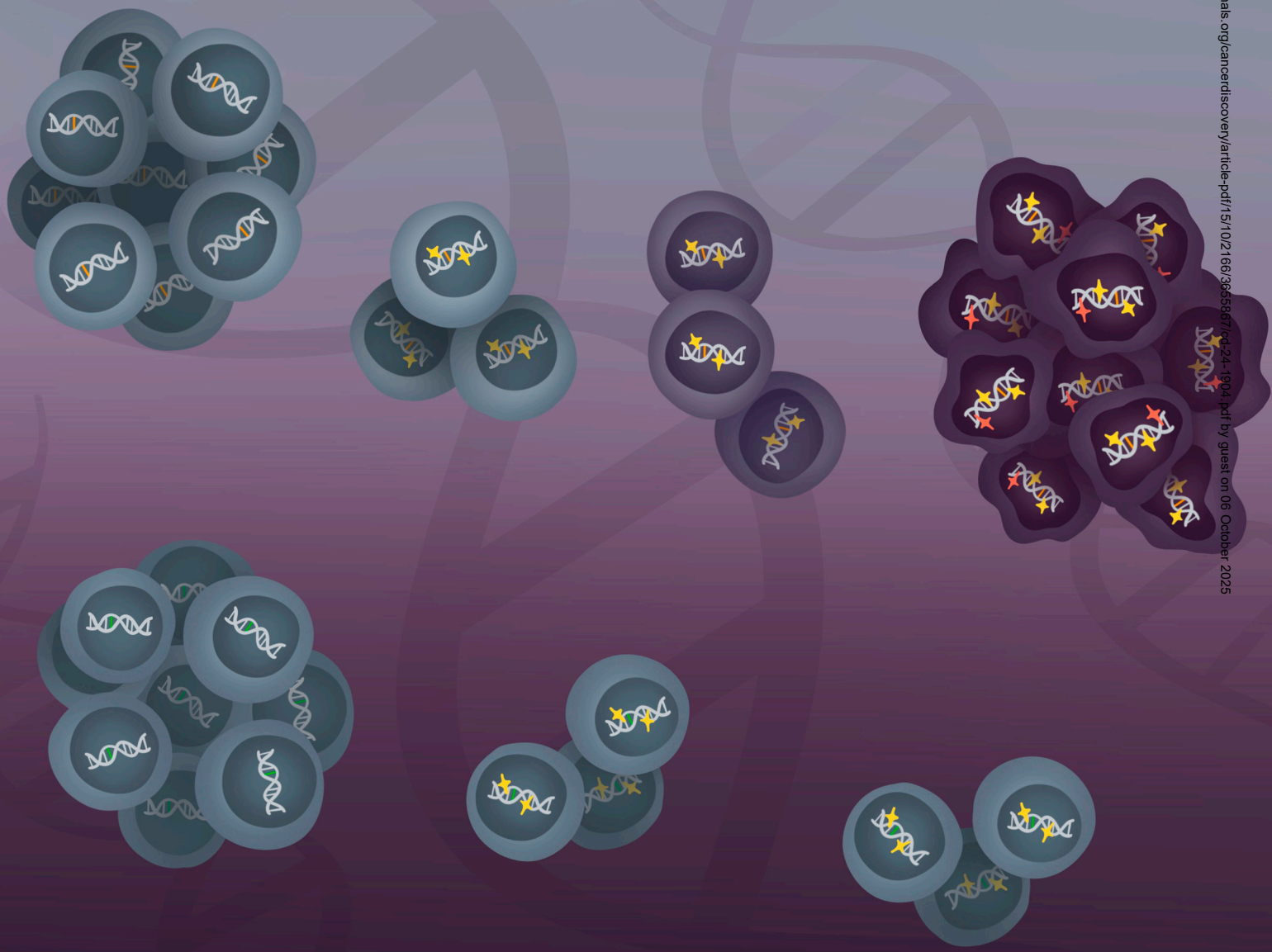
RESEARCH ARTICLE

Germline–Somatic Liaison Dictates Cancer Subtypes via *de novo* Steroid Biosynthesis

Paola Gasperini[1], Alessandro Alaimo[1], Blerta Stringa[1], Yoon-Mi Chung[2], Yari Ciani[1], Francesca Lorenzin[1], Giulia Fracassi[1], Yanis Zekri[3,4], Francesco Orlando[1], Orsetta Quaini[1], Sebastian Gregoricchio[3,4], Gianluca Petris[1,5,6], Antonio Casini[7], Christopher E. Barbieri[8,9,10], Wilbert Zwart[3,4,11], Anna Cereseto[1], Nima Sharifi[2], Andrea Lunardi[1], and Francesca Demichelis[1]



ABSTRACT The biological mechanisms underlying the cooperation between germline genetic variants and somatic mutations during carcinogenesis are rarely elucidated. In this study, characterizing isogenic prostate cancer cell lines, we dissected the interplay between a germline variant at the *7p14.3* locus (rs1376350, G>A) and early recurrent prostate cancer–specific mutation in the speckle-type POZ protein (*SPOP*) gene across human prostate adenocarcinomas. The transcriptomes of multiple edited models pointed to GLI3 and the Hedgehog signaling pathway in a genotype-specific manner, whereas *SPOP* mutation and androgen receptor stimulation promote GLI3 accumulation in the full-length, transcriptionally active form. This, in turn, triggers the cell-autonomous production of steroids that prostate cancer relies on, in line with the enhanced responsiveness of *SPOP*-mutated prostate cancer to androgen deprivation therapy. These data demonstrate that germline variants dictate prostate cancer somatic evolution and suggest opportunities to jointly model germline–somatic relationship to help untangle the complexity of human cancer.

SIGNIFICANCE: Significant heritability is observed for common cancer types worldwide. The molecular mechanisms by which inherited genetics facilitate cancer initiation might transit through its cooperation with specific somatic events that then dictate the tumor features. Through a germline–somatic tandem leading to steroid biosynthesis, we suggest a paradigm to study cancer initiation.

INTRODUCTION

Cancer hallmarks and the abundance of polymorphisms of the germline genome shape the complexity of human cancer and augment the variability in the treatment response of individuals with the same clinical features at diagnosis (1, 2). Estimates of familial cancer risk from population-based studies, such as large twin cohorts, inform cancer risk prediction (3). Significant heritability (i.e., the proportion of variance in cancer risk due to inter-individual genetic differences) has been observed for two of the five most common cancer types worldwide (4), i.e., prostate [57%; 95% confidence interval (CI), 51%–63%] and breast (31%; 95% CI, 11%–51%) cancers, with more than 1.4 and 2.3 million new cases diagnosed in 2022, respectively. Despite strong evidence of this genetic component and the identification of hundreds of germline risk variants, there is still a substantial knowledge gap between an individual's genetic makeup and the full understanding of how risk variants contribute to human diseases (5–7).

Prostate cancer, a leading cause of cancer death among men (responsible for 4.1% of all cancer deaths; ref. 4), characterized by diverse clinical outcomes and unknown biological mechanisms behind its etiology, represents an enigmatic example of this unresolved germline–somatic dilemma. The etiology of prostate cancer is multifactorial, including aging, environmental risk factors, and genetics (3). Large association studies provide evidence for hundreds of genetically inherited variants linked with prostate cancer risk with modest size effects, including SNPs, insertions/deletions, and structural changes (8–10), followed by the more recent discovery of penetrant rare pathogenic variants in DNA damage repair genes such as *BRCA2*, *BRCA1*, *ATM*, *MLH1*, *MSH2*, and *MSH6* (11) and the pioneering transcription factor *HOXB13* (12). Additional germline variations, including a common missense-encoding variant in *HSD3B1,* which encodes for the 3β-hydroxysteroid dehydrogenase-1 (3βHSD1) enzyme, confers poor outcomes after hormonal therapy and increases prostate cancer–specific mortality (13, 14). At diagnosis, prostate cancer presents with somatic, noninherited, aberrations that define specific distinct molecular subclasses (15, 16), which still hold in the advanced metastatic setting (17). We thus hypothesized a cooperative effect between inherited variants and somatic prostate cancer–specific genomic events that arise early during carcinogenesis (18), which helps explain the molecular pathogenetic processes behind the disease etiology. Through large human data–driven analyses, we identified a germline–somatic mutation association between an inherited noncoding

[1]Department of Cellular, Computational and Integrative Biology, University of Trento, Trento, Italy. [2]Desai Sethi Urology Institute and Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, Florida. [3]Division of Oncogenomics, The Netherlands Cancer Institute, Amsterdam, the Netherlands. [4]Oncode Institute, Utrecht, the Netherlands. [5]Genome Engineering and Biotechnology Unit, Fondazione Italiana Fegato, AREA Science Park, Trieste, Italy. [6]Department of Medicine, University of Udine, Udine, Italy. [7]Alia Therapeutics, Trento, Italy. [8]Department of Urology, Weill Cornell Medicine, New York, New York. [9]Sandra and Edward Meyer Cancer Center, Weill Cornell Medicine, New York, New York. [10]Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, New York. [11]Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands.

P. Gasperini, A. Alaimo, and B. Stringa are co-first authors and contributed equally to this article.

A. Lunardi and F. Demichelis are co-last authors and contributed equally to this article.

Corresponding Author: Francesca Demichelis, Department of Cellular, Computational, and Integrative Biology, University of Trento, Via Sommarive 9, Trento 38123, Italy. E-mail: f.demichelis@unitn.it

Cancer Discov 2025;15:2166–84

doi: 10.1158/2159-8290.CD-24-1904

polymorphic regulatory element at *7p14.3* (rs1376350) and early recurrent prostate cancer–specific somatic mutations in the speckle-type POZ protein (*SPOP*) gene (Fig. 1A; ref. 19). SPOP, a cullin 3–based ubiquitin ligase, is mutated in about 10% of patients with prostate cancer across multiple ethnicities (16, 20, 21). SPOP is responsible for the ubiquitination of several proteins, including the key androgen receptor (AR) coactivator steroid receptor coactivator-3 and AR itself (22–24), in line with the observed enhanced AR signaling in *SPOP*-mutant prostate cancer. This well-defined subclass of primary prostate cancer is characterized by expanded genomic instability (25), worse prognosis (26), and increased dependency on AR-mediated oncogenic signaling (27), in line with enhanced responsiveness to androgen deprivation therapy (ADT; refs. 28–30).

In this study, we sought to investigate the liaison between the *7p14.3* variant and the *SPOP*-mutant prostate cancer phenotype and identified the transcription factor GLI3 as the intermediate. GLI3 is an effector of the Hedgehog (Hh) signaling pathway. Whereas most noted for its involvement in embryogenesis, inappropriate Hh signaling is increasingly viewed as a factor in the development of human malignancies and a key driver of prostate cancer cell growth by controlling the expression of limiting enzymes of steroid metabolism. Specifically, GLI3 promotes *CYP11A1* gene transcription, whereas GLI2 mediates 3βHSD1 and aromatase expression (31, 32). Interestingly, functional AR physically binds the transcriptionally active form of GLI3 (full-length GLI3, GLI3-FL; refs. 33–35), protecting it from SPOP-mediated ubiquitination and proteasomal cleavage (36, 37). The intermediates between the *7p14.3* variant and the *SPOP*-mutant prostate cancer phenotype rely on a GLI3-mediated steroid hormone cascade that, in return, promotes a hyperactive GLI3–SPOP–AR axis that sustains AR-dependent tumorigenesis.

These data demonstrate that a germline-inherited variant dictates a prostate cancer somatic subclass by increasing the cell-autonomous fitness of mutated cells. This work provides a detailed proof-of-concept as to how the germline genome can be interrogated for associations with somatic endpoints that characterize human cancers as cancer-specific subclasses to understand the biological mechanisms of tumor initiation and complexity.

RESULTS

The Nonancestral Allele at *7p14.3* Favors Locus Enhancer Activity Independently of AR Binding

In silico correlative investigations of germline variants and early prostate cancer–specific somatic aberrations nominated a SNP in a noncoding region of *7p14.3* associated with a recurrent mutation in the *SPOP* gene (OR = 5.54; P = 1.22e−08; ref. 19), harbored in about 10% of the cases (Supplementary Fig. S1A). The inherited variant (rs1376350, G>A) has a minor allelic frequency of 1.59% in the total population (ALFA allele frequency, release version: 20201027095038) and resides within a noncoding regulatory element between two genes with no known biological connection to the prostate gland (~500 kb upstream by *BMPER* and ~250 kb downstream by *BBS9*; Supplementary Fig. S1B). A noncanonical half-site AR-binding motif overlapping the inherited variant locus (M00962) with a predicted higher affinity for the alternative allele A (0.82 vs. 0.18 as La/Lm ratio, P value = 0.0009) and a second noncanonical AR half-binding site at 76 bp downstream were previously reported (19). To functionally characterize the polymorphic *7p14.3* locus, CRISPR–Cas9 technology was applied in four sets of DNA editing experiments to disrupt or alter the locus (Fig. 1B) through the following: (i) a macrodeletion of 731 bp (Supplementary Fig. S1C) and (ii) a microdeletion of 50 bp spanning the locus in PC3 (Supplementary Fig. S1D) and LNCaP prostate cancer cell lines; (iii) the disruption of the *in silico* predicted AR-binding motif around the locus in PC3 cells (Supplementary Fig. S1E), and (iv) the generation of isogenic cells harboring the heterozygous (AG) or the homozygous (AA) genotype for the alternative allele at rs1376350 in PC3 cells (Fig. 1C; Supplementary Fig. S1F–S1J; Supplementary Tables S1–S3).

