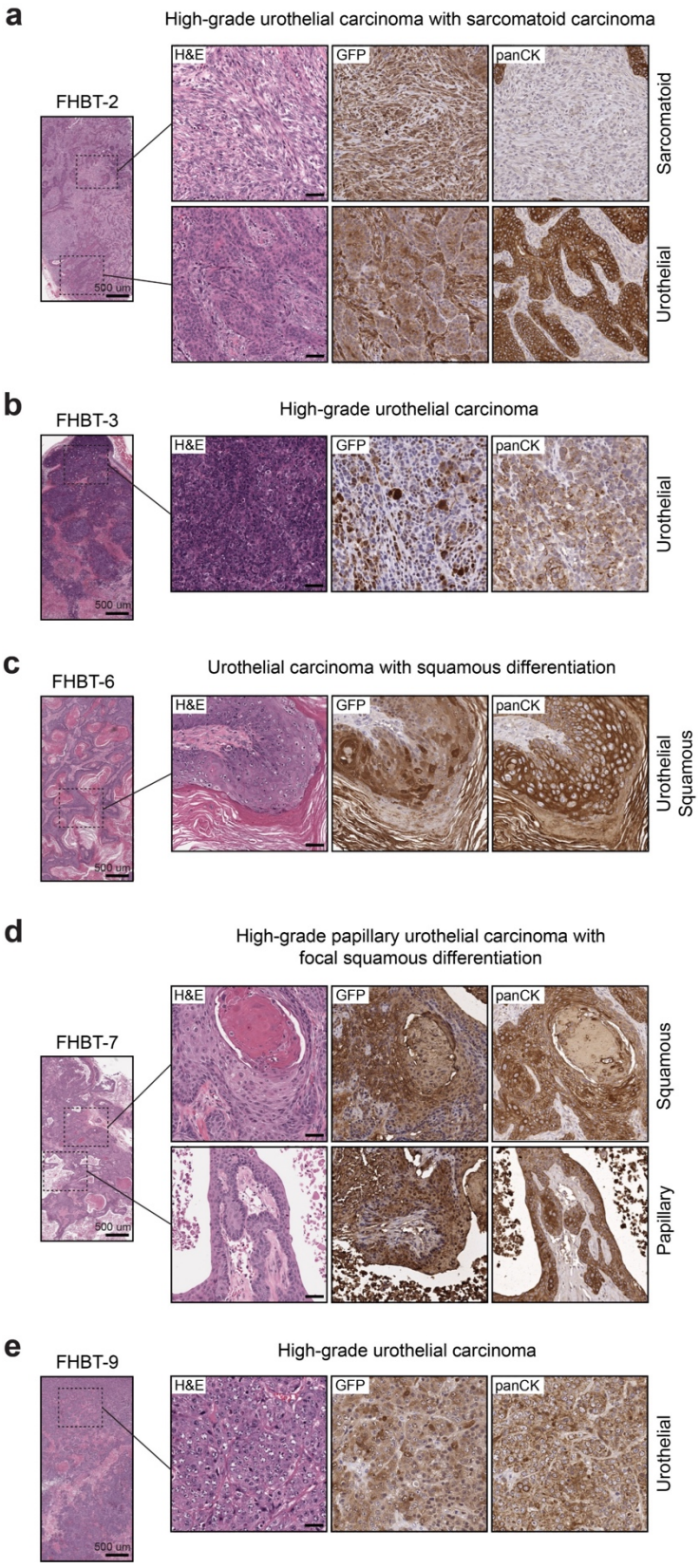
2.3.5. Extended Data Fig. 3. **Generation of barcoded lentiviral libraries and normalization of library representation.**

Schema showing the (**a**) generation of individual lentiviruses from the library in arrayed format with subsequent pooling and concentration by ultracentrifugation and (**b**) transduction of respective mouse bladder urothelial (mBU) or prostate epithelial (mPE) cells in organoid culture with concentrated lentiviral libraries to determine lentiviral barcode representation by bulk amplicon sequencing of genomic DNA. (**c**) Representative distribution of barcoded lentiviruses within a library with skewed enrichment of shRNA relative to ORF lentiviruses. (**d**) Representative distribution of barcoded lentiviruses within a library after applying information from **c** to adjust producer cell surface area in **a** for the generation of the lentiviral library.



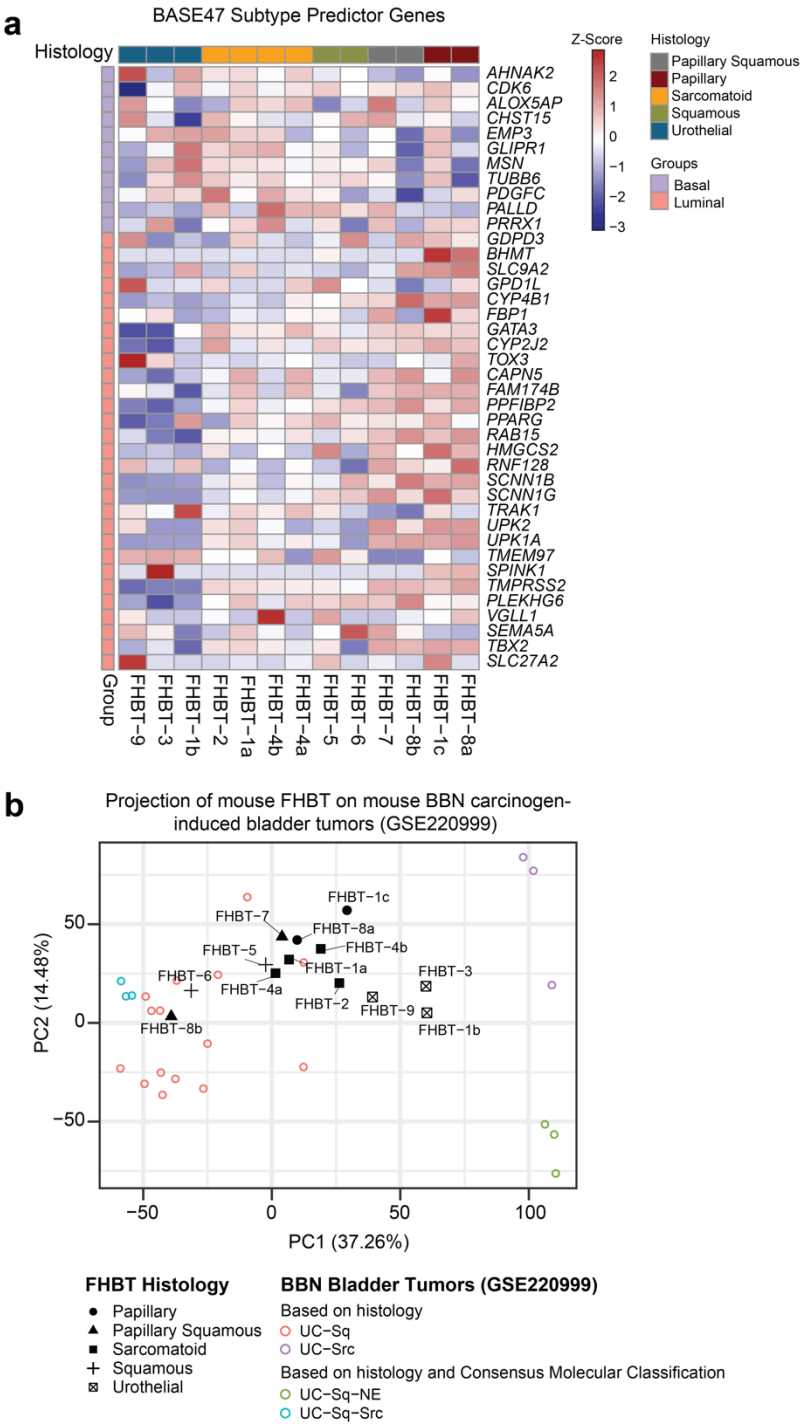
2.3.6. Extended Data Fig. 4. **Active mutant Fgfr3 S243C cooperates with other oncogenic factors in mouse bladder urothelial cells to drive papillary urothelial carcinoma with inverted growth pattern.**

(**a**) Scheme of the mBU organoid transformation assay using a defined lentiviral library to confirm functional genotype-phenotype associations. (**b**) High-magnification images of H&E- and IHC-stained sections of a resultant tumor of the experiment in **a** with histologic features consistent with papillary urothelial carcinoma with inverted growth pattern.



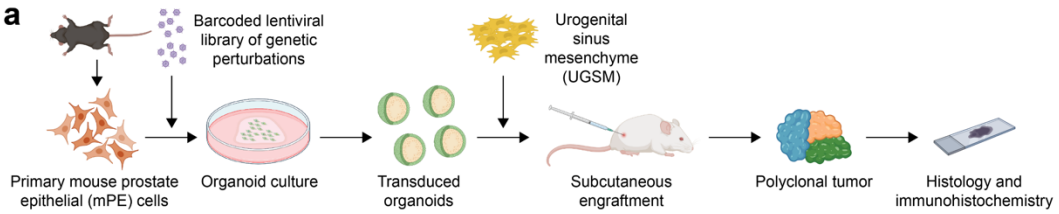
2.3.7. Extended Data Fig. 5. **FHBT models demonstrate diverse cancer histologies.**

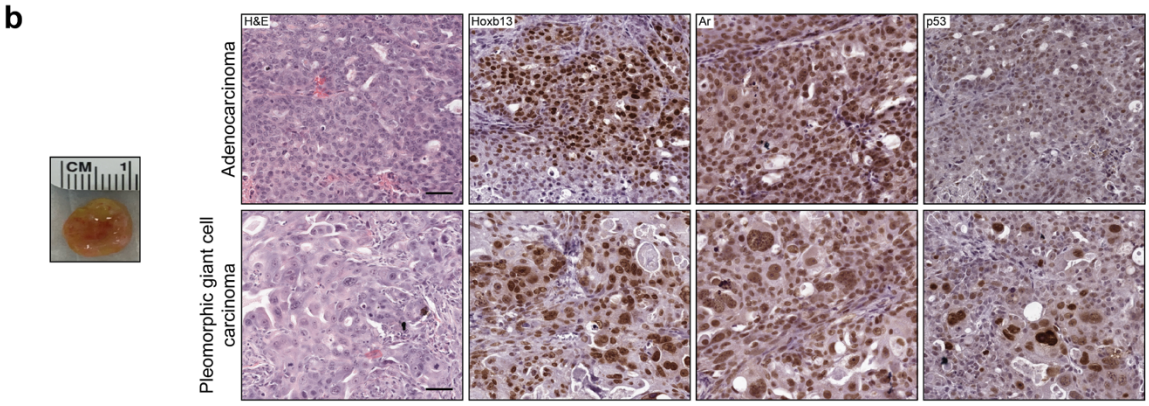
(**a-e**) Low- and high-magnification images of H&E-stained sections and high-magnification images of IHC-stained sections for GFP and pan-cytokeratin (panCK) expression depicting characteristic histologies. Scale bars=50 μm.

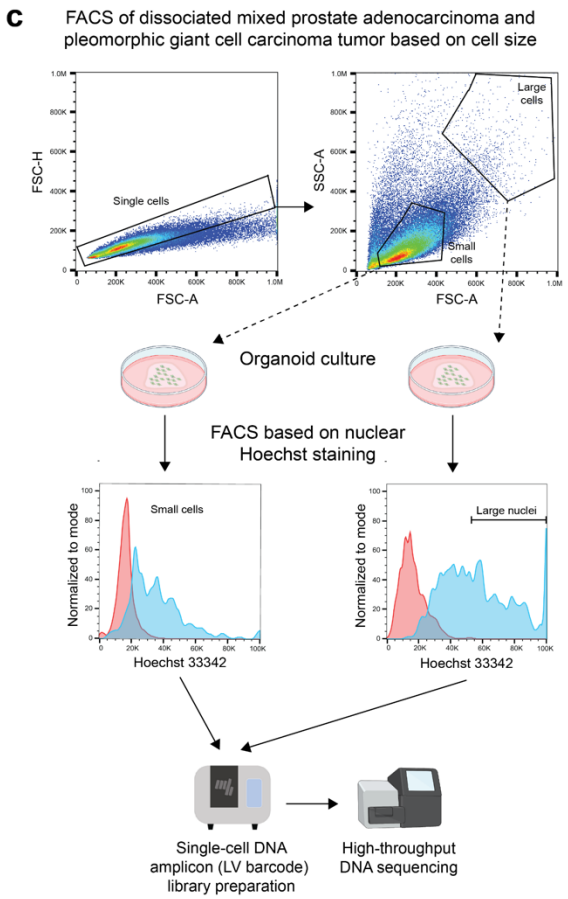


2.3.8. Extended Data Fig. 6. **Phenotypic diversity and relevance of FHBT models.**

(**a**) Heatmap showing the histologies of the FHBT series relative to expressions of genes that constitute basal and luminal signatures for the BASE47 subtype predictor. (**b**) PCA projection plot of FHBT samples over BBN carcinogen-induced mouse bladder tumors color-coded based on histology or histology and Consensus Molecular Classification.







d

Single-cell clonality analysis of small cells/nuclei

Genes encoded by LV	Clonal frequency (%)
Etv1	34.3
Spop F133C, Etv1	26.1
shZfhx3	5.0
Spop F133C, shZfhx3	2.9
Myc, shZfhx3	2.6
Myc, Spop F133C, shZfhx3	2.2
Etv1, shZfhx3	1.9
Myc, Etv1	1.9
Spop F133C, Etv1, shZfhx3	1.6
Trp53 R245Q, Etv1, shZfhx3	1.5

Single-cell clonality analysis of large cells/nuclei

Genes encoded by LV	Clonal frequency (%)
shKmt2c, shTtn, Etv1	11.8
shKmt2c, shTtn, Spop F133C, Etv1	5.5
shKmt2c, shTtn, Erg, Etv1	5.1
Myc, Etv1	4.7
Spop F133C, Etv1	4.5
shKmt2c, Etv1	4.1
shKmt2c, shTtn, Spop F133C, Erg, Etv1	2.7
shKmt2c, Spop F133C, Etv1	2.7
shKmt2c, shTtn, Myc, Etv1	2.7
Myc, Spop F133C, Etv1	2.3
shKmt2c, Spop F133C, Erg, Etv1	2.1
shKmt2c, Myc, Etv1	1.9
shKmt2c, Pik3ca E545K, Spop F133C, Erg, Etv1	1.9
shKmt2c, Erg, Etv1	1.8
shKmt2c, Myc, Spop F133C, Etv1	1.7
shKmt2c, Pik3ca E545K, shTtn, Erg, Etv1	1.7
shKmt2c, Pik3ca E545K, shTtn, Etv1	1.5

2.3.9. Extended Data Fig. 7. **Association of adenocarcinoma with polymorphic giant cell carcinoma of the prostate with perturbation of *Kmt2c*.**

(**a**) Scheme of the mouse prostate epithelial (mPE) organoid transformation assay to uncover functional genotype-phenotype associations in prostate cancer. (**b**) *Left*, Gross image of a tumor arising from mPE transformed with PE-LVp. *Right*, high-magnification images of H&E- and IHC-stained sections of regions with high-grade adenocarcinoma and pleomorphic giant cell carcinoma. Scale bar=50 μm. (**c**) Overview of the experimental approach to enrich for prostate adenocarcinoma and pleomorphic giant cell carcinoma based on cell size and nuclear DNA content followed by single-cell lentiviral barcode enumeration. (**d**) Tables showing single-cell clonality analysis of: *Top*, tumor cells enriched for "small cells/nuclei." *Bottom*, tumor cells enriched for "large cells/nuclei." Highlighted in red is sh*Kmt2c* based on the enumeration of the associated lentiviral barcode.

2.4 TABLES

Extended Data Table 1. **Summary of barcoded lentiviral ORF and shRNA constructs.**

Barcode (BC) #	***Barcode sequence***	***Encoded gene***	***TRC shRNA clone ID***
01	GCAGATTGTA	*Yap1*	-
02	CAACGTCGAC	ORF control	-
03	TGTTGATTCA	*E2f3*	-
04	TCGCACACAC	*Pparg*	-
05	TAGTTGGCTT	sh*Kdm6a*	TRCN0000331919
06	GGGATCGATC	*Mdm2*	-
07	AGCGTGTACA	*Ccnd1*	-
08	TCGTCCTTGA	sh*Stag2*	TRCN0000295340
09	GAGTCATGTC	sh*Ncor1*	TRCN0000350169
10	GGCCATTCAG	*Pvrl4*	-
11	ATCTACTCGC	*Ywhaz*	-
12	CAAACTACGT	sh*Kmt2d*	TRCN0000239233
13	ACGATATTAG	sh*Rb1*	TRCN0000235830
14	CAAACAATGA	sh*Crebbp*	TRCN0000012725
15	TCGGGACAGA	sh*Ep300*	TRCN0000071207
16	ACCGTTAGAG	*Ccne1*	-
18	ACCCACATGC	sh*Sptan1*	TRCN0000090595
20	CTTTGACTAT	*Zfp703*	-
22	GGTCAAATCG	*Fgfr3* S243C	-
23	CCCAAATGAT	sh*Atm*	TRCN0000360328
24	AGGCCTATCG	*Myc*	-
25	AGACTCGATG	*Erbb2* S311Y	-
26	CTATGCGTCC	sh*Kmt2a*	TRCN0000034426
27	TCGTACGGTT	sh*Kmt2c*	TRCN0000238934
28	GCTGGGAGTA	sh*Pten*	TRCN0000355842
29	GAGTTCTCTA	*Egfr*	-
30	CCTATGAGTT	sh*Cdkn2a*	TRCN0000231227
31	CCAATCTGGG	sh*Arid1a*	TRCN0000071396
32	GAACTGAACT	TRC control	-
33	TGCATCGTTT	*Erbb3* V104L	-
34	GAAATACCTG	*Pik3ca* E545K	-
35	TTTGTTAGGT	sh*Spen*	TRCN0000226288
36	AAACCATACT	sh*Brca2*	TRCN0000349771
37	AACGTGTCGT	sh*Ttn*	TRCN0000362964
38	CGCATAAGCA	sh*Cdk12*	TRCN0000361778
39	GGTCGCAATT	shChd1	TRCN0000096526
40	GGTGGAACTG	sh*Cdkn1b*	TRCN0000294885
41	TGGGTCAGCA	sh*Apc*	TRCN0000244294
43	ATGCTGTAGG	*Prex2*	-
44	TTAAGCGCCT	*Rspo2*	-
45	ATCGGTTCAC	*Foxa1* R261C	-
46	CATGTGGACC	*Spop* F133C	-
47	TGCATGTCCT	*Ctnnb* D32A	-
48	CATTACCAAG	*Erg*	-
49	GGTCACCACA	*Ar*	-
51	ATAGAAACTC	*Etv1*	-
52	AAATGGGTAT	*Braf* G469A	-
54	CATGTCAGTC	sh*Zfhx3*	TRCN0000075411
59	GCGTGGTTCT	*Trp53* R245Q	-

Extended Data Table 2. **Primers used for qPCR studies to quantify relative expression of target genes.**

Target gene	***Primer 1 (5' - 3')***	***Primer 2 (5' - 3')***
Kdm6a	ACCTAGTCCTCAGATCATACCA	GCTACTATTAGACAGGCCGTT
Stag2	GGCAGATAAATTTAACCGGCTTC	GCATTATGAAAAGCAGTGATCCTC
Ncor1	ATCCAGCTATGCCCTTTCAC	TGTCTGCCTTGTATTCTCCATT
Kmt2d	TCAGTGCTATCACCCGTACT	CACCTCACACACGATACACTC
Rb1	CCTCAGCCTTCCATACTCAG	CGGAGATATGCTAGACGGTACA
Crebbp	TCACAATCAACATCTCCTTCCC	TGTCGATAGAGTGCTTCTAGAGT
Ep300	CTCAGAAACTGTATGCCACCA	GCATCTCTACCGTCCATCAG
Sptan1	GGAGGTGTATGGTGCGATG	TTGATGGAGTTGAAGGTAGCC
Atm	GTCACAAAGAACCATGCTTGC	CACCTTCGCAACCTCAAGA
Kmt2a	GAAGATGCCTGGAAGTCACT	TGCTCAATCAGAAACACAACG
Kmt2c	AGAAGGATGAAGAGGAAAAGCA	GGTGGTGTAGGAGGAAGAGAG
Pten	CACTGCTGTTTCACAAGATGATG	TTCACCTTTAGCTGGCAGAC
Cdkn2a	GTGCGATATTTGCGTTCCG	CTCTGCTCTTGGGATTGGC
Arid1a	CCTCTATCGCCTCTATGTGTCT	GCACTGCTTGATGTACCCA
Spen	CGACTACTTACCACGACCTTC	CACACACTAGCGATATCACAGT
Brca2	GATGCCTAAACCCAGAAAGAGT	TGTGTCATCCCTCTCCAGTATC
Ttn	CAATGGATCTGGACAAGCGA	CACTCTCACTTGGAGTCTCAC
Cdk12	AACAGACCCTACAGAGTGACT	TCGACGTTTCTTACTCCACAA
Chd1	CGCCCAGCTTCATCTAATAGTG	CATCATTGTGCTTCTTCCTCTTG
Cdkn1b	GAGCAGACGCCCAAGAAG	GCAGTGATGTATCTAATAAACAAGGA
Apc	AGAATGAAGGTCAAGGAGTGG	TACTAGAACTCAAAACACTGGCT
Actb	GATTACTGCTCTGGCTCCTAG	GACTCATCGTACTCCTGCTTG
Zfhx3	AACAACAAGATCCACCTCCAG	ACTAGGCATAACCATCTCAGGA
Ubc	AACATCCAGAAAGAGTCCACC	CATTCTCTATGGTGTCACTGGG

Chapter 3. DEFINED CELLULAR REPROGRAMMING OF ANDROGEN RECEPTOR-ACTIVE PROSTATE CANCER TO NEUROENDOCRINE PROSTATE CANCER

This work has been submitted for publication.

3.1 INTRODUCTION

Prostate cancer (PC) is a hormonally-driven disease in which androgen receptor (AR) signaling plays a fundamental role in defining a luminal epithelial cancer lineage and promoting cell survival and proliferation. Thus, androgen deprivation therapy (ADT) and AR signaling inhibitors (ARSIs) represent standard and effective frontline treatments for advanced PC. While the majority of PCs initially respond to combined treatment with ADT and an ARSI, therapeutic resistance ultimately develops over the course of months to years and leads to a disease state known as castration-resistant PC (CRPC) [112, 113]. The mechanisms by which CRPC subvert ADT and ARSIs are numerous, but most are centered on aberrant reactivation of AR signaling through AR gene mutations, amplification, alternative splicing, and enhancer duplication [43, 113-115]. Beyond reactivation of the targeted AR pathway, clinical evidence implicates lineage plasticity or a shift in cell identity as a more extreme adaptive mechanism of resistance in PC and other epithelial malignancies. A prime example is treatment-related transdifferentiation of AR-active PC (ARPC) to neuroendocrine PC (NEPC) in which the AR program is extinguished and supplanted with a NE program (NEtD) [53, 116], thereby leading to a cancer lineage state that no longer depends on AR signaling and is resistant to ADT and ARSIs.

Multiple genetic and molecular events have been associated with the process of NEtD, including amplification of *MYCN*, loss of RE1-silencing transcription factor (REST) activity, and loss of *PTEN*, *TP53*, and *RB1* [69, 70, 117-120]. While our prior studies using human prostate epithelial transformation models have demonstrated that overexpression of *MYCN* and constitutively active myristoylated *AKT1* [69] or the PARCB factors (dominant-negative *TP53*, myristoylated *AKT1*, short hairpin targeting *RB1*, *MYC*, and *BCL2*) can initiate NEPC [70], these experimental systems start with AR-null basal prostate epithelial cells. Thus, they do not address how AR-enforced luminal epithelial lineage commitment may be repressed and bypassed which is a critical step in the NEtD of PC. The leading biological model currently available to investigate NEtD is the LTL-331 ARPC patient-derived xenograft (PDX) model which reproducibly undergoes transdifferentiation after passage in castrated mice to the LTL-331R NEPC tumor model [67]. Repeated biopsies and molecular analyses of tumors from this system have provided insights into molecular programs that are associated with NEtD [121]. Yet, this model remains experimentally unwieldy for functional studies, and other more tractable approaches have been lacking.

Another challenge has been the inconsistent definition of NEPC applied in the field. Clinically, NEPC consists of a heterogeneous group of neuroendocrine tumors primarily defined by histologic morphology that includes a range of subtypes from well-differentiated carcinoid to aggressive large-cell and small-cell carcinomas [122]. The expression of NE markers like chromogranin A (CHGA), synaptophysin (SYP), or CD56 (neural cell adhesion molecule 1 or NCAM1) are often assayed by immunohistochemistry (IHC) to support or confirm a diagnosis of NEPC. In general, individual NE markers expression can be heterogeneous, and combined marker scores are more

robust [123]. In experimental studies, phenotypic features such as NE marker expression or the emergence of dendrite-like projections from PC cells are broadly used as indicators of NE transdifferentiation of PC. The use of these sparse phenotypic features to define NEPC may be problematic as they may not reflect a *bona fide* cancer lineage state divergent from ARPC. A prime example is amphicrine prostate cancer (AMPC) which co-expresses AR and NE markers [124]. Despite the expression of NE markers, AMPC models such as VCaP remain sensitive to the ARSI enzalutamide, indicating a functional dependence on the AR-driven lineage program.

Lineage-defining transcription factors and transcription factor networks specify cell states by regulating chromatin accessibility and gene expression programs. The basic helix-loop-helix (bHLH) proteins Achaete-scute homolog 1 (ASCL1) and neurogenic differentiation 1 (NeuroD1) are pioneer transcriptional factors that are necessary for lineage specification during neuronal development [125, 126]. In cancer, ASCL1 and NeuroD1 define distinct subtypes of small cell lung cancer (SCLC) [127] and NEPC [128] with differential transcriptional, epigenetic, cell surface proteomic, and drug sensitivity profiles [129-131]. The role of ASCL1 as a driver of neuroendocrine transformation has been supported by the genetically engineered *TCKO* mouse model of high-grade neuroendocrine lung cancer where *Ascl1* knockout prevents tumor formation [127]. Further, knockdown of *ASCL1* or *NEUROD1* in established human SCLC or NEPC cell line models significantly impairs cell viability [132, 133], indicating that these factors are needed for sustained survival and growth. Insights into the functional role of these pioneer transcription factors in mediating the acute lineage conversion from ARPC to NEPC are limited. Recent work has shown that ASCL1 expression may be induced by ARSI therapy in PC and ASCL1 activates

a neuronal stem cell-like lineage program through chromatin remodeling mediated by the polycomb repressive complex 2 (PRC2) [134].

In this study, we developed a genetically-defined strategy adapted from the field of cellular reprogramming to directly convert ARPC to NEPC. We identified key functional roles for ASCL1 and NeuroD1 in NE transdifferentiation, specifically new findings uncovering their ability to silence the AR program through chromatin remodeling of the AR enhancer which is somatically amplified in PC and global AR binding sites with enhancer activity. In contrast, we found that loss of REST activity induces the expression of NE markers but is insufficient to drive a NE cancer lineage state. Repeated sampling over the reprogramming time course revealed a wide range of temporal changes in the transcriptomic and epigenomic landscapes associated with the lineage conversion from ARPC to NEPC. Lastly, we found that ASCL1 and NeuroD1 expression, but not the loss of REST activity, induces the downregulation of class I major histocompatibility complex (MHC) expression in PC, providing a basis for the lineage-specific potential for immune evasion associated with NEPC.

3.2 RESULTS

Candidate factors directly reprogram ARPC to NEPC and bypass the dependence on AR signaling

To functionally define factors required for the conversion of ARPC to NEPC, we established a NEtD assay wherein candidate factors were introduced into the human ARPC cell lines LNCaP and C4-2B after which they were propagated in media conditions permissive to the growth of

NEPC and analyzed for a set of phenotypic markers of ARPC and NEPC after two weeks (**Figure 1A**). Initially, we introduced a pool of lentiviruses (LVs) expressing eight factors associated with NEPC including dominant-negative *TP53* R175H, a short hairpin targeting *RB1* (shRB1), *MYCN*, *ASCL1*, *SRRM4*, *NR0B2*, *BCL2*, and *KRAS* G12V. Loss of *TP53* and *RB1* as well as amplification of *MYCN* have been identified as common genetic features associated with NEPC [69-71, 135]. SRRM4 is a splicing regulator that has been shown to inactivate the RE-1 silencing transcription factor (REST) which functions to suppress a major neural gene program. Loss of REST has been implicated as a common feature in NEPC that has been reported to functionally drive NEtD of PC [119, 136]. In addition, the pioneer neural transcription factor ASCL1 is expressed in most NEPC and appears to define a transcriptional and epigenetic subtype of the disease [128]. Lastly, BCL2 and KRAS G12V were included to potentially reduce apoptosis and enhance reprogramming efficiencies [137, 138].

The introduction of this candidate factor LV pool at a multiplicity-of-infection (MOI) of four for each factor in this assay led to a significant downregulation of AR and the AR target NK3 homeobox 1 (NKX3-1) as well as the upregulation of NE markers including forkhead box A2 (FOXA2), SYP, NCAM1, insulinoma-associated protein 1 (INSM1), and POU class 3 homeobox 2 (POU3F2 or BRN2) by immunoblot analysis (**Figure 1B**). Furthermore, immunocytochemical analysis of the C4-2B and LNCaP cell lines transduced with the candidate factor LV pool demonstrated a prominent loss of nuclear AR expression and increased SYP expression (**Figure 1C**). The $AR^{low/-}NE^{+}$ cancer phenotype induced by the candidate factor LV pool was durable based on IHC studies in tumors obtained after subcutaneous xenografting in non-castrate male NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl} *SzJ* (NSG) mice (**Figure S1A**).

We next sought to determine whether the PC cells reprogrammed with the candidate factor LV pool were functionally independent of AR signaling. We implemented a potent AR-dependent negative selection strategy (**Figure 1D**) wherein C4-2B cells were transduced with LV encoding AR response elements (ARE) upstream of a FK506 binding protein-caspase 8 (FKBP-Casp8) fusion. A clonal C4-2B ARE-FKBP-Casp8 line was established and transduced with either a control LV expressing green fluorescent protein (GFP) or the candidate factor LV pool and subjected to selection with 1 nM R1881 (to induce AR signaling and expression of FKBP-Casp8 in AR^+ cells) and AP20187 which is a synthetic FK506 analog (to activate FKBP-Casp8 and induce apoptosis). We identified no viable, adherent cells in the control LV condition after five days of selection. In contrast, we appreciated viable, adherent cells in the candidate factor LV pool condition and ~70% of cells demonstrated functional AR bypass as they could survive stringent selection against AR signaling (**Figures 1E and S1B**). These finding indicated that a discrete set of factors is capable of reprogramming human ARPC, in the context of negative selection for non-reprogrammed cells (modeling highly stringent AR inhibitor therapy), to a cellular state that recapitulates critical phenotypic and functional features of NEPC.

The pioneer neural transcription factor ASCL1 suppresses the AR program during NEtD but requires additional genetic alterations to maintain a proliferative phenotype

To specify the functional contribution of each candidate factor in the LV pool to NE reprogramming, we employed a leave-one-out approach in which each factor was iteratively left out from the LV pool in the C4-2B (**Figure 2A**) and LNCaP **(Figure S2)** cell lines. Based on

immunoblot analyses, we observed several established regulatory interactions that have been previously identified in cancer. These included 1) ectopic expression of N-Myc leading to elevated levels of NCAM1 in neuroblastoma [139], 2) repression of POU3F2 expression by Rb1 in retinoblastoma [140], 3) and enhanced SOX2 expression with Rb1 inactivation in PC [120]. Most strikingly, we observed that the exclusion of ASCL1 diminished the decrease in AR and NKX3-1 expression and likewise abrogated the otherwise induced FOXA2, NCAM1, and INSM1 expression, implicating ASCL1 as a critical player in NE transdifferentiation. In contrast, the omission of SRRM4 restored REST expression and suppressed SYP but had little effect on AR or other NE markers. These data are consistent with prior reports in which ectopic introduction of SRRM4 in PC cell lines induced SYP and chromogranin B expression but did not directly affect AR expression [136, 141]. Removal of NR0B2, BCL2, and KRAS G12V had no significant effects on AR and NE marker expression (**Figures 2A and S2A**).

Next, we employed an orthogonal factor reconstitution strategy to identify the minimal set of factors needed to induce NEtD in the ARPC lines. Immunoblot analyses showed that ASCL1 expression alone downregulated AR markers and induced NE marker expression **(Figure 2B)**. On the other hand, the PRNB factors (dominant-negative TP53 H175R, shRB1, MYCN, and BCL2) were insufficient to enforce NEtD (**Figure 2B**). SRRM4 alone imparted a partial effect on phenotype as it led to a loss of REST expression and upregulated SYP without reducing AR markers (**Figure 2B)**. As *TP53*, *RB1*, *MYCN*, and *BCL2* alterations are common in NEPC and other aggressive NE carcinomas we questioned why they appeared to be dispensable in our NEtD assay. We measured cell proliferation for each of the factor reconstitution conditions as an indicator of proliferative fitness in our experimental system. PRNB significantly enhanced

proliferation while ASCL1 was detrimental to cell viability over time in the ARPC lines (**Figure 2C and Figure S2B**). This effect of ASCL1 is consistent with prior reports that ASCL1 promotes cell cycle exit and terminal differentiation when overexpressed in neural progenitor cells and glioblastoma stem cells [142]. Notably, SRRM4 promoted proliferation but alone was insufficient to rescue the effects of ASCL1 without PRNB. These data support the concept that the appropriate genetic and proliferative context may be an important determinant of ASCL1's impact on the emergence of NEPC and highlight the idea that overall fitness may be a combination of proliferation- and lineage-linked benefit. In other words, this mechanism of therapeutic escape may require a change in lineage without a substantial decrease in proliferative fitness.

NeuroD1 is competent to induce neuroendocrine lineage reprogramming

We have previously observed that most NEPCs can be assigned to ASCL1^{high}, NeuroD1^{high}, or mixed ASCL1/NeuroD1 groups based on transcriptional profiling [132]. These findings have also been described by other groups wherein ASCL1 and NeuroD1 appear to define molecular subtypes of NEPC [128] and the neuroendocrine subtypes of small cell lung cancer [127]. To determine whether NeuroD1 can also drive lineage conversion from ARPC to NEPC, we performed the reprogramming assay using the C4-2B and LNCaP ARPC lines by replacing ASCL1 with NeuroD1 or combining NeuroD1 with ASCL1. The addition of NeuroD1 to PRNB and SRRM4 substantially diminished AR markers and increased NE marker expression (**Figure 3A**). We discovered that NeuroD1 induced ASCL1 expression and heightened overall NE marker panel expression. Bulk RNA-seq gene expression analysis of the reprogrammed cell lines with respect to established 22-gene AR signature and NE signature scores revealed a significant de-enrichment

of the AR gene signature and enrichment of the NE gene signature in the PRNBSA (PRNB, SRRM4, and ASCL1) and PRNBSN (PRNB, SRRM4, and NeuroD1) conditions relative to the GFP control (**Figure 3B**). As C4-2B is derived from the LNCaP cell line, we also confirmed the reproducibility of these findings in the independent ARPC cell line MDA PCa 2b using this cellular reprogramming assay (**Figure S3A**).

Several neural transcription factors including FOXA2, INSM1, POU3F2, and SOX2 have also been described as documented drivers or regulators of NE transdifferentiation in PC and other cancer types [75, 120, 143]. We sought to determine whether these factors are independently competent to induce NE lineage reprogramming. We introduced PRNB, SRRM4, and each of the above listed neural transcription factors individually to the C4-2B line in the NEtD assay. In contrast to ASCL1 and NeuroD1, these factors were by themselves incapable of NE lineage reprogramming (**Figure S3B**). As ASCL1 and NeuroD1 also appear to regulate the expression of FOXA2 and INSM1, ASCL1 and NeuroD1 may sit atop a hierarchy of multiple NE-associated transcription factors required in combination to establish and maintain a NEPC transcriptional regulatory network.

To confirm the relevance of our NE transdifferentiation models in PC and to use omic-scale data to minimize the effect of individual marker heterogeneity, we performed partial least squares-discriminant analysis (PLS-DA) of RNA-seq gene expression data from previously characterized human metastatic CRPC patient specimens annotated as ARPC or NEPC [62] and projected the C4-2B factor-reconstituted samples onto them (**Figure 3C**). This analysis showed that PRNB alone did not induce a substantial shift away from the ARPC phenotype. Further, the addition of

PRNB and SRRM4 (PRNBS) resulted in a minor, intermediate shift toward the NEPC phenotype. Only upon the addition of ASCL1 or NeuroD1 to PRNB (PRNBA and PRNBN) or PRNBS (PRNBSA and PRNBSN) did we observe a dramatic gene expression program-based shift from ARPC to NEPC. These data highlight the relative lineage impacting importance of these factors in NEtD of PC.

SRRM4 promotes alternative splicing and inactivates REST through the inclusion of a neural exon that generates the REST-4 isoform [136, 144]. We confirmed the presence of alternatively spliced *REST* mRNA in C4-2B reprogramming conditions that included SRRM4 (**Figure S4A**). Further, we performed PLS-DA analysis on alternative splicing data from Beltran et al., 2016 and projected the C4-2B reprogrammed samples onto them. This analysis indicated that SRRM4 drives the alternative splicing phenotype associated with NEPC with a marginal contribution from ASCL1 or NeuroD1 (**Figure S4B**).

Dynamic changes in cancer phenotype, transcriptional and epigenetic landscapes during NE reprogramming

Having established a defined and tractable system that recapitulates NE transdifferentiation of PC, we sought to investigate the dynamic molecular and regulatory changes underlying this shift in cancer lineage state. We applied the NE reprogramming assay to the ARPC C4-2B line and temporally sampled multiple timepoints within the 14-day reprogramming period (D2, D3, D4, D5, D8, D11, and D14) for downstream analyses (**Figure 4A**). Immunoblot analyses revealed a reduced expression of AR markers (AR and NKX3-1) as early as D2, with expression of the NE

markers INSM1 and FOXA2 evident shortly thereafter on D4-5 (**Figure 4B**). These findings indicate that downregulation of the AR program precedes the establishment of the NE program in our experimental system. Transcriptome profiling by bulk RNA-seq of the C4-2B cells reprogrammed with PRNBSA or PRNBSN recapitulated the findings from the immunoblot analyses where downregulation of *AR* expression preceded the induction of *INSM1* expression (**Figures 4C and D**). Notably, PCA analysis of the time course RNA-seq gene expression data from C4-2B cells reprogrammed with PRNBSA or PRNBSN, but not PRNBS, showed an arc-like trajectory during NEtD, indicating dynamic changes in gene expression programs during ARPC to NEPC reprogramming (**Figures S5A-C**). This arc-like trajectory is also broadly observed in normal and cancer development and cellular differentiation [145-148].

We next used PLS-DA projections from the time course RNA-seq gene expression data to further investigate the temporal nature of the phenotypic shift from ARPC to NEPC during NEtD. We distinguished three general phases of reprogramming in the C4-2B cells modified with PRNBSA and PRNBSN that could be defined as “early” on D2, “intermediate” from D3-5, and “late” from D8-14 (**Figure 4E**). In contrast, these phases were not apparent in C4-2B cells transduced with PRNB and SRRM4 (**Figure S5D**), further supporting the previous finding that ASCL1 or NEUROD1 is needed for full lineage shift via NEtD.

Lineage reprogramming of ARPC to NEPC is largely driven by epigenetic dysregulation and reprogramming as few genetic alterations apart from the enrichment of *TP53* and *RB1* mutations and genomic amplification of *MYCN* have been identified in NEPC [62, 71, 149, 150]. To investigate global changes in the epigenetic landscape during the transition from ARPC to NEPC,

we performed Cleavage Under Targets & Release Using Nuclease (CUT&RUN) analysis on PRNBSA reprogrammed C4-2B cells over the 14-day time course to examine genome-wide changes in histone modifications, including monomethylation of histone H3 at lysine 4 (H3K4me1) which is commonly associated with gene enhancers, acetylation of histone H3 on lysine 27 (H3K27ac) which indicates both active promoters and enhancers [151, 152], and trimethylation of histone H3 at lysine 4 (H3K4me3) which marks active gene promoters. Unsupervised principal component analysis (PCA) showed that these histone modifications shift significantly during the time course (**Figure 4F**), indicating that dynamic epigenetic changes occur during NEtD.

As ASCL1 is a pioneer transcription factor, we queried whether there may be time-dependent changes in putative ASCL1-mediated gene regulation during the acute conversion of ARPC to NEPC. We integrated the RNA-seq gene expression data and ASCL1 CUT&RUN genome occupancy data over the 14-day reprogramming period to identify genes both in proximity to ASCL1 binding and upregulated in PRNBSA compared to PRNBS. Gene set enrichment analysis (GSEA) revealed several pathways enriched early during NE reprogramming including the CDC25 pathway, TGFβ1 pathway, and estradiol response pathway (**Figure 5A**). In contrast, ASCL1 targets, PRC2 targets including those of the core subunits embryonic ectoderm development (EED) and SUZ12, and neuronal fate commitment pathways were significantly enriched later in the NE transdifferentiation process. The pathways enriched early may represent components of an unstable transcriptional state prior to the establishment of a stable NEPC lineage, consistent with our observation of an arc-like trajectory.

Next, we evaluated the dynamic activities of active enhancers, super-enhancers, and active promoters in the NE reprogramming assay. We defined active enhancers based on overlapping H3K4me1 and H3K27ac peaks and characterized super-enhancers by applying the Rank Ordering of Super-Enhancer (ROSE) algorithm to H3K27ac peaks over the reprogramming time course [153]. This analysis revealed a substantial temporal shift in active enhancers (**Figure S6A**) and super-enhancers (**Figure 5B**). Gene ontology (GO) analysis of downregulated active enhancer- and super-enhancer-associated genes over time revealed a highly significant enrichment of androgen response genes including the Hallmark Androgen Response gene set (**Figures S6B and 5C**). In contrast, genes associated with upregulated active enhancers over time were enriched for endocrine therapy resistance pathways and targets of the PRC2 complex core subunits SUZ12 and enhancer of zeste homolog 2 (EZH2) (**Figure S6C**). We also examined active promoters by determining the overlap of H3K4me3 and H3K27ac peaks over the reprogramming time course (**Figure S7A**). Genes associated with downregulated active promoters over time were enriched in androgen response and AR regulated targets (**Figure S7B**). On the other hand, GO analysis of upregulated active promoters over time did not highlight any gene sets of statistical significance.

ASCL1/NeuroD1 inhibits AR expression by remodeling the somatically acquired AR enhancer

To further investigate how ASCL1 and NeuroD1 may suppress AR expression and signaling during NE reprogramming, we performed single-cell (sc) multiome RNA-seq and ATAC-seq analysis on D14 after reprogramming of C4-2B cells with PRNB, PRNBS, PRNBSA, and PRNBSN. The PRNB and PRNBS conditions were included as controls as these factor

combinations have no apparent effect on the AR program. Pseudotime analysis revealed a progression from PRNB to PRNBS to PRNBSA to PRNBSN which was associated with a heightened NEPC score and reduced *AR* expression (**Figure S8A-D**). In addition, we observed two distinct clusters within the PRNBSA and PRNBSN conditions indicating heterogeneous cell populations likely due to differences in expression levels of *ASCL1* and *NEUROD1* (**Figures S8E, F**).

Most strikingly, the integration analysis of RNA-seq and ATAC-seq data identified a significant reduction in chromatin accessibility of the somatically acquired *AR* enhancer, but not the *AR* gene body, that was associated with reduced *AR* expression in the PRNBSA and PRNBSN conditions (**Figure 6A**). In contrast, the reduction in chromatin accessibility of the somatically acquired *AR* enhancer was not seen in the control PRNB and PRNBS conditions (**Figure 6A**). The somatically acquired enhancer of *AR* and its amplification independent of the *AR* gene body has previously been shown to be a noncoding driver of CRPC [43]. Attesting to the clinical relevance of this finding, we identified a hepatic metastasis from a patient with lethal mCRPC who underwent rapid autopsy as part of the University of Washington Tissue Acquisition Necropsy (UW TAN) program that showed a mixed population of $AR^{+}/ASCL1^{-}$ and $AR^{-}/ASCL1^{+}$ tumors cells by immunofluorescence staining (**Figure 6B**). Similar to the findings from our NE reprogramming studies, scATAC-seq analysis of this metastatic tumor showed loss of chromatin accessibility at the enhancer of *AR* but not the *AR* gene body in $AR^{-}/ASCL1^{+}$ tumor cells compared to $AR^{+}/ASCL1^{-}$ cells (**Figure 6C**).

We next sought to investigate the chromatin accessibility of genome-wide AR binding sites (ARBS) with enhancer activity. ARBS with enhancer activity have previously been characterized using Self-Transcribing Active Regulatory Regions sequencing (STARR-seq) in the LNCaP cell line and classified as constitutively active enhancers, inducible enhancers, and inactive enhancers based on their responsiveness to the androgen dihydrotestosterone [154]. We evaluated the chromatin accessibility at these ARBS and identified a global reduction in the accessibility of ARBS with enhancer activity from least reprogrammed (PRNB) to most reprogrammed (PRNBSA or PRNBSN) cell lines (**Figure 6D**). In contrast, a random selection of genomic sites of similar intervals across all chromosomes, as a control, did not show reduction in chromatin accessibility. These results further demonstrate the acute epigenetic rewiring of PC during NEtD to silence an AR-driven PC lineage program.

ASCL1/NeuroD1 but not loss of REST via SRRM4 activity establish a NE transcriptional network

Lineage-defining transcription factors including neural lineage transcriptions factors have been reported to auto-activate their own expression by binding to super-enhancers to establish a positive feedback loop required to sustain a lineage program [155]. We investigated the potential for self-regulation of ASCL1 and NeuroD1 in our NE reprogramming assay based on differential RNA-seq read alignment of exogenous proviral transcripts and endogenous transcripts (**Figure S9A**). We specifically leveraged the presence of the synonymous single-nucleotide variant NM_004316.4(ASCL1);c.627C>G encoded in the ASCL1 lentivirus to distinguish exogenous and endogenous reads. The introduction of either exogenous ASCL1 or NeuroD1 to PRNB or PRNBS

resulted in the induction of endogenous *ASCL1* expression (**Figure S9B**). This was also confirmed by the presence of reads mapping to the 5' and 3' untranslated regions (UTRs) which are present in endogenous transcripts but not encoded in exogenous proviral transcripts. Using a similar strategy of mapping reads to the 5' and 3' UTR, we also found that exogeneous NeuroD1 could activate endogenous *NEUROD1* expression (**Figure S9C**).

We then investigated how the different conditions in our NE reprogramming assay may impact the establishment of a transcriptional regulatory network (TRN) underlying an NEPC lineage state. We first focused on the differential expression of transcription factors due to SRRM4 (PRNBS vs. PRNB), ASCL1 (PRNBSA vs. PRNBS), and NeuroD1 (PRNBSN vs. PRNBS). We observed a significant induction of NE-associated transcription factors including INSM1, SOX2, SOX4, NK2 homeobox 2 (NKX2-2), and FOXA2 related to ASCL1 and NeuroD1 (**Figure 7A**). In stark contrast, not a single transcription factor demonstrated statistically significant differential expression upon the addition of SRRM4 to PRNB. Consistent with these findings, footprinting analysis of ATAC-seq data identified a significant enrichment in binding motifs associated with these NE-associated transcription factors in conditions of enforced ASCL1 and NeuroD1 expression but not SRRM4 expression (**Figure 7B**). Global chromatin accessibility profiles are also informative as they can differentiate lineage states in normal tissues and cancer. When we projected pseudo-bulk scATAC-seq data of reprogrammed C4-2B lines on the PCA space defined by ATAC-seq data of previously published ARPC and NEPC PDX models [128], we found that C4-2B lines reprogrammed with PRNBSA and PRNBSN clustered with the NEPC PDXs while those reprogrammed with PRNB and PRNBS were more similar to ARPC PDXs (**Figure 7C**). Further, we found that the expression of NEPC-associated cell surface targets (*DLL3*, *L1CAM*,

CEACAM5, *RET* and *SEZ6*) were significantly increased and the ARPC-associated cell surface targets (*FOLH1, PSCA, STEAP1*, and *TACSTD2*) were significantly decreased upon the addition of ASCL1 or NeuroD1. In contrast, SRRM4 alone did not have a prominent effect on the expression of these cell surface targets (**Figure 7D**). Taken together, these findings indicate that, unlike ASCL1 and NeuroD1, SRRM4 activity and the resultant loss of REST repressor activity alone may be insufficient to drive a NEPC lineage due to the absence of active transcriptional and epigenetic reprogramming.

Downregulation of major histocompatibility complex class I (MHC I) in NEPC is functionally attributable to ASCL1/NeuroD1-driven programs

The downregulation of MHC I antigen presentation machinery by cancer to overcome immune surveillance and avoid immune destruction is an established hallmark of cancer. In particular, this phenomenon has been associated with cancers with neuroendocrine differentiation [156]. In a recent study, examination of SCLC data sets showed that tumors with low NE scores demonstrated higher expression of MHC I genes while those with high NE scores showed decreased MHC I and reduced immune infiltration [157]. Further, cancer cells are capable of co-opting PRC2 in a lineage-specific fashion to silence the MHC I antigen processing and presentation pathway [158]. Mechanistically, mutation of the catalytic domain or pharmacologic inhibition of the PRC2 core subunit EZH2 was shown to impair MHC I repression, highlighting the critical role of PRC2 in this process.

We therefore asked whether a similar phenomenon may be operative in NEPC and could be functionally attributable to lineage-defining ASCL1 and/or NeuroD1 expression. We first examined a published RNA-seq gene expression dataset of mCRPC [62] including CRPC-Ad (adenocarcinoma) and CRPC-NE (NEPC) patient specimens and found that NEPC was associated with reduced MHC I antigen processing and presentation gene signature scores as well as lower *B2M* (beta-2-microglobulin, a critical component of MHC I) expression (**Figures 8A and S10**). A similar analysis using published data from tumors in the UW TAN PC rapid autopsy cohort also showed a reduced overall expression of *B2M* in the NEPC samples (**Figure 8B**). We then evaluated RNA-seq gene expression data from our reprogrammed C4-2B lines and correlated each condition and NE signature score with MHC I antigen processing and presentation gene signature scores. Consistent with prior literature, we discovered a reduction in MHC I gene expression with increasing NE differentiation (**Figure 8C**). Downregulation of MHC I gene expression started as early as D2 along with AR pathway repression (**Figure S11**). Furthermore, the addition of SRRM4 to PRNB did not have an appreciable effect on MHC I expression, indicating that loss of REST does not itself contribute to the downregulation of MHC I. In contrast, conditions where ASCL1 or NeuroD1 were added did demonstrate reduced expression of MHC I gene sets. We found that *B2M* expression was especially downregulated in conditions reprogrammed with NeuroD1 which were also associated with prominent NE signature scores (**Figure 8C**). These findings were in line with integrated scRNA-seq and scATAC-seq data from reprogrammed C4-2B cells showing an inverse correlation between NEPC score and *B2M* expression (**Figure 8D-F**) To confirm these findings at the protein level, we performed flow cytometry on reprogrammed C4-2B cells and showed that the PRNBSN condition was associated with a significant reduction in B2M positivity relative to the GFP control condition (29.2% vs. 60.9%) (**Figure 8G**). Together, these findings

functionally attribute the lineage-specific downregulation of MHC I in NEPC (relative to ARPC) to the lineage-defining transcription factors ASCL1 and NeuroD1.

3.3 DISCUSSION

AR signaling is pivotal in defining the luminal epithelial lineage of most PCs. However, potent therapies have been developed that apply stringent selective pressure against AR signaling, leading to the increased therapy-driven clinical presentation of lineage plasticity in PC. One such example is the conversion of ARPC to NEPC where the epithelial cancer state enforced by the AR program is replaced by a neural/NE cancer lineage [116]. Alternative AR-null PC lineages have also been identified including double-negative prostate cancer (DNPC) which lacks both AR and NE marker expression [124, 159]. Further, this lineage plasticity in the setting of targeted therapies has also been observed in other epithelial cancer types such as epidermal growth factor receptor (EGFR)-mutant lung adenocarcinoma in which transformation to SCLC occurs in up to 15% of patients treated with EGFR tyrosine kinase inhibitors [160]. However, the mechanisms by which AR (or EGFR) expression and signaling are effectively silenced during this transition have been poorly understood. The ability to functionally characterize these dynamic processes in a defined manner by reverse genetics has been limited by the lack of available model systems allowing for rapid manipulation and repeated sampling over time. Furthermore, due to documented heterogeneity, the application of narrow phenotypic markers to define NEPC has also been problematic and complicated our understanding of this PC lineage state.

We established an *in vitro* system in which we can reproducibly and durably reprogram established human ARPC cell lines to NEPC with defined genetic factors over a specified time course. This

tractable system recapitulates key features of the process of NEtD including global, dynamic transcriptional and epigenetic reprogramming. Importantly, we uncovered the ability of the pioneer neural transcription factors ASCL1 and NeuroD1 to suppress AR expression and signaling by chromatin remodeling of the somatically acquired enhancer of AR and global ARBS with enhancer activity. These findings show that the activity of a new lineage-defining transcription factor can concurrently silence the prior cancer lineage program while establishing a divergent lineage. An immediate question that emerges from these findings is whether AR and ASCL1 or NeuroD1 activity is mutually exclusive in individual PC cells. Emerging single-cell sequencing data sets of human and mouse models of NEPC indicate that this may be the case [161, 162]. The competence of ASCL1 and/or NeuroD1 to silence the AR program in prostate cancer may be a major constraint to NEtD, providing a potential basis for why NEPC only occurs in 15% of therapy-driven mCRPC cases. How the expression of these neural/NE transcription factors is initially induced in PC also remains to be determined and warrants further investigation.

Our studies have also distinguished the relative contributions of SRRM4 and ASCL1 or NeuroD1 to NE reprogramming of PC, and yield insight into the problematic nature of using sparse protein (IHC) or RNA analysis based on only a few markers to define NEPC lineage states. In our reprogramming studies, we find that SRRM4 activity results in the loss of REST which activates neural/NE gene expression including SYP and CHGA. However, unlike ASCL1 or NeuroD1, SRRM4 does not induce a new transcriptional regulatory network based on transcription factor expression and inferred transcription factor activity analyses, does not elicit epigenomic shifts consistent with NEtD, does not affect AR expression and signaling, and does not alter the expression of ARPC-associated cell surface targets. Thus, our findings indicate that a reliance on

a sparse set of NE markers, including highly canonical markers, to define NEPC can be flawed as their expression may be independent of true NEPC lineage states. A major clinical consideration is that mCRPCs that express select NE markers but have not undergone NEtD may still be responsive to therapies intended for ARPC including ARSIs and PSMA radioligand therapy.

In our NE reprogramming assay, we have found that ASCL1 and NeuroD1 but not other NE-associated transcription factors are competent to induce NEtD of PC. This supports a transcription regulatory network model where ASCL1 and NeuroD1 sit atop a hierarchy of NE-associated transcription factors necessary to establish self-sustaining NEPC transcriptional convergence. We have shown that ASCL1 and NeuroD1 are capable of self-regulation but the full set of regulatory interactions between these and other NE-associated transcription factors are unknown in NEPC. Future studies to define these relationships through genetic loss-of-function screening using this and other NEtD assays elucidate these functional interactions, identify potential vulnerabilities, and promote therapeutic development to inhibit NEtD under the pressures of otherwise effective targeted therapies.

Prior work identified suppression of MHC I antigen processing and presentation in neuroendocrine cancers and a functional association with lineage-specific PRC2 activity [156-158]. Because our studies directly implicated ASCL1 and NeuroD1 in lineage reprogramming of ARPC to NEPC as well as the induction of PRC2 activity, we examined the expression of MHC I pathway gene sets and discovered that MHC I pathway genes were suppressed in the effective ASCL1- or NeuroD1-driven reprogramming conditions. In contrast, the addition of SRRM4 and subsequent loss of REST activity had little to no effect on the MHC I pathway. These findings support that

downregulation of the MHC I antigen processing and presentation pathway in NEPC is functionally attributable to ASCL1 and NeuroD1 activity. Immune checkpoint inhibitors have demonstrated limited utility in mCRPC except in the case of DNA mismatch repair deficiency [163]. Thus, our results support therapeutic efforts to induce immunogenic responses in NEPC using combinations with epigenetic inhibitors such as those targeting EZH2, histone deacetylases, or demethylases to enhance MHC I expression [164, 165].