We first tested AR binding to the rs1376350 variant site in edited clones. Chromatin immunoprecipitation (ChIP) experiments for AR in GG, AG, and AA isogenic PC3 clones expressing exogenous AR and stimulated with dihydrotestosterone (DHT) demonstrated no binding within a 103-bp region around the variant (rs1376350_1, chr7:33692717–33692819, GRChg38; Fig. 1D, top). No marked signal was observed in a wider 160-bp stretch of DNA, including the second AR-binding motif (rs1376350_2, chr7:33692714–33692873; Supplementary Fig. S1K, left). As expected, AR bound the *FKBP5* enhancer when stimulated with DHT (Fig. 1D, bottom) but not a desertic region lacking AR consensus sequences (Supplementary Fig. S1K, right). To exclude that chromatin organization was limiting binding of AR to rs1376350, we used a biotinylated DNA pull-down assay to verify AR binding to a DNA probe carrying the sequence of the noncanonical AR half-site overlapping the inherited variant locus (M00962) with the alternative allele A (7p14.3_A) or with the ancestral allele G (7p14.3_G). Consistent with the ChIP data, both probes carrying either the A or G variant (position 15, corresponding to rs1376350) showed a lack of AR affinity, whereas the AR efficiently bound to 28 nucleotide probes carrying the canonical AR half-site motif AR M08908 (AR_hs_A). In this study, a single substitution from A to G of nucleotide 15 abrogated AR binding (AR_hs_G; Fig. 1E, lanes 7p14.3_G and 7p14.3_A vs. AR_fs used as a positive control; Supplementary Table S1). As the *in silico* prediction identified a CEBP-binding motif overlapping to the rs1376350 locus (M00770), we mutated two additional nucleotides (7p14.3_G* and 7p14.3_A*) to prevent the potential interference of CEBPβ on AR binding. Nevertheless, AR's affinity for the region was not changed. Altogether, these *in vitro* results indicate that AR does not directly bind the rs1376350 locus. Finally, interrogation of AR ChIP sequencing (ChIP-seq) data from human patients with prostate cancer (n = 167) in the DARANA (38), POMERANTZ (39), and PORTO (40) cohorts further supported the absence of AR binding at the rs1376350 locus (Fig. 1F). Investigations intersecting the human variant map and the Encyclopedia of DNA Elements histone mark ChIP-seq data (41) nominated the *7p14.3* variant locus as a putative functionally active region. ChIP studies for the monomethylation and trimethylation of H3K4 showed high and low levels, respectively, around the locus in wild-type (WT) PC3 cells (Fig. 1G), a feature typical of enhancer regions. To further characterize the epigenetic

status of the locus and the possible impact of the nucleotide variant, we analyzed acetylation and trimethylation of H3K27 in the PC3-edited clones expressing exogenous WT AR. Indeed, whereas H3K4me1 marks both active and poised/neutral enhancers, H3K27Ac only marks active ones and serves to distinguish active from poised enhancers.

A significant increase in H3K27Ac levels (P value ≤ 0.05) was observed in AA_1 compared with GG_1 (Fig. 1H, left). Conversely, the signal for H3K27me3 was significantly lower in AA_1 compared with GG_1 (P value ≤ 0.001; Fig. 1H, right), suggesting that the region is more active with A allele than G. These results indicate that the *7p14.3* locus harbors epigenetic marks characteristic of enhancers, and its functional activity is higher for the alternative allele A compared with the ancestral allele G, independently from AR activation.

7p14.3 Locus Controls the Expression of Hormone Metabolism Genes

To investigate the role of the *7p14.3* locus as a transcriptional enhancer and the impact of the variant on its activity, we first queried the transcriptome of bulk PC3 and LNCaP edited cells via RNA sequencing (RNA-seq). Global analysis demonstrated that, the wider the genomic editing, the larger the set of differentially expressed genes (DEG). For instance, eight times more DEGs (n = 800; FDR < 0.05) were observed in macrodeleted PC3 cells with respect to single-nucleotide changes in both alleles (AA), an effect also observed with the microdeletion (Supplementary Fig. S2A; Supplementary Table S4).

Next, we profiled multiple isogenic clones to specifically nominate the transcriptional signal associated with the variant. Regardless of the expected high correlation of the RNA-seq profiles (Fig. 2A; Supplementary Fig. S2B), reflected in the clustering of the clones irrespective of the edited single nucleotide, the isogenic clone analysis allowed for the identification of pronounced deregulation when comparing AA or AG versus GG clones (1,062 and 642 genes, FDR = 0.05, respectively; Fig. 2B; Supplementary Tables S5 and S6). Genotype-specific differential expression of selected transcripts was confirmed by RT-qPCR, including *NT5E*, *CLDN3*, and *TGFA* (Fig. 2C). The concordance between AA and AG clones' fold change values is high (genes with FDR <0.05 in AG vs. GG comparison, Spearman correlation test correlation = 0.51 P value = 2.2e−16), overall suggesting a coherent transcriptional activity influenced by the presence of at least one adenine at the rs1376350 locus (Fig. 2B).

To assess this transcriptional signal in the context of the human prostate cancer tissue transcriptome, we evaluated the joint set of AA and AG DEGs (a total of 1,516 unique DEGs) in a cohort of prostate cancer data (n = 285, The Cancer Genome Atlas + Weill Cornell Medicine (WCM)), stratifying patients based on *SPOP* mutation status and identified a significant (FDR <0.05) difference in expression levels for 218 genes (Fig. 2D; Supplementary Fig. S2C; Supplementary Table S7). Gene ontology analysis suggested that genes deregulated by the variant and *SPOP* status are enriched for progesterone and hormone metabolic processes (Fig. 2E; Supplementary Table S8).

7p14.3 Variant and SPOP Mutant Functions Converge Toward the Hh Pathway

To nominate the core player of the *7p14.3* locus biological function, we queried which genes were commonly affected by the different degrees of genomic manipulations across the cell lines (i.e., macrodeletion in PC3 and LNCaP, microdeletion in PC3, and single-nucleotide editing in PC3 clones; Fig. 3A). The intersection of the DEGs affected by the diverse editing strategies pointed to a single gene, *GLI3*, a transcription factor of the GLI family with crucial functions in the Hh signaling pathway (Fig. 3A; ref. 42). The *GLI3* gene is located on chromosome 7p14.1 (chr7: 41,960,949–42,264,268; GRCh38) at a distance of 8 Mb from the variant. GLI3 protein is a known SPOP target (36, 37), and it has been involved in prostate cancer progression (33, 35, 43). *GLI3* mRNA expression in PC3-edited clones by RT-qPCR confirmed the RNA-seq data (Fig. 3B, left). The G to A substitution increased *GLI3* mRNA levels in both AG and AA clones compared with clones homozygous for the ancestral allele G. We then assessed *GLI3* mRNA levels in PC3 and LNCaP pools and clones harboring different sizes of deletions encompassing the variant of interest. We found that the deletion negatively affected the level of *GLI3* expression and that the smaller the deletion, the stronger the effect on the transcript expression in both PC3 and LNCaP cells (Fig. 3B, middle and right).

The Hh signaling pathway is primarily known for its role in the development and regeneration of a wide spectrum of human and animal tissues (44, 45). The transcriptional machinery of Hh signaling relies on the transcription factors of the GLI family. GLI3 plays a substantial role, acting as a transcriptional activator when the protein is full length (GLI3-FL) and as a transcriptional repressor (GLI3-R) in its shorter cleaved form (46). In prostate cancer cells, both AR and SPOP interact with GLI3: active AR protects GLI3-FL from cleavage, whereas WT SPOP promotes the formation of the shorter isoform of GLI3 (GLI3-R; Fig. 3C; refs. 33–35). To elucidate the molecular circuit connecting GLI3 to AR and SPOP in prostate cancer, AR^+ LNCaP and 22Rv1 and isogenic AR^- and AR^+ PC3 cell lines were transduced with lentiviral vectors expressing SPOP WT and SPOP mutant (F133V). Consistent with its role as a transcriptional activator, GLI3-FL accumulation increased the expression of its downstream targets GLI1 and PTCH1 (Fig. 3D). GLI3 is highly expressed in all tested prostate cancer cell lines (Fig. 3D), with GLI3-FL being stabilized by (i) DHT-stimulated AR or constitutively active AR splicing variant 7 (as in 22Rv1 cells), (ii) expression of SPOP F133V, and (iii) the combination of active AR and SPOP mutant with partial additive effect, whereas, as expected, the overexpression of SPOP WT resulted in the accumulation of GLI3-R.

Genomic and transcriptomic analyses of human prostate cancer The Cancer Genome Atlas datasets have identified the nucleotide variant at the *7p14.3* locus associated with *SPOP* mutations and AR activity (19). To decipher the role of this association, SPOP WT and SPOP F133V were stably expressed in GG, AG, and AA PC3 clones. Both SPOP F133V and active AR increased the GLI3-FL content in a *7p14.3* allele–specific manner, with AG and AA clones expressing the highest levels of GLI1 and PTCH1 (Fig. 3E and F; Supplementary Fig. S3A). As the macrodeletions and microdeletions encompassing the

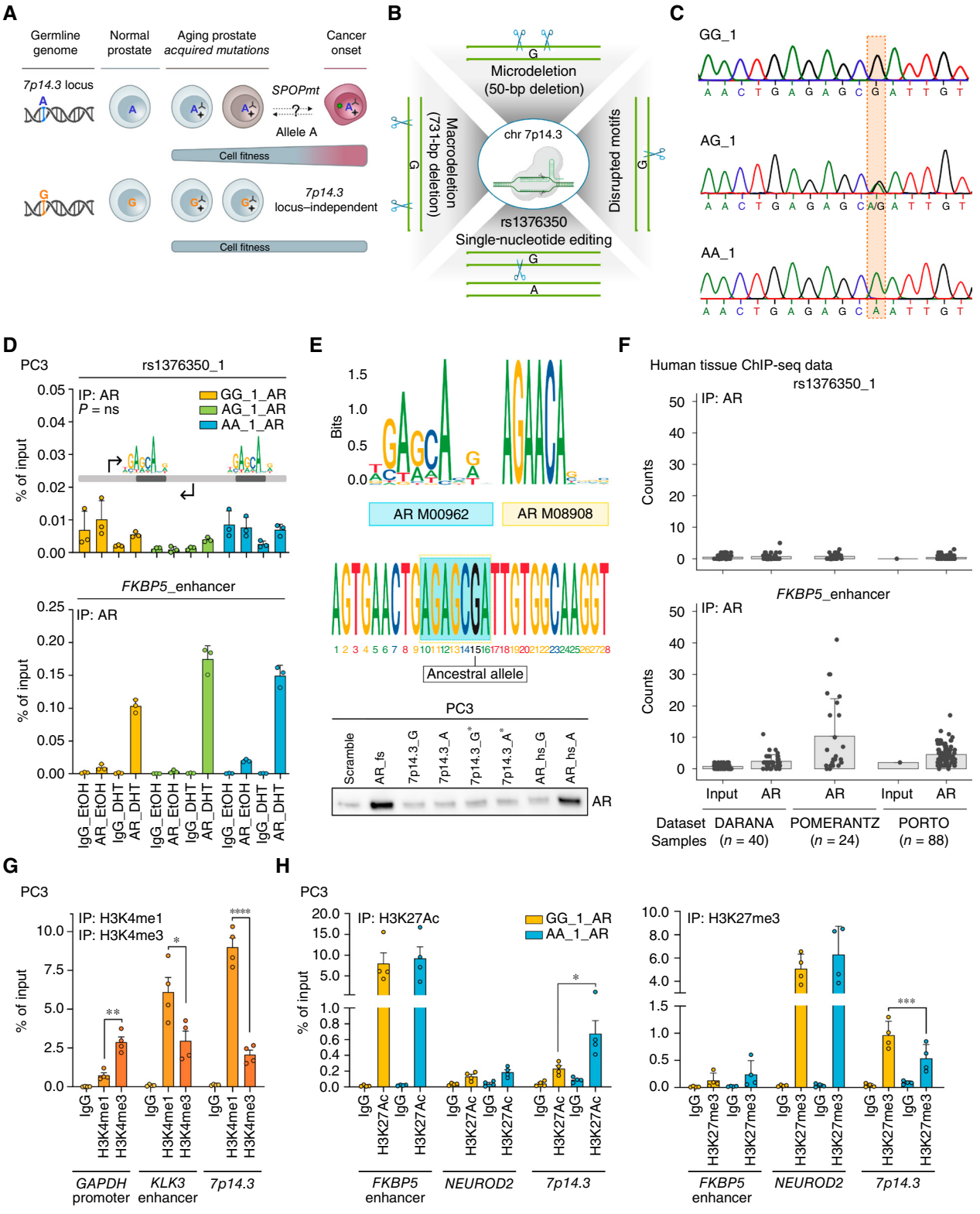


Figure 1. *7p14.3* locus editing and characterization. **A,** Human data–driven analysis identified a germline–somatic mutation association between an inherited noncoding polymorphic regulatory element at *7p14.3* (rs1376350) and early recurrent prostate cancer–specific somatic mutations in the *SPOP* gene. Schematic representation of the working hypothesis regarding the molecular nature of this pathologic liaison in prostate cancer onset. **B,** Schematic representation of the implemented editing solutions for generating *7p14.3* edited cell pools and/or single clones. **C,** Chromatograms of three representative prostate cancer cell clones are shown. The frame indicates the rs1376350 nucleotide of interest. **D,** ChIP-qPCR results for AR in PC3 representative edited clones (GG, AG, and AA) overexpressing AR upon EtOH or DHT treatment. (*continued on following page*)

locus negatively affected *GLI3* transcription in LNCaP cells (Fig. 3B), we overexpressed SPOP WT and SPOP F133V in LNCaP with microdeletion and the corresponding control to verify the effect on GLI3 and its targets. As expected, SPOP F133V and activated AR were less effective in stabilizing GLI3-FL and its targets in microdeleted LNCaP cells than in LNCaP cells with a WT *7p14* locus (Fig. 3G). Finally, to confirm the direct involvement of the Hh signaling pathway, *SPOP*-mutated single-nucleotide edited PC3 clones were incubated with cyclopamine, a transmembrane protein Smoothened-specific inhibitor. Treatment with cyclopamine triggered the processing of GLI3-FL to GLI3-R form and the reduction of GLI1 and PTCH1 levels in an allele-specific manner (Supplementary Fig. S3B). These results confirm a functional link between the *7p14.3* rs1376350 variant, the SPOP F133V mutation, and the AR, with the inherited variant increasing *GLI3* expression and *SPOP* mutation and AR activation preventing GLI3 cleavage across models.