Overall, this study provides important mechanistic details of NEtD of PC by highlighting key lineage-defining activities of neural pioneer transcription factors in this process. These findings have important implications on the molecular definition of NEPC and therapeutic considerations related to the application of targeted agents and immune-based therapies.

3.4 MATERIALS AND METHODS

Cell lines

LNCaP (Cat# CRL-1740, RRID: CVCL_0395), C4-2B (Cat# CRL-3315, RRID: CVCL_4784), NCI-H660 (Cat# CRL-5813, RRID: CVCL_1576), MDA PCa 2b (Cat# CRL-2422, RRID: CVCL_4748) cell lines were purchased from ATCC. MSKCC EF1 were derived from the MSKCC PCa4 organoid line [166], as previously described [167]. All cell lines were validated by short tandem repeat analysis after receipt and on a yearly basis. LNCaP, C4-2B, and MSKCC EF1 were maintained in RPMI medium supplemented with 10% FBS, 100 U/mL penicillin and 100 µg/mL streptomycin, and 4 mmol/L GlutaMAX. NCI-H660 cells were maintained in Advanced

DMEM/F12 medium supplemented with 1X serum-free B27, 10 ng/mL recombinant human basic FGF (bFGF), 10 ng/mL recombinant human EGF, 100 U/mL penicillin and 100 μg/mL streptomycin, and 4 mmol/L GlutaMAX. LAPC4 was maintained in IMDM medium supplemented with 10% FBS, 100 U/mL penicillin and 100 μg/mL streptomycin, and 4 mmol/L GlutaMAX. MDA PCa 2b was maintained in F-12K medium supplemented with 20% FBS, 25 ng/ml cholera toxin, 10 ng/ml mouse epidermal growth factor, 0.005 mM phosphoethanolamine, 100 pg/ml hydrocortisone, 1X insulin-transferrin-selenium (ITS-G), 100 U/mL penicillin and 100 μg/mL streptomycin, and 4 mmol/L GlutaMAX.

Lentiviral constructs and lentivirus production

Lentiviral vectors expressing shRNA targeting RB1, dominant-negative TP53, BCL2, and MYCN have been previously described [69, 70]. Lentiviral vectors expressing SRRM4, NR0B2, KRAS, ASCL1 and NeuroD1 were generated by PCR amplification of the open reading frames and subcloning into the EcoRI site of the FU-CGW lentiviral backbone by NEBuilder HiFi DNA Assembly (New England Biolabs). Lentiviruses were prepared and titered as previously described [69].

Immunoblotting

Whole cell extracts were collected in 9 M urea lysis buffer and quantified using the Pierce Rapid Gold BCA Protein Assay Kit (Thermo Fisher Scientific). Protein samples were fractionated by SDS-PAGE using Bolt Bis-Tris Plus gels and transferred to a nitrocellulose membrane using Bolt

Transfer Buffer according to the manufacturer's instructions (Invitrogen). Membranes were blocked with 5% non-fat milk in PBST (DPBS + 0.5% Tween 20) for 30 minutes while shaking at room temperature, then incubated with primary antibodies at 4°C overnight. Membranes were washed three times for 5 minutes with PBST and incubated with horseradish peroxidase (HRP) conjugated anti-mouse or anti-rabbit secondary antibodies for 1 hour at room temperature. Blots were washed three times for 5 minutes each with PBST and developed with Immobilon Western Chemiluminescent HRP Substrate (EMD Millipore) for 3 minutes at room temperature. Blot images were acquired using a ChemiDoc Imaging System (Bio-Rad).

Immunohistochemistry (IHC) and immunofluorescence (IF) studies

Cells and xenografts were fixed in 10% buffered formalin for 12 hours and then embedded in HistoGel (Thermo Fisher Scientific) and paraffin, sectioned, and mounted on glass slides (Thermo Fisher Scientific). IHC was performed as previously described [132]. In brief, slides containing formalin-fixed paraffin-embedded sections were deparaffinized in xylene and rehydrated in 100%, 95%, and 70% ethanol. Slides were then heated in antigen retrieval buffer (0.2 mol/L citric acid and 0.2 mol/L sodium citrate) within a pressure cooker followed by PBS wash. Slides were blocked with 2.5% horse serum for 30 minutes and then incubated with primary antibody diluted in 2.5% horse serum overnight at 4°C. HRP was detected with ImmPRESS-HRP anti-mouse or anti-rabbit IgG peroxidase detection kits (Vector Laboratories) and staining was visualized with DAB peroxidase substrate (Dako). Slides were counterstained with hematoxylin and dehydrated for mounting.

Immunofluorescence studies using specific antibodies were carried out on archival formalin fixed paraffin embedded tissues. In brief, 5 μm paraffin sections were de-waxed and rehydrated following standard protocols. Antigen retrieval consisted of steaming for 40 minutes in Target Retrieval Solution (Agilent). Slides were then washed and equilibrated in TBS-Tween buffer (Sigma) for 10 minutes. Primary antibodies were applied at at 37□C for 60 minutes. Sequential dual-immunofluorescence labeling was carried out using Tyramide SuperBoost kits (Thermo Fisher).

Neuroendocrine transdifferentiation assay

Cells were seeded in 6-well tissue culture plates at a density of 3 x 105 cells per mL in 3 mL of RPMI medium supplemented with 10% FBS, 100 U/mL penicillin and 100 μg/mL streptomycin, and 4 mmol/L GlutaMAX. Cells were transduced approximately 4-6 hours after seeding at a defined multiplicity of infection (MOI) of 4 for each lentivirus. 72 hours after transduction, cells were trypsinized, washed, and transferred to 100 mm tissue culture plates in 15 mL of neural stem cell media (N-SCM) consisting of Advanced DMEM/F12 medium supplemented with 1X serum-free B27, 10 ng/mL recombinant human bFGF, 10 ng/mL recombinant human EGF, 100 U/mL penicillin and 100 μg/mL streptomycin, and 4 mmol/L GlutaMAX. Media were replenished every 3-4 days. Cells were collected 14 days post-transduction for analysis.

Mouse studies

All animal care and studies were performed in accordance with an approved Fred Hutchinson Cancer Center Institutional Animal Care and Use Committee protocol and Comparative Medicine regulations. Eight-week-old male NSG (NOD-SCID-IL2Rγ-null, RRID:BCBC_4142) mice were obtained from the Jackson Laboratory. A total of 1 x 10^6 cells from each C4-2B reprogramming condition were suspended in 100 μL of cold Matrigel (Corning) and implanted by injection subcutaneously into NSG mice.

Cell proliferation assay

For cell proliferation assays, the CellTiter-Glo® 2.0 Cell Viability Assay system (Promega) was used according to the manufacturer's instructions. 5×10^3 cells were plated into each well in a 96-well and 25 μL per well of CellTiter-Glo® 2.0 reagent was added. After a 30 minute incubation at room temperature, luminescence was measured every 24 hours using a BioTek Synergy H1 Multimode Reader (Agilent). Six replicate wells per time point were used to obtain measures of cell proliferation.

RNA-seq

Total RNA was extracted with a Purelink™ RNA mini kit (Invitrogen). cDNA libraries were prepared from isolated RNA using the Illumina Stranded mRNA prep kit. High-throughput sequencing with 50 bp paired-end reads was performed using an Illumina NovaSeq 6000 SP. Sequencing reads were mapped to human genome reference hg38 and gene expression was quantified and normalized using the Toil RNA-seq pipeline (v4.1.2). Sequence adapters were

trimmed using CutAdapt (v1.9) [168]. Trimmed sequences were then aligned to human reference genome GRCh38 using STAR (v2.4.2a) [169] and gene expression quantification was performed using RSEM v1.2.25 [170]. RSEM expected counts were subsequently normalized by upper-quartile (UQ) normalization for downstream analysis.

The enrichment scores were calculated in R using the GSVA package using a previously published set of 8 AR-regulated genes for AR-activity scores and the 14 genes in the NEURO I and NEURO II gene sets for NEURO scores[124]. Differential gene expression analysis was performed using the R package DEseq2 (v1.36.0) RNA-seq data from Beltran et al., 2016 was processed through the Toil RNA-seq pipeline (v4.1.2). Gene expression data was projected onto the Beltran et al., 2016 PCA space to examine reprogramming towards the NEPC phenotype. Gene signature scores were calculated based on RNA-seq expression data using the R package GSVA [171]. The neuroendocrine score was previously described based on CRPC patient data [62]. Briefly, 70 genes were used to construct the score based on differential expression or methylation between CRPC-NE and CRPC-Ad or relevance to CRPC-NE based on the literature. The mean expression of the 70 genes was calculated to form the NE score. The scores were visualized in heatmap format using ComplexHeatmap (v2.12.0) [172].

Differential RNA alternative splicing analysis

Differential RNA alternative splicing (AS) between vector transduction conditions was calculated from bulk RNA-sequencing using the rMATS (version 4.1.1)[173]. Gene and transcript annotations were sourced from Ensembl GRCH38 release 104[174]. The program was run under

the statoff mode to obtain the percent spliced in (PSI) matrix for all C42B samples. AS events that were defined by both junction counts and exon counts were used for downstream analysis. The PSI values of all five AS events: skipped exon (SE), alternative 5' splice site (A5SS), alternative 3' splice site (A3SS), mutually exclusive exons (MXE) and retained intron (RI) were aggregated together to form a large matrix for PCA. CRPC-Adeno and CRPC-NE Beltran et al 2016 patient RNAseq data was processed using the same rMATS workflow for AS calling. C42B was projected onto Betlran et al 2016 PSI PCA space for examining the remodeling of AS by oncogene transduction towards the NE direction.

CUT&RUN

Cells were harvested by centrifugation at 1200 rpm for 3 minutes in a swinging bucket rotor and washed twice in ice cold wash buffer (20 mM HEPES pH 7.5; 150 mM NaCl; 0.5 mM Spermidine; 1 tablet of Roche complete EDTA-free protease inhibitor). Concanavalin A-coated beads (Bangs Laboratories) were activated by resuspension and washed in binding buffer (20 mM HEPES-KOH at pH 7.9; 10 mM KCl, 1 mM CaCl2; 1 mM MnCl2). Cells were mixed with activated beads at a ratio of 9:1 at room temperature for 10 minutes. The cell and bead mixtures were resuspended in antibody solution (2 mM EDTA; 0.05% digitonin in wash buffer). Corresponding antibodies were added at desired concentrations to 150 µl of cell and bead mixture and rotated overnight at 4oC. Libraries were prepared using the AutoCUT&RUN protocol (https://www.protocols.io/view/autocut-run-genome-wide-profiling-of-chromatin-pro-

6qpvre6zblmk/v1). High-throughput sequencing with 50 bp paired-end reads was performed on CUT&RUN libraries using an Illumina NextSeq.

Single-cell multiome RNA-seq and ATAC-seq

Cells were collected 14 days post transduction and processed per manufacturer's instructions using the 10x Chromium Next GEM Single Cell Multiome ATAC + Gene Expression platform to generate libraries. Single-cell RNA-seq libraries and single-cell ATAC-seq libraries were sequenced on an Illumina NovaSeq 6000 targeting 35,000 read pairs/nuclei for RNA-seq and 40,000 read pairs/nucleus for ATAC-seq.

Single-cell RNA-seq and ATAC-seq analysis

Single-cell multiome ATAC- and RNA-sequencing reads were processed through the CellRanger ARC analysis pipeline. RNA-sequencing filtered feature count matrices and ATAC-sequencing fragment paths from each sample were loaded individually loaded into R with Signac (version 1.7.0)[175]. The data was filtered for total RNA count between 1000 and 50000, total feature RNA count between 500 and 10000, percent mitochondria RNA less than 20%, total ATAC fragment count between 1000 and 50000, nucleosome signal score less than 2, and a TSS enrichment score above 1.5. ATACseq peaks were called using MACS2 [176]. Peaks on nonstandard chromosomes and in genomic blacklist regions of hg38 were removed. ATAC counts were quantified for all peaks across all samples.

Single-cell trajectory analysis

Filtered Seurat object was loaded into Monocle3 (version 1.2.9) to calculate Single-cell trajectory [177, 178]. First 50 PCA dimensions were used in the preprocessing step. Cells were aligned by regressing on percent mitochondria RNA.

CUT&RUN data processing and analyses

Peak calls for H3K4me1, H3K4me3 and H3K27ac from CUT&RUN data at different timepoints were performed using SEACR (v1.3) [179]. Bedgraph files were prepared using bedtools (v2.27.1) to serve as the inputs for SEACR. IgG control conditions at each timepoint were used to define empirical thresholds for peak calling. The "norm" parameter was used in SEACR to normalize the data. Peaks were called in "stringent" mode. Active promoters were defined by checking if H3K4me3 peaks overlap with H3K27ac signals using the bedtools intersect -wa option. The resulting regions within +/- 5 kb of transcription start sites (TSS) were retained. TSS regions were calculated by applying GTFtools (v0.8.5) [180] to the gencode v39 (GRCh38.p13) GTF file. Active enhancers were obtained by overlapping H3K4me1 peaks with H3K27ac peaks, followed by subtracting the +/- 5 kb TSS regions using the bedtools subtract –A option. Super-enhancers were called using ROSE (v0.1) [153, 181] on H3K27ac CUT&RUN data. STITCHING_DISTANCE was set to 12.5 kb. TSS_EXCLUSION_ZONE_SIZE was set to 2.5 kb. hg38 was used as the reference genome. The resulting genomic regions of active promoters, active enhancers and super-enhancers were loaded into DiffBind [182] for differential binding analysis. Counts that mapped to the genomic regions were normalized using the Relative Log Expression

(RLE) method. For super-enhancers, the counts of member regions that belong to each parental broad genomic region were aggregated together to form the overall count. The correlations between time and genomic regions were calculated using Kendall test followed by Benjamini-Hochberg P-value adjustment. Genomic regions were mapped to genes using ChIPseeker (v1.32.0) [183]. Genomic regions and genes demonstrating significant changes along with time were visualized in heatmap format using ComplexHeatmaps (v2.12.0)[172] and analyzed by Gene Ontology using hypergeometric test with MSigDB gene sets [184].

AR binding site analyses

Inducible, constitutively active, and inactive AR binding sites (ARBS) were obtained from a previously published study [154]. A random set of genomic regions with the same length of ARBS (700 bp) was generated by randomly selecting 50 start sites for each chromosome. Fragments of scATAC-seq data were quantified against these sites. The significance of the difference between conditions was calculated based on the log transformed counts using Welch's t-test.

AR enhancer analyses

The genomic coordinates of the somatically acquired AR enhancer were obtained from a previously published study [43]. scATAC-seq data were examined at the same overall genomic region to determine the location of the AR enhancer in our model system and in the prostate cancer hepatic metastasis. The correlation between scATAC chromatin accessibility and scRNA gene

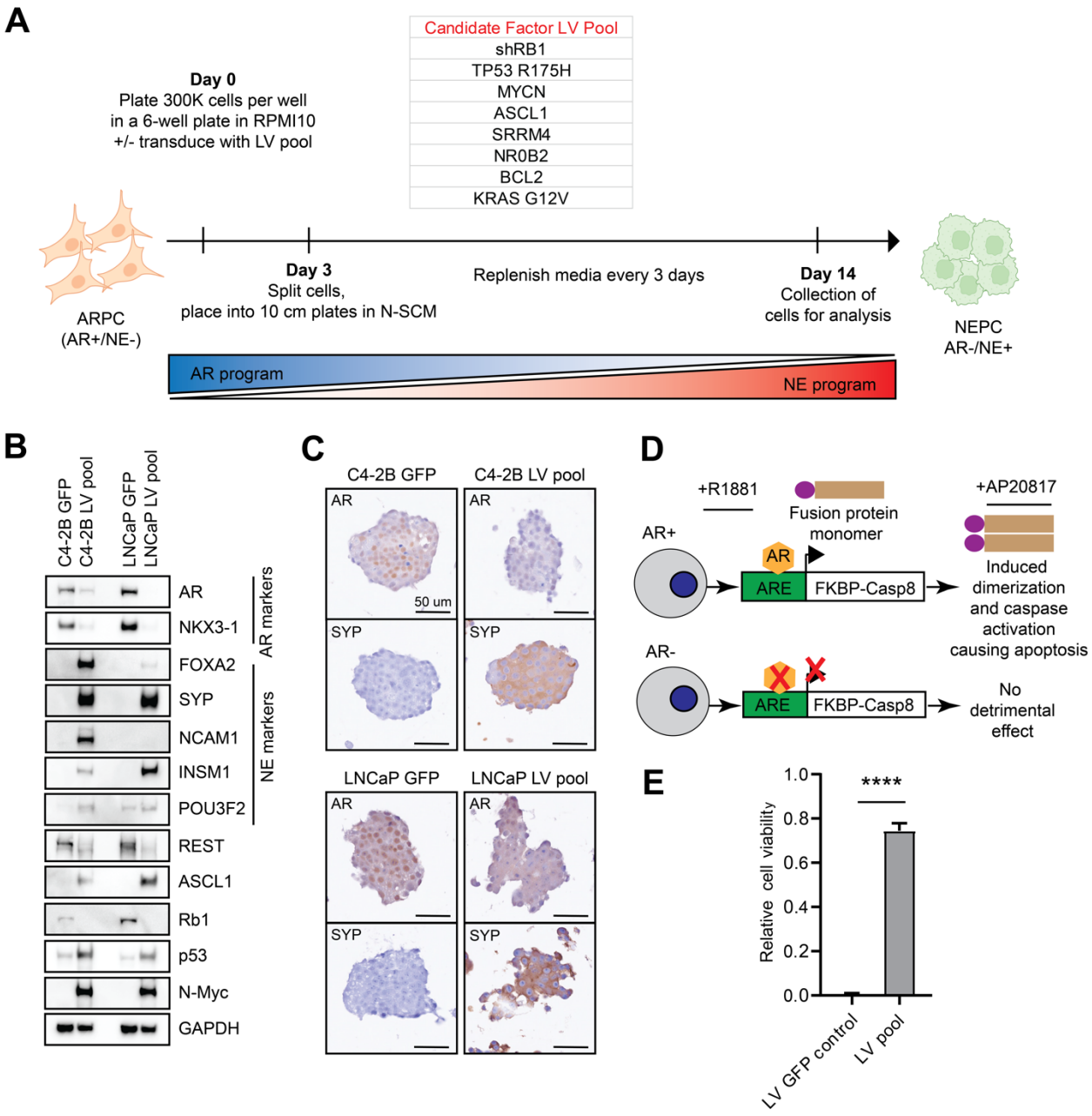
expression in reprogrammed cells was visualized using CoveragePlot in the R package Signac (version 1.7.0) [175].

Statistical analyses

Data analysis was performed on GraphPad Prism 9 (GraphPad Software, Inc.). Quantitative PCR results were analyzed in Excel. Statistical significance was determined using the unpaired two-tailed Student t test, unpaired two-tailed Welch t test where the variances are shown to be different via F-test, one-way ANOVA, or two-way ANOVA. Only two-tailed tests were used. Results are depicted as mean + SD unless stated otherwise. All P values of <0.05, <0.01, <0.001, and <0.0001 were considered significant. Pearson correlation coefficient was used to determine correlation between genes. The symbols used to represent the P values were: ns, nonsignificant for $P > 0.05$; *, $P \leqslant 0.05$; **, $P \leqslant 0.01$; ***, $P \leqslant 0.001$; ****, $P \leqslant 0.0001$. The test used in each statistical analysis is specified in the figure legends.

3.5 FIGURES

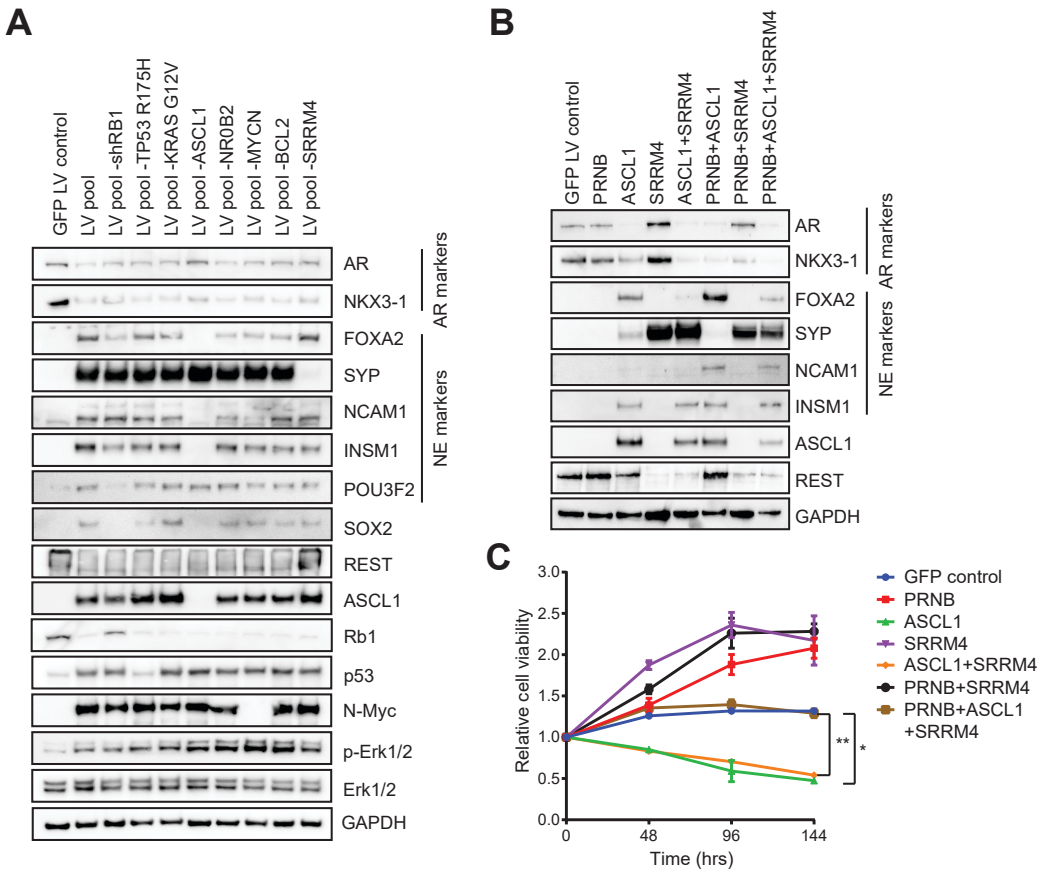
Figure 1



3.5.1. Figure 1. **Direct reprogramming of ARPC to NEPC with functional bypass of a dependence on AR signaling.**

(**A**) Experimental schema of the reprogramming of ARPC cells to NEPC through the introduction of a candidate factor lentiviral (LV) pool and propagation in neural stem cell media (N-SCM) permissive to NEPC growth. (**B**) Immunoblot analysis of the C4-2B and LNCaP cell lines transduced with either green fluorescent protein (GFP) or the LV pool and subjected to the reprogramming assay. Lysates were collected on day 14. (**C**) Immunocytochemical analysis of the C4-2B and LNCaP lines at day 14. (**D**) Experimental schema for the stringent negative selection of AR^+ cell populations based on the AR-dependent expression of an inducible FKBP-Casp8 fusion protein that dimerizes and induces caspase activation and apoptosis in the presence of the FK506 analog AP20817. (**E**) Relative cell viability determined by CellTiter-Glo assay of treated versus untreated groups after five days (n=4 each). P-value was assessed by Student's t-test. **** denotes $p < 0.0001$.

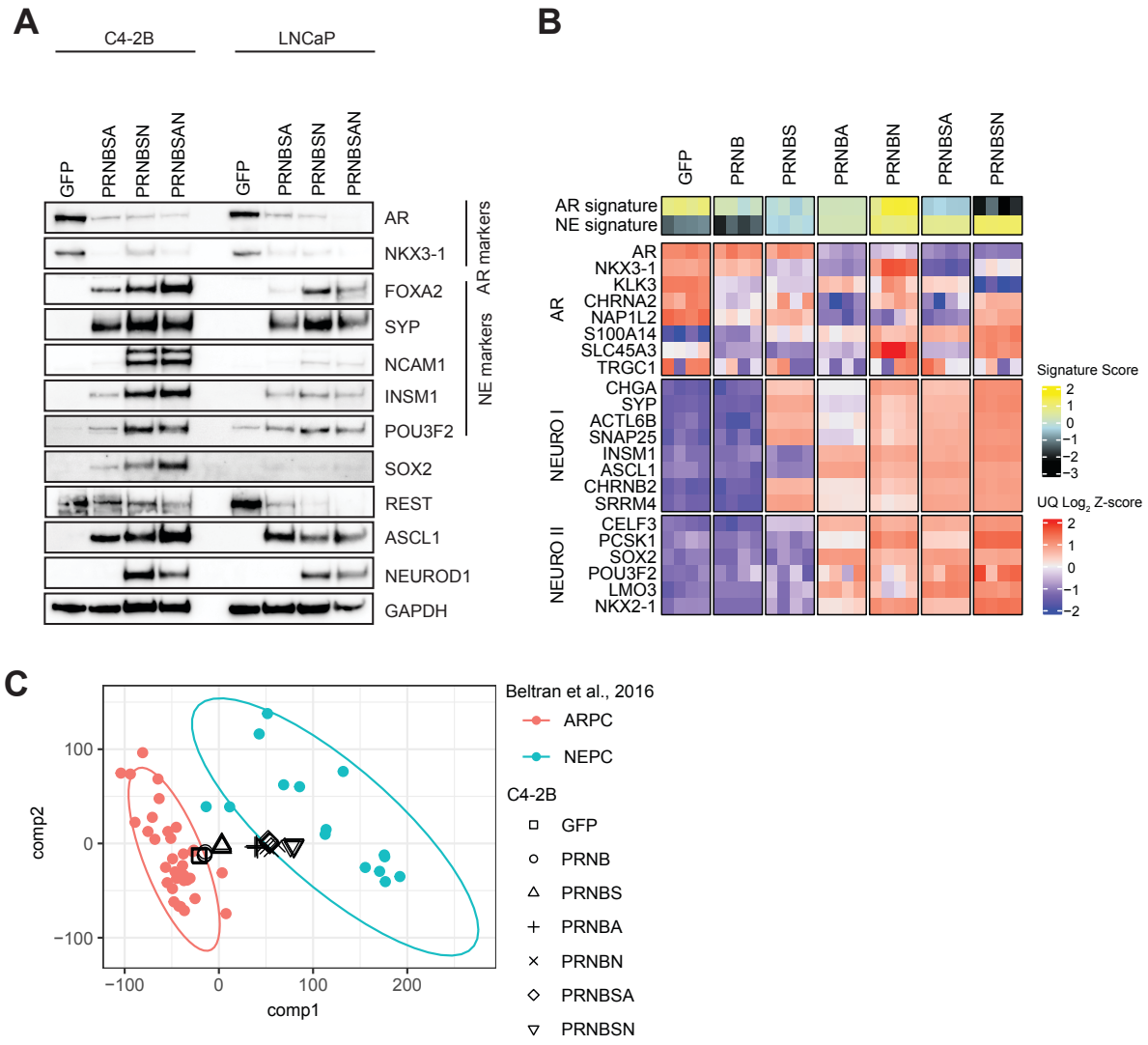
Figure 2



3.5.2. Figure 2. **The pioneer neural transcription factor ASCL1 suppresses AR expression and drives NE transdifferentiation in prostate cancer.**

Immunoblot analysis of (**A**) leave-one-out conditions and (**B**) factor reconstitution conditions in reprogramming studies using the C4-2B cell line. PRNB represents dominant-negative TP53 H175R, shRB1, MYCN, and BCL2. Lysates were collected on day 14. (**C**) Relative cell viability over time determined by CellTiter-Glo assay of C4-2B cells reprogrammed with various factor combinations (n=4 per condition). * denotes $p < 0.05$, ** denotes $p < 0.01$.