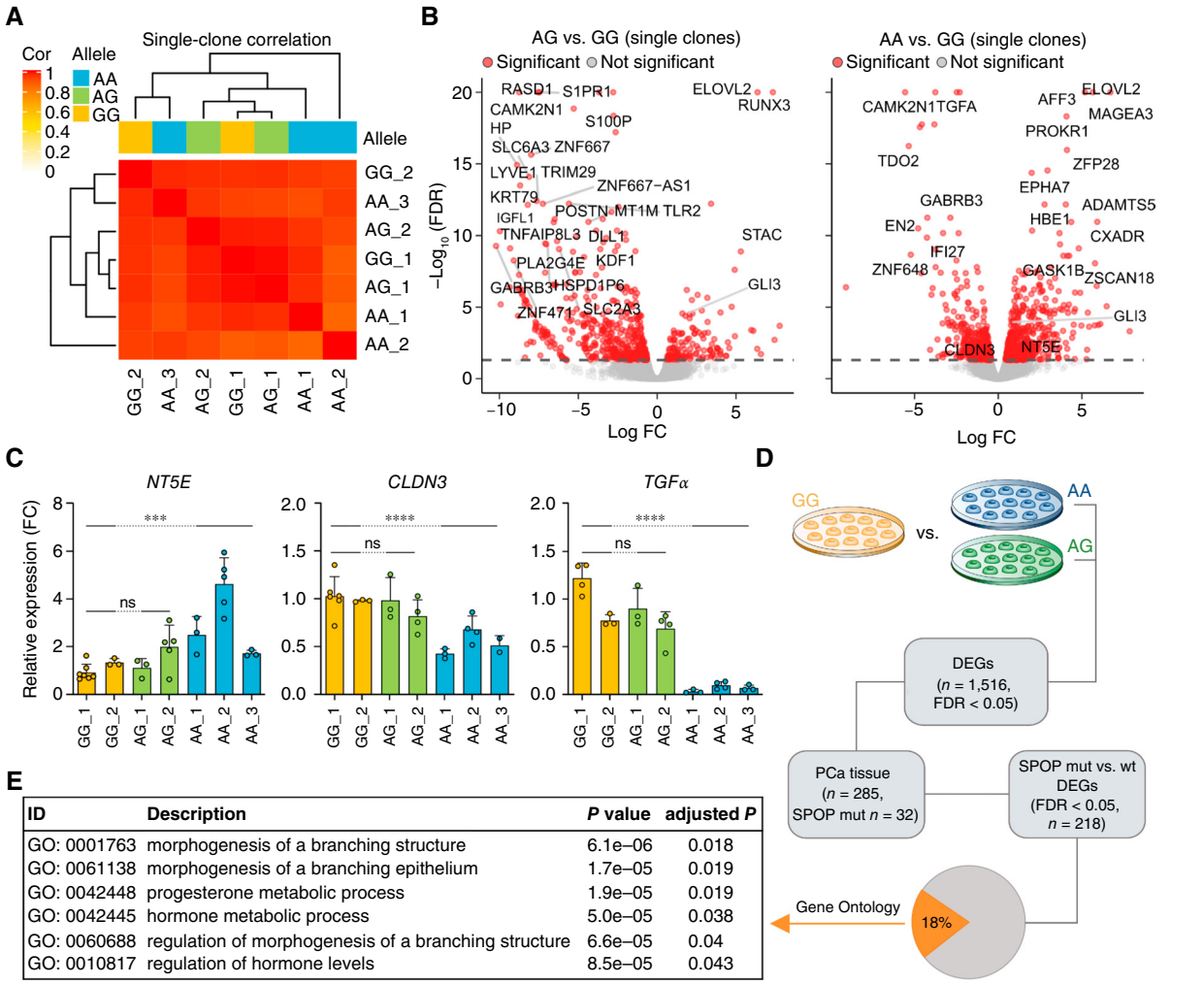
Allele-Enhanced Pathologically Stabilized GLI3-FL Promotes Cholesterol Metabolism and *de novo* Steroidogenesis

Hh signaling activity regulates cholesterol metabolism and is important in steroidogenesis (47). In response to the activation of the Hh pathway, GLI3-FL induces the expression of the *CYP11A1* gene encoding for the P450 cholesterol side-chain cleavage enzyme (P450scc), which is essential for the conversion of cholesterol to pregnenolone (Fig. 4A; refs. 48, 49). Moreover, Hh signaling is involved in the regulation of the *HSD3B1* gene encoding for the 3βHSD1 isoenzyme dominantly expressed in the prostate that converts pregnenolone to progesterone (31). To test whether GLI3-FL increased activity induced by the *7p14.3* locus together with mutant *SPOP* and AR activation affected steroidogenesis in our cell models, we first analyzed P450scc and 3βHSD1 protein levels by Western blot analysis. As shown in Fig. 4B, *CYP11A1*, generally undetectable in PC3 cells (50), was induced by both overexpression of SPOP F133V and DHT stimulation with a synergistic effect of the two signals, consisten with GLI3-FL stabilization. The same pattern was observed in LNCaP and 22Rv1 cells stimulated with DHT and expressing SPOP F133V (Supplementary Fig. S4). Strikingly, expression of SPOP F133V in single nucleotide–edited PC3 cells enhanced the *CYP11A1*-encoded P450scc protein levels in an allele-specific manner, with AG and AA clones characterized by increasing levels of the enzyme compared with GG clones (Fig. 4C). Even more interesting, 3βHSD1 protein—undetectable in PC3 cells bearing the G variant regardless of either AR stimulation or SPOP F133V expression (Fig. 4B)—became detectable in AG and AA PC3 clones expressing SPOP F133V (Fig. 4C). Because microdeletion encompassing the variant of interest negatively affected *GLI3* mRNA and protein expression in LNCaP cells (Fig. 3B and G), we tested whether the same was true for GLI3-FL targets P450scc and 3βHSD1. P450scc protein expression was negatively affected by the microdeletion, consistent with the expected correlation between GLI3-FL stabilization and the expression of its targeted genes, whereas 3βHSD1 was unaffected (Fig. 4D). This result is consistent with the observation that LNCaP cells express a mutated form of 3βHSD1 resistant to the proteasome and AMFR-meditated degradation (51) that can determine the accumulation of the protein.

To verify that increased expression of P450scc protein had a biological impact on the accumulation of its substrate, potentially affecting androgen steroids, we measured the levels of pregnenolone released by PC3 in cell-conditioned media by ELISA. As expected, cells with GLI3-FL accumulation—by mutant SPOP and/or activated overexpressed AR—released more pregnenolone into the culture supernatant, confirming the previous finding (Fig. 4E). On the contrary, ELISA measurement of pregnenolone levels in single nucleotide–edited PC3 clone conditioned medium showed that AG and AA clones released less pregnenolone than GG cells in the supernatant (Fig. 4F). In line with previous observations on high cellular retention of pregnenolone (52), we hypothesized that pregnenolone was not released in the cell media either because it was quickly transformed in progesterone and/or immediately used inside the cell.

We performed LC-MS/MS analysis and quantification of steroids in the media of our edited clones expressing either SPOP WT or SPOP F133V following the pregnenolone spike and observed progesterone accumulation in AG edited clones peaking at 12 hours after spiking in the presence of SPOP F133V and at 24 hours in the presence of SPOP WT. We did

Figure 1. ***(Continued)*** The inset at the top is a schematic representation of putative AR binding sites at the variant of interest; rs1376350_1 (106 bp) encompasses one AR-binding site generated by the presence of the alternative allele A. The *FKBP5* enhancer (bottom) is shown as a positive control and IgG as a negative control. Data are represented as the mean + SD. ($n = 3$ technical replicates). The *P* value for rs1376350_1 was determined using an unpaired *t* test on IgG_EtOH (GG_1, AG_1, and AA_1 clones) and AR_DHT (GG_1, AG_1, and AA_1 clones). ns, not significant. **E,** Schematic representation of half-site binding motifs for AR in a 28-oligo stretch centered around rs1376350 in the *7p14.3* locus. DNA pull-down assays followed by Western blot analysis for AR. Nuclear proteins used for AR pull-down were extracted from the PC3 cell line overexpressing AR and immunoprecipitated with biotinylated 28-nucleotide oligo probes (Supplementary Table S1) carrying the sequence of the noncanonical AR half-site overlapping the inherited variant locus (M00962) with the alternative allele A (7p14.3_A) or with the ancestral allele G (7p14.3_G) or carrying the canonical AR half-site motif AR M08908 (AR_hs_A). Single substitution from (A to G) of nucleotide 15 abrogated AR binding (AR_hs_G), whereas the substitution of two additional nucleotides (7p14.3_G* and 7p14.3_A*) to prevent the potential interference of CEBPβ does not affect AR affinity for the sequence. Scramble probe was used as a negative control; AR_fs containing the canonical DNA full-site binding motif M08907 was used as a positive control. **F,** Read counts at the rs1376350_1 (top) and *FKBP5* enhancer (bottom) regions from AR ChIP-seq in human prostate samples obtained from the DARANA, POMERANTZ, and PORTO datasets. The *FKBP5* enhancer is shown as a positive control. Barplots represent data as the mean + SD. **G,** ChIP-qPCR results for H3K4me1 and H3K4me3 in PC3. *GAPDH* promoter and *KLK3* enhancer are shown as promoter and enhancer region-positive controls, respectively. Data are represented as the mean + SEM ($n = 4$ biological replicates). *P* values were determined using an unpaired *t* test. **H,** ChIP-qPCR results for H3K27Ac and H3K27me3 as markers of activation and repression in GG and AA representative PC3 clones overexpressing AR. *FKBP5* enhancer was selected as positive and negative controls for H3K27Ac and H3K27me3 and *NEUROD2* as negative and positive controls for H3K27Ac and H3K27me3, respectively. Data are represented as the mean + SEM ($n = 4$ biological replicates). *P* values were determined using an unpaired *t* test on fold change (GG vs. AA clones). *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$. [**A,** Partially created in BioRender. Demichelis, F. (2025) https://BioRender.com/hdkj918 and **B,** Partially created in BioRender. Demichelis, F. (2025) https://BioRender.com/ldo56cd.]



E

ID	Description	*P* value	adjusted *P*
GO: 0001763	morphogenesis of a branching structure	6.1e−06	0.018
GO: 0061138	morphogenesis of a branching epithelium	1.7e−05	0.019
GO: 0042448	progesterone metabolic process	1.9e−05	0.019
GO: 0042445	hormone metabolic process	5.0e−05	0.038
GO: 0060688	regulation of morphogenesis of a branching structure	6.6e−05	0.04
GO: 0010817	regulation of hormone levels	8.5e−05	0.043

Figure 2. Transcriptome analysis of PC3 edited cells. **A,** RNA-seq data correlation (Cor) map of single-nucleotide edited clones, evidencing strong correlations. **B,** Volcano plots representing the DEGs in AG (left) and AA (right) vs. GG clones. **C,** Real-time validation of selected RNA-seq upregulated and downregulated transcripts in two GG, two AG, and three AA PC3 clones. Data are represented as fold changes (FC) compared with the mean of GG clones and expressed as the mean + SD. ($n \geq 3$ biological replicates). *P* values were determined using one-way ANOVA followed by Bonferroni's multiple comparisons test, comparing AG and AA samples with GG samples. ***, $P \leq 0.001$; ****, $P \leq 0.0001$. ns, not significant. **D,** Comparative analysis of differential genes from AG vs. GG and AA vs. GG clones in a human prostate cancer (PCa) tissue samples cohort upon *SPOP* mutation stratification. **E,** Enriched Gene Ontology terms from the comparative clones/tissue analysis genes per (**D**) workflow [DEGs in AG vs. GG or AA vs. GG and differential in *SPOP* mutant (mut) vs. WT].

not see progesterone accumulation for GG or AA edited clones (Fig. 5A). The results observed in GG and AG are consistent with 3βHSD1 detection: its absence in GG clones prevented the transformation in progesterone, whereas its presence in AG clones allowed the efficient conversion of pregnenolone to progesterone. The absence of progesterone in the media of AA clones might suggest that, in this setting, progesterone was more efficiently used inside the cell or metabolized into other androgen steroids.

We first verified progesterone receptor (PR) expression and activation in prostate cancer cell lines to test the hypothesis of differential use of intracellular progesterone. Although WT LNCaP and PC3 cell lines do not express PR at a detectable level, it is known from the literature that PR is an SPOP target (53). Therefore, we reasoned that SPOP F133V overexpression could sustain PR accumulation, as confirmed by Western blot analysis (Supplementary Fig. S5A; ref. 53). We then tested progesterone biological activity in engaging mutant SPOP–stabilized PR and inducing its translocation to the nucleus of PC3 edited cells. A higher amount of PR was detected in the nuclear fraction of AG and AA edited cells than in GG nonedited cells (Fig. 5B), suggesting that PR can be more active in these conditions.

To test the hypothesis that progesterone, like pregnenolone, might not be released in the cell media as a result of quick conversion into other androgen steroids, we used mass spectrometry to analyze the cellular fraction of the edited clones expressing either SPOP WT or mutant SPOP F133V spiked with pregnenolone. Whereas the level of pregnenolone observed in these fractions after 72 hours was comparable between edited clones, the levels of the other androgen steroids were not. Specifically, as previously observed in the cell media,

progesterone accumulated in AG clones while showing the lowest level in AA clones, in which it was quickly transformed into DHEA, androstenediol (A5-diol), and DHT (Fig. 5C and D), all detectable in edited clones only. Accordingly, A5-diol and DHT were detected only in AA clones. No marked differences were observed in steroid levels with respect to *SPOP* status.

These observations suggested that the *7p14.3* genotype and *SPOP* status could affect enzymes involved in *de novo* androgen biosynthesis, previously not linked to GLI3-FL stabilization, such as (i) CYP17A1, responsible for pregnenolone conversion to DHEA; (ii) 3βHSD1, mainly expressed in peripheral tissues and responsible for androstenediol conversion in testosterone; and (iii) SRD5A1, which converts testosterone to DHT and also converts progesterone to downstream metabolites (54). In our setting, the assessment of those enzymes in SPOP-overexpressing PC3 edited clones showed CYP17A1 expression in AA only (Fig. 5E) together with a modest SRD5A1 enhancement by both the presence of the allele A and SPOP F133V (Supplementary Fig. S5B). Whereas 3βHSD1 was affected both by the genotype and the mutation (Fig. 4C), 3βHSD2 expression was unaltered (Supplementary Fig. S5B). Finally, nucleocytoplasmic fractionation coupled with Western blot analysis showed nuclear AR and concomitant expression of FKBP5 and NKX3.1, but not kallikrein-2 and PSA, proteins only in AA clones (Fig. 5F), thus supporting the biological relevance of intratumor steroids biosynthesis. Although we have direct evidence that single-nucleotide substitution affects CYP17A1 expression, we recognize that the intricate map of androgen biosynthesis likely includes other players not controlled by the variant genotype but involved in sustaining the production of different hormones.

DISCUSSION

With the current knowledge of the genomic landscape of human tumors and the full understanding of the germline human genome, the discovery of the interactions of inherited polymorphisms and genomic somatic events is desirable not only to uncover the whole picture but also to guide tumor prevention research. GWAS and familial studies nominate inherited variants associated with cancer risk. However, there is still a substantial knowledge gap in understanding how risk genetic variants contribute to human diseases (5–7). Only a handful of pan-cancer studies addressed the concept of germline–somatic links. Carter and colleagues (55) proposed a general investigation framework and qualified a few interactions *in vitro*, including between the *19p13.3* allele that magnifies the effect of *PTEN* alteration on mTOR signaling. Most recently, the role of germline immunoediting in sculpting subtypes of breast cancer (56), a cancer type that presents the most emblematic germline–somatic association between inherited *BRCA1* mutations and triple-negative breast cancer, has been proposed.

Whereas age, family history, and ancestry have been reported as the prominent prostate cancer risk factors by epidemiologic studies, the genetic component remains the leading player of prostate cancer etiology, as evidenced by familial linkage studies, twin-based studies, and GWAS efforts (3, 57). More than 160 common polymorphic loci, mainly residing in non–protein-coding regions, have been reported to be associated with prostate cancer susceptibility, in principle, implicated these in biological processes required to initiate carcinogenesis. A few studies have investigated the mechanisms involved in the progression of prostate cancer in the presence of risk variants, as in the case of the rs11672691 polymorphism within an intron of the *PCAT19* noncoding RNA associated with aggressive prostate cancer (58) and of rs1047303 in *HSD3B1* (59, 60).

Whereas the association between germline variants and prostate cancer susceptibility has been studied extensively, the connection of germline variants with specific somatic alterations has rarely been reported. Upon its discovery in 2005, the community investigated the associations between the *TMPRSS2–ERG* gene fusion molecular subclass and inherited variants (61, 62), including rare variants in DNA repair genes, such as *ESCO1*^{N191S} and *POLI*^{F532S} (63), and signal emerging from the enrichment of *TMPRSS2–ERG* fusion in familial prostate cancer studies (64, 65). In 2017, Taylor and colleagues showed that localized castration-sensitive tumors with inactivating *BRCA2* germline variants, a disease that arises at a younger age of onset and is characterized by dismal outcomes, present with somatic aberrations commonly associated with metastatic disease, including epigenomic dysregulation of the MED12/MED12L axis (66). In 2022, Giannareas and colleagues reported mechanistic insights into how the *17q12/HNF1B* locus, a pleiotropic cancer risk locus, contributes to prostate cancer progression through the interplay between *HNF1B* and *TMPRSS2-ERG* (67). Still, a handful of studies investigated the molecular mechanism behind germline–somatic associations.

We analyzed the association between germline variants and prostate cancer somatic lesions. Our results revealed the existence of a noncoding polymorphic regulatory element specifically located at the *7p14.3* locus (rs1376350, SNP, G>A), which is associated with an early recurrent prostate cancer–specific somatic mutation in the *SPOP* gene. Using CRISPR–Cas-edited prostate cancer cell lines to disrupt or alter this locus, we observed an enrichment in histone markers typical of enhancer regions and higher activity in the presence of the alternative allele AA compared with the parental allele GG (Fig. 1H). The transcriptional profiling of a series of PC3 and LNCaP cells subjected to varying degrees of editing at the *7p14.3* locus revealed *GLI3* as the only gene whose transcription was consistently affected. Interestingly, the *GLI3* gene is located on the same chromosome (*7p14.1*) at a distance of 8 Mb from the variant. In a classical view, enhancers are regulatory elements of the genome located no more than hundreds of kilobases from the controlled locus and capable of regulating gene expression through physical interaction (68), Recently, ultra-long-range (tens of megabases) enhancers have been identified (69, 70), and more sophisticated molecular mechanisms than simple direct contact with the cognate gene have been suggested to control gene expression (68, 71, 72).

Although how exactly the *7p14.3* allelic variant affects *GLI3* transcription remains unclear, the mechanism that links germline and somatic association to *SPOP* has been thoroughly elucidated in the current work. The transcription factor GLI3 is a known target of SPOP-mediated ubiquitination (36, 37) and has been reported to play a role in prostate cancer progression (33, 35, 43). Accumulating evidence indicates that

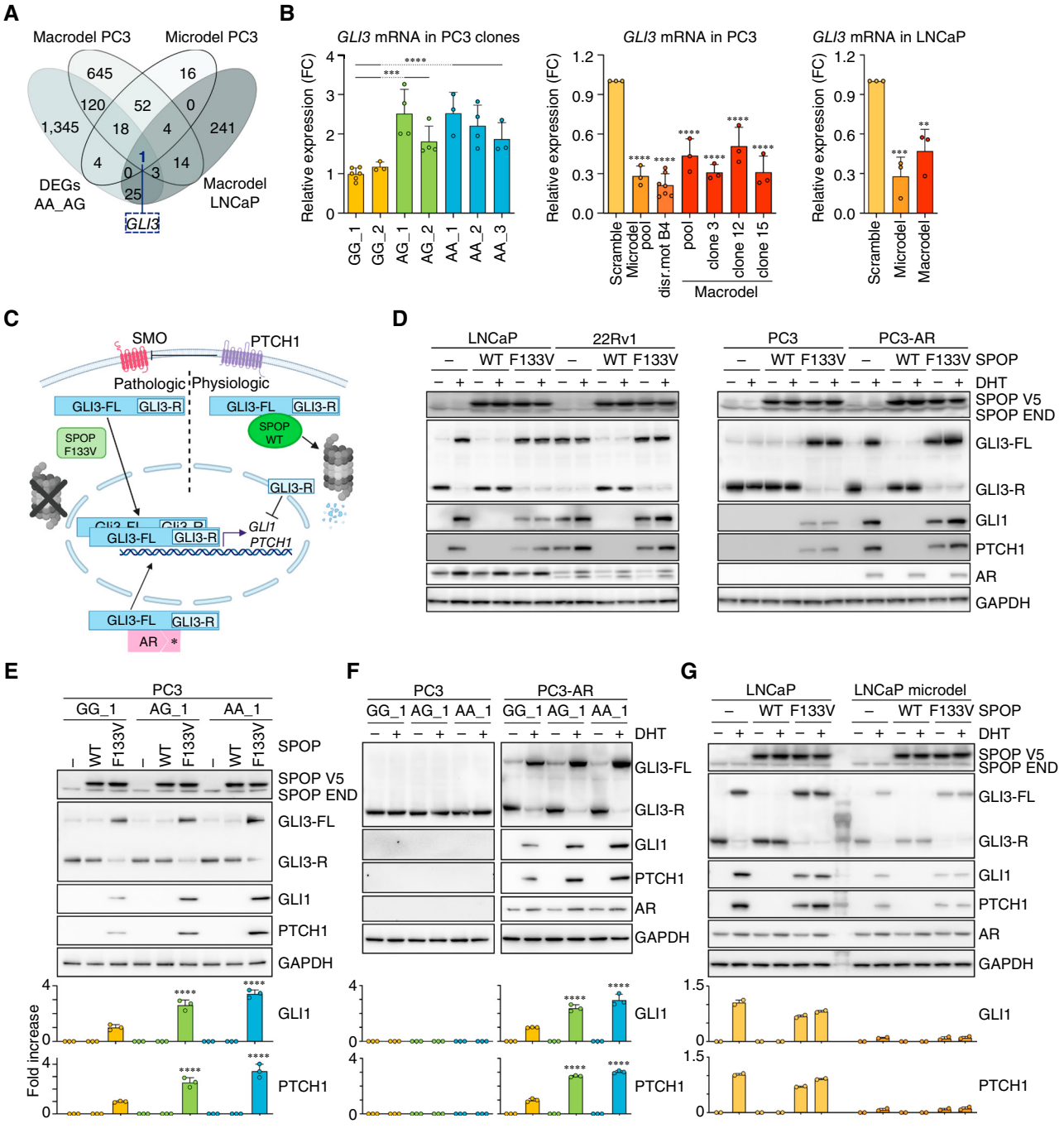


Figure 3. Impact of the *7p14.3* locus, AR activation, and *SPOP* mutation on *GLI3* transcription and stabilization. **A,** The Venn diagram depicts the overlap among DEGs as a function of the extent of the deletion surrounding the variant of interest in PC3 and LNCaP cells and of the presence of the ancestral (G) vs. the alternative allele (A) in the PC3 clones. **B,** *GLI3* mRNA levels measured by RT-qPCR are shown for single-nucleotide edited PC3 clones (left), for PC3 microdeletion and macrodeletion pools (microdel pool and macrodel pool, respectively), along with single clones harboring a disruption motif (disr.mot B4) or macrodeletion (clones 3, 12, 15; middle), and for the pools of LNCaP cells with microdeletion and macrodeletion (right). Data are represented as the mean + SD. ($n \geq 3$ biological replicates). *P* values were determined using one-way ANOVA and the Bonferroni multiple comparisons test. **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$. FC, fold change. **C,** Schematic representation of the GLI3 pathway in the presence of *SPOP* somatic mutation F133V (pathologic scenario, left) and WT functional SPOP (physiologic scenario, right). **D,** LNCaP, 22Rv1, PC3, and AR-overexpressing PC3 (PC3-AR) cells were transduced with LV vectors expressing SPOP WT and mutant SPOP (F133V) with a V5 tag. LV empty was used as control (marked as "-"). Cells were androgen-deprived for 5 days and then stimulated with 10 nmol/L DHT for 18 hours. The whole-cell lysate was analyzed by Western blot with the indicated antibodies. GAPDH was used as a loading control. **E,** Single-nucleotide edited PC3 clones with ancestral alleles GG, heterozygous alleles AG, or homozygous alternative allele AA were transduced with LV vectors expressing SPOP WT and mutant SPOP (F133V) with a V5 tag. Cells were androgen-deprived, and total cell lysates were processed by Western blot analysis as in (**D**) and quantified ($n = 3$ biological replicates). *P* values were determined using the Spearman correlation test ($r = 0.95$, ****, *P* value = 9.585e−05). **F,** Western blot analysis and its quantification ($n = 3$ biological replicates) of the indicated proteins from single-nucleotide edited PC3 and PC3-AR clones. The average expression of the protein in the GG_1 clone was taken as a reference. (*continued on following page*)

improper Hh signaling promotes malignant transformation (73, 74). The Hh and androgenic signaling pathways are intimately linked in driving the growth of prostate cancer cells through AR and GLI family transcription factors, which functionally interact with each other to enhance the transcription of their targets (35, 37, 43). Kaushal and colleagues (75) recently suggested that most prostate cancer cell lines predominantly express the repressor form of GLI3 (GLI3-R). However, we and others (35) observed massive stabilization of GLI3-FL in AR^+ prostate cancer cells either by DHT stimulation or by the expression of intrinsically active isoforms of AR (e.g., ARv7; Fig. 3D), supporting the notion that active AR protects the transcriptionally active GLI3-FL from its proteolytic processing into the transcriptionally repressive shorter form (GLI3-R). The AR–GLIs molecular circuit can be further integrated by the presence of an *SPOP* somatic mutation that also sustains GLI3-FL transcriptional activity by preventing its proteasome-mediated degradation (33–35).

In this study, we showed that the single-nucleotide substitution in the *7p14.3* locus from G to A in the AR^- PC3 cells favors *GLI3* gene expression with no effect on GLI3 protein stabilization in the transcriptionally active form (i.e., GLI3-FL), which is instead regulated by mutant SPOP or AR activation. In line with this, our transcriptomic analysis of the *7p14.3* edited cells did not indicate GLI3 target genes or Hh pathway as differentially expressed. Whereas the impact is insignificant in AR^-/SPOP WT cells, both AR activation and mutant SPOP expression consistently changed the GLI3-FL/GLI3-R ratio, promoting GLI3-FL form stabilization, which results in allele-dependent amplification of the GLI3 signaling cascade.

A positive correlation of SPOP signature score with AR activity score has been reported in prostate cancer patient cohorts (15). Furthermore, it has been found that the gene signature of all SPOP mutants exhibited high overlap with the gene signature of androgen-treated prostate cancer cells *in vitro* (24). Interestingly, several studies linked Hh signaling pathway activation to significantly increased expression of key steroidogenic enzymes and testosterone production through the posttranslational processing of GLI transcription factors (43, 76). In this context, the role of the transcription factors GLI1 and GLI2 has been extensively studied in prostate cancer, whereas the role of GLI3 has not been considered.

In prostate cancer cells, we demonstrate that the combination of aberrant GLI3 expression due to the *7p14.3* locus with the stabilization of the FL form of GLI3 by mutated SPOP determines the allele-dependent accumulation of the key steroidogenic enzyme P450scc, encoded by the *CYP11A1* gene. Furthermore, the concomitance of the *7p14.3* allele variant A and SPOP mutant expression increases the levels of 3βHSD1, an additional crucial enzyme for the synthesis of testosterone, although no data in the literature report a direct role of SPOP in 3βHSD1 ubiquitination and degradation. At present, although interesting, how the combination of the allelic variant in *7p14.3* and the expression of mutant SPOP can promote the accumulation of 3βHSD1 remains unclear.