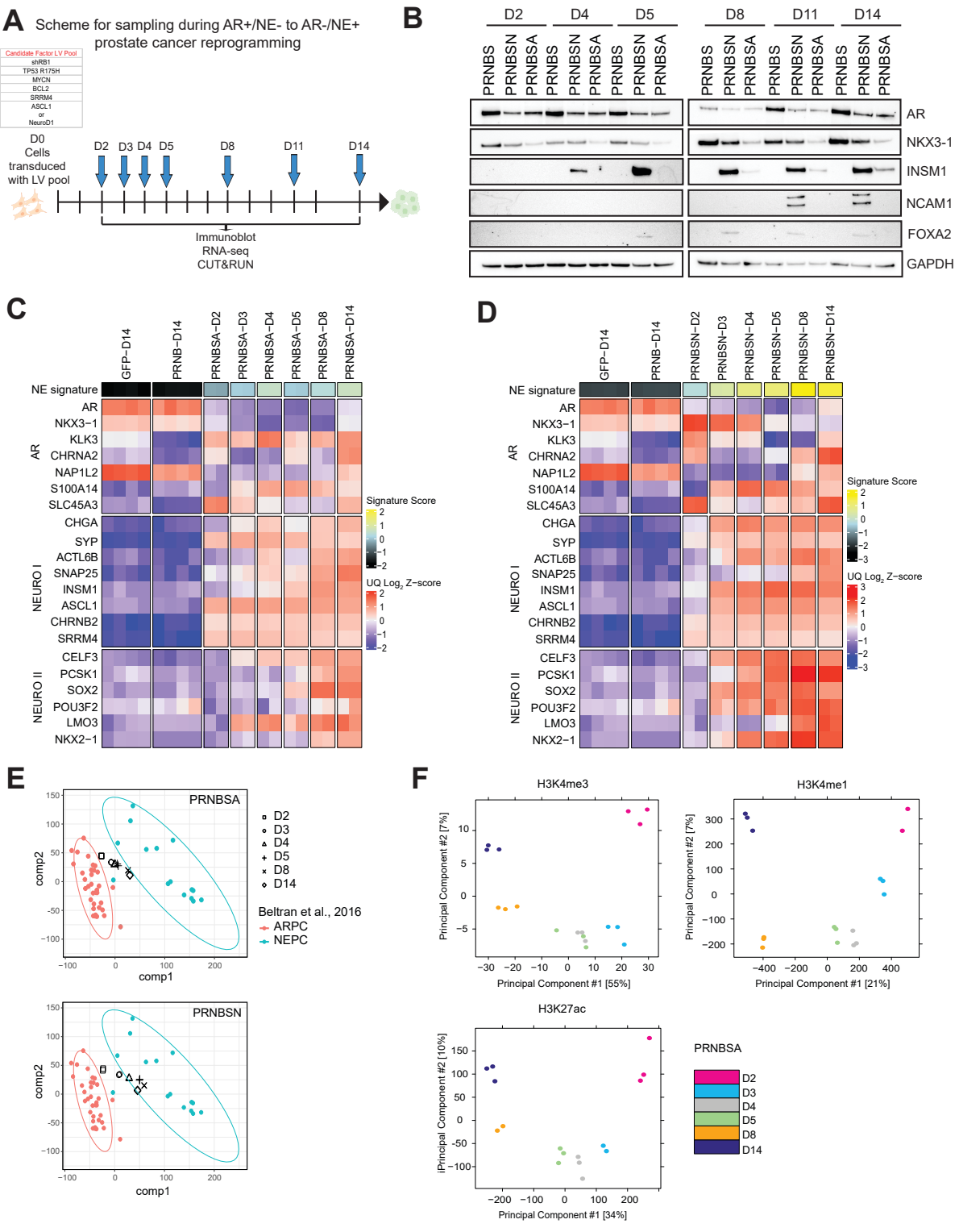
Figure 3



3.5.3. Figure 3. **ASCL1 and NeuroD1 are competent to induce NE lineage reprogramming of prostate cancer.**

(A) Immunoblot analysis of C4-2B and LNCaP cells reprogrammed with GFP or PRNB, SRRM4, and ASCL1 and/or NeuroD1. Lysates were collected on day 14. (**B**) Heatmap of RNA-seq gene expression from reprogrammed C4-2B cell line conditions showing 22-gene AR and NE signature scores and select genes associated with the AR program and NE programs (NEURO I and NEURO II). UQ: upper quartile normalization. (**C**) Partial least squares-discriminant analysis (PLS-DA) plot based on RNA-seq gene expression of reprogrammed C4-2B cell line conditions (black shapes) projected onto human ARPC (red dots) and NEPC (blue dots) samples from Beltran et al., 2016. Ellipses represent 95% confidence level for multivariate t-distributions defined by ARPC (red) and NEPC (blue) data.

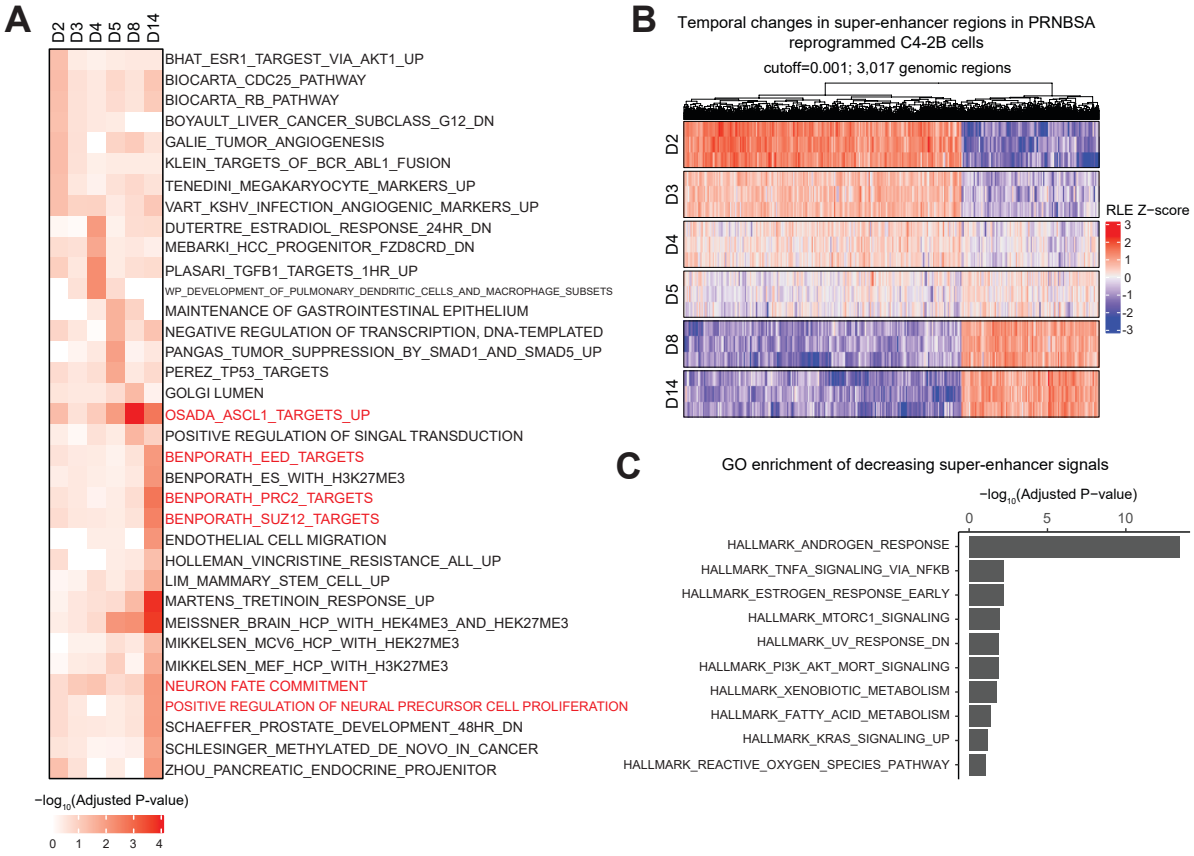
Figure 4



3.5.4. Figure 4. **Dynamic changes in cancer phenotype and transcriptional and epigenetic landscapes during acute reprogramming of ARPC to NEPC.**

(**A**) Experimental schema for the temporal investigation of the reprogramming of ARPC to NEPC by immunoblot, RNA-seq, and CUT&RUN analyses. (**B**) Immunoblot analysis of C4-2B cells modified with PRNB and SRRM4 (PRNBS), PRNBS and NeuroD1 (PRNBSN), or PRNBS and ASCL1 (PRNBSA) that were collected at various timepoints during the 14-day reprogramming period. Heatmaps of RNA-seq gene expression from C4-2B cells reprogrammed with (**C**) PRNBSA or (**D**) PRNBSN over the 14-day reprogramming period showing 22-gene NE signature scores and select genes associated with the AR program and NE programs (NEURO I and NEURO II). (**E**) PLS-DA plot based on RNA-seq gene expression of C4-2B cells reprogrammed with PRNBSA (top) or PRNBSN (bottom) at different timepoints (black shapes) projected onto human ARPC (red dots) and NEPC (blue dots) samples from Beltran et al., 2016. Ellipses represent 95% confidence of t-distribution defined by ARPC (red) and NEPC (blue) data. (**F**) Principal component analysis (PCA) of CUT&RUN enrichment for H3K4me3, H3K4me1 and H3K27Ac signal in C4-2B cells reprogrammed with PRNBSA over time.

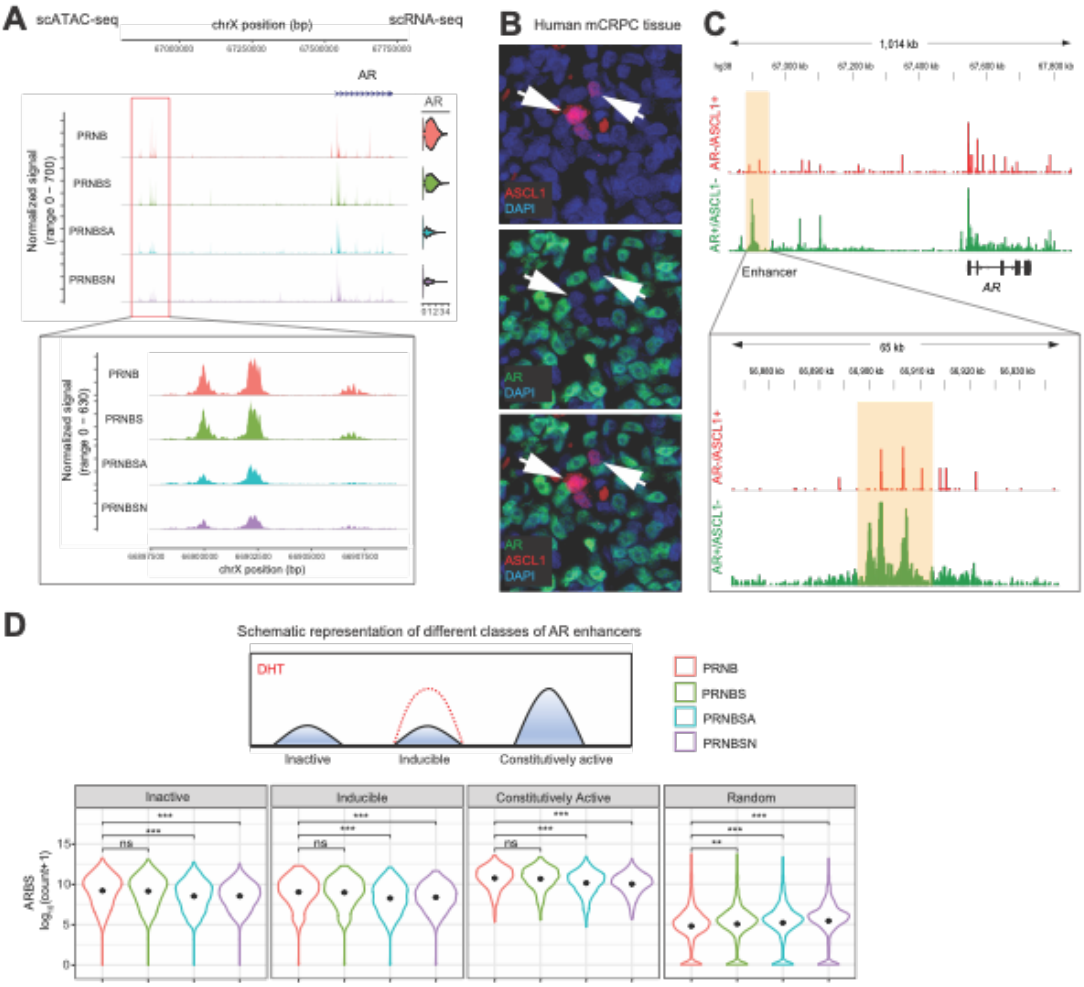
Figure 5



3.5.5. Figure 5. **Time-dependent changes in ASCL1-regulated gene programs and super-enhancer organization during NE transdifferentiation of prostate cancer.**

(A) Heatmap of MSigDB gene sets enriched over time with ASCL1-regulated genes identified by integrating CUT&RUN and RNA-seq data from C4-2B cells reprogrammed with PRNBSA over the 14-day reprogramming period. (**B**) Heatmap showing dynamic changes in super-enhancer regions in C4-2B cells reprogrammed with PRNBSA over the 14-day reprogramming period. Genomic regions shown were obtained by setting a p-value cutoff of 1e-03 on Kendall correlation between day and peak activity. RLE: Relative Log Expression normalization. (**C**) Plot showing Gene Ontology (GO) enrichment of Hallmark gene sets associated with decreasing super enhancer signals from D2 to D14 in C4-2B cells reprogrammed with PRNBSA.

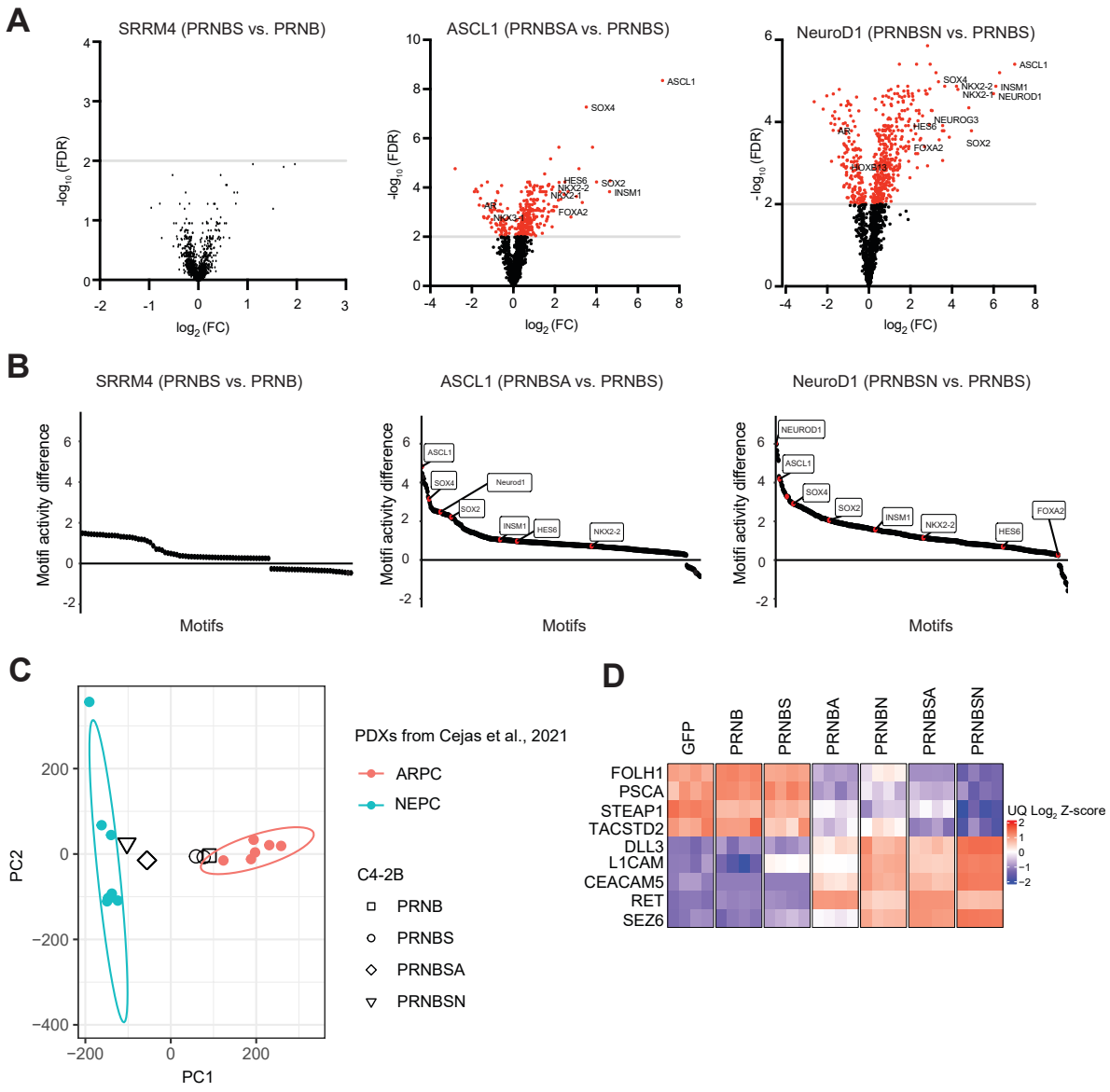
Figure 6



3.5.6. Figure 6. **ASCL1/NeuroD1 inhibit AR expression by remodeling chromatin accessibility at the somatically acquired AR enhancer and global AR binding sites with enhancer activity.**

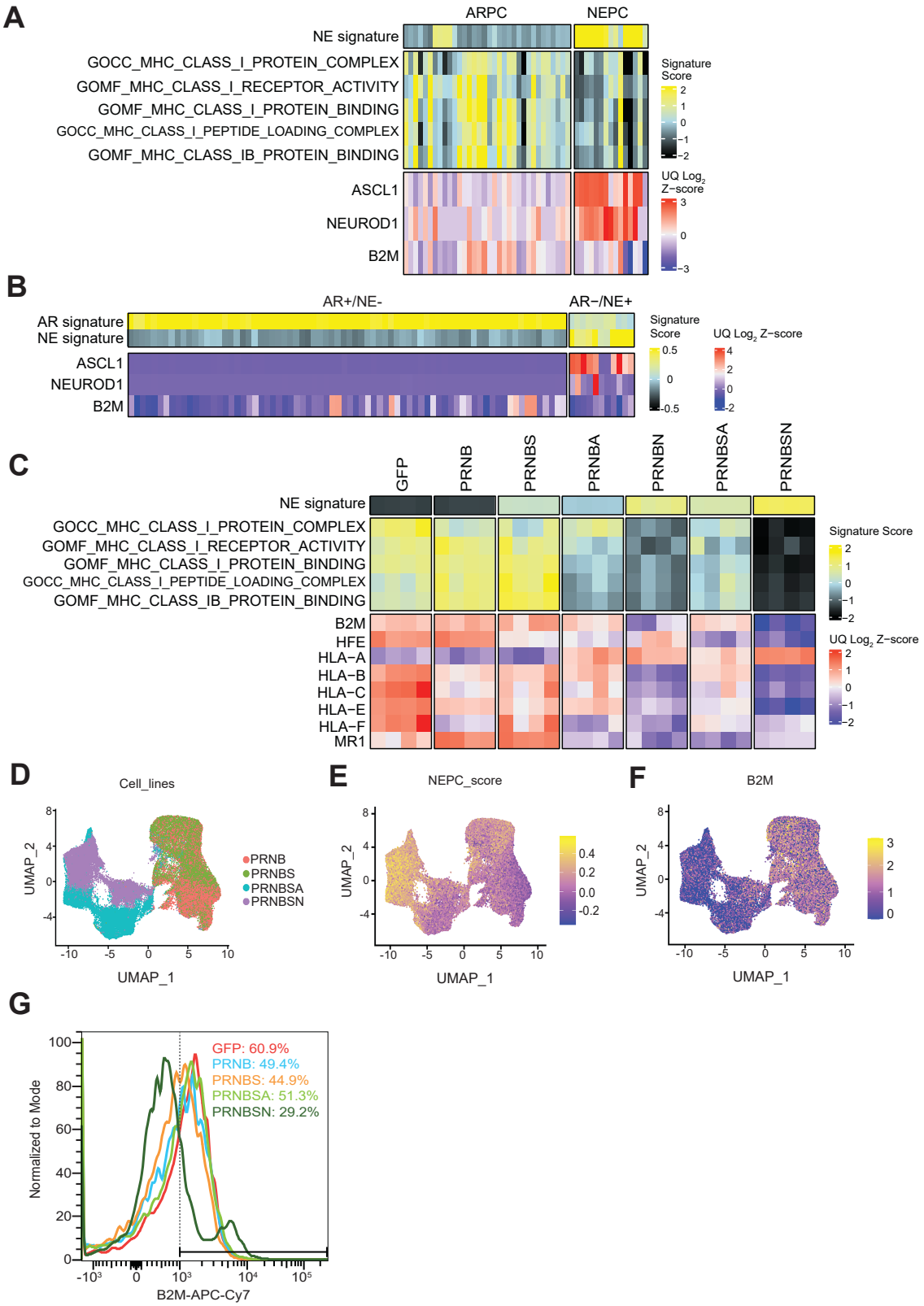
(**A**) Tracks showing chromatin accessibility peaks at the AR locus and AR gene expression (violin plot) from pseudo-bulk analysis of scATAC-seq and scRNA-seq data from C4-2B cells reprogrammed with PRNB, PRNBS, PRNBSA, and PRNBSN. Tracks in the region of the somatically acquired AR enhancer are magnified. (**B**) Immunofluorescence analysis of a human prostate cancer hepatic metastasis with ASCL1 staining (top), AR staining (middle), and an overlay of ASCL1 and AR staining (bottom) shown. DAPI was used as a nuclear counterstain. (**C**) Tracks showing chromatin accessibility at the AR locus by pseudo-bulk analysis of scATAC-seq data from the $AR^{-}/ASCL1^{+}$ and $AR^{+}/ASCL1^{-}$ tumor cell populations present in the human prostate cancer hepatic metastasis shown in **B**. Tracks in the region of the somatically acquired AR enhancer are magnified. (**D**) Schematic representation of different classes of AR enhancers (top) is shown. The distribution of chromatin accessibility at AR bindings sites (ARBS) of different enhancer activity and at random genomic regions of the same interval are shown. Black dots represent the means. ns denotes $p > 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

Figure 7



3.5.7. Figure 7. **ASCL1/NeuroD1 induce the expression and activity of NE-associated transcription factors.**

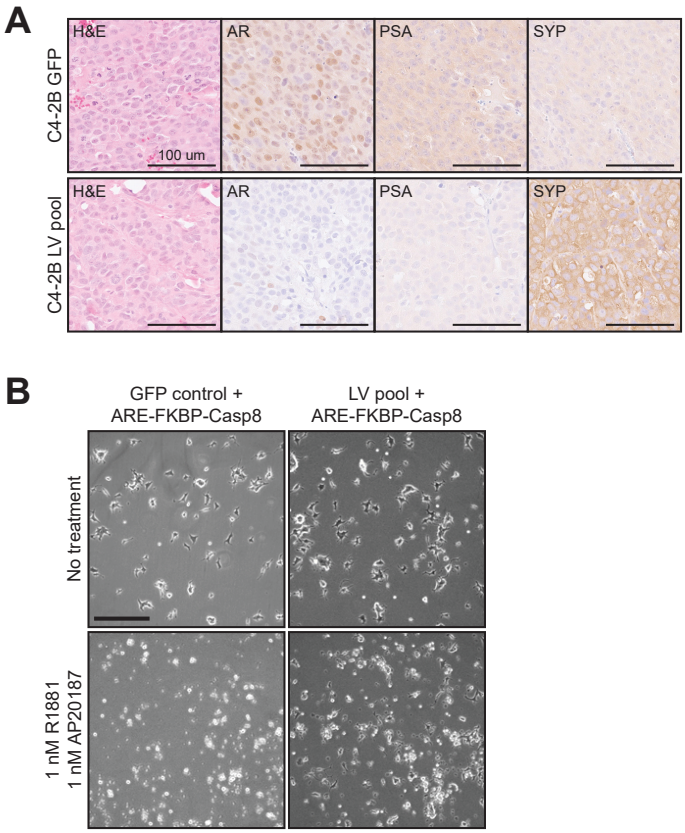
(**A**) Volcano plots showing the differential expression of genes encoding transcriptions factors in pairwise comparisons of RNA-seq gene expression data of C4-2B cells reprogrammed with PRNBS vs. PRNB, PRNBSA vs. PRNBS, and PRNBSN vs. PRNBS. Red dots represent genes with a $-\log_{10}(\text{FDR}) > 2$. FC represents fold-change. (**B**) Plots of differential transcription factor motif activity from footprinting analysis of pseudo-bulk scATAC-seq data from C4-2B cells reprogrammed with PRNBS vs. PRNB, PRNBSA vs. PRNBS, and PRNBSN vs. PRNBS. Note that labelled motifs in the second and third panel do not show up in the first panel. (**C**) PCA of chromatin accessibility signals from pseudo-bulk scATAC-seq of reprogrammed C4-2B cell line conditions projected on previously published ATAC-seq data from NEPC and ARPC PDX models. (**D**) Heatmap of RNA-seq gene expression data showing cell surface targets relevant to ARPC and NEPC in reprogrammed C4-2B cell conditions.



3.5.8. Figure 8. **Downregulation of MHC class I antigen processing and presentation genes in NEPC is functionally attributable to ASCL1/NeuroD1.**

(**A**) Heatmap of RNA-seq gene expression data showing the NE signature and MHC class I pathway scores (top) and select genes (bottom) in ARPC and NEPC from Beltran et al., 2016. (**B**) Heatmap of RNA-seq gene expression data showing the AR and NE signature scores and select genes in AR^{+}/NE^{-} and AR^{-}/NE^{+} metastatic prostate cancers from the University of Washington Tissue Acquisition Necropsy (UW TAN) program. (**C**) Heatmap of RNA-seq gene expression data from reprogrammed C4-2B cell line conditions showing the NE signature and MHC class I pathway scores (top) and select MHC class I genes including B2M (bottom). UQ: upper quartile normalization. (**D**) UMAP analysis of scRNA-seq and scATAC-seq data from C4-2B cells reprogrammed with PRNB, PRNBS, PRNBSA or PRNBSN using the Seurat package. UMAP plots of reprogrammed cells colored based on (**E**) NEPC score and (**F**) B2M gene expression. (**G**) Flow cytometry histogram plots showing the percentage of B2M positive cells in reprogrammed C4-2B cells.

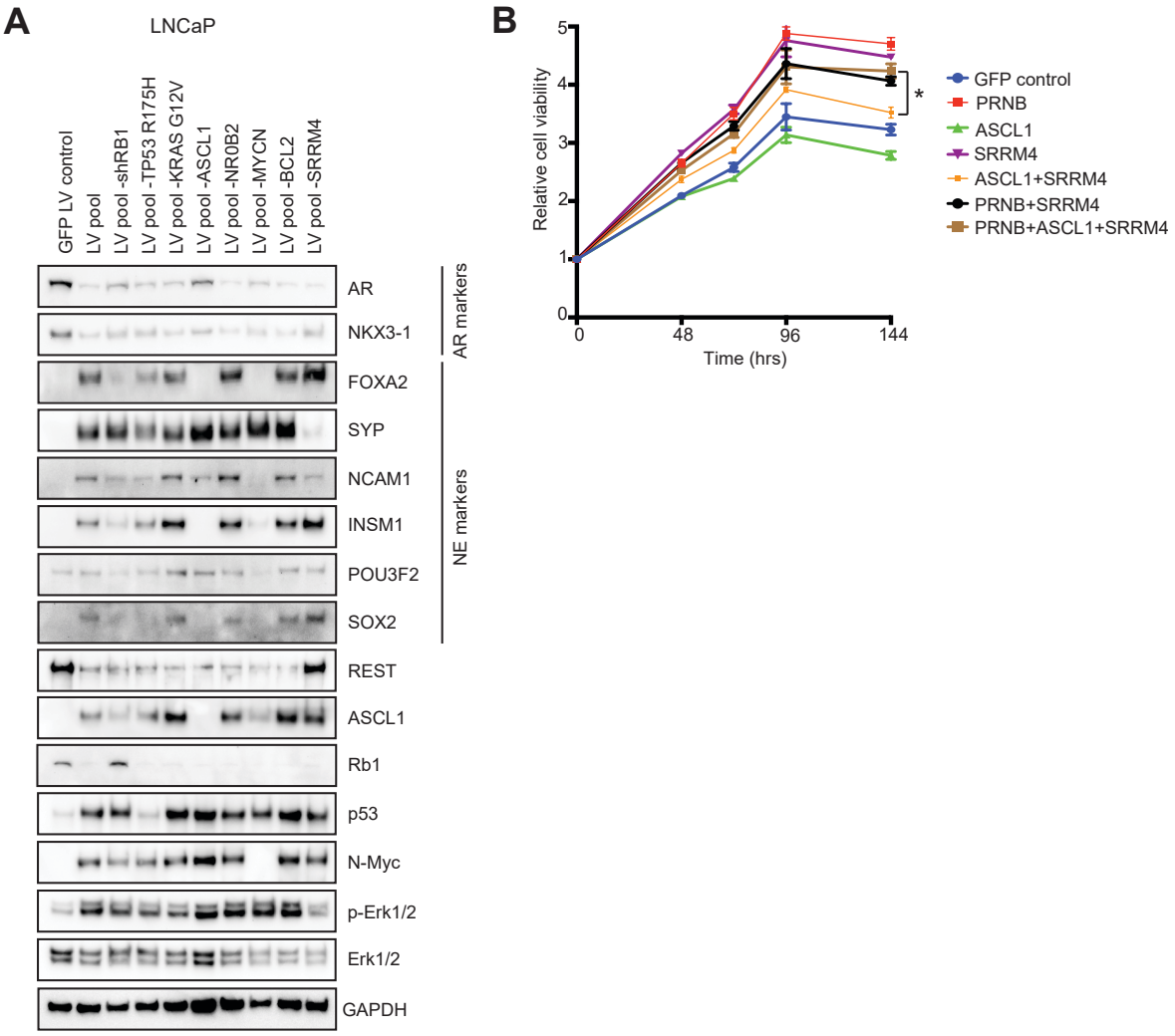
Figure S1



3.5.9. Figure S1. **Sustained reprogramming of ARPC to NEPC and functional bypass of AR signaling.**

(**A**) Photomicrographs of H&E- and IHC-stained tissue sections of C4-2B cell line xenograft tumors established subcutaneously in non-castrate adult male NGS mice. (**B**) Photomicrographs of clonal C4-2B ARE-FKBP-Casp8 cells transduced with either GFP control or the candidate factor LV pool and subjected to no treatment or treatment with R1881 and AP20187 for five days. Scale bar = 100 μm.

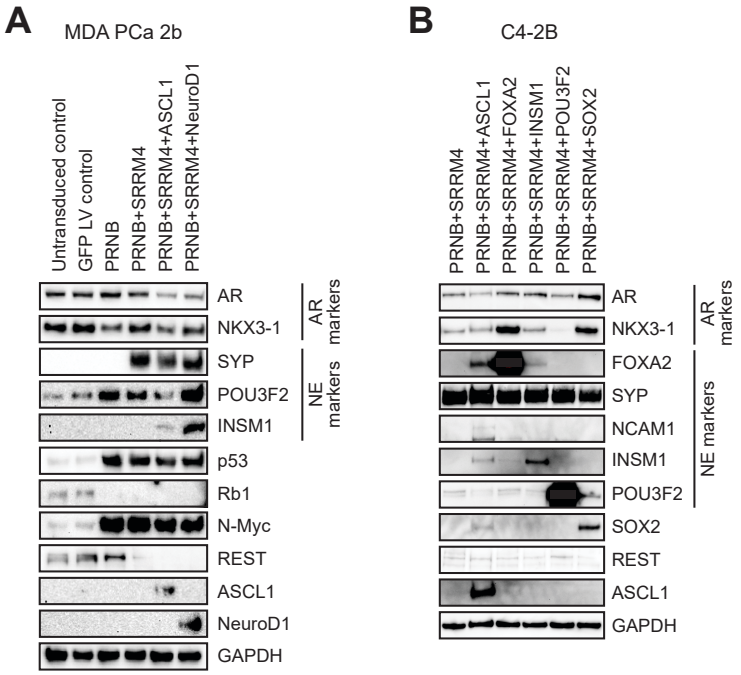
Figure S2



3.5.10. Figure S2. **ASCL1 drives NEtD and suppresses AR expression in LNCaP cells.**

(A) Immunoblot analysis of leave-one-out conditions in reprogramming studies using the C4-2B cell line. Lysates were collected on day 14. **(B)** Relative cell viability over time determined by CellTiter-Glo assay of LNCaP cells reprogrammed with various factor combinations. * denotes $p < 0.05$.

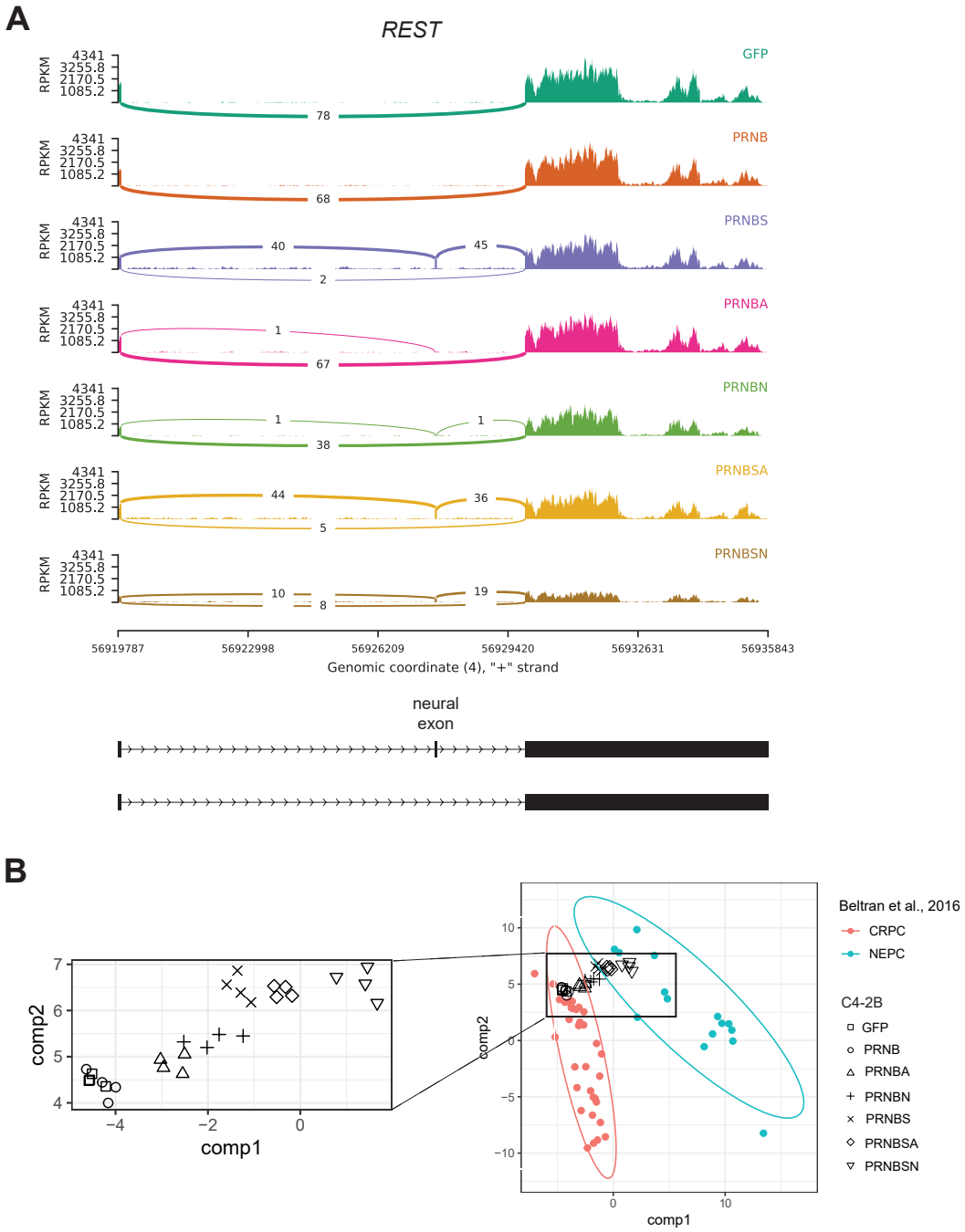
Figure S3



3.5.11. Figure S3. **NEtD of MDA PCa 2b cells and evaluation of the competence of NE-associated transcription factors in driving NE reprogramming.**

(**A**) Immunoblot analysis of MDA PCa 2b cell line conditions subjected to the NE transdifferentiation assay. (**B**) Immunoblot analysis of reprogrammed C4-2B cell line conditions in which NE-associated transcription factors were added to PRNB (dominant-negative TP53 H175R, shRB1, MYCN, and BCL2) and SRRM4 to evaluate their effects on markers of the AR and NE programs.

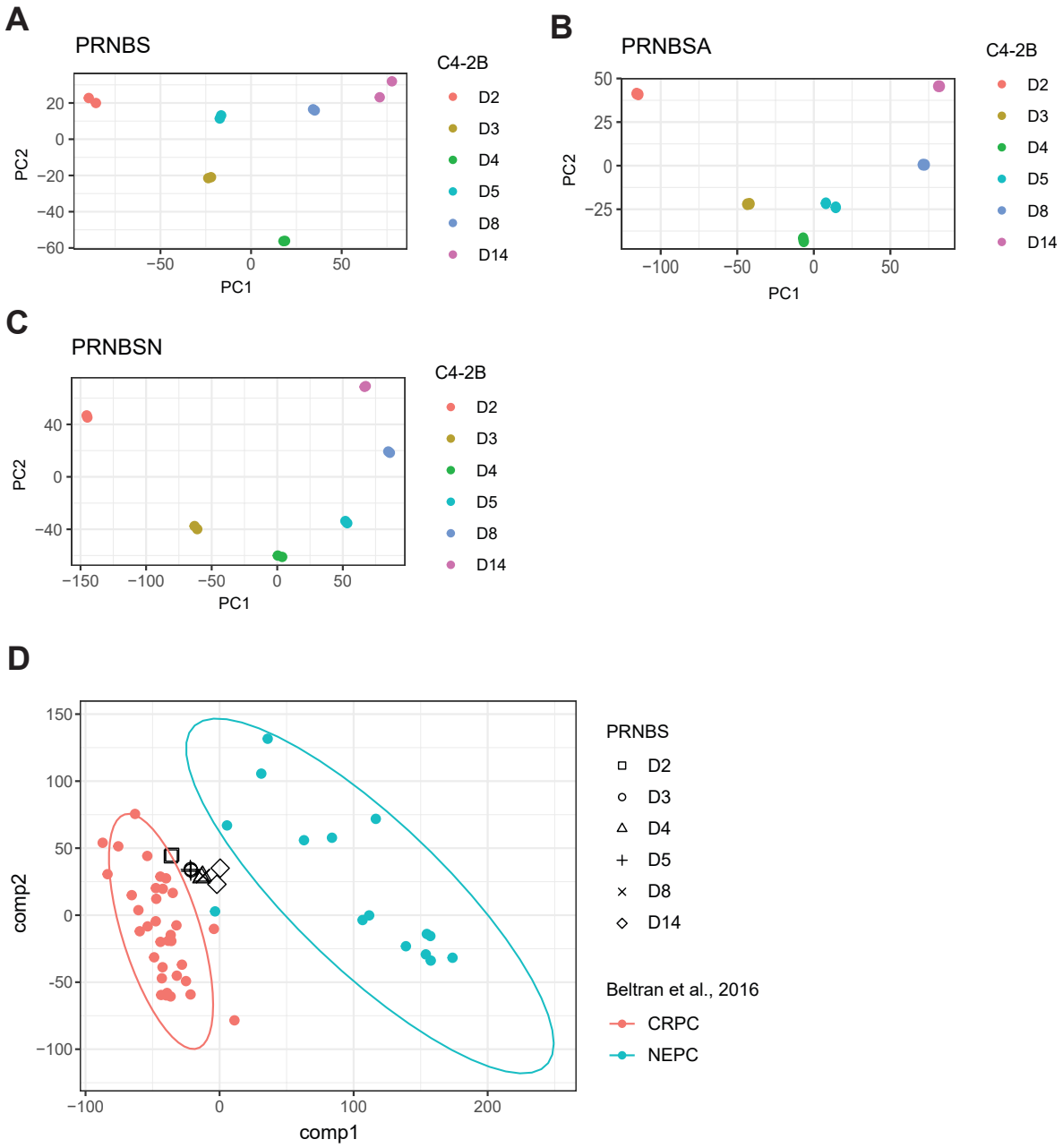
Figure S4



3.5.12. Figure S4. **SRRM4 expression leads to alternative splicing of REST in C4-2B reprogrammed cells.**

(**A**) Sashimi plots of *REST* RNA-seq reads from C4-2B cells modified with PRNB and SRRM4 (PRNBS), PRNBS and NeuroD1 (PRNBSN), or PRNBS and ASCL1 (PRNBSA). Conditions with SRRM4 demonstrate inclusion of a neural exon leading to the REST-4 isoform. (**B**) PLS-DA plot based on alternative splicing analysis of C4-2B cells line conditions projected onto human ARPC (red dots) and NEPC (blue dots) samples from Beltran et al., 2016. Ellipses represent 95% confidence of t-distribution defined by ARPC (red) and NEPC (blue) data.

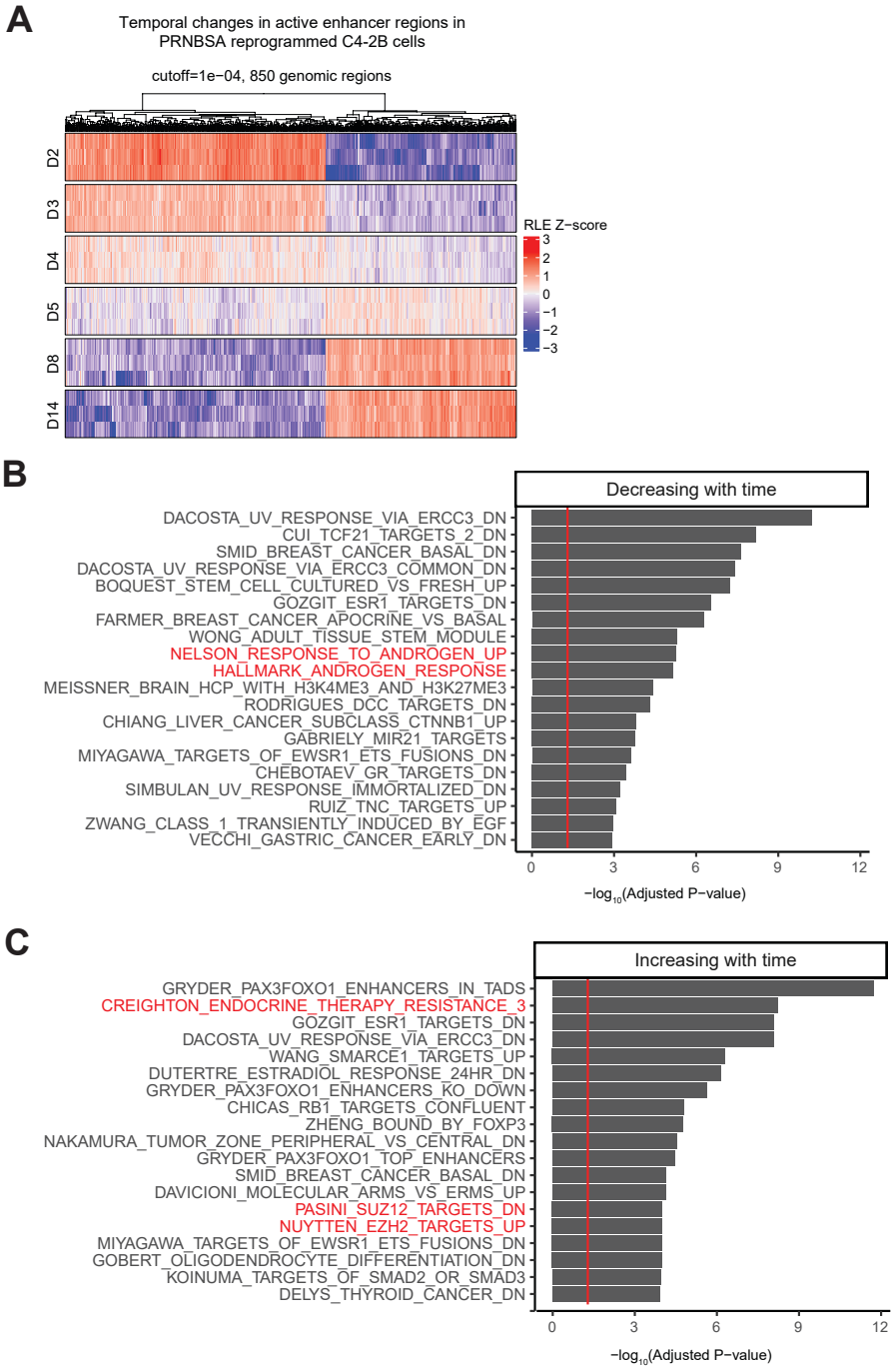
Figure S5



3.5.13. Figure S5. **ASCL1 or NeuroD1 but not SRRM4 in combination with PRNB is sufficient to reprogram ARPC to NEPC.**

PCA analyses of bulk RNA-seq data from C4-2B cells reprogrammed with (**A**) PRNBS, (**B**) PRNBSA, and (**C**) PRNBSN over time. (**D**) PLS-DA plot based on RNA-seq gene expression of C4-2B cells reprogrammed with PRNBS over time (black shapes) projected onto human ARPC (red dots) and NEPC (blue dots) samples from Beltran et al., 2016. Ellipses represent 90% confidence for ARPC (red) and NEPC (blue).

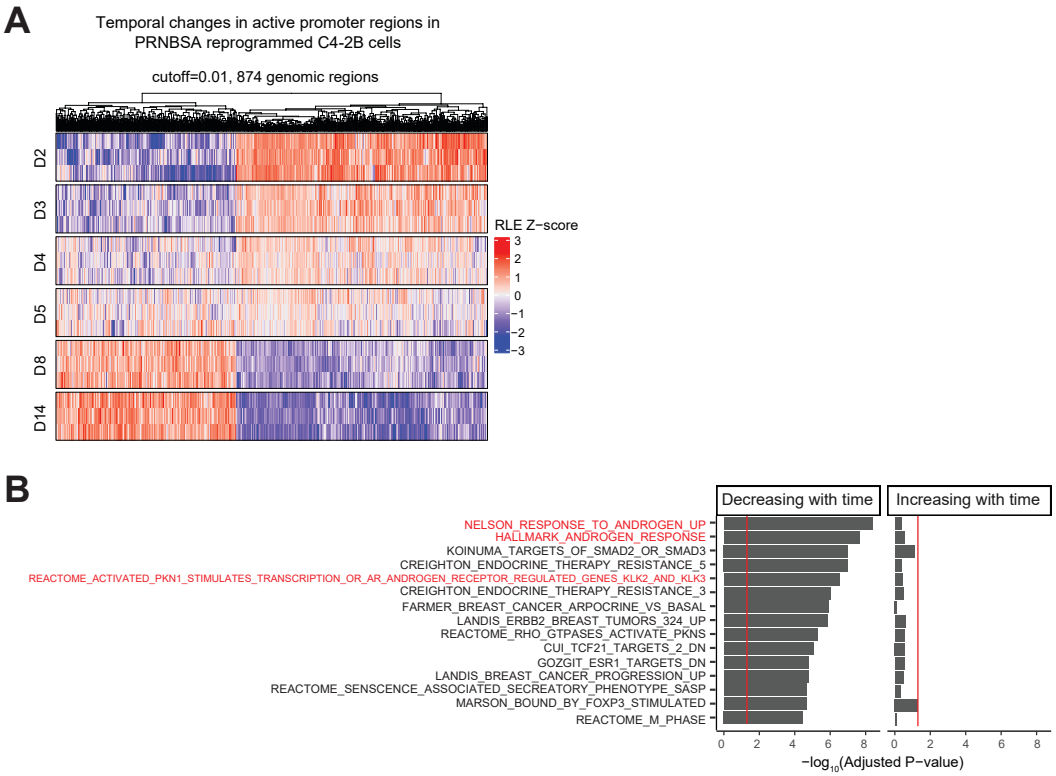
Figure S6



3.5.14. Figure S6. **Temporal shift in global active enhancers in PRNBSA reprogrammed C4-2B cells.**

(**A**) Heatmap showing dynamic changes in active enhancer regions defined by overlapping H3K4me1 and H3K27ac signals in C4-2B cells reprogrammed with PRNBSA over the 14-day reprogramming period. Plots showing Gene Ontology (GO) enrichments of MSigDB gene sets associated with (**B**) decreasing and (**C**) increasing active enhancer regions from D2 to D14 in C4-2B cells reprogrammed with PRNBSA. RLE: Relative Log Expression normalization. Genomics regions in (A) were obtained by setting a p-value cutoff of 1e-04 on Kendall correlation between day and peak activity. Red lines in (B) and (C) represent adjusted p-value = 0.05.

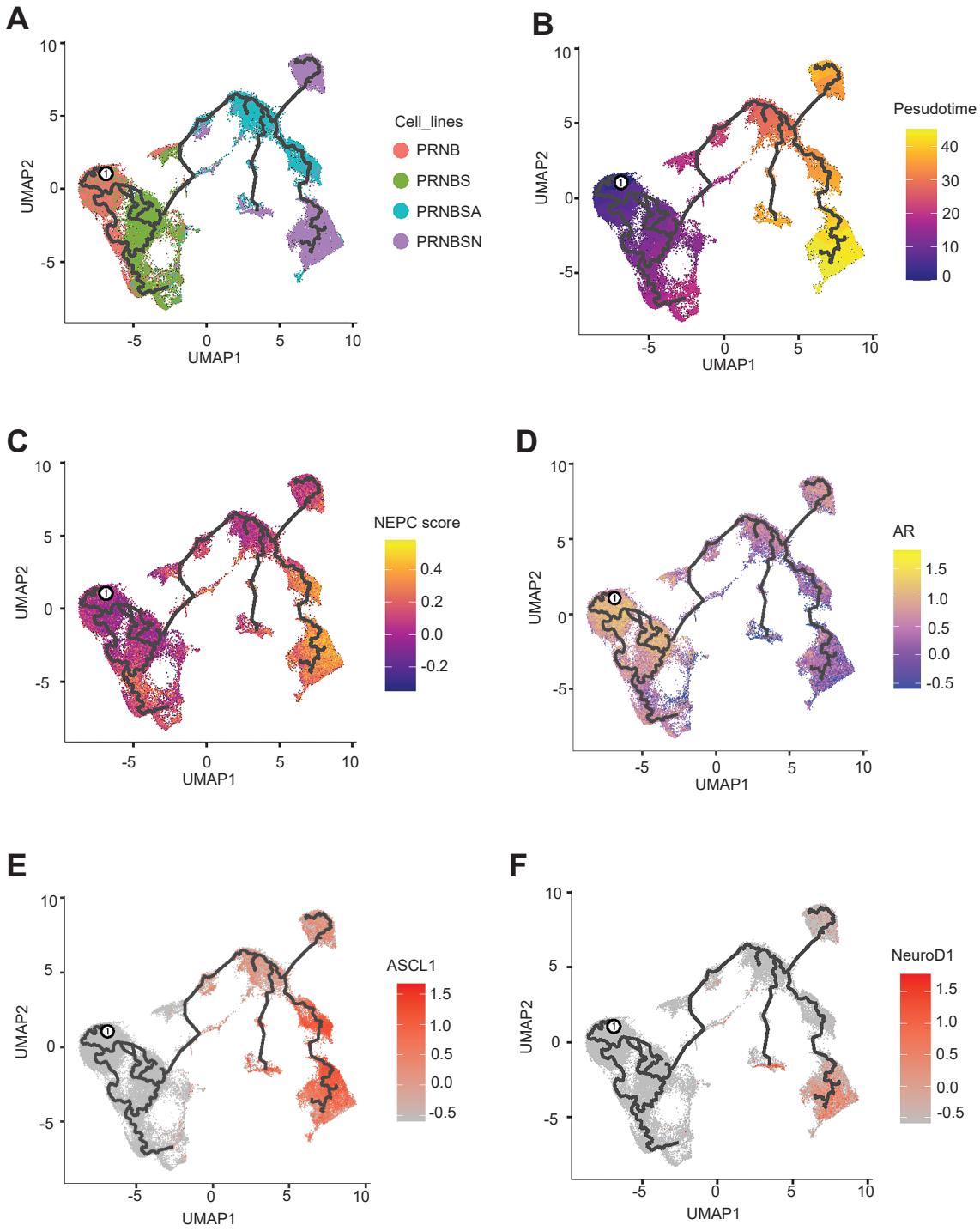
Figure S7



3.5.15. Figure S7. **Temporal shift in global active promoters in PRNBSA reprogrammed C4-2B cells.**

(**A**) Heatmap showing significant changes in active promoter regions defined by overlapping H3K4me3 and H3K27ac signals in C4-2B cells reprogrammed with PRNBSA over the 14-day reprogramming period. RLE: Relative Log Expression normalization. Genomics regions were obtained by setting a p-value cutoff of 1e-02 on Kendall correlation between day and peak activity. (**B**) Plots showing GO enrichment of MSigDB gene sets associated with decreasing and increasing active promoter regions from D2 to D14 in in C4-2B cells reprogrammed with PRNBSA. Red lines represent adjusted p-value = 0.05.

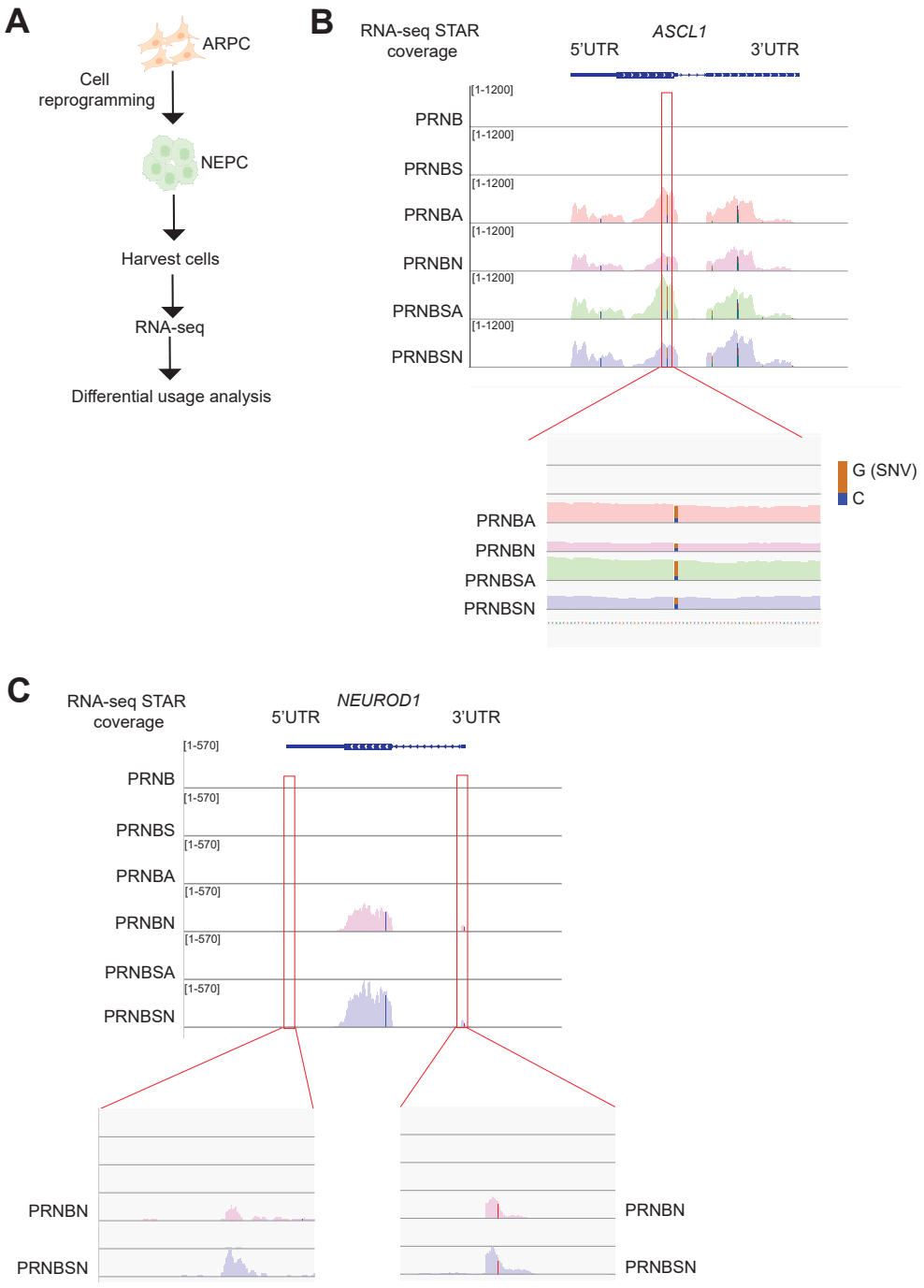
Figure S8



3.5.16. Figure S8. **scRNA-seq profiles distinct cell states induced by NE reprogramming of C4-2B cells.**

Monocle pseudotime trajectory plots of C4-2B cells reprogramming conditions. Cell populations were colored by (**A**) conditions, (**B**) pseudotime, (**C**), NE score, (**D**) *AR* expression, (**E**) *ASCL1* expression, and (**F**) *NEUROD1* expression.