The rise of steroidogenic enzymes prompted our interest in the capability of the edited cells to synthesize *de novo* steroid precursors starting from cholesterol. LC-MS/MS analysis of steroids in the media and at the intracellular level of our single-nucleotide edited cell model with the overexpression of either WT or mutated SPOP highlighted differences in the pattern of conversion of cholesterol in progesterone or DHT that result from a complex interplay between germline (single-nucleotide editing) and somatic mutation (SPOP F133V overexpression). Our findings show that pathologically stabilized GLI3 sustains intratumoral steroidogenesis and that the relative levels of GLI3-FL are instrumental in regulating the relative abundance/availability of steroids (i.e., progesterone, DHEA, A5-diol, and DHT) in the tumor (Fig. 6A). This explains the strong association between the variant of interest and *SPOP* mutation. In fact, by preventing GLI3-FL degradation, *SPOP* mutation enhanced GLI3-mediated intratumoral steroidogenesis in an allele-dependent manner.

The role of the GLI family of transcription factors in prostate cancer has been extensively studied in the context of androgen deprivation that mimics treatment in advanced cancer stages (33, 35, 43, 76). The translational potential of this study stands on the nomination of the mechanisms by which patients with *SPOP*-mutant prostate cancer better respond to ADT. To fully test its applicability (77) and given the current standard-of-care for high-risk localized and metastatic patients with prostate cancer, it becomes key to interrogate for differential response with upfront hormonal agents (e.g., NHA, abiraterone, enzalutamide, apalutamide, and darolutamide) added to ADT.

Whereas we do not expect a diverse effect of abiraterone versus enzalutamide treatment on the inhibition of the androgen pathway in prostate cancer characterized by WT AR and the co-occurrence of SPOP mutation and the *7p14.3* variant, it should be noted that abiraterone inhibition of CYP17 can lead to the accumulation of progesterone in prostate cancer cells, whose implication in resistance and tumor progression has been recently described (78). This condition could be more exacerbated in prostate cancer with SPOP mutation and the *7p14.3* allelic variant. These considerations suggest the testing of a different androgen pathway inhibitor, potentially combined with the concomitant inhibition of 3βHSD1.

Single risk variants might not be able to disrupt the normal equilibrium of a cell, but the combination of more than one SNP may influence cancer predisposition. Although common SNPs only confer weak to modest risk to prostate cancer, some studies indicate that if an individual carries multiple risk alleles, the risk increases (79). This study has important implications for males whose testosterone levels change with advancing age, in which a subtle differential effect might become significant to the cell and facilitate or accelerate the initiation of tumorigenesis in hormone-sensitive tissues.

Figure 3. ***(Continued)*** GAPDH was used as a loading control. *P* values were determined using the Spearman correlation test ($r = 0.95$, ****, *P* value = 9.585e−05). **G,** LNCaP cells with *7p14.3* WT locus and microdeletion were transduced with LV vectors expressing SPOP WT and mutant SPOP (F133V) with a V5 tag. Cells were androgen-deprived, and total cell lysates were processed by Western blot analysis as in **D** and quantified ($n = 2$ biological replicates). GAPDH was used as a loading control. [**C,** Partially created in BioRender. Demichelis, F. (2025) https://BioRender.com/27382r1.]

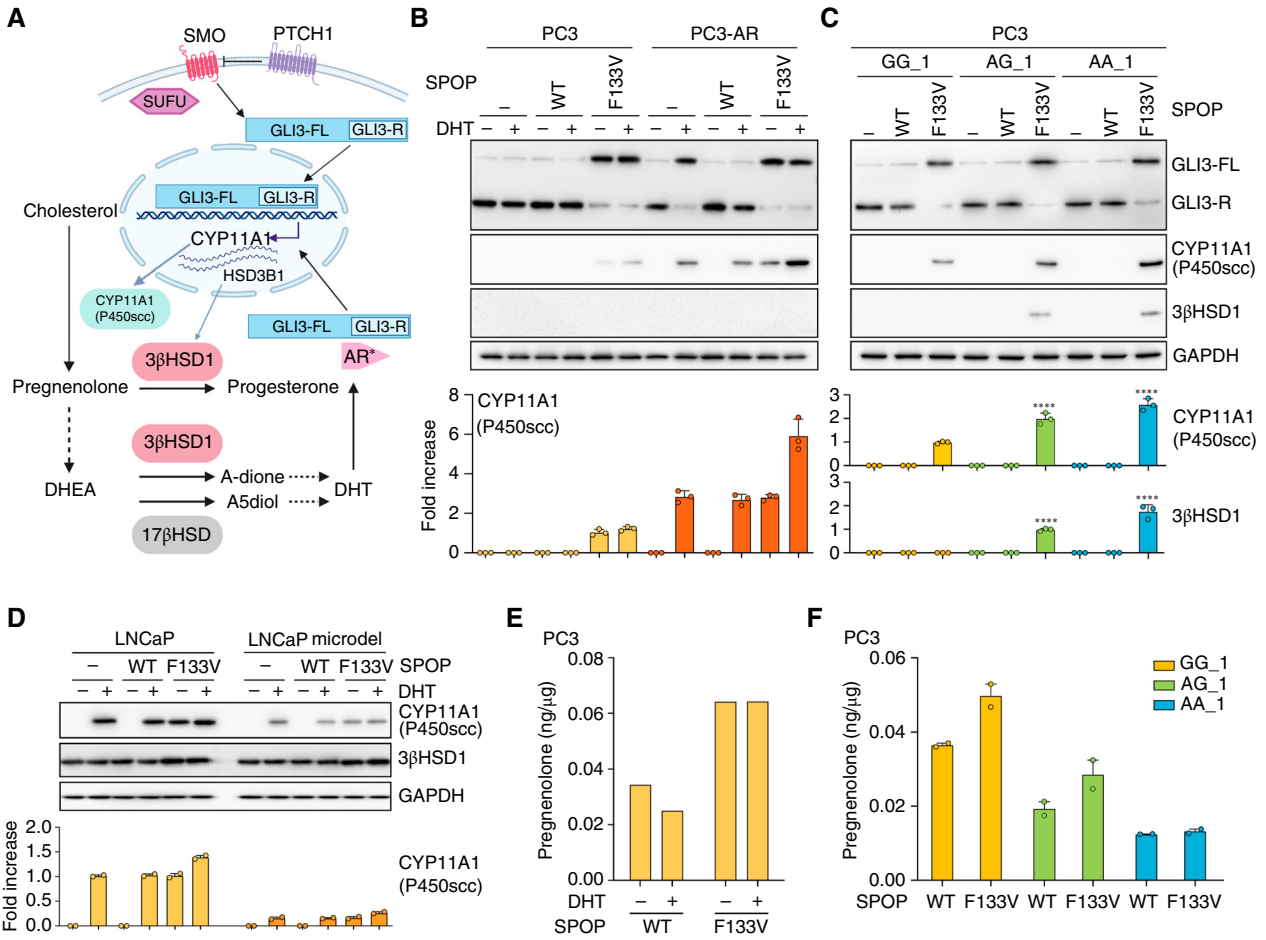


Figure 4. Effects of pathologically stabilized GLI3 on cholesterol metabolism and *de novo* pregnenolone synthesis. **A,** Schematic representation of the effect of Hh signaling on cholesterol metabolism. Activation of Hh signaling induces the conversion of cholesterol to progesterone and estradiol through the upregulation of steroidogenic enzymes, including P450scc and 3βHSD1. GLI3 is required for Hh-induced *CYP11A1* expression. **B,** Western blot analysis of the expression and quantification ($n = 3$ biological replicates) of P450scc and 3βHSD1 in PC3 and PC3-AR overexpressing SPOP WT and mutant SPOP (F133V). LV empty was used as control (marked as "-"). Cells were androgen-deprived for 5 days and then stimulated with 10 nmol/L DHT for 18 hours. GAPDH was used as a loading control. **C,** Western blot analysis and quantification ($n = 3$ biological replicates) of P450scc and 3βHSD1 expression of single-nucleotide edited PC3 clones expressing SPOP WT and mutant SPOP F133V. LV empty was used as control. GAPDH was used as a loading control. *P* values were determined using the Spearman correlation test ($r = 0.95$, ****, *P* value = 9.585e−05). **D,** LNCaP cells with *7p14.3* WT locus and microdeletion (microdel) were transduced with LV vectors expressing SPOP WT and mutant SPOP (F133V) with a V5 tag. LV empty was used as control. Cells were androgen-deprived, and total cell lysates were processed by Western blot analysis as in Fig. 3G and quantified ($n = 2$ biological replicates). GAPDH was used as loading control. **E,** Pregnenolone accumulation in the media of PC3 cells overexpressing WT SPOP and F133V mutated SPOP as measured by ELISA. **F,** Pregnenolone accumulation in the media of single-nucleotide edited PC3 clones expressing SPOP WT and mutant SPOP F133V measured by ELISA. Data are represented as the mean + SEM ($n = 2$ biological replicates). [**A,** Partially created in BioRender. Demichelis, F. (2025) https://BioRender.com/icp0sun.]

Altogether, our study fuels the need to query and model the complexity of the germline human genome in conjunction with the somatic characteristics of tumors to push the understanding of cancer initiation and disease evolution (80). This is especially relevant to all tumor types with high heritability rates and diverse tumor lineages. Cancer hallmarks can be acquired when an oncogenic alteration presents in a specific genetic background determined by the germline genotypes at one or more inherited polymorphisms of an individual (Fig. 6B). Germline–somatic interplays can impinge on metabolism, immune response, and inflammation and jointly facilitate an advantageous cell or oncogenic microenvironment state that presents at a certain point in time during individual's lifetime at or before cancer onset. Whereas implicated germline variants might be known cancer risk loci, either tumor type–specific or with pleiotropic signals (i.e., pan-cancer or multiple cancer risk variants), they can also be missed by GWAS studies and only emerge when somatic subclasses or lineages are interrogated. Hypothesis-driven investigations of germline–somatic interactions informed by functional genomics studies, eQTL signals, or *in silico* predictions can overcome the power limitations dictated by the extent of genetic variability, the low minor allele frequency of variants of interest, and the intrinsic inter-individual tumor diversity. One can speculate that certain germline–somatic interactions also inform individuals' treatment responses and that the upcoming integration of genomics into learning health systems (e.g., Genomics-enabled Learning Health System supported by the NIH) will eventually translate these interactions into improvements in medical care.

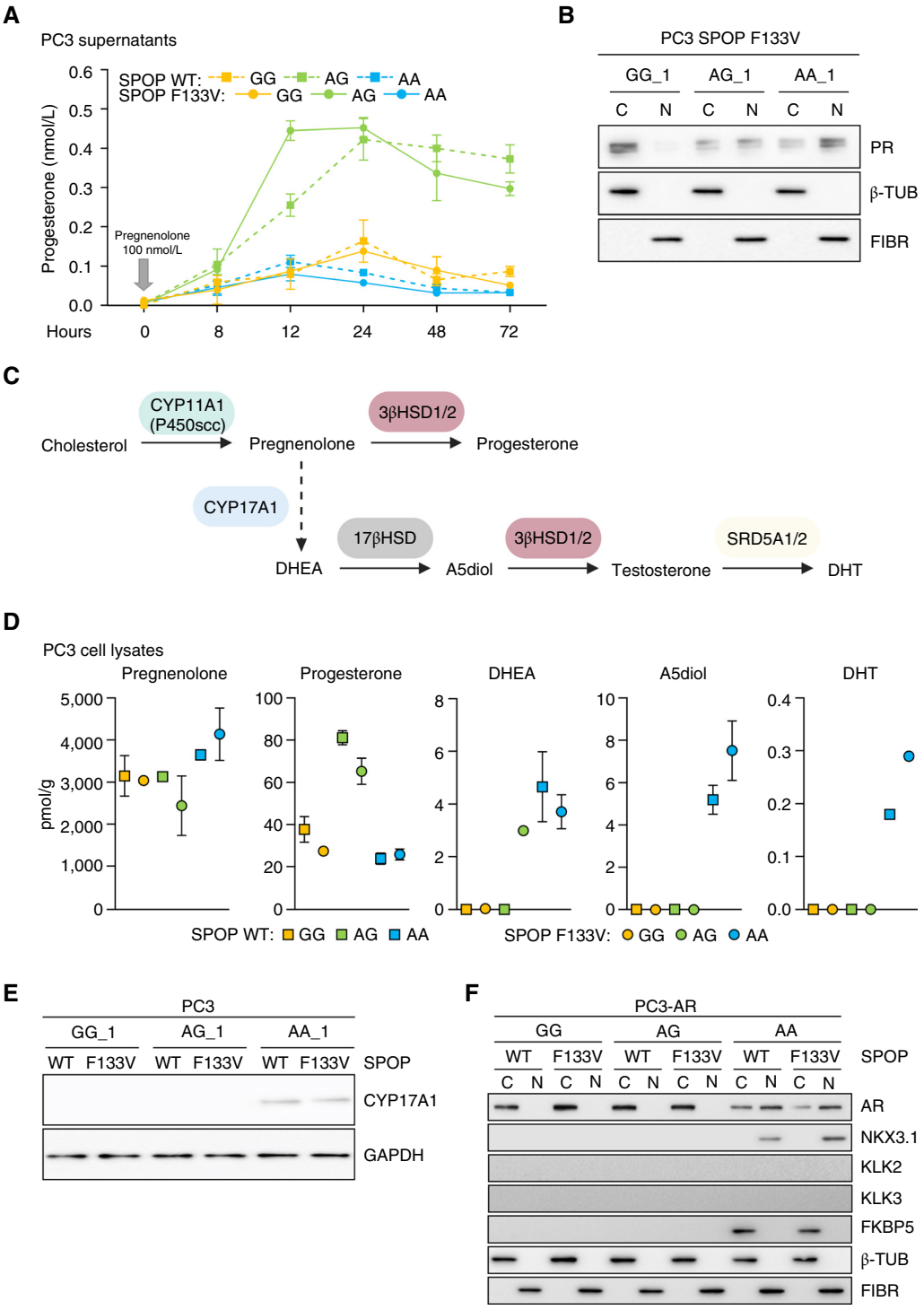


Figure 5. Steroid analysis reveals *de novo* progesterone and DHT synthesis. **A,** Progesterone accumulation in the media of single-nucleotide edited PC3 clones expressing SPOP WT and mutant SPOP F133V after pregnenolone spike-in (100 nmol/L), detected by mass spectrometry analysis at 8, 12, 24, 48, and 72 hours after spike. **B,** Western blot analysis of PR in cytosolic and nuclear fractions of single-nucleotide edited PC3 clones expressing mutant SPOP (F133V). β-TUB and FIBR were used as cytoplasmic and nuclear markers, respectively. **C,** Schematic representation of enzymatic reactions and steroid products required for DHT synthesis from cholesterol. **D,** LC-MS/MS analysis and quantification of steroid in the cell lysates of single-nucleotide edited PC3 clones expressing SPOP WT and mutant SPOP F133V androgen-deprived for 5 days and then stimulated with 100 nmol/L pregnenolone for 72 hours. **E,** Western blot analysis of CYP17A1 expression of single-nucleotide edited PC3 clones expressing SPOP WT and mutant SPOP F133V. GAPDH was used as a loading control. **F,** Western blot analysis after nuclear/cytoplasmic fractionation of single-nucleotide edited PC3-AR clones expressing SPOP WT and mutant SPOP F133V. Cells were androgen-deprived for 5 days and then stimulated with 100 nmol/L pregnenolone for 72 hours. β-TUB and FIBR were used as cytoplasmic and nuclear markers, respectively. β-TUB, β-tubulin; C, cytoplasmic; FIBR, fibrillarin; N, nuclear. [**C,** Partially created in BioRender. Demichelis, F. (2025) https://BioRender.com/es6qhxh.]