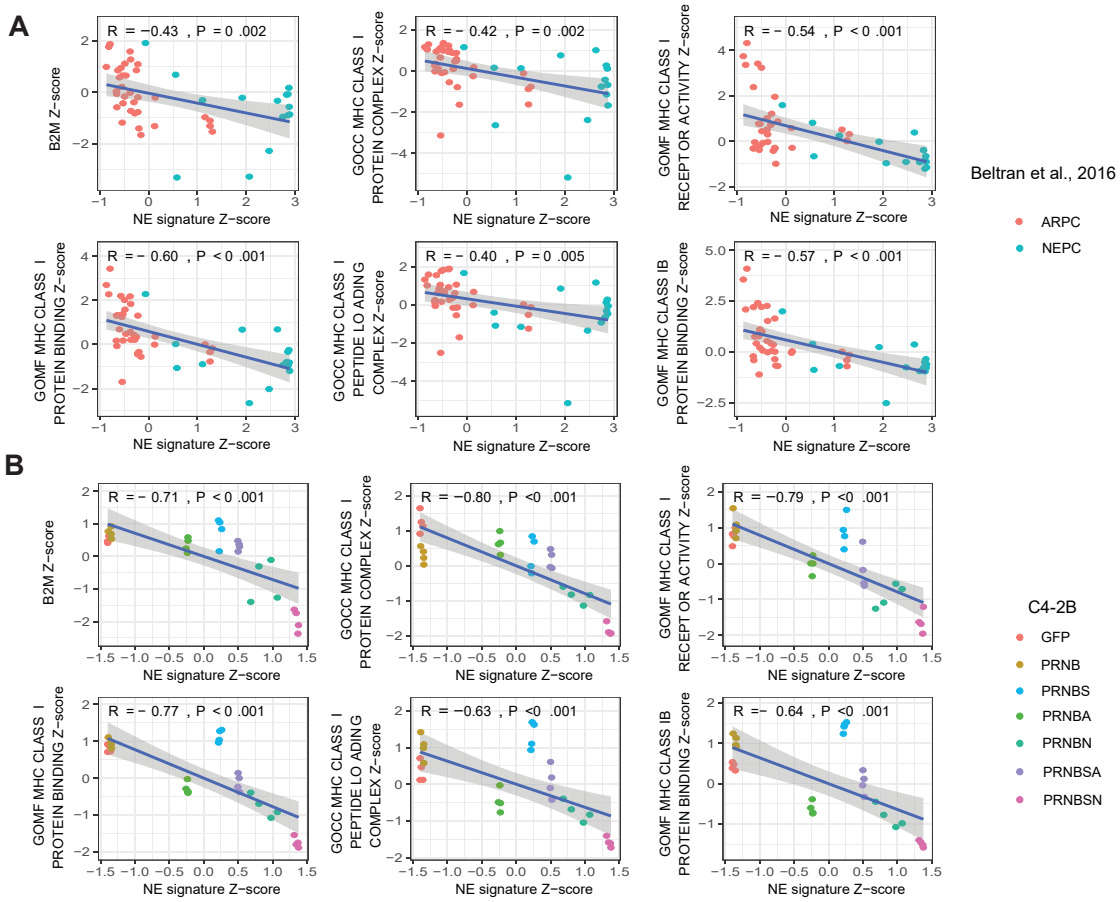
Figure S9



3.5.17. Figure S9. **ASCL1 and NeuroD1 engage in self-regulation after NE reprogramming of PC.**

(**A**) Experimental schema to evaluate exogenous and endogenous ASCL1 or NeuroD1 gene expression using differential usage analysis. (**B**) Integrated Genomics Viewer (IGV) tracks of RNA-seq reads from reprogrammed C4-2B cell line conditions mapping to *ASCL1* or the untranslated regions (UTRs) of *ASCL1*. The relative distribution of reads with the synonymous single-nucleotide variant NM_004316.4(ASCL1);c.627C>G encoded in the ASCL1 lentivirus is shown in the magnified image. (**C**) IGV tracks of RNA-seq reads from reprogrammed C4-2B cell line conditions mapping to *NEUROD1* or *NEUROD1* UTRs. The magnified tracks below show reads mapping to the *NEUROD1* UTRs associated with conditions where exogenous NeuroD1 was introduced.

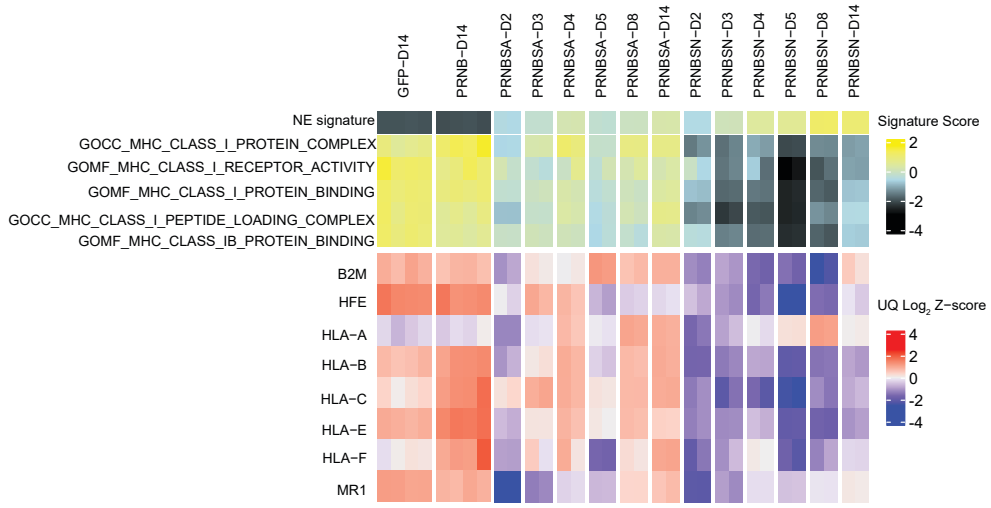
Figure S10



3.5.18. Figure S10. **NE signature scores are associated with reduced MHC I antigen processing and presentation signature scores in PC.**

(**A**) Correlations of gene expression and various MHC signature scores to NE signature scores in the ARPC and NEPC samples from Beltran et al., 2016. (**B**) Correlations of gene expression and various MHC signature scores to NE signature scores in the reprogrammed C4-2B cell lines. Correlation coefficients (R) and P-values (P) were derived from Pearson correlations.

Figure S11



3.5.19. Figure S11. **Downregulation of MHC class I antigen processing and presentation genes starts early during reprogramming of ARPC to NEPC.**

Heatmap of RNA-seq gene expression data from C4-2B cells reprogrammed with PRNBSA or PRNBSN over the 14-day reprogramming period showing the NE signature and MHC class I pathway scores (top) and select MHC class I genes including *B2M* (bottom). UQ: upper quartile normalization.

3.6 TABLES

Table 1. **List of antibodies used for Western Blotting (WB), Immunohistochemistry (IHC), Immunocytochemistry (ICC), Immunofluorescence (IF), and Cleavage Under Targets and Release Using Nuclease (CUT&RUN).**

Antibodies	Application	Source	Catalog number	Dilution	Species
AR	WB, ICC,IHC	Santa Cruz	sc-816	1:1000	mouse
AR	IF	Cell Signaling	5153T	1:100	rabbit
ASCL1 (24B72D11.1)	WB	BD Pharmingen	556604	1:1000	mouse
ASCL1 (24B72D11.1)	IF	BD Pharmingen	556604	1:25	mouse
N-MYC	WB	Santa Cruz	sc-53993	1:1000	mouse
RB1	WB	Santa Cruz	sc-74562	1:500	mouse
P53 (D0-1)	WB	Santa Cruz	sc-126	1:500	mouse
BCL2 (124)	WB	Cell Signaling	15071S	1:1000	mouse
NeuroD1 (A-10)	WB	Santa Cruz	sc-46684	1:500	mouse
NKX3-1	WB	Cell Signaling	83700S	1:2000	rabbit
INSM1	WB	Santa Cruz	sc-271408	1:500	mouse
NCAM	WB	Santa Cruz	sc-374289	1:500	mouse
FOXA2	WB	Santa Cruz	sc-377033	1:500	mouse
SYP	WB, ICC, IHC	Santa Cruz	sc-17750	1:500	mouse
PSA	IHC	Santa Cruz	sc-7306	1:500	mouse
POU3F2	WB	Cell Signaling	12137	1:1000	rabbit
REST	WB	Proteintech	22242-1-AP	1:250	mouse
SOX2 (E-4)	WB	Santa Cruz	sc-365823	1:1000	mouse
Erk1/2	WB	Cell Signaling	4695	1:1000	rabbit
Phospho Erk1/2	WB	Cell Signaling	9101	1:1000	rabbit
GAPDH-HRP	WB	GeneTex	GTX627408-01	1:10,000	mouse
Goat anti-rabbit-HRP Conjugate	WB	BioRad	1706515	1:10,000	mouse
Goat anti-mouse-HRP conjugate	WB	BioRad	1706516	1:10,000	mouse
ASCL1	CUT&RUN	Abcam	ab74065	1:50	rabbit
H3K4me1	CUT&RUN	Abcam	ab8895	1:100	rabbit
H3K4me3 (C42D8)	CUT&RUN	Cell Signaling	9751	1:100	rabbit
H3K27ac	CUT&RUN	EMD Millipore	MABE647	1:50	rabbit
Rabbit IgG	CUT&RUN	EpiCypher	13-0042	1:10	rabbit

Chapter 4. POU2F3 IS A DRIVER OF A VARIANT MOLECULAR SUBTYPE CASTRATION-RESISTANT PROSTATE CANCER

4.1 INTRODUCTION

With poor prognosis, overall survival, and very limited therapeutic options, castration-resistant prostate cancer (CRPC) is a highly aggressive and lethal form of prostate cancer. The acquired castration resistance is attributed to reactivation of androgen receptor (AR) signaling via genomic mutations, AR amplification and alternative splicing variants of AR [37]. Nevertheless, an estimated 15% to 20% of CRPC tumors is shown to display loss of AR expression and/or downstream signaling [185, 186]. Loss of the AR lineage program in response to stringent therapeutic pressure leads to enrichment for adaptive genetic and epigenetic events causing re-programming to alternate prostate cancer cell states that are resistant to AR pathway inhibition. A recent study revealed distinct cell populations with differential expression of Pou2f3 in the loss of Pten, Rb1 and Nmyc+ mouse model with AR-negative and poorly differentiated prostate tumors [162]. They also identified POU2F3 expression in clinical CRPC samples, indicating the clinical relevance of POU2F3 during the progression to advanced AR-negative prostate cancer. Another single-cell based study showed Pou2f3+ subpopulation emergence during the progression from adenocarcinoma to neuroendocrine in genetically engineered mouse models of prostate cancer.

POU2F3 is a POU homeodomain transcription factor, which is required for the generation of chemosensory tuft cells in multiple digestive tracts tissues [187]. POU2F3 has not been associated with cancer progression until Huang et al found POU2F3 is expressed in the subset of small-cell lung cancer (SCLC) tumors that lack ASCL1, NEUROD1, and other neuroendocrine lineage markers and overexpress chemosensory tuft cells markers [188]. Mechanism of how POU2F3 is involved in the progression of advanced prostate cancer is unknown.

In this study, we showed cell dependency of POU2F3 in POU2F3 expressing small-cell prostate cancer and neuroendocrine prostate cancer cell line. Functional study also revealed POU2F3 is not responsible for inducing and maintaining NE reprogramming of advanced prostate cancer, it is important for lineage transition which involves downregulating AR downstream signaling in CRPC.

4.2 RESULTS AND DISCUSSION

POU2F3 is a lineage oncogene in castration-resistant prostate cancer

The currently established models to study POU2F3 function in prostate cancer is lacking. There is no information about POU2F3 expression in the established prostate cancer cell lines. We screened a series of existing NEPC cell lines and publicly available RNA-seq data and found two models have high POU2F3 expression at transcription level, which are studies involved in transdifferentiation human basal prostate cells into NEPC [69, 70]. LASCPC-01 was a NEPC line generated by overexpression of N-Myc and myrAKT1 in human primary prostate basal epithelial cells [69], and PARCB lines were small cell prostate cancer lines established from combination of

expression of dominant negative p53 (TP53DN) (P), myrAKT1 (A), RB1 (R), c-Myc (C), and BCL2 (B) in human primary prostate basal epithelial cells from different donors [70]. RNA-seq heatmaps (Figure 1a) showed high transcriptional level of POU2F3 and previously reported POU2F3 regulating genes, SOX9, ASCL2, AVIL, GFI1B, CHAT and TRPM5 [188, 189] in the LASCPC-01 and PARCBs originated tumor samples compared to the primary prostate basal samples. Western blot analysis found LASCPC-01 and PARCB5 were positively expressing POU2F3 protein, in a comparable level compared to SCLC-POU2F3 cell lines NCI-H1048 and NCI-H211 (Figure 1b).

To determine the cell dependency of POU2F3 in LASCPC-01 and PARCB5, shRNA was transduced in LASCPC-01 and PARCB5 to determine if proliferation of the cells could be affected by POU2F3 knock-down. In both cell lines, downregulation of POU2F3 resulted in a significant decrease in cell proliferation compared to the shRNA scramble control (Figure 1c). Similar results were seen in small cell lung cancer POU2F3 subtype cell lines (Figure 1c). On the other hand, knocking down expression of other tuft cell markers, SOX9, AVIL, CHAT, and TRPM5 were not able to affect the proliferation of LASCPC-01 cells (Supplementary Figure 1). These findings indicate that POU2F3, not its regulating genes, acts as an oncogenic driver and its acute perturbation leads to a significant disruption of cancer cell viability.

POU2F3 can down-regulate AR signaling pathway under low androgen condition

In SCLC, POU2F3 often is associated with a NE-low subtype, which lacks neuroendocrine lineage markers (41). Therefore, we did not expect POU2F3 to stere NE reprogramming in prostate cancer.

Instead, we want to examine whether POU2F3 is involved in lineage transition which alter AR signaling pathways. Overexpressed POU2F3 in various CRPC and NEPC cell lines was performed to address this question. For the ARPC cells, overexpression of POU2F3 showed no effect on downregulation of AR and AR downstream genes with androgen present in the media (Supplementary Figure 1a). However, in cell lines of C4-2B, LNCaP, LAPC4, POU2F3 overexpression reduced AR regulating gene NKX3.1 at protein level and other subsets of AR downstream signaling at transcriptional level, such as NKX3.1 and SPON2, in low presence of androgen which mimics androgen depletion circumstances in the clinical settings. (Figure 2a, 2b). This effect was not seen in 22Rv1 and MDA PCa 2b cells (Supplementary Figure 2a). One possible explanation could be that 22Rv1 cells exhibit expression of the AR-v7 splice variant, leading to a disregarding of the reduction in androgen levels. Interestingly, replenish androgen in the charcoal-strip serum condition revealed rescue expression of NKX3.1 level in the three cell lines (Supplementary Figure 2b), suggesting the effect of POU2F3 overexpression on the AR downstream signaling is mainly dependent on AR activation. Analysis of neuroendocrine marker expression upon POU2F3 expression demonstrated that POU2F3 alone was not able to induce any NE marker expression in CRPC cells (Figure 2a), which is expected since POU2F3 in small cell lung cancer is not a NE-phenotype driver.

RNA-seq to assess global transcriptional landscape alteration affected by POU2F3 overexpression showed attenuated AR downstream signaling pathways (Figure 2c). The volcano plot demonstrated genes enriched in POU2F3 overexpression group showed overlapping enriched gene expression, such as MUL1 and ABCA12 (Figure 2d), which are both associated with AKT signaling pathway dysregulation in many cancer types [4,5] and play a role in the castration

resistance []. The Gene Ontology (GO) analysis displayed limited enrichment of pathways (data not shown), indicating moderate transcriptional landscape upon POU2F3 activation. Together, our work demonstrated POU2F3 effect in driving downregulation of AR signaling pathway, suggesting potential role of POU2F3 in progression of castration resistant prostate cancer to acquire castration resistance.

4.3 MATERIALS AND METHODS

Cell lines

LNCAP (catalog No. CRL-1740, RRID: CVCL_0395), C4-2B (catalog No. CRL-3315, RRID: CVCL_4784), LAPC4 (catalog No. CRL-13009, RRID: CVCL_4744), LASCPC-01 (catalog No. CRL-3356, RRID: CVCL_UE17), NCI-H1048 (catalog No. CRL-5853, RRID: CVCL_1453) and NCI-H211 (catalog No. CRL-5824, RRID: CVCL_1529) cell lines were purchased from the ATCC. All cell lines were validated by short tandem repeat analysis after receipt. PARCB cell lines were gift from Jung Wook Park, (Duke University, Durham, NC). were kindly gifted by David Macpherson (Fred Hutchinson Cancer Center, Seattle, WA). LNCAP, 22Rv1, C4-2B, and MSKCC EF1 (derived from the organoid line MSKCC-CaP4) were maintained in RPMI medium supplemented with 10% FBS, 100 U/mL penicillin and 100 μg/mL streptomycin, and 4 mmol/L GlutaMAX. NCI-H660 and PARCB cells were maintained in Advanced DMEM/F12 medium supplemented with B27, 4 mmol/L GlutaMAX, and 10 ng/mL recombinant human basic FGF (bFGF) and EGF. LAPC4 was maintained in IMDM medium supplemented with 10% FBS, 100 U/mL penicillin and 100 μg/mL streptomycin, and 4 mmol/L GlutaMAX. LASCPC-01 was

maintained in RPMI medium supplemented with 10% FBS, 10 pM Hydrocortisone, 10 pM Beta-estradiol, 100 U/mL penicillin and 100 μg/mL streptomycin, and 4 mmol/L GlutaMAX. Cell lines were cultured no more than 3 weeks after thawing prior to use in described experiments.

Lentiviral constructs and lentivirus preparation

shRNA targeting POU2F3 were generated by cloning the top 3 shRNA from Broad Institute Genetic Perturbation Platform into site between ECORI and AGEI in pLKO.1-puro (Addgene). The shRNA with highest knock down efficiency was used in the experiment. Lentiviral vectors of POU2F3 were cloned by PCR amplification of plasmids with corresponding open read frame sequences. The amplified PCR products were then cloned into the EcoRI site of the FU-CGW lentiviral backbone by HiFi DNA Assembly. The lentivirus was prepared as described in (45,46). Lentivirus titter was determined by infecting target cells with serially diluted virus followed by FACS analysis for GFP expression.

Immunoblotting

Whole cell extracts were fractionated by SDS-PAGE and transferred to a nitrocellulose membrane using a transfer apparatus according to the manufacturer's instructions (Invitrogen). Membranes were blocked with 5% nonfat milk in PBST (DPBS + 0.5% Tween 20) for 30 minutes while shaking, then incubated with primary antibodies at 4°C for 16 hours. Membranes were washed three times for 5 minutes with PBST and incubated with horseradish peroxidase (HRP)-conjugated anti-mouse or anti-rabbit secondary antibody for 1 hour at room temperature. Blots were washed

three times for 5 minutes each with PBST and developed with Immobilon Western Chemiluminescent HRP Substrate (MilliporeSigma) for 3 minutes at room temperature. Blot images were acquired with a ChemiDoc Imaging System (Bio-Rad) or autoradiography film.

RNA-seq

Total RNA was extracted with a PurelinkTM RNA mini kit (Invotrogen). cDNA libraries were prepared from isolated RNA by the Fred Hutchinson Genomics core facility (Fred Hutchinson Cancer Center, Seattle, WA). High-throughput sequencing with 50 bp paired-end reads was performed using an Illumina NovaSeq 6000 SP100 mode. Sequencing reads were mapped to human genome reference HG38 and gene expression was quantified and normalized using Toil RNA-seq workflow (47). Gene set enrichment analysis was performed using GSEA software from Broad Institute (Subramanian et al., 2005) with a pre-ranked list of genes differentially expressed (>4-fold) between control and experimental samples.

Quantitative RT-qPCR

RNA was extracted from each cell line using $0.5x10^6$ cells with a PurelinkTM RNA mini kit (Invotrogen). RNA was converted to cDNA using the SuperScript™ IV Reverse Transcriptase (Invitrogen). SYBR Green RT-PCR was performed using the PowerUp™ SYBR™ Green Master Mix (Applied Biosystems) and the following gene expression probes: AR; NKX3-1; SPON2; TMPRSS2; FKBP5; ABCC4; KLK3; NDRG1; POU2F3; RPL13A for a normalization control. qPCR was performed on a QuantStudio 6 Real-Time PCR System Instrument (Applied Biosystems). Data was analyzed using the comparative Ct method (fold change = $2^{-\Delta\Delta Ct}$(48)).

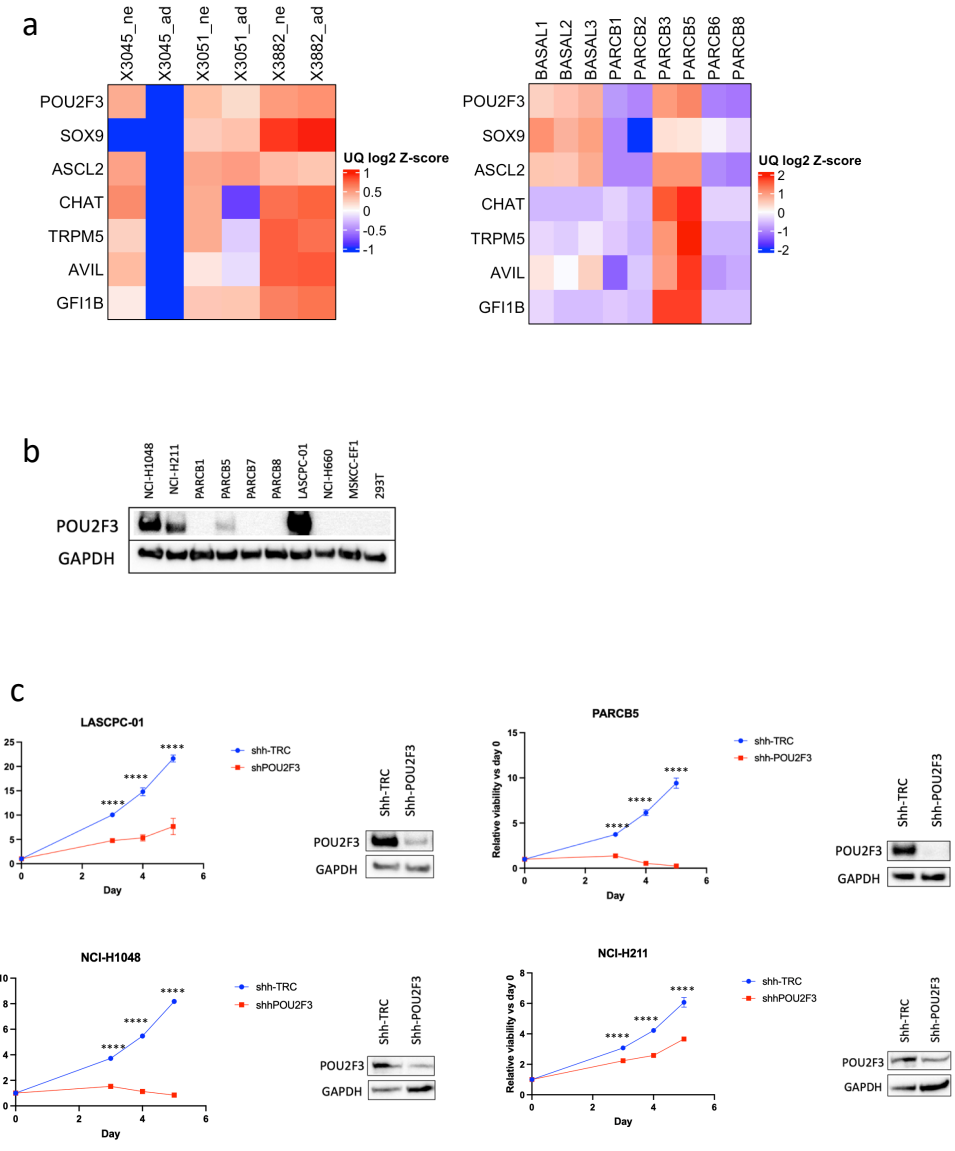
Cell proliferation assay

For cell proliferation assays, the CellTiter-Glo® 2.0 Cell Viability Assay system (Promega) was used according to the manufacturer's instructions. 5×10^3 cells were plated into each well in a 96-well and 25 μL per well of CellTiter-Glo® 2.0 reagent was added. After 0.5h incubation at room temperature, luminescence was measured every 24 h using a BioTek Synergy H1 Multimode Reader reader (Agilent). Six replicate wells per time point were used to obtain measures of cell proliferation.

Statistical analysis

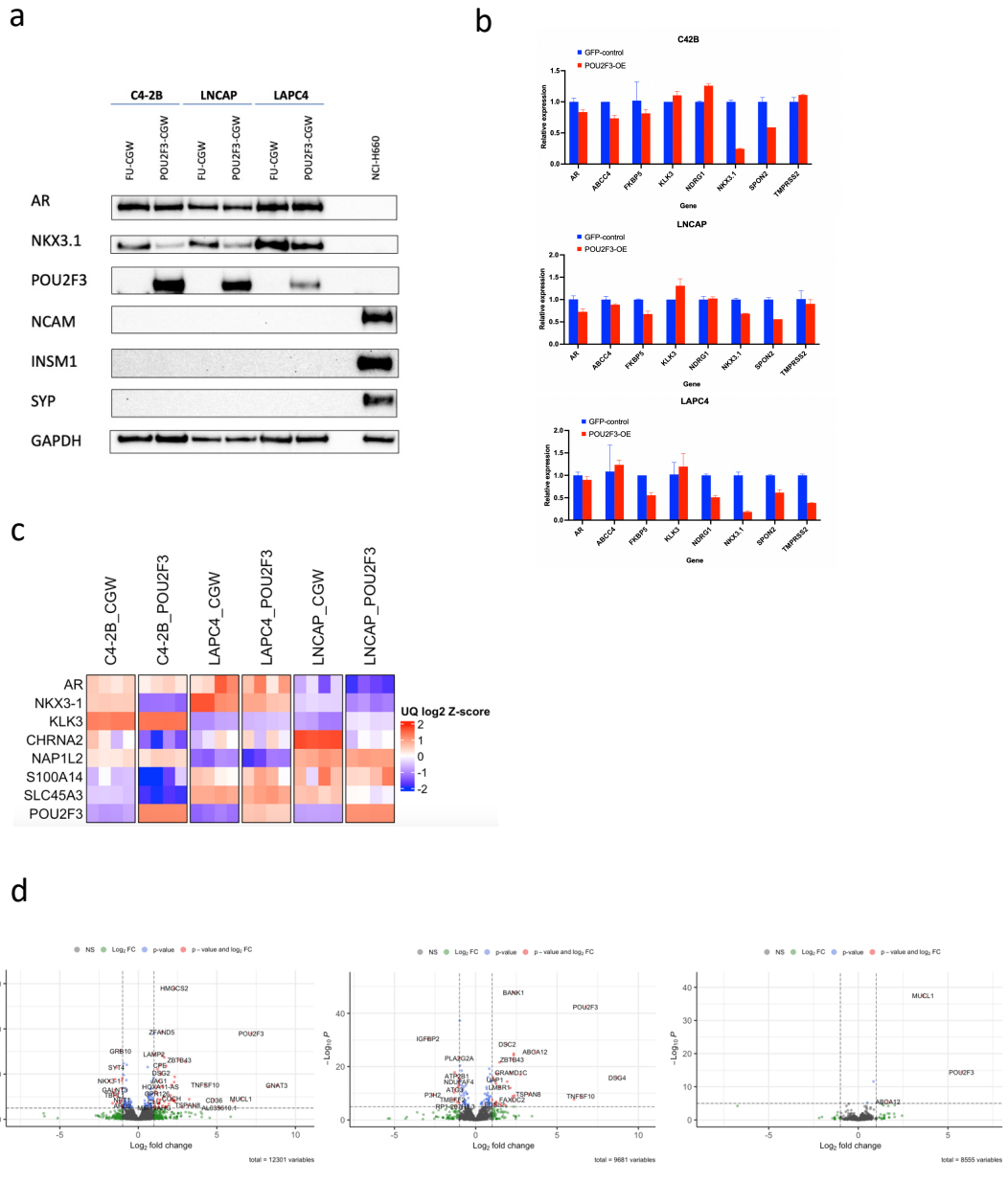
Data analysis was performed on GraphPad Prism 9 (GraphPad Software Inc.). Quantitative PCR results were analyzed in Excel. Statistical significance was determined using the unpaired two-tailed Student t test, unpaired two-tailed Welch t test where the variances are shown to be different via F-test, one-way ANOVA, or two-way ANOVA. Only two-tailed tests were used. Results are depicted as mean + SD unless stated otherwise. All P values of <0.05, <0.01, <0.001, and <0.0001 were considered significant. Pearson correlation coefficient was used to determine correlation between genes (Pearson ρ). The symbols used to represent the P values were: ns, nonsignificant, $P > 0.05$; *, $P \leqslant 0.05$; **, $P \leqslant 0.01$; ***, $P \leqslant 0.001$; ****, $P \leqslant 0.0001$. The test used in each statistical analysis is specified in the figure legends.