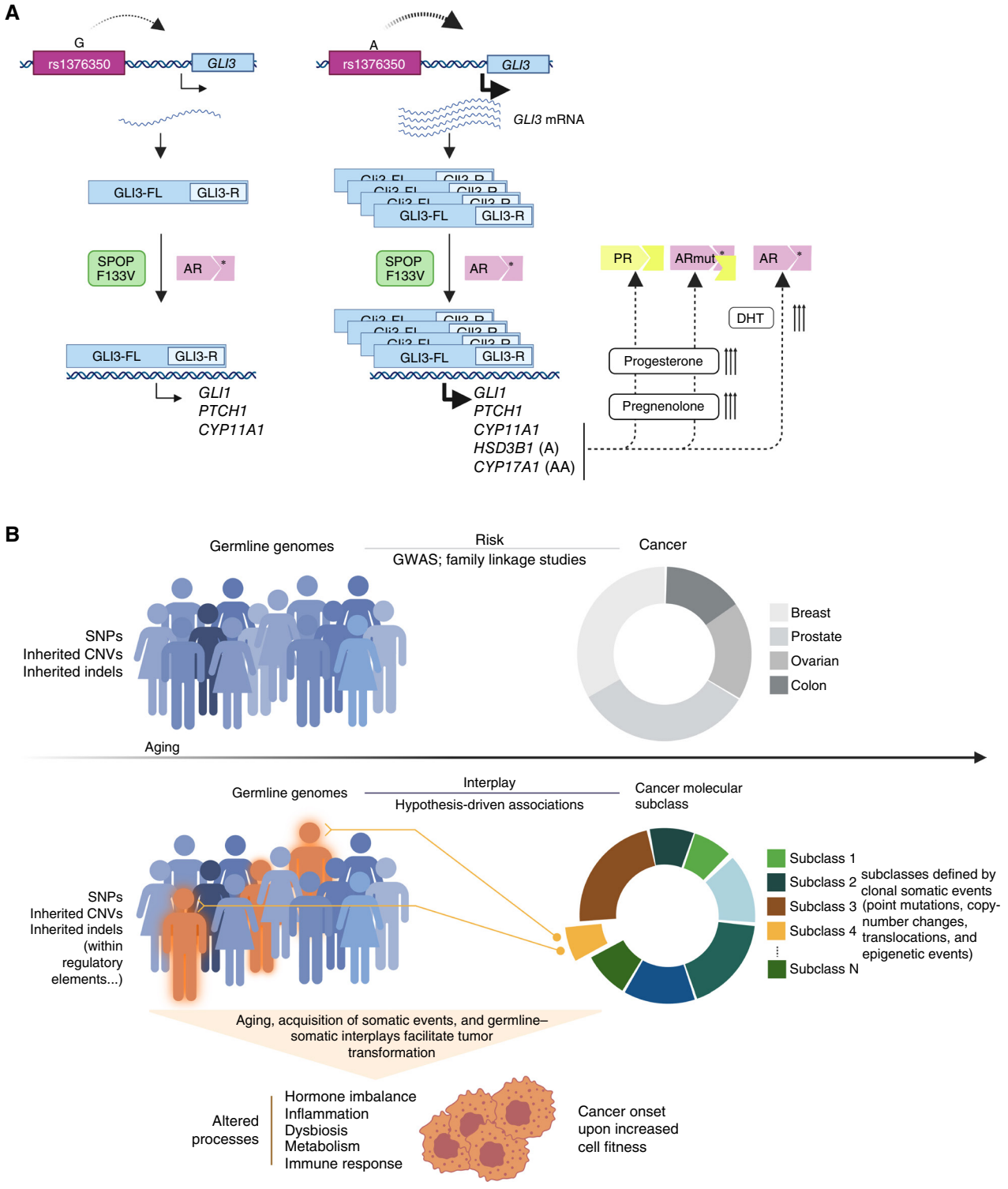


Figure 6. Germline–somatic interactions increase cell fitness favoring cell transformation. **A,** Schematic representation of the impact of genetic and somatic interplay on activation of Hh signaling and steroid biosynthesis via single-nucleotide substitution effect on GLI3 expression and stabilization in the condition of *SPOP* mutation and AR activation. Activation of Hh signaling induces the upregulation of steroidogenic enzymes, including CYP11A1 (P450scc) and 3βHSD1, that promote the conversion of cholesterol to progesterone and/or DHT that can activate the AR and PR. **B,** Inherited germline variants in the general population predispose individuals to cancer types (top), conventionally detected via cancer risk genome-wide association studies (GWAS). With time, aging and the accumulation of somatic mutations may cooperate with inherited variants [(SNPs, copy-number variations (CNV), or insertions/deletions (indels)] to trigger cell processes that increase cell fitness and promote the onset of specific cancer molecular subclasses (bottom). This type of germline–somatic tandem is detectable through association studies testing hypothesis-constrained germline variants and molecular subclasses of human tumors. [**A,** Partially created in BioRender. Demichelis, F. (2025) https://BioRender.com/q4ntg0g and **B,** Created in BioRender. Demichelis, F. (2025) https://BioRender.com/e407ze1.]

METHODS

Cell Culture

22Rv1 (RRID: CVCL_1045), PC3 (RRID: CVCL_0035), LNCaP (RRID: CVCL_0395), VCaP (RRID: CVCL_2235), T47D (RRID: CVCL_0553), and HEK293T (RRID: CVCL_0063) cell lines were purchased from the ATCC (LGC Standards) and cultured in a humidified incubator at 37°C and 5% CO_2 in conformity with the manufacturer's protocols. Sex hormone depletion (androgens and estrogens), before DHT (Sigma-Aldrich) treatments, was achieved by growing the cells in medium without phenol red (Euroclone), supplemented with 10% charcoal/dextran treated FBS (Hyclone, Celbio) for 5 days. All prostate cell lines were GG at rs1376350 (GSM888588 and GSM888346) and WT for SPOP status. All cell lines used in this study were routinely confirmed to be *Mycoplasma*-free using PlasmoTest Mycoplasma Detection Kit (Invivogen) and authenticated with DNA markers used by the ATCC.

Transfection and Plasmids

For the transfection of plasmids, 3×10^5 PC3 cells were seeded on six-well plates and transfected with 2.5 μg DNA and FuGENE HD (cat. # E2311, Promega). LNCaP transfection was performed with Lipo2000 (cat. # 11668019, Invitrogen). Single-guide RNA (sgRNA) expressing plasmid under the control of U6 promoter harboring a neomycin resistance cassette pU19_NeoR was derived by Addgene 137776 (RRID: Addgene_137776). High-fidelity Cas9-expressing vector pX330-eSpCas9(1.1) was a gift from Yuichiro Miyaoka (Addgene, RRID: Addgene_108301). Cas9-expressing vector harboring a blasticidin resistance cassette lentiCas9-Blast plasmid was a gift from Feng Zhang (RRID: Addgene_52962). AR-expressing lentiviral backbone pAIP_AR was cloned in house using pAIP from Jeremy Luban (RRID: Addgene_74171). VSV-G envelope expressing plasmid pMD2G (RRID: Addgene_12259) and second-generation packaging vector psPAX2 were a gift from Didier Trono (RRID: Addgene_12260). V5-tagged lentiviral vector for the overexpression of WT SPOP, mutant SPOP F133V, and the empty vector used as control (LvLacz 6.3/V5 empty, SPOP WT, and SPOP F133V) were a kind gift from Mark A Rubin. Lentiviruses were generated by transfection of HEK293T cells with packaging and envelope plasmids pMD2G and psPAX2 with the vector of interest or its control vector. They were then used to stably transduce prostate cancer cell lines.

sgRNA Design

sgRNAs used for the macrodeletion (731 bp) spanning the *7p14.3* locus were selected using the GPP Web Portal CRISPick (RRID: SCR_025148 at http://portals.broadinstitute.org/gpp/public/). sgRNAs for the microdeletion (50 bp) and disruption of DNA-binding motifs around the SNP of interest were designed based on protospacer adjacent motif 5′-NGG-3′ sequences nearby the locus. (Supplementary Tables S1–S3). All sgRNAs utilized in this work were characterized for predicted off-targets using Cas OFFinder (RRID: SCR_023390; Supplementary Table S2; ref. 81). All sgRNAs were cloned into pUC19 and pX330 plasmids.

Assessment of Genome Editing by TOPO TA Cloning of PCR Products

The Sanger sequencing of PCR products encompassing our region of interest performed on pool of edited cells often showed a mix of different genomic events. To confirm the insertions/deletions occurring on either alleles of the *7p14.3* locus, PCRs from selected clones were performed using primers encompassing the edited locus (FBS005-RBS005, Supplementary Table S1), ligated into the pCR-4 TOPO TA vector and cloned using a TOPO TA cloning sequencing kit (K450641, Invitrogen Co.), and then amplified in competent DH5-α (Thermo Fisher Scientific, cat. # 18265017). Isolated plasmids were sent for Sanger sequencing. At least 10 colonies were tested for each cellular clone.

Generation of Cell Lines with Homozygous Macrodeletion, Microdeletion, and Disruption Motifs

LNCaP and PC3 cells were transfected with combos of Cas9-expressing vectors and sgRNA guides to target the locus surrounding as described in Supplementary Table S1 to obtain pools of cells with macrodeletion, microdeletion, and disrupted motifs. The efficiency of genome editing evaluated by PCR (Supplementary Fig. S1C and S1D), (as described in "Assessment of Genome Editing with PCR or ddPCR"), revealed that—whereas the efficiency of microdeletion was high, and single-clone isolation was not necessary—limited dilution and isolation of single clones were instead necessary to obtain cells with homozygous macrodeletion or disruption motif resulting in 10 (14,28%) positive clones in at least one allele and 2 (2.8%) homozygous for a deletion (deletion of 14 and 24 bp for clones B4 and deletion of 8 bp for B5; Supplementary Fig. S1E). Each model/sample used in the study is listed in Supplementary Table S9.

Generation of PC3 Cell Line with Single-Nucleotide Substitution

PC3 cells were electroporated with Cas9–RNP complexes (prepared using SF cell line 4D-Nucleofector X, Kit S (Lonza, cat. # CHV4XC-1032) recommended for PC3 cells according to the manufacturer's instructions). Briefly, RNP complexes were obtained by mixing: previously annealed tracrRNAs and crRNAs (from Integrated DNA Technologies), donor DNAs (PS-ssDNA or ssDNA sequence reported in Supplementary Table S1), and recombinant Cas9 protein and electroporation enhancer (from Integrated DNA Technologies), as listed in Supplementary Fig. S1F. PC3 cells, resuspended in Nucleofector solution buffer, were then combined with RNPs complex, electroporated, and then plated in medium supplemented with HDR enhancer. Once enough cells were present, DNA was extracted, and editing efficiency was assessed.

Assessment of Genome Editing with PCR or Droplet Digital PCR

To evaluate the editing in pools of edited cells, PCR reactions were performed using Platinum SuperFi Green PCR Master Mix (Invitrogen), with primers spanning di expected deletion (Supplementary Table S1). Amplified PCR products were run on 1% to 2% agarose gel, purified, and sent for Sanger sequencing. For clones screening, the single-cell dilution method was used to have one cell per well in a 96-well plate as previously described (82). DNA was extracted with QuickExtract DNA extraction solution (Lucigen) and used as template for PCR or droplet digital PCR (ddPCR). For an appropriate design of primers and probes for rare event detection by ddPCR, we used Primer3Plus (RRID: SCR_003081; http://primer3plus.com). ddPCR reactions were assembled using primers and allele-specific probes (Supplementary Table S1) with ddPCR SuperMix for Probes (Bio-Rad) and 20 ng of DNA. All ddPCR assays were analyzed using the QX200 droplet reader and Bio-Rad QuantaSoft Analysis Software (RRID: SCR_025696) version 1.7.4.

Assessment of Genome Editing with Deep Target Sequencing

Pools with a higher probability of HDR events as determined by ddPCR (Supplementary Fig. S1G) were sequenced at and around the locus of interest with the following approach: three sets of primers were designed using Primer3 to cover 500 bp around the cutting site of Cas9. PCR products were purified with AMPure XP paramagnetic

beads (Beckman Coulter). A second PCR was performed to index PCR products (Nextera, Illumina). DNA was pooled before MiSeq 2 × 250 bp sequencing (Illumina). PCR primer sequences are reported in Supplementary Table S1. Pools with higher probability of HDR events (Supplementary Fig. S1H) were limited-diluted to isolated single clones that were then tested by ddPCR (Supplementary Fig. S1I) and, when positive, confirmed by Sanger sequencing (Fig. 1C). From a total of 140 screened clones, we found three homozygous genotypes for the alternative allele (2.14%; AA_1, AA_2, and AA_3) and 17 heterozygous genotypes (12.14%; AG; Supplementary Fig. S1J). Clones in which the editing did not occur were selected to represent the GG ancestral allele (GG_1, GG_2, and GG_3).