4.4 FIGURES



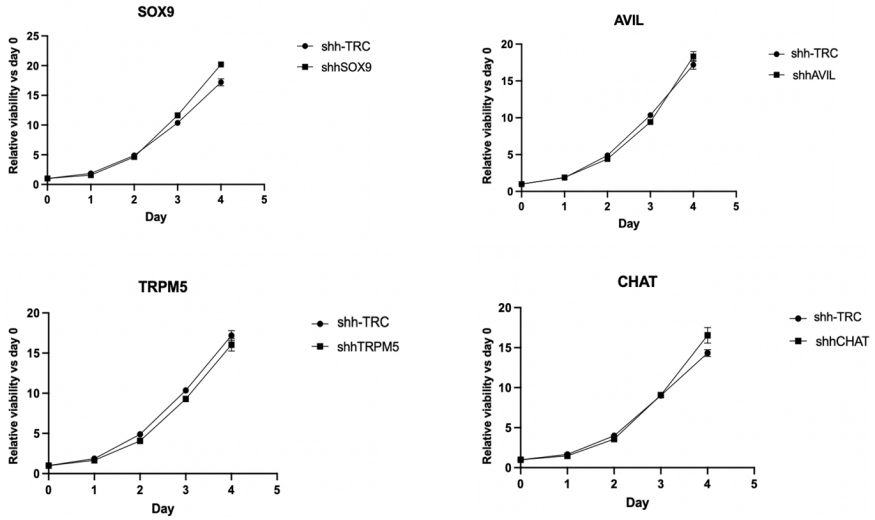
4.4.1. Figure 1. POU2F3 is an oncogenic driver in POU2F3 overexpressing prostate cancer cell line.

(a) Heatmap of RNA-seq gene expression data showing POU2F3 and POU2F3 regulating genes in adenocarcinoma and NEPC from [69](left), in basal cells and generated PARCB tumors from [70](right). UQ: upper quartile normalization. (b) Immunoblot analysis of NCI-H1048, NCI-H211, PARCB1, PARCB5, PARCB7, PACRB8, LASCPC-01, NCI-H660, MSKCC-EF1, 293T cells. (c) Relative cell viability (left) determined by CellTiter-Glo assay over time of LASCPC-01, PACRB5, NCI-H1048, HCI-H211 cells transduced with TRC shRNA control or shPOU2F3 (n=6 each). P-value was assessed by Student's t-test. **** denotes p <0.0001. Immunoblot analysis (right) of LASCPC-01, PACRB5, NCI-H1048, HCI-H211 cells transduced with TRC shRNA control or shPOU2F3. Lysates were collected at day 5.



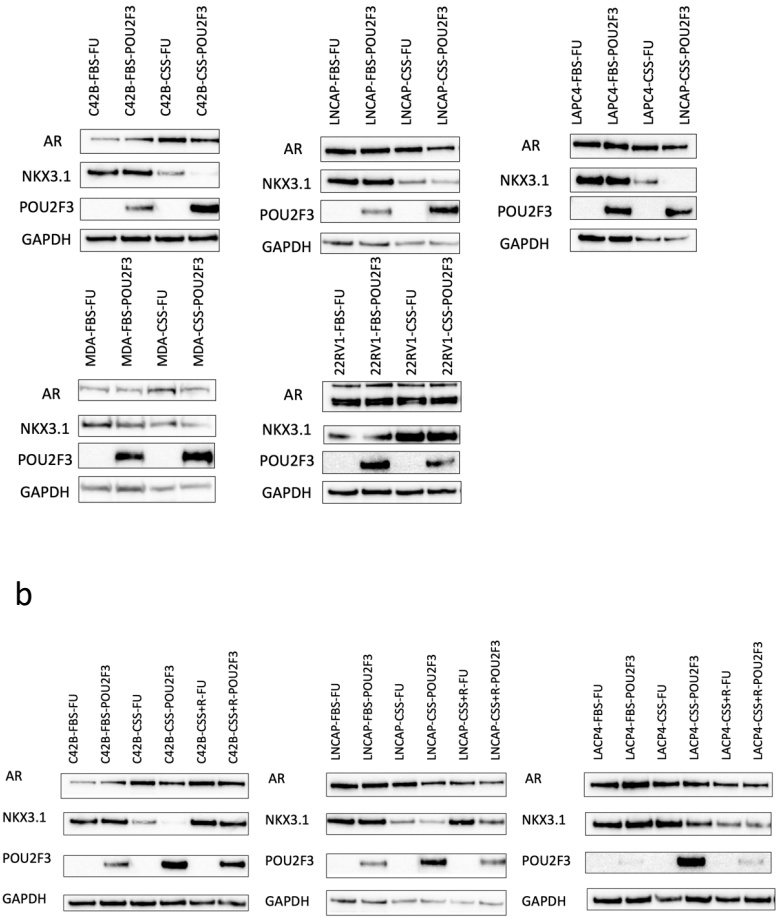
4.4.2. Figure 2. POU2F3 overexpression in CRPC cells decrease the AR signaling.

(a) Immunoblot analysis of C4-2B, LNCAP, LAPC4 cells transduced with POU2F3 overexpression lentivirus under low androgen condition. Lysates were collected on day 14. (b) Quantitative PCR analysis of C4-2B, LNCAP, LAPC4 cells transduced with POU2F3 overexpression under low androgen condition. RNA was collected on day 14. (c) Heatmap of RNA-seq gene expression from POU2F3 overexpressing C4-2B, LNCAP, LAPC4 cells showing genes associated with the AR program. (d) Volcano plots showing the differential expression of genes encoding transcriptions factors in pairwise comparisons of RNA-seq gene expression data of C4-2B, LNCAP, LAPC4 cells transduced with CGW-control vs. POU2F3. Red dots represent genes with a $-\log_2(FC) > 2$. FC represents fold-change.



4.4.3. Supplementary Figure 1. POU2F3 regulating gene is not a cell dependency in LASCPC-01 cells.

Relative cell viability determined by CellTiter-Glo assay over time of LASCPC-01 cells transduced with TRC shRNA control or shSOX9, shAVIL, shTRPM5, shCHAT (n=6 each). P-value was assessed by Student's t-test.



b

4.4.4. Supplementary Figure 2. Downregulation of AR signaling by POU2F3 only under low androgen condition.

(a) Immunoblot analysis of C4-2B, LNCAP, LAPC4, MDA PCa 2b, and 22RV1 cells transduced with POU2F3 overexpression lentivirus under 10% FBS and 10% CSS. Lysates were collected on day 14. (b) Immunoblot analysis of C4-2B, LNCAP, and LAPC4cells transduced with POU2F3 overexpression lentivirus under 10% FBS, 10% CSS, and 10% CSS +10nM R1881. Lysates were collected on day 14.

Chapter 5. CONCLUSIONS AND FUTURE STUDIES

5.1 CONCLUSIONS

The aims of the research presented here is to identify and characterize genetic determinants involved in prostate cancer initiation and progression. In the previous chapters, we have built a novel functional genomic strategy to interrogate genotype-phenotype relationships in cancer initiation/progression from primary epithelial cells (Chapter 2), defined critical functional roles of genetic factors ASCL1 and NEUROD1 in neuroendocrine transdifferentiation in prostate cancer (Chapter 3), identified POU2F3 as an oncogenic driver in CRPC progression (Chapter 4).

We have developed a multiplexed genetic strategy using an *in vivo* organoid transformation assay to efficiently generate diverse, clinically relevant bladder and prostate cancer models. Incorporating single-cell and spatially resolved next-generation sequencing to the assay, we are able to deconvolute the oncogenic drivers associated with the different cancer phenotypes in the resultant tumors. We have used this assay to successfully establish first-in-field genetically-defined bladder cancers with non-urothelial histologic subtypes and prostate cancer with pleomorphic giant cells. In the generated model, we identified Fgfr3 as a driver of a papillary urothelial carcinoma lineage and reduced Kmt2c expression associated with pleomorphic giant cell carcinoma phenotype in prostate cancer. This work describes a methodology that can be applied to rapidly investigate the extensive genomic abnormalities discovered in cancer genome sequencing studies and their functional implications in the development of cancer.

We have adapted a genetically-defined assay to identify a set of candidate factors that could reprogram ARPC cell lines to NEPC. With this model of NEPC transdifferentiation, we found the pioneer neural transcription factors ASCL1 and NEUROD1 are responsible to induce a neuroendocrine lineage and silence AR signaling in the transdifferentiation of adenocarcinoma to neuroendocrine prostate cancer. Additional genetic alterations, loss of TP53 and RB1, overexpression of NMYC, BCL2, and SRRM4 are required to overcome the acute growth senescence induced by ASCL1 and NEUROD1. Suppression of AR activity by ASCL1 and NEUROD1 is mediated by reducing chromatin accessibility of somatically acquired AR enhancer. We show the ASCL1 and NEUROD1 are driving downregulation of major histocompatibility complex class I (MHC I) in NEPC, resulting in an immunoevasive state. In summary, we successfully recapitulate ARPC to NEPC transidfferentiation in prostate cancer cell model, and uncover key mechanisms implications in the NEPC progression.

With the increasing recognition of POU2F3 in the maintenance of lineage state in small cell lung cancer, there are evidence of POU2F3 expression in the castration resistance progression in prostate cancer. We have identified POU2F3 expression in two genetically modified small cell prostate cancer and neuroendocrine prostate cancer cell lines, and demonstrated cell dependency on POU2F3 expression in the cells, suggesting that POU2F3 is oncogenic driver in prostate cancer and is essential for the proliferation the cells. Overexpression of POU2F3 in AR positive CRPC cell lines leads to the suppression of AR downstream signaling pathways in an androgen depletion manner. In this work, we elaborate mechanism insights of how POU2F3 might be involved in the AR lineage transition in prostate cancer, and reveal POU2F3 as a potential driver in prostate cancer progression.

5.2 FUTURE DIRECTIONS

Improving and extending applications of the combinatorial genetic strategy

The recent great advance of high-throughput next generation sequencing technology allows for deep understanding of the human genome. The massive candidate genetic abnormalities were uncovered in human diseases. Our combinatorial assay by multiplexing candidate genetic factors to study the disease progression are capable to close the gap between the large scale of genomic aberrations and our understanding of their functional consequences. The initial studies demonstrate the potency of the approach but further optimization is still necessary. The tumors were generated in an immunocompromised mouse model, whereas tumorigenesis occurs within a complex microenvironment. A more ideal way to study tumorigenesis is to perform the transformation assay in an immunocompetent mouse model to recapitulate tumor initiation in a more native tissue microenvironment. As previously mentioned, the process of single-cell dissociation during single-cell DNA sequencing results in the loss of spatial resolution, thus disconnecting the identified tumor clonality from the corresponding phenotype/histology within the tumors. Instead of laser capture microdissection, recent development of spatial sequencing provides an easier and efficient way of assignment of specific locations to transcriptomic or genomic variation. Slide-DNA-seq [190], a spatially resolved DNA sequencing enabled detection of spatial distribution of copy number alterations, suggesting capability of spatial sequencing being used for recovering our barcoded oncogenic events associated with different histologies in the generated tumors.

Besides the biology of disease progression, there are vast biological process which have not yet been fully understand. One important biological question is to understand cell programming and differentiation. The potential for somatic cells to be reprogrammed into a pluripotent state opens up new avenues for treating diseases. The famous Nobel Prize-winning research that identified critical factors capable of inducing pluripotent stem cells from mouse embryonic and adult fibroblast cultures laid a strong foundation for functional studies in cell reprogramming [191]. Yamanaka et al. selected more than 20 candidate factors and tested various combinations of genetic factors that were essential for inducing pluripotency in mouse embryonic fibroblasts, narrowing down the pivotal transcriptional factors to be OCT4, SOX2, KLF4, and MYC [191]. The work of identifying 4 factors out of 24 candidates are massive and time-consuming. Our combinatorial genetic assay provides a new approach to functionally interrogate the induction of pluripotency in somatic cells. Including way fewer steps, it is possible to deconvolute pivotal factors and clonal architecture in the induced pluripotent stem cells (iPSC), which are developed by transducing mouse embryonic fibroblasts with a combination of 24 candidate factors, thereby greatly reducing the experimental time frame. Our strategy has demonstrated its efficacy by being successfully applied to two different cancer types, suggesting its potential for broader applications. As long as there are two distinct phenotypes, genetic factors involved in the transition from one phenotype to the other, can be uncovered via our assay.

Potential therapeutic strategies of NEPC

Understanding the mechanisms of neuroendocrine transdifferentiation will provide insights into developing inhibitors or targeted therapy for the pathways that govern and sustain neuroendocrine

lineage in prostate cancer. We have successfully generated an *in vitro* model that recapitulate transdifferentiation of adenocarcinoma to NEPC, which allows us to study the potential therapeutic vulnerabilities of NEPC. In our work, we demonstrate that ASCL1 or NEUROD1-driven neuroendocrine program displays an immunoevasive state, where the MHC I MHC I antigen processing and presentation pathway is suppressed. Due to the immune evasion of NECP, though NEPC patients were predicted to have better response to immunotherapy than patients with adenocarcinoma CRPC [192], immune check inhibitors have demonstrated limited clinical efficacy in men with metastatic NEPC [193]. Potential strategy for NEPC could be combination of immune checkpoint inhibitors with epigenetic inhibitors such as those targeting EZH2, histone deacetylases, or demethylases, which might induce immunogenic response by enhancing MHC I expression and sensitize the cells for immunotherapy [194, 195].

The cell surface protein repertoire (surfaceome) study has revealed that unique cell identities exhibit quantitative and qualitative differences in cell surface protein expression [196]. We have identified cell surface targets that are differentially expressed in ASCL1 or NEUROD1-driven NE lineage (DLL3, L1CAM, CEACAM5, RET and SEZ6), which could become potential targets for antibody conjugated therapy or chimeric antigen receptor T-cell therapies. There are supporting evidences that AMG 757, an IgG-like T cell-engaging bispecific antibody that redirects T cells to lyse DLL3-expressing cells, and rova-IR700, an anti-DLL3 antibody conjugated to an IR700 photosensitizer, have shown high efficacy and sensitivity in preclinical models of SCLC [197, 198]. Targeted immune therapies that for the other surface markers enriched in NEPC will provide more potential options for NEPC treatment.

Functional mechanism of POU2F3 in lineage plasticity

POU2F3, acting as a transcription factor, exerts regulatory control over gene expression through its binding to specific DNA sequences within the promoter or enhancer regions of target genes. Although the precise mechanism by which POU2F3 influences AR downstream gene expression remains unknown, a possible explanation could involve the interference of POU2F3 binding with the binding of AR to its target genes, resulting in decrease expression of those genes. To elucidate the relationship between POU2F3 and AR and gain insights into how POU2F3 affects the expression of AR downstream genes, one approach is to perform DNA binding site profiling of POU2F3 and AR using sequencing techniques such as ChIP-Seq (chromatin immunoprecipitation followed by sequencing) [199] or CUT&RUN (cleavage under targets and release using nuclease) [200]. These methods would allow for the identification of the genomic locations where POU2F3 and AR bind and potentially overlap. By analyzing the overlapping binding sites, valuable insights can be obtained regarding the potential regulatory interactions between POU2F3 and AR and their impact on the expression of AR downstream genes.

Given the limited models available for studying the role of POU2F3 in prostate cancer, additional evidence of POU2F3 expression in CRPC patients, cell lines, or PDX models is necessary to investigate its functional role in the progression of CRPC. We can infer from our study that POU2F3 is not a transcriptional factor that can induce acute transcriptional/epigenetic landscape alteration and there is possibility that POU2F3 being a driver for a short-term intermediate state of ARPC to NEPC or DNPC progression or POU2F3 expression is a resultant of other genetic factors associated with CPRC progression. In the two cell line models with POU2F3 expression, they are

both generated by activation of AKT and MYC signaling pathways. Therefore, the two pathways might be involved in inducing POU2F3 expression in prostate cancer progression. The interaction of POU2F3 with the AKT1 and N/C-MYC can be explored to understand their relationship in driving AR-null lineage. Since POU2F3 alone is not able to induce AR downregulation in AR positive cell lines, combination of POU2F3 with activation of AKT1, N/C-MYC, and other oncogenic events might have a more extensive effect on AR signaling pathway and initiate an AR-independent phenotype in prostate cancer.

As our understanding of POU2F3 in prostate cancer continues to advance, comparison based on transcriptomic and epigenomics profile with small cell lung cancer with POU2F3 subtype can help identify common or unique therapeutic vulnerabilities in these cancers. By examining the similarities and differences in the role and function of POU2F3 in prostate cancer and small cell lung cancer, we can gain insights into potential shared molecular pathways or distinct mechanisms that contribute to the development and progression of these malignancies. Understanding these similarities and differences may open up new avenues for targeted therapeutic interventions tailored to each cancer type.

VITA

2015	B.S., Biological Sciences Nanjing University, Nanjing, China
2017	M.S., Microbial and Cellular Biology Emporia State University, Emporia, Kansas
2016	Award Recipient, Harold Durst Graduate Student Research Award Emporia State University, Emporia, Kansas
2017	Award Recipient, Robert F. Clarke Outstanding Biology Graduate Student Award in Research Emporia State University, Emporia, Kansas
2017	Award Recipient, Kansas Academy of Science Graduate Research Grant Kansas Academy of Science, Lawrence, Kansas
2017-2023	PhD trainee, Molecular Engineering and Sciences University of Washington, Seattle, Washington

PUBLICATIONS

Bhatia, V., Kamat, N., Pariva, T., Wu, L., Tsao, A., **Sun, H**., et al, (2023) Targeting advanced prostate cancer with STEAP1 chimeric antigen receptor T cell and tumor-localized IL-12 immunotherapy. Nature communications, 14(1), 2041

Hu, C., Zhao, Y., **Sun, H**., & Yang, Y. (2017) Synergism of Dam, MutH, and MutS in methylation-directed mismatch repair in Escherichia coli. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 795, 31-33.

BIBLIOGRAPHY

1. Siegel, R.L., et al., *Cancer statistics, 2023.* CA Cancer J Clin, 2023. **73**(1): p. 17-48.
2. Sandhu, S., et al., *Prostate cancer.* Lancet, 2021. **398**(10305): p. 1075-1090.
3. de Bono, J.S., et al., *Abiraterone and increased survival in metastatic prostate cancer.* N Engl J Med, 2011. **364**(21): p. 1995-2005.
4. Tran, C., et al., *Development of a second-generation antiandrogen for treatment of advanced prostate cancer.* Science, 2009. **324**(5928): p. 787-90.
5. James, N.D., et al., *Abiraterone for Prostate Cancer Not Previously Treated with Hormone Therapy.* N Engl J Med, 2017. **377**(4): p. 338-351.
6. Robinson, D., et al., *Integrative clinical genomics of advanced prostate cancer.* Cell, 2015. **161**(5): p. 1215-1228.
7. Taylor, B.S., et al., *Integrative genomic profiling of human prostate cancer.* Cancer Cell, 2010. **18**(1): p. 11-22.
8. Quigley, D.A., et al., *Genomic Hallmarks and Structural Variation in Metastatic Prostate Cancer.* Cell, 2018. **175**(3): p. 889.
9. Tomlins, S.A., et al., *Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer.* Science, 2005. **310**(5748): p. 644-8.
10. Demichelis, F., et al., *TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort.* Oncogene, 2007. **26**(31): p. 4596-9.
11. Taplin, M.E., et al., *Mutation of the androgen-receptor gene in metastatic androgen-independent prostate cancer.* N Engl J Med, 1995. **332**(21): p. 1393-8.

12. Lei, Q., et al., *NKX3.1 stabilizes p53, inhibits AKT activation, and blocks prostate cancer initiation caused by PTEN loss.* Cancer Cell, 2006. **9**(5): p. 367-78.

13. He, W.W., et al., *A novel human prostate-specific, androgen-regulated homeobox gene (NKX3.1) that maps to 8p21, a region frequently deleted in prostate cancer.* Genomics, 1997. **43**(1): p. 69-77.

14. Barbieri, C.E., et al., *Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer.* Nat Genet, 2012. **44**(6): p. 685-9.

15. Wang, Z., et al., *The diverse roles of SPOP in prostate cancer and kidney cancer.* Nat Rev Urol, 2020. **17**(6): p. 339-350.

16. Grasso, C.S., et al., *The mutational landscape of lethal castration-resistant prostate cancer.* Nature, 2012. **487**(7406): p. 239-43.

17. Visakorpi, T., et al., *In vivo amplification of the androgen receptor gene and progression of human prostate cancer.* Nat Genet, 1995. **9**(4): p. 401-6.

18. Bryce, A.H. and E.S. Antonarakis, *Androgen receptor splice variant 7 in castration-resistant prostate cancer: Clinical considerations.* Int J Urol, 2016. **23**(8): p. 646-53.

19. Liu, P., et al., *Targeting the phosphoinositide 3-kinase pathway in cancer.* Nat Rev Drug Discov, 2009. **8**(8): p. 627-44.

20. Reid, A.H., et al., *Molecular characterisation of ERG, ETV1 and PTEN gene loci identifies patients at low and high risk of death from prostate cancer.* Br J Cancer, 2010. **102**(4): p. 678-84.

21. Pearson, H.B., et al., *Identification of Pik3ca Mutation as a Genetic Driver of Prostate Cancer That Cooperates with Pten Loss to Accelerate Progression and Castration-Resistant Growth.* Cancer Discov, 2018. **8**(6): p. 764-779.

22. Armenia, J., et al., *The long tail of oncogenic drivers in prostate cancer.* Nat Genet, 2018. **50**(5): p. 645-651.

23. Mulholland, D.J., et al., *Pten loss and RAS/MAPK activation cooperate to promote EMT and metastasis initiated from prostate cancer stem/progenitor cells.* Cancer Res, 2012. **72**(7): p. 1878-89.

24. Nickols, N.G., et al., *MEK-ERK signaling is a therapeutic target in metastatic castration resistant prostate cancer.* Prostate Cancer Prostatic Dis, 2019. **22**(4): p. 531-538.

25. Cho, N.Y., et al., *BRAF and KRAS mutations in prostatic adenocarcinoma.* Int J Cancer, 2006. **119**(8): p. 1858-62.

26. Angers, S. and R.T. Moon, *Proximal events in Wnt signal transduction.* Nat Rev Mol Cell Biol, 2009. **10**(7): p. 468-77.

27. Isaacsson Velho, P., et al., *Wnt-pathway Activating Mutations Are Associated with Resistance to First-line Abiraterone and Enzalutamide in Castration-resistant Prostate Cancer.* Eur Urol, 2020. **77**(1): p. 14-21.

28. Murillo-Garzon, V. and R. Kypta, *WNT signalling in prostate cancer.* Nat Rev Urol, 2017. **14**(11): p. 683-696.

29. Lozano, R., et al., *Genetic aberrations in DNA repair pathways: a cornerstone of precision oncology in prostate cancer.* Br J Cancer, 2021. **124**(3): p. 552-563.

30. Fraser, M., et al., *Genomic hallmarks of localized, non-indolent prostate cancer.* Nature, 2017. **541**(7637): p. 359-364.

31. Nyquist, M.D., et al., *Combined TP53 and RB1 Loss Promotes Prostate Cancer Resistance to a Spectrum of Therapeutics and Confers Vulnerability to Replication Stress.* Cell Rep, 2020. **31**(8): p. 107669.

32. Fleming, W.H., et al., *Expression of the c-myc protooncogene in human prostatic carcinoma and benign prostatic hyperplasia.* Cancer Res, 1986. **46**(3): p. 1535-8.

33. Bernard, D., et al., *Myc confers androgen-independent prostate cancer cell growth.* J Clin Invest, 2003. **112**(11): p. 1724-31.

34. Hellerstedt, B.A. and K.J. Pienta, *The current state of hormonal therapy for prostate cancer.* CA Cancer J Clin, 2002. **52**(3): p. 154-79.

35. Scher, H.I. and C.L. Sawyers, *Biology of progressive, castration-resistant prostate cancer: directed therapies targeting the androgen-receptor signaling axis.* J Clin Oncol, 2005. **23**(32): p. 8253-61.

36. Bluemn, E.G., et al., *Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained through FGF Signaling.* Cancer Cell, 2017. **32**(4): p. 474-489 e6.

37. Zong, Y. and A.S. Goldstein, *Adaptation or selection--mechanisms of castration-resistant prostate cancer.* Nat Rev Urol, 2013. **10**(2): p. 90-8.

38. Veldscholte, J., et al., *A mutation in the ligand binding domain of the androgen receptor of human LNCaP cells affects steroid binding characteristics and response to anti-androgens.* Biochem Biophys Res Commun, 1990. **173**(2): p. 534-40.

39. Culig, Z., et al., *Mutant androgen receptor detected in an advanced-stage prostatic carcinoma is activated by adrenal androgens and progesterone.* Mol Endocrinol, 1993. **7**(12): p. 1541-50.

40. Karantanos, T., P.G. Corn, and T.C. Thompson, *Prostate cancer progression after androgen deprivation therapy: mechanisms of castrate resistance and novel therapeutic approaches.* Oncogene, 2013. **32**(49): p. 5501-11.

41. Watson, P.A., V.K. Arora, and C.L. Sawyers, *Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer.* Nat Rev Cancer, 2015. **15**(12): p. 701-11.

42. Lu, N.Z., et al., *International Union of Pharmacology. LXV. The pharmacology and classification of the nuclear receptor superfamily: glucocorticoid, mineralocorticoid, progesterone, and androgen receptors.* Pharmacol Rev, 2006. **58**(4): p. 782-97.

43. Takeda, D.Y., et al., *A Somatically Acquired Enhancer of the Androgen Receptor Is a Noncoding Driver in Advanced Prostate Cancer.* Cell, 2018. **174**(2): p. 422-432 e13.

44. Viswanathan, S.R., et al., *Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing.* Cell, 2018. **174**(2): p. 433-447 e19.

45. Arora, V.K., et al., *Glucocorticoid receptor confers resistance to antiandrogens by bypassing androgen receptor blockade.* Cell, 2013. **155**(6): p. 1309-22.

46. Zaffuto, E., et al., *Contemporary Incidence and Cancer Control Outcomes of Primary Neuroendocrine Prostate Cancer: A SEER Database Analysis.* Clin Genitourin Cancer, 2017. **15**(5): p. e793-e800.