ChIP Assay

PC3 cells with and without AR stable overexpression were androgen-deprived for 4 days and then treated with EtOH or DHT (100 nmol/L) for 16 h before harvesting. ChIP was performed using the following antibodies: anti-histone H3 monomethyl K4 (Abcam, cat. # ab8895, RRID: AB_306847), anti-histone H3 trimethyl K4 (Abcam, cat. # ab8580, RRID: AB_306649), anti-histone H3 trimethyl K27 (Millipore, cat. # 07-449, RRID: AB_310624), anti-histone H3 acetyl K27 (Abcam, cat. # ab4729, RRID: AB_2118291), anti-AR (part number CS207339 of Millipore #17-10489), or a normal IgG (part number PP64B of Millipore #17-10489). Briefly, the experiment procedure included chromatin crosslinking with formaldehyde, nuclei isolation, and lysis followed by chromatin shearing. Magnetic beads (Dynabeads, Invitrogen) were incubated with each antibody overnight at 4°C. DNA was eluted, purified with classic phenol/chloroform protocol, and then analyzed by RT-qPCR. When possible, we used the same control for each histone modification. When the controls did not show the expected pattern, we selected more appropriate controls. This was the case for *GAPDH* promoter, and *KLK3* enhancer used as a positive control for H3K4me3 and H3K4me1 enrichment, respectively, instead of *FKBP5* enhancer.

FKBP5 enhancer was used as a positive control for AR and H3K27Ac enrichment (83). *NeuroD2* enhancer–promoter region was used as a positive control for H3K27m3 and as a negative control for the H3K27Ac enrichment (40), and the desertic region was used as negative copromoter control for the AR. Enrichments of these regions were compared with enrichments of the region of interest surrounding the *7p14.3* variant. The primers used are reported in Supplementary Table S1. Antibodies-specific recruitment was calculated as enrichment to the IgG according to the ΔCt method.

AR ChIP-Seq Data of Human Patients with Prostate Cancer

AR ChIP-Seq datasets of three human patients with prostate cancer were accessed, namely the DARANA (N.PCA.pts = 24; ref. 38), PORTO (N.PCA.pts = 94; ref. 40), and POMERANTZ (N.PCA.pts = 49; ref. 39) cohorts. The locus of interest (rs1376350) was queried for read counts, together with the *FKBP5* enhancer, as a positive control. When multiple samples per patient were available, all counts were retrieved to inspect the presence of AR binding.

Biotinylated DNA Pull-down Assay

Nuclear protein extraction was performed using EpiQuik Nuclear Extraction Kit I (EpigenTek) according to the manufacturer's instructions. Twenty-eight–nucleotide-long oligos were designed to harbor the canonical DNA full-site binding sequence for the AR (AR_fs; ref. 84), a canonical DNA half-site binding sequence for the AR (M08908, AR_hs_A), and the sequences that encompass our SNP of interest carrying the major allele G (7p14.3_G) or the minor allele A (7p14.3_A; DNA-binding sequence for AR M00962). As the *in silico* forecast identified a CEBP-binding motif overlapping to the rs1376350 (M00770), oligoprobes mutated at nucleotides 18 and 24 (7p14.3_G* and 7p14.3_A*) were used to test the potential interference of CEBP β on AR binding. A sequence lacking AR-binding sites with a 5′ biotin residue was used as negative control (scramble; All oligo sequences are listed in Supplementary Table S1). Oligos were ordered as single strands, hybridized, and mixed with nuclear extracts and magnetic beads (Dynabeads M-280 Streptavidin, Invitrogen). After overnight incubation, the precipitated matrix was washed, and proteins pulled down were probed by Western blot for the AR.

Library Preparation for RNA-seq

RNA integrity was assessed through electrophoresis with Bioanalyzer RNA 6000 Nano Kit (Agilent) for library preparation. All samples showed an RNA integrity number ≥9.90. RNA libraries were prepared according to the manufacturer's protocol using TruSeq Stranded mRNA Kit (Illumina). Single-end (100 bp) sequencing was performed on a HiSeq 2500 (Illumina). RNA-seq statistics are reported in Supplementary Table S10.

RNA Isolation and RT-qPCR

Total RNA was isolated from cultured cells using RNeasy Kit (Qiagen, cat 74104) following the manufacturer's instructions. cDNAs were synthesized from 1 μg of total RNA using a cDNA synthesis kit (K1622 RevertAid First Strand cDNA Synthesis Kit, Thermo Fisher Scientific). RT-qPCR was performed with KapaSybr (KK4601 – KAPA SYBR FAST Universal Kit, Sigma-Aldrich). Analysis of relative mRNA expression was performed using the ΔΔCt method with GAPDH as the reference gene (Supplementary Table S1). Antibody-specific recruitment was calculated as enrichment to the IgG according to the ΔCt method.

Analysis of RNA-seq Data

Data Preprocessing and Differential Analysis. Preprocessing and differential analysis of RNA-seq data was performed as follows: raw data quality was controlled using fastQC (RRID: SCR_014583; https://www.bioinformatics.babraham.ac.uk/projects/fastqc) and visualized using MultiQC (RRID: SCR_014982; ref. 85). Trimmomatic (RRID: SCR_011848; ref. 86) was used to trim adapters from sequences to filter out low-quality sequences (<15 in a sliding window of 4 bp) and sequences that resulted too short after the trimming (<50 bp). Trimmed and filtered fastQ files were aligned to the hg38 reference human genome using STAR (RRID: SCR_004463; ref. 87). Read counts were generated through the STAR quantMode function and rearranged in a single expression matrix through a custom R script. Differential expression analysis was performed using edgeR (RRID: SCR_012802; ref. 88), which includes an internal normalization step to generate the pseudo counts used in the actual differential expression analysis.

DEG analysis of edited PC3 cells (i.e., macrodeletion, microdeletion, and rs1376350 homozygous and heterozygous genotypes, with or without exogenous AR expression) and LNCaP cells (i.e., macrodeletion and microdeletion; Supplementary Table S4) was performed using WT parental cells as a reference.

Single-clone DEGs were obtained using edgeR (RRID: SCR_012802), analyzing all single-clone samples and setting clone of origin and treatment (no treatment, AR-DHT, or AR-EtOH) as covariates; DEGs are filtered based on a FDR < 0.05, without any fold change threshold. Normalized counts per million reads (cpm) were used exclusively for visualization. Functional characterization of DEG lists was performed using the enrichGO function of clusterProfiler (RRID: SCR_016884; ref. 89).

Three genes with no established links between *SPOP* mutations or to the regulatory element at *7p14.3* were selected for validation. The *NT5E* and *CLDN3* genes were chosen because close to the

significance threshold in AA versus GG PC3 edited clones (Fig 2B, right), whereas *TGFα* was chosen as highly deregulate in AA versus GG in the same setting.

Human Transcript Data and Genotypes. Human prostate data queried for genotype–transcript association include $N = 63$ benign prostate tissues and $N = 319$ prostate cancer tissues. Briefly, benign ($N = 63$) and tumor ($N = 319$) prostate tissue RNA-seq data with available FASTA files (15, 20, 90) and matched genotype data were aligned to the reference genome hg19 using STAR aligner (RRID: SCR_004463), and logarithm-transformed RPKM+1 of each gene (UCSCknownGenes) were computed using mrf Quantifier (91) and were quantile-normalized. Data processing for transcript and genotype was performed as previously (19). Differential expression analysis was performed using edgeR (starting from published read counts data).

Code and Manipulation of Sequencing Data. Statistical analyses, data manipulation, and data visualization were performed using R R Core Team (2018). R Project for Statistical Computing (RRID: SCR_001905): A language and environment for statistical computing, R Foundation for Statistical Computing. URL https://www.R-project.org/). FastQ and alignment files were manipulated using SAMTOOLS (RRID: SCR_002105; ref. 92). All the analyses were implemented using custom shell scripts.

Western Blotting

Cell lysates were prepared from cells kept in androgen-deprived media (RPMI1640 without phenol red supplemented with 10% charcoal-stripped FBS (Hyclone, Celbio) for 5 days. When DHT (10 nmol/L) was used, it was added at day 4 per 24 hours. Nuclear/cytoplasmic fractionation was achieved using the NE-PER Nuclear and Cytoplasmic Extraction Kit (Life Technologies) following the manufacturer's instructions. Western blot analysis was performed as previously described (93, 94). Briefly, cells were lysed in RIPA buffer supplemented with protease inhibitors (Halt protease inhibitor cocktail, Life Technologies). Equal amounts of protein were separated by SDS-PAGE and transferred onto a polyvinylidene difluoride membrane (Hybond, Thermo Fisher Scientific). After blocking, the membranes were incubated O/N at 4°C with specific primary antibodies (see below) and then with horseradish peroxidase–conjugated anti-rabbit (Cell Signaling Technology, cat. # 7074, RRID: AB_2099233) or horseradish peroxidase–linked anti-mouse (Cell Signaling Technology, cat. # 7076, RRID: AB_330924) secondary antibodies for 1 hour at room temperature. The immunoreactive bands were detected using ECL Select Western Blotting Detection Reagent (GEHRPN2232, Thermo Fisher Scientific) with an Alliance LD2-77WL system and software (UVITEC). Densitometric quantification of Western blot bands normalized to loading controls was performed using ImageJ (RRID: SCR_003070). The following primary antibodies were used: AR (Cell Signaling Technology, cat. # 5153, RRID: AB_10691711), GLI1 (Cell Signaling Technology, cat. # 3538, RRID: AB_1903989), GLI3 (Abcam, cat. # ab181130, RRID: AB_2732891), PTCH1 (Abcam, cat. # ab53715, RRID: AB_882208), SPOP (Proteintech, cat. # 16750-1-AP, RRID: AB_2756394), CYP11A1 (Abcam, cat. # ab272494, RRID: AB_2916032), CYP17A1 (Abcam, cat. # ab125022, RRID: AB_10975095), 3βHSD1 (Abcam, cat. # ab55268, RRID: AB_942015), 3βHSD2 (Abcam, cat. # ab75710, RRID: AB_2631971), PR (Santa Cruz Biotechnology, cat. # sc-539, RRID: AB_632264), SRD5A1 (Proteintech, cat. # 66329-1-Ig, RRID: AB_2881710), KLK2 (Origene, cat. # 802077, RRID: AB_2626200), KLK3 (Agilent, cat. # M0750, RRID: AB 2281105), NKX3.1 (Millipore, cat. # AB5983, RRID: AB_ 92188), FKBP5 (Cell Signaling Technology, cat. # 12210, RRID: AB_ 2797846), FIBRILLARIN (Abcam, cat. # ab4566, RRID: AB_304523), β-TUBULIN (Santa Cruz Biotechnology, cat. # sc-5274, RRID: AB_2288090), and GAPDH (Thermo Fisher Scientific, cat. # MA5-15738, RRID: AB_10977387). All antibodies are listed in Supplementary Table S11.

ELISA

For the measurement of pregnenolone, PC3 and LNCaP cells were seeded in quadruplicate at the density of 60,000 and 120,000 per well of a 24-well plate in androgen-deprived media (RPMI1640 without phenol red supplemented with 10% charcoal-stripped FBS) for 5 days. When DHT (10 nmol/L) was used, it was added at day 4 per 24 hours. Culture supernatants and cell lysates were harvested and tested by ELISA (K1912 Abnova) following the manufacturer's instructions.

LC-MS/MS Analysis of Steroid

Freshly frozen media samples were thawed from storage at −80°C, and 200 μL of each sample were transferred to a 13 × 100 glass tube. To ensure accurate quantification, samples were spiked with 10 μL of mixed internal standards solution containing androstene-3, 17-dione-2,3,4-13C3 (Cerilliant), 5α-dihydrotestosterone-16,17,17-d3 (Cambridge Isotope Laboratories, Inc.), and pregnenolone-17α,21,21,21-d4 (CDN Isotopes). The samples were briefly vortexed to ensure homogeneity. Subsequently, 2 mL of methyl-tert-butyl ether (Across) was added to each tube. The tubes were then vortexed for 5 minutes using a multi-tube vortexer (Thermo Fisher Scientific) and centrifuged for 5 minutes at 3,000 RPM at 4°C. Following centrifugation, the top organic solvent layer containing extracted steroids was carefully transferred into a new glass tube. The solvent was evaporated under nitrogen gas at 40°C and reconstituted in 120 μL of 50% methanol/water (v/v), followed by 1 minute of vortexing. The reconstituted samples were transferred into 1.5 mL microcentrifuge tubes and centrifuged at 14,000 RPM for 15 minutes at 4°C. The supernatants, containing the extracted steroid compounds, were collected in inserts for LC-MS/MS analysis. For cell pellets, samples were prepared by subjecting them to five freeze–thaw cycles. Specifically, the cell pellets, stored at −80°C, were thawed at room temperature for each cycle. Following the freeze–thaw cycles, the extraction procedure for cell pellets followed the same protocol as described above for media samples. The extracted steroids from the samples were quantified using LC-MS/MS. The LC-MS/MS system utilized an ultra-pressure liquid chromatography system (Shimadzu Corporation, Japan), which included two LC-40AD pumps, a DGU-405 degasser, a CTO-40C column oven, a NEXERA SIL-40CX3 autosampler, and a system controller SCL-40. This ultra-pressure liquid chromatography system was coupled with a QTRAP 6500+ mass spectrometer (AB Sciex). The gradient started from 20% solvent B [acetonitrile/methanol (80/20, v/v) with 0.1% formic acid] over 4 minutes, followed by an increase to 52% solvent B over 4 minutes, 75% by 20 min, and finally to 100% solvent B for 5 minutes. Steroids were detected by a QTRAP 6500+ mass spectrometer. Data acquisition and processing were performed using Sciex OS-MQ software (AB Sciex, version 3.3.1.43). All samples were analyzed in triplicate and represented as the mean of the three values obtained with exception of intracellular pregnenolone were the outlier AG WT_R3 was removed from the analysis (Supplementary Table S12).

Statistical Analysis

Unpaired *t* tests were used to determine the statistical significance of the ChIP-qPCR experiments. One-way ANOVA test, followed by the Bonferroni correction, was used to assess the statistical significance of intergroup differences in the analysis of gene expression through RT-qPCR. Spearman correlation test was used to assess the statistical significance in the Western blot analyses. *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$. The results with $P \leq 0.05$ were

considered statistically significant. Other statistics are included in the specific paragraphs. R Software (version 3.6) and Prism version 10 (GraphPad Software, software RRID: SCR_002798) were used for data analysis.