47. Chedgy, E.C., et al., *Biallelic tumour suppressor loss and DNA repair defects in de novo small-cell prostate carcinoma.* J Pathol, 2018. **246**(2): p. 244-253.

48. Shah, R.B., et al., *Androgen-independent prostate cancer is a heterogeneous group of diseases: lessons from a rapid autopsy program.* Cancer Res, 2004. **64**(24): p. 9209-16.

49. Amato, R.J., et al., *Chemotherapy for small cell carcinoma of prostatic origin.* J Urol, 1992. **147**(3 Pt 2): p. 935-7.

50. Aparicio, A.M., et al., *Platinum-based chemotherapy for variant castrate-resistant prostate cancer.* Clin Cancer Res, 2013. **19**(13): p. 3621-30.

51. Epstein, J.I., et al., *Proposed morphologic classification of prostate cancer with neuroendocrine differentiation.* Am J Surg Pathol, 2014. **38**(6): p. 756-67.

52. Spiess, P.E., et al., *Treatment outcomes of small cell carcinoma of the prostate: a single-center study.* Cancer, 2007. **110**(8): p. 1729-37.

53. Beltran, H., et al., *Molecular characterization of neuroendocrine prostate cancer and identification of new drug targets.* Cancer Discov, 2011. **1**(6): p. 487-95.

54. Prendeville, S., et al., *Prostate carcinoma with amphicrine features: further refining the spectrum of neuroendocrine differentiation in tumours of primary prostatic origin?* Histopathology, 2017. **71**(6): p. 926-933.

55. Graham, L.S., et al., *Clinical, pathologic, and molecular features of amphicrine prostate cancer.* Prostate, 2023. **83**(7): p. 641-648.

56. Abrahamsson, P.A., *Neuroendocrine cells in tumour growth of the prostate.* Endocr Relat Cancer, 1999. **6**(4): p. 503-19.

57. Hu, Y., et al., *Molecular characterization of a metastatic neuroendocrine cell cancer arising in the prostates of transgenic mice.* J Biol Chem, 2002. **277**(46): p. 44462-74.

58. Beltran, H., et al., *The Role of Lineage Plasticity in Prostate Cancer Therapy Resistance.* Clin Cancer Res, 2019. **25**(23): p. 6916-6924.

59. Lotan, T.L., et al., *ERG gene rearrangements are common in prostatic small cell carcinomas.* Mod Pathol, 2011. **24**(6): p. 820-8.

60. Williamson, S.R., et al., *ERG-TMPRSS2 rearrangement is shared by concurrent prostatic adenocarcinoma and prostatic small cell carcinoma and absent in small cell carcinoma of the urinary bladder: evidence supporting monoclonal origin.* Mod Pathol, 2011. **24**(8): p. 1120-7.

61. Hansel, D.E., et al., *Shared TP53 gene mutation in morphologically and phenotypically distinct concurrent primary small cell neuroendocrine carcinoma and adenocarcinoma of the prostate.* Prostate, 2009. **69**(6): p. 603-9.

62. Beltran, H., et al., *Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer.* Nat Med, 2016. **22**(3): p. 298-305.

63. Dong, B., et al., *Single-cell analysis supports a luminal-neuroendocrine transdifferentiation in human prostate cancer.* Commun Biol, 2020. **3**(1): p. 778.

64. Bang, Y.J., et al., *Terminal neuroendocrine differentiation of human prostate carcinoma cells in response to increased intracellular cyclic AMP.* Proc Natl Acad Sci U S A, 1994. **91**(12): p. 5330-4.

65. Yuan, T.C., et al., *Androgen deprivation induces human prostate epithelial neuroendocrine differentiation of androgen-sensitive LNCaP cells.* Endocr Relat Cancer, 2006. **13**(1): p. 151-67.

66. Zou, M., et al., *Transdifferentiation as a Mechanism of Treatment Resistance in a Mouse Model of Castration-Resistant Prostate Cancer.* Cancer Discov, 2017. **7**(7): p. 736-749.

67. Lin, D., et al., *High fidelity patient-derived xenografts for accelerating prostate cancer discovery and drug development.* Cancer Res, 2014. **74**(4): p. 1272-83.

68. Zhou, Z., et al., *Synergy of p53 and Rb deficiency in a conditional mouse model for metastatic prostate cancer.* Cancer Res, 2006. **66**(16): p. 7889-98.

69. Lee, J.K., et al., *N-Myc Drives Neuroendocrine Prostate Cancer Initiated from Human Prostate Epithelial Cells.* Cancer Cell, 2016. **29**(4): p. 536-547.

70. Park, J.W., et al., *Reprogramming normal human epithelial tissues to a common, lethal neuroendocrine cancer lineage.* Science, 2018. **362**(6410): p. 91-95.

71. Dardenne, E., et al., *N-Myc Induces an EZH2-Mediated Transcriptional Program Driving Neuroendocrine Prostate Cancer.* Cancer Cell, 2016. **30**(4): p. 563-577.

72. Clermont, P.L., et al., *Polycomb-mediated silencing in neuroendocrine prostate cancer.* Clin Epigenetics, 2015. **7**(1): p. 40.

73. Lapuk, A.V., et al., *From sequence to molecular pathology, and a mechanism driving the neuroendocrine phenotype in prostate cancer.* J Pathol, 2012. **227**(3): p. 286-97.

74. Chang, Y.T., et al., *REST is a crucial regulator for acquiring EMT-like and stemness phenotypes in hormone-refractory prostate cancer.* Sci Rep, 2017. **7**: p. 42795.

75. Bishop, J.L., et al., *The Master Neural Transcription Factor BRN2 Is an Androgen Receptor-Suppressed Driver of Neuroendocrine Differentiation in Prostate Cancer.* Cancer Discov, 2017. **7**(1): p. 54-71.

76. Russo, M.V., et al., *SOX2 boosts major tumor progression genes in prostate cancer and is a functional biomarker of lymph node metastasis.* Oncotarget, 2016. **7**(11): p. 12372-85.

77. Guo, H., et al., *ONECUT2 is a driver of neuroendocrine prostate cancer.* Nat Commun, 2019. **10**(1): p. 278.

78. Chen, S., et al., *Identifying non-muscle-invasive and muscle-invasive bladder cancer based on blood serum surface-enhanced Raman spectroscopy.* Biomed Opt Express, 2019. **10**(7): p. 3533-3544.

79. Knowles, M.A. and C.D. Hurst, *Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity.* Nat Rev Cancer, 2015. **15**(1): p. 25-41.

80. Rachakonda, P.S., et al., *TERT promoter mutations in bladder cancer affect patient survival and disease recurrence through modification by a common polymorphism.* Proc Natl Acad Sci U S A, 2013. **110**(43): p. 17426-31.

81. Hernandez, S., et al., *Prospective study of FGFR3 mutations as a prognostic factor in nonmuscle invasive urothelial bladder carcinomas.* J Clin Oncol, 2006. **24**(22): p. 3664-71.

82. Biton, A., et al., *Independent component analysis uncovers the landscape of the bladder tumor transcriptome and reveals insights into luminal and basal subtypes.* Cell Rep, 2014. **9**(4): p. 1235-45.

83. Cairns, P., M.E. Shaw, and M.A. Knowles, *Initiation of bladder cancer may involve deletion of a tumour-suppressor gene on chromosome 9.* Oncogene, 1993. **8**(4): p. 1083-5.

84. Lopez-Knowles, E., et al., *PIK3CA mutations are an early genetic alteration associated with FGFR3 mutations in superficial papillary bladder tumors.* Cancer Res, 2006. **66**(15): p. 7401-4.

85. Cancer Genome Atlas Research, N., *Comprehensive molecular characterization of urothelial bladder carcinoma.* Nature, 2014. **507**(7492): p. 315-22.

86. Fujimoto, K., et al., *Frequent association of p53 gene mutation in invasive bladder cancer.* Cancer Res, 1992. **52**(6): p. 1393-8.

87. Kamoun, A., et al., *A Consensus Molecular Classification of Muscle-invasive Bladder Cancer.* Eur Urol, 2020. **77**(4): p. 420-433.

88. Wolff, E.M., et al., *Unique DNA methylation patterns distinguish noninvasive and invasive urothelial cancers and establish an epigenetic field defect in premalignant tissue.* Cancer Res, 2010. **70**(20): p. 8169-78.

89. Robertson, A.G., et al., *Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer.* Cell, 2017. **171**(3): p. 540-556 e25.

90. Santos, C.P., et al., *Urothelial organoids originating from Cd49f(high) mouse stem cells display Notch-dependent differentiation capacity.* Nat Commun, 2019. **10**(1): p. 4407.

91. Drost, J., et al., *Organoid culture systems for prostate epithelial and cancer tissue.* Nature Protocols, 2016. **11**(2): p. 347-358.

92. Winters, B.R., et al., *Genomic distinctions between metastatic lower and upper tract urothelial carcinoma revealed through rapid autopsy.* JCI Insight, 2019. **5**(13).

93. Kumar, M., et al., *Systematic determination of the packaging limit of lentiviral vectors.* Hum Gene Ther, 2001. **12**(15): p. 1893-905.

94. Oottamasathien, S., et al., *Bladder tissue formation from cultured bladder urothelium.* Dev Dyn, 2006. **235**(10): p. 2795-801.

95. Zong, Y., A.S. Goldstein, and O.N. Witte, *Preparation of Urogenital Sinus Mesenchymal Cells for Prostate Tissue Recombination Models.* Cold Spring Harb Protoc, 2015. **2015**(11): p. 988-90.

96. Marzouka, N.A., et al., *The Lund Molecular Taxonomy Applied to Non-Muscle-Invasive Urothelial Carcinoma.* J Mol Diagn, 2022. **24**(9): p. 992-1008.

97. Damrauer, J.S., et al., *Intrinsic subtypes of high-grade bladder cancer reflect the hallmarks of breast cancer biology.* Proc Natl Acad Sci U S A, 2014. **111**(8): p. 3110-5.

98. Guo, C.C., et al., *Dysregulation of EMT Drives the Progression to Clinically Aggressive Sarcomatoid Bladder Cancer.* Cell Rep, 2019. **27**(6): p. 1781-1793 e4.

99. Shah, S.D., et al., *Syngeneic model of carcinogen-induced tumor mimics basal/squamous, stromal-rich, and neuroendocrine molecular and immunological features of muscle-invasive bladder cancer.* Front Oncol, 2023. **13**: p. 1120329.

100. Abdullah, M.A., et al., *Prostatic Adenocarcinoma With Focal Pleomorphic Giant Cell Features.* Surg Pathol, 2018. **42**: p. 1286-1296.

101. Pienta, K.J., et al., *Cancer recurrence and lethality are enabled by enhanced survival and reversible cell cycle arrest of polyaneuploid cells.* Proc Natl Acad Sci U S A, 2021. **118**(7).

102. Rampias, T., et al., *The lysine-specific methyltransferase KMT2C/MLL3 regulates DNA repair components in cancer.* EMBO Rep, 2019. **20**(3).

103. Chang, A., et al., *Recruitment of KMT2C/MLL3 to DNA Damage Sites Mediates DNA Damage Responses and Regulates PARP Inhibitor Sensitivity in Cancer.* Cancer Res, 2021. **81**(12): p. 3358-3373.

104. Lotan, T.L., et al., *DNA damage repair alterations are frequent in prostatic adenocarcinomas with focal pleomorphic giant-cell features.* Histopathology, 2019. **74**(6): p. 836-843.

105. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012. **9**(4): p. 357-9.

106. Martin, M., *Cutadapt removes adapter sequences from high-throughput sequencing reads.* 2011, 2011. **17**(1): p. 3.

107. Danecek, P., et al., *Twelve years of SAMtools and BCFtools.* Gigascience, 2021. **10**(2).

108. Vivian, J., et al., *Toil enables reproducible, open source, big biomedical data analyses.* Nat Biotechnol, 2017. **35**(4): p. 314-316.

109. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.

110. Smith, T., A. Heger, and I. Sudbery, *UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy.* Genome Res, 2017. **27**(3): p. 491-499.

111. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows–Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-1760.

112. Chen, C.D., et al., *Molecular determinants of resistance to antiandrogen therapy.* Nature Medicine, 2004. **10**(1): p. 33-39.

113. Watson, P.A., V.K. Arora, and C.L. Sawyers, *Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer.* Nature Reviews Cancer, 2015. **15**(12): p. 701-711.

114. Hu, R., et al., *Ligand-Independent Androgen Receptor Variants Derived from Splicing of Cryptic Exons Signify Hormone-Refractory Prostate Cancer.* Cancer Research, 2009. **69**(1): p. 16-22.

115. Viswanathan, S.R., et al., *Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing.* Cell, 2018. **174**(2): p. 433-+.

116. Beltran, H., et al., *The Role of Lineage Plasticity in Prostate Cancer Therapy Resistance.* Clinical Cancer Research, 2019. **25**(23): p. 6916-6924.

117. Nyquist, M.D., et al., *Combined TP53 and RB1 Loss Promotes Prostate Cancer Resistance to a Spectrum of Therapeutics and Confers Vulnerability to Replication Stress.* Cell Reports, 2020. **31**(8).

118. Berger, A., et al., *N-Myc mediated epigenetic reprogramming drives lineage plasticity in advanced prostate cancer.* Journal of Clinical Investigation, 2019. **129**(9): p. 3924-3940.

119. Zhang, X.T., et al., *SRRM4 Expression and the Loss of REST Activity May Promote the Emergence of the Neuroendocrine Phenotype in Castration-Resistant Prostate Cancer.* Clinical Cancer Research, 2015. **21**(20): p. 4698-4708.

120. Mu, P., et al., *SOX2 promotes lineage plasticity and antiandrogen resistance in TP53- and RB1-deficient prostate cancer.* Science, 2017. **355**(6320): p. 84-88.

121. Akamatsu, S., et al., *The Placental Gene PEG10 Promotes Progression of Neuroendocrine Prostate Cancer.* Cell Reports, 2015. **12**(6): p. 922-936.

122. Epstein, J.I., et al., *Proposed Morphologic Classification of Prostate Cancer With Neuroendocrine Differentiation.* American Journal of Surgical Pathology, 2014. **38**(6): p. 756-767.

123. Balanis, N.G., et al., *Pan-cancer Convergence to a Small-Cell Neuroendocrine Phenotype that Shares Susceptibilities with Hematological Malignancies.* Cancer Cell, 2019. **36**(1): p. 17-+.

124. Labrecque, M.P., et al., *Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer.* Journal of Clinical Investigation, 2019. **129**(10): p. 4492-4505.

125. Merkens, L., et al., *Aggressive variants of prostate cancer: underlying mechanisms of neuroendocrine transdifferentiation.* Journal of Experimental & Clinical Cancer Research, 2022. **41**(1).

126. Raposo, A.A.S.F., et al., *Ascl1 Coordinately Regulates Gene Expression and the Chromatin Landscape during Neurogenesis.* Cell Reports, 2015. **10**(9): p. 1544-1556.

127. Borromeo, M.D., et al., *ASCL1 and NEUROD1 Reveal Heterogeneity in Pulmonary Neuroendocrine Tumors and Regulate Distinct Genetic Programs.* Cell Reports, 2016. **16**(5): p. 1259-1272.

128. Cejas, P., et al., *Subtype heterogeneity and epigenetic convergence in neuroendocrine prostate cancer.* Nature Communications, 2021. **12**(1).

129. Costanzo, F., et al., *Promoters of ASCL1-and NEUROD1-dependent genes are specific targets of lurbinectedin in SCLC cells.* Embo Molecular Medicine, 2022. **14**(4).

130. Baine, M.K., et al., *SCLC Subtypes Defined by ASCL1, NEUROD1, POU2F3, and YAP1: A Comprehensive Immunohistochemical and Histopathologic Characterization.* Journal of Thoracic Oncology, 2020. **15**(12): p. 1823-1835.

131. Szeitz, B., et al., *In-depth proteomic analysis reveals unique subtype-specific signatures in human small-cell lung cancer.* Clinical and Translational Medicine, 2022. **12**(9).

132. DeLucia, D.C., et al., *Regulation of CEACAM5 and Therapeutic Efficacy of an Anti-CEACAM5-SN38 Antibody-drug Conjugate in Neuroendocrine Prostate Cancer.* Clin Cancer Res, 2021. **27**(3): p. 759-774.

133. Augustyn, A., et al., *ASCL1 is a lineage oncogene providing therapeutic targets for high-grade neuroendocrine lung cancers.* Proc Natl Acad Sci U S A, 2014. **111**(41): p. 14788-93.

134. Nouruzi, S., et al., *ASCL1 activates neuronal stem cell-like lineage programming through remodeling of the chromatin landscape in prostate cancer.* Nature Communications, 2022. **13**(1).

135. Ku, S.Y., et al., *Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance.* Science, 2017. **355**(6320): p. 78-83.

136. Li, Y.A., et al., *SRRM4 Drives Neuroendocrine Transdifferentiation of Prostate Adenocarcinoma Under Androgen Receptor Pathway Inhibition.* European Urology, 2017. **71**(1): p. 68-78.

137. Babos, K.N., et al., *Mitigating Antagonism between Transcription and Proliferation Allows Near-Deterministic Cellular Reprogramming.* Cell Stem Cell, 2019. **25**(4): p. 486-+.

138. Jouravel, N., et al., *Interaction between the androgen receptor and a segment of its corepressor SHP.* Acta Crystallographica Section D-Structural Biology, 2007. **63**: p. 1198-1200.

139. Judware, R. and L.A. Culp, *Over-expression of transfected N-myc oncogene in human SKNSH neuroblastoma cells down-regulates expression of beta 1 integrin subunit.* Oncogene, 1995. **11**(12): p. 2599-607.

140. Cobrinik, D., et al., *Rb induces a proliferative arrest and curtails Brn-2 expression in retinoblastoma cells.* Mol Cancer, 2006. **5**: p. 72.

141. Li, Y.N., et al., *Establishment of a neuroendocrine prostate cancer model driven by the RNA splicing factor SRRM4.* Oncotarget, 2017. **8**(40): p. 66878-66888.

142. Vue, T.Y., et al., *ASCL1 regulates neurodevelopmental transcription factors and cell cycle genes in brain tumors of glioma mouse models.* Glia, 2020. **68**(12): p. 2613-2630.

143. Han, M., et al., *FOXA2 drives lineage plasticity and KIT pathway activation in neuroendocrine prostate cancer.* Cancer Cell, 2022. **40**(11): p. 1306-1323.e8.

144. Ohnishi, T., M. Shirane, and K.I. Nakayama, *SRRM4-dependent neuron-specific alternative splicing of protrudin transcripts regulates neurite outgrowth.* Scientific Reports, 2017. **7**.

145. Tsoi, J., et al., *Multi-stage Differentiation Defines Melanoma Subtypes with Differential Vulnerability to Drug-Induced Iron-Dependent Oxidative Stress.* Cancer Cell, 2018. **33**(5): p. 890-+.

146. Richard, A., et al., *Single-Cell-Based Analysis Highlights a Surge in Cell-to-Cell Molecular Variability Preceding Irreversible Commitment in a Differentiation Process.* Plos Biology, 2016. **14**(12).

147. Basili, D., et al., *In Silico Computational Transcriptomics Reveals Novel Endocrine Disruptors in Largemouth Bass (Micropterus salmoides).* Environmental Science & Technology, 2018. **52**(13): p. 7553-7565.

148. Jiang, S., et al., *Single-nucleus RNA-seq identifies divergent populations of FSHD2 myotube nuclei.* Plos Genetics, 2020. **16**(5).

149. Karanikolas, B.D.W., M.L. Figueiredo, and L. Wu, *Comprehensive Evaluation of the Role of EZH2 in the Growth, Invasion, and Aggression of a Panel of Prostate Cancer Cell Lines.* Prostate, 2010. **70**(6): p. 675-688.

150. Cyrta, J., et al., *Role of specialized composition of SWI/SNF complexes in prostate cancer lineage plasticity.* Nature Communications, 2020. **11**(1).

151. Crispatzu, G., et al., *The chromatin, topological and regulatory properties of pluripotency-associated poised enhancers are conserved in vivo.* Nature Communications, 2021. **12**(1).

152. Stepniak, K., et al., *Mapping chromatin accessibility and active regulatory elements reveals pathological mechanisms in human gliomas (vol 12, 3621, 2021).* Nature Communications, 2021. **12**(1).

153. Whyte, W.A., et al., *Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes.* Cell, 2013. **153**(2): p. 307-319.

154. Huang, C.C.F., et al., *Functional mapping of androgen receptor enhancer activity.* Genome Biology, 2021. **22**(1).

155. Saint-Andre, V., et al., *Models of human core transcriptional regulatory circuitries.* Genome Research, 2016. **26**(3): p. 385-396.

156. Dhatchinamoorthy, K., J.D. Colbert, and K.L. Rock, *Cancer Immune Evasion Through Loss of MHC Class I Antigen Presentation.* Frontiers in Immunology, 2021. **12**.

157. Cai, L., et al., *Cell-autonomous immune gene expression is repressed in pulmonary neuroendocrine cells and small cell lung cancer.* Communications Biology, 2021. **4**(1).

158. Burr, M.L., et al., *An Evolutionarily Conserved Function of Polycomb Silences the MHC Class I Antigen Presentation Pathway and Enables Immune Evasion in Cancer.* Cancer Cell, 2019. **36**(4): p. 385-+.

159. Bluemn, E.G., et al., *Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained through FGF Signaling.* Cancer Cell, 2017. **32**(4): p. 474-+.

160. Sequist, L.V., et al., *Genotypic and Histological Evolution of Lung Cancers Acquiring Resistance to EGFR Inhibitors.* Science Translational Medicine, 2011. **3**(75).

161. Wang, Z.W., et al., *Single-cell transcriptional regulation and genetic evolution of neuroendocrine prostate cancer.* Iscience, 2022. **25**(7).

162. Brady, N.J., et al., *Temporal evolution of cellular heterogeneity during the progression to advanced AR-negative prostate cancer.* Nat Commun, 2021. **12**(1): p. 3372.

163. Abida, W., et al., *Analysis of the Prevalence of Microsatellite Instability in Prostate Cancer and Response to Immune Checkpoint Blockade.* JAMA Oncol, 2019. **5**(4): p. 471-478.

164. Hiatt, J.B., et al., *Inhibition of LSD1 with Bomedemstat Sensitizes Small Cell Lung Cancer to Immune Checkpoint Blockade and T-Cell Killing.* Clinical Cancer Research, 2022. **28**(20): p. 4551-4564.

165. Nguyen, E.M., et al., *Targeting Lysine-Specific Demethylase 1 Rescues Major Histocompatibility Complex Class I Antigen Presentation and Overcomes Programmed Death- Ligand 1 Blockade Resistance in SCLC.* Journal of Thoracic Oncology, 2022. **17**(8): p. 1014-1031.

166. Gao, D., et al., *Organoid Cultures Derived from Patients with Advanced Prostate Cancer.* Cell, 2014. **159**(1): p. 176-187.

167. Lee, J.K., et al., *Systemic surfaceome profiling identifies target antigens for immune-based therapy in subtypes of advanced prostate cancer.* Proc Natl Acad Sci U S A, 2018. **115**(19): p. E4473-E4482.

168. Martin, M., *Cutadapt removes adapter sequences from high-throughput sequencing reads.* EMBnet.journal, 2011. **17**: p. 10-12.

169. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.

170. Li, B. and C.N. Dewey, *RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome.* Bmc Bioinformatics, 2011. **12**.

171. Hanzelmann, S., R. Castelo, and J. Guinney, *GSVA: gene set variation analysis for microarray and RNA-Seq data.* Bmc Bioinformatics, 2013. **14**.

172. Gu, Z.G., R. Eils, and M. Schlesner, *Complex heatmaps reveal patterns and correlations in multidimensional genomic data.* Bioinformatics, 2016. **32**(18): p. 2847-2849.

173. Shen, S.H., et al., *rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data.* Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(51): p. E5593-E5601.

174. Howe, K.L., et al., *Ensembl 2021.* Nucleic Acids Research, 2021. **49**(D1): p. D884-D891.

175. Stuart, T., et al., *Single-cell chromatin state analysis with Signac.* Nature Methods, 2021. **18**(11): p. 1333-+.

176. Zhang, Y., et al., *Model-based Analysis of ChIP-Seq (MACS).* Genome Biology, 2008. **9**(9).

177. Cao, J.Y., et al., *The single-cell transcriptional landscape of mammalian organogenesis.* Nature, 2019. **566**(7745): p. 496-+.

178. Qiu, X.J., et al., *Reversed graph embedding resolves complex single-cell trajectories.* Nature Methods, 2017. **14**(10): p. 979-+.

179. Meers, M.P., D. Tenenbaum, and S. Henikoff, *Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling.* Epigenetics & Chromatin, 2019. **12**.

180. Li, H.-D., C.-X. Lin, and J. Zheng, *GTFtools: a software package for analyzing various features of gene models.* Bioinformatics, 2022. **38**(20): p. 4806-4808.

181. Loven, J., et al., *Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers.* Cell, 2013. **153**(2): p. 320-334.

182. Ross-Innes, C.S., et al., *Differential oestrogen receptor binding is associated with clinical outcome in breast cancer.* Nature, 2012. **481**(7381): p. 389-U177.

183. Yu, G.C., L.G. Wang, and Q.Y. He, *ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization.* Bioinformatics, 2015. **31**(14): p. 2382-2383.

184. Subramanian, A., et al., *Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(43): p. 15545-15550.

185. Aggarwal, R., et al., *Clinical and Genomic Characterization of Treatment-Emergent Small-Cell Neuroendocrine Prostate Cancer: A Multi-institutional Prospective Study.* J Clin Oncol, 2018. **36**(24): p. 2492-2503.

186. Abida, W., et al., *Genomic correlates of clinical outcome in advanced prostate cancer.* Proc Natl Acad Sci U S A, 2019. **116**(23): p. 11428-11436.

187. Matsumoto, I., et al., *Skn-1a (Pou2f3) specifies taste receptor cell lineage.* Nat Neurosci, 2011. **14**(6): p. 685-7.

188. Huang, Y.H., et al., *POU2F3 is a master regulator of a tuft cell-like variant of small cell lung cancer.* Genes Dev, 2018. **32**(13-14): p. 915-928.

189. Yamaguchi, T., et al., *Skn-1a/Pou2f3 is required for the generation of Trpm5-expressing microvillous cells in the mouse main olfactory epithelium.* BMC Neurosci, 2014. **15**: p. 13.

190. Zhao, T., et al., *Spatial genomics enables multi-modal study of clonal heterogeneity in tissues.* Nature, 2022. **601**(7891): p. 85-91.

191. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.

192. Nava Rodrigues, D., et al., *Immunogenomic analyses associate immunological alterations with mismatch repair defects in prostate cancer.* J Clin Invest, 2018. **128**(10): p. 4441-4453.

193. Brown, L.C., et al., *A phase 2 trial of avelumab in men with aggressive-variant or neuroendocrine prostate cancer.* Prostate Cancer Prostatic Dis, 2022. **25**(4): p. 762-769.

194. Hiatt, J.B., et al., *Inhibition of LSD1 with Bomedemstat Sensitizes Small Cell Lung Cancer to Immune Checkpoint Blockade and T-Cell Killing.* Clin Cancer Res, 2022. **28**(20): p. 4551-4564.

195. Nguyen, E.M., et al., *Targeting Lysine-Specific Demethylase 1 Rescues Major Histocompatibility Complex Class I Antigen Presentation and Overcomes Programmed Death-Ligand 1 Blockade Resistance in SCLC.* J Thorac Oncol, 2022. **17**(8): p. 1014-1031.

196. Bausch-Fluck, D., et al., *A mass spectrometric-derived cell surface protein atlas.* PLoS One, 2015. **10**(3): p. e0121314.

197. Hipp, S., et al., *A Bispecific DLL3/CD3 IgG-Like T-Cell Engaging Antibody Induces Antitumor Responses in Small Cell Lung Cancer.* Clin Cancer Res, 2020. **26**(19): p. 5258-5268.

198. Isobe, Y., et al., *Near infrared photoimmunotherapy targeting DLL3 for small cell lung cancer.* EBioMedicine, 2020. **52**: p. 102632.

199. Johnson, D.S., et al., *Genome-wide mapping of in vivo protein-DNA interactions.* Science, 2007. **316**(5830): p. 1497-502.

200. Skene, P.J. and S. Henikoff, *An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites.* Elife, 2017. **6**.