Data Availability

Raw and processed data are deposited at the Gene Expression Omnibus (RRID: SCR_005012) database with the accession number GSE293478.

Authors' Disclosures

A. Casini reports being an employee and a quotaholder of Alia Therapeutics, a genome editing company. C.E. Barbieri reports personal fees from Pfizer outside the submitted work; in addition, C.E. Barbieri has a patent for US20130331279A1 issued. N. Sharifi reports being a coinventor on a Cleveland Clinic patent on HSD3B1. A. Cereseto reports being a quotaholder of Alia Therapeutics, a genome editing company. No disclosures were reported by the other authors.

Authors' Contributions

P. Gasperini: Formal analysis, investigation, visualization, writing–original draft, writing–review and editing. **A. Alaimo:** Formal analysis, investigation, visualization, writing–original draft, writing–review and editing. **B. Stringa:** Formal analysis, investigation, writing–original draft. **Y.-M. Chung:** Data curation, formal analysis, investigation, writing–original draft. **Y. Ciani:** Data curation, formal analysis, visualization. **F. Lorenzin:** Data curation, formal analysis, investigation, writing–original draft, writing–review and editing. **G. Fracassi:** Data curation, formal analysis, investigation, visualization, writing–review and editing. **Y. Zekri:** Data curation, formal analysis. **F. Orlando:** Data curation, formal analysis. **O. Quaini:** Investigation. **S. Gregoricchio:** Formal analysis. **G. Petris:** Methodology. **A. Casini:** Methodology. **C.E. Barbieri:** Methodology. **W. Zwart:** Supervision, methodology. **A. Cereseto:** Supervision, methodology. **N. Sharifi:** Supervision, methodology. **A. Lunardi:** Conceptualization, formal analysis, supervision, writing–original draft, writing–review and editing. **F. Demichelis:** Conceptualization, resources, formal analysis, supervision, funding acquisition, writing–original draft, writing–review and editing.

Acknowledgments

The authors like to thank Ilaria Cherchi for helpful insights into the analysis of the *7p14.3* locus genomic sequence, Mark A. Rubin for sharing the SPOP constructs and the Next Generation Sequencing Core Facility, part of the Department CIBIO Core Facilities (IRBIO) for their technical support. F. Demichelis is supported by Fondazione AIRC Investigator Grant 19221, 29370. A. Lunardi is supported by Fondazione AIRC Investigator Grant 27893 and The Giovanni Armenise-Harvard Foundation CDA Grant. N. Sharifi is supported by grants from the NCI, NIH (R01CA172382, R01CA236780, R01CA261995, and R01CA249279). F. Lorenzin is supported by the Prostate Cancer Foundation Young Investigator Award 19YOUN16. G. Fracassi is supported by an AIRC fellowship for Italy (29926-2023). W. Zwart is funded by the Oncode Institute, partly financed by the Dutch Cancer Society KWF. Research at the Netherlands Cancer Institute is supported by institutional grants from the Dutch Cancer Society and the Dutch Ministry of Health, Welfare, and Sport. Department CIBIO Core Facilities (IRBIO) are supported by the European Regional Development Fund (ERDF) 2014–2020 and 2021-2027. This work has been supported by the initiative "Dipartimenti di Eccellenza 2023–2027 (Legge 232/2016)" funded by the Italian Ministry of University and Research.

Note

Supplementary data for this article are available at Cancer Discovery Online (http://cancerdiscovery.aacrjournals.org/).

Received December 31, 2024; revised April 10, 2025; accepted June 11, 2025; posted first June 13, 2025.

REFERENCES

1. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434–43.
2. Swanton C, Bernard E, Abbosh C, André F, Auwerx J, Balmain A, et al. Embracing cancer complexity: hallmarks of systemic disease. Cell 2024;187:1589–616.
3. Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, et al. Familial risk and heritability of cancer among twins in nordic countries. JAMA 2016;315:68–76.
4. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74:229–63.
5. Chatrath A, Przanowska R, Kiran S, Su Z, Saha S, Wilson B, et al. The pan-cancer landscape of prognostic germline variants in 10,582 patients. Genome Med 2020;12:15.
6. Cheng DT, Prasad M, Chekaluk Y, Benayed R, Sadowska J, Zehir A, et al. Comprehensive detection of germline variants by MSK-IMPACT, a clinical diagnostic platform for solid tumor molecular oncology and concurrent cancer predisposition testing. BMC Med Genomics 2017;10:33.
7. Huang K-L, Mashl RJ, Wu Y, Ritter DI, Wang J, Oh C, et al. Pathogenic germline variants in 10,389 adult cancers. Cell 2018;173:355–70.e14.
8. Conti DV, Darst BF, Moss LC, Saunders EJ, Sheng X, Chou A, et al. Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. Nat Genet 2021;53:65–75.
9. Demichelis F, Setlur SR, Banerjee S, Chakravarty D, Chen JYH, Chen CX, et al. Identification of functionally active, low frequency copy number variants at 15q21.3 and 12q21.31 associated with prostate cancer risk. Proc Natl Acad Sci U S A 2012;109:6686–91.
10. Seibert TM, Garraway IP, Plym A, Mahal BA, Giri V, Jacobs MF, et al. Genetic risk prediction for prostate cancer: implications for early detection and prevention. Eur Urol 2023;83:241–8.
11. Pritchard CC, Mateo J, Walsh MF, De Sarkar N, Abida W, Beltran H, et al. Inherited DNA-repair gene mutations in men with metastatic prostate cancer. N Engl J Med 2016;375:443–53.
12. Ewing CM, Ray AM, Lange EM, Zuhlke KA, Robbins CM, Tembe WD, et al. Germline mutations in HOXB13 and prostate-cancer risk. N Engl J Med 2012;366:141–9.
13. Hearn JWD, Sweeney CJ, Almassi N, Reichard CA, Reddy CA, Li H, et al. HSD3B1 genotype and clinical outcomes in metastatic castration-sensitive prostate cancer. JAMA Oncol 2020;6:e196496.
14. McKay RR, Nelson TJ, Pagadala MS, Teerlink CC, Gao A, Bryant AK, et al. Adrenal-Permissive germline HSD3B1 allele and prostate cancer outcomes. JAMA Netw Open 2024;7:e242976.
15. Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. Cell 2015;163:1011–25.
16. Li J, Xu C, Lee HJ, Ren S, Zi X, Zhang Z, et al. A genomic and epigenomic atlas of prostate cancer in Asian populations. Nature 2020;580:93–9.
17. Abida W, Cyrta J, Heller G, Prandi D, Armenia J, Coleman I, et al. Genomic correlates of clinical outcome in advanced prostate cancer. Proc Natl Acad Sci U S A 2019;116:11428–36.
18. Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, et al. Punctuated evolution of prostate cancer genomes. Cell 2013;153:666–77.

19. Romanel A, Garritano S, Stringa B, Blattner M, Dalfovo D, Chakravarty D, et al. Inherited determinants of early recurrent somatic mutations in prostate cancer. Nat Commun 2017;8:48.
20. Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, et al. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. Nat Genet 2012;44:685–9.
21. Jaratlerdsiri W, Jiang J, Gong T, Patrick SM, Willet C, Chew T, et al. African-specific molecular taxonomy of prostate cancer. Nature 2022; 609:552–9.
22. Dai X, Gan W, Li X, Wang S, Zhang W, Huang L, et al. Prostate cancer-associated SPOP mutations confer resistance to BET inhibitors through stabilization of BRD4. Nat Med 2017;23:1063–71.
23. Geng C, He B, Xu L, Barbieri CE, Eedunuri VK, Chew SA, et al. Prostate cancer-associated mutations in speckle-type POZ protein (SPOP) regulate steroid receptor coactivator 3 protein turnover. Proc Natl Acad Sci U S A 2013;110:6997–7002.
24. Geng C, Rajapakshe K, Shah SS, Shou J, Eedunuri VK, Foley C, et al. Androgen receptor is the key transcriptional mediator of the tumor suppressor SPOP in prostate cancer. Cancer Res 2014;74: 5631–43.
25. Boysen G, Barbieri CE, Prandi D, Blattner M, Chae SS, Dahija A, et al. SPOP mutation leads to genomic instability in prostate cancer. Elife 2015;4:e09207.
26. García-Flores M, Casanova-Salas I, Rubio-Briones J, Calatrava A, Domínguez-Escrig J, Rubio L, et al. Clinico-pathological significance of the molecular alterations of the SPOP gene in prostate cancer. Eur J Cancer 2014;50:2994–3002.
27. Stangl A, Wilner C, Li P, Maahs L, Hwang C, Pilling A. Molecular features and race-associated outcomes of SPOP-mutant metastatic castration-resistant prostate cancer. Prostate 2023;83:524–33.
28. Boysen G, Rodrigues DN, Rescigno P, Seed G, Dolling D, Riisnaes R, et al. SPOP-mutated/CHD1-deleted lethal prostate cancer and abiraterone sensitivity. Clin Cancer Res 2018;24:5585–93.
29. Nakazawa M, Fang M, H Marshall C, Lotan TL, Isaacsson Velho P, Antonarakis ES. Clinical and genomic features of SPOP-mutant prostate cancer. Prostate 2022;82:260–8.
30. Swami U, Graf RP, Nussenzveig RH, Fisher V, Tukachinsky H, Schrock AB, et al. SPOP mutations as a predictive biomarker for androgen receptor axis-targeted therapy in de novo metastatic castration-sensitive prostate cancer. Clin Cancer Res 2022;28: 4917–25.
31. Tang C, Pan Y, Luo H, Xiong W, Zhu H, Ruan H, et al. Hedgehog signaling stimulates the conversion of cholesterol to steroids. Cell Signal 2015;27:487–97.
32. Kothandapani A, Lewis SR, Noel JL, Zacharski A, Krellwitz K, Baines A, et al. GLI3 resides at the intersection of hedgehog and androgen action to promote male sex differentiation. PLoS Genet 2020;16: e1008810.
33. Burleson M, Deng JJ, Qin T, Duong TM, Yan Y, Gu X, et al. GLI3 is stabilized by SPOP mutations and promotes castration resistance via functional cooperation with androgen receptor in prostate cancer. Mol Cancer Res 2022;20:62–76.
34. Chen G, Goto Y, Sakamoto R, Tanaka K, Matsubara E, Nakamura M, et al. GLI1, a crucial mediator of sonic hedgehog signaling in prostate cancer, functions as a negative modulator for androgen receptor. Biochem Biophys Res Commun 2011;404:809–15.
35. Li N, Truong S, Nouri M, Moore J, Al Nakouzi N, Lubik AA, et al. Non-canonical activation of hedgehog in prostate cancer cells mediated by the interaction of transcriptionally active androgen receptor proteins with Gli3. Oncogene 2018;37:2313–25.
36. Wang C, Pan Y, Wang B. Suppressor of fused and Spop regulate the stability, processing and function of Gli2 and Gli3 full-length activators but not their repressors. Development 2010;137: 2001–9.
37. Zhang Q, Shi Q, Chen Y, Yue T, Li S, Wang B, et al. Multiple Ser/Thr-rich degrons mediate the degradation of Ci/Gli by the Cul3-HIB/SPOP E3 ubiquitin ligase. Proc Natl Acad Sci U S A 2009;106: 21191–6.
38. Linder S, Hoogstraat M, Stelloo S, Eickhoff N, Schuurman K, de Barros H, et al. Drug-induced epigenomic plasticity reprograms circadian rhythm regulation to drive prostate cancer toward androgen independence. Cancer Discov 2022;12:2074–97.
39. Pomerantz MM, Qiu X, Zhu Y, Takeda DY, Pan W, Baca SC, et al. Prostate cancer reactivates developmental epigenomic programs during metastatic progression. Nat Genet 2020;52:790–9.
40. Stelloo S, Nevedomskaya E, Kim Y, Schuurman K, Valle-Encinas E, Lobo J, et al. Integrative epigenetic taxonomy of primary prostate cancer. Nat Commun 2018;9:4900.
41. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57–74.
42. Kalderon D. Transducing the hedgehog signal. Cell 2000;103: 371–4.
43. Chen M, Feuerstein MA, Levina E, Baghel PS, Carkner RD, Tanner MJ, et al. Hedgehog/Gli supports androgen signaling in androgen deprived and androgen independent prostate cancer cells. Mol Cancer 2010; 9:89.
44. Smelkinson MG. The hedgehog signaling pathway emerges as a pathogenic target. J Dev Biol 2017;5:14.
45. Lee RTH, Zhao Z, Ingham PW. Hedgehog signalling. Development 2016;143:367–72.
46. Infante P, Alfonsi R, Botta B, Mori M, Di Marcotullio L. Targeting GLI factors to inhibit the Hedgehog pathway. Trends Pharmacol Sci 2015;36:547–58.
47. Xu S, Tang C. Cholesterol and hedgehog signaling: mutual regulation and beyond. Front Cell Dev Biol 2022;10:774291.
48. Bose HS, Whittal RM, Marshall B, Rajapaksha M, Wang NP, Bose M, et al. A novel mitochondrial complex of aldosterone synthase, steroidogenic acute regulatory protein, and Tom22 synthesizes aldosterone in the rat heart. J Pharmacol Exp Ther 2021;377:108–20.
49. Finco I, LaPensee CR, Krill KT, Hammer GD. Hedgehog signaling and steroidogenesis. Annu Rev Physiol 2015;77:105–29.
50. Sakai M, Martinez-Arguelles DB, Aprikian AG, Magliocco AM, Papadopoulos V. De novo steroid biosynthesis in human prostate cell lines and biopsies. Prostate 2016;76:575–87.
51. Chang K-H, Li R, Kuri B, Lotan Y, Roehrborn CG, Liu J, et al. A gain-of-function mutation in DHT synthesis in castration-resistant prostate cancer. Cell 2013;154:1074–84.
52. McManus JM, Bohn K, Alyamani M, Chung Y-M, Klein EA, Sharifi N. Rapid and structure-specific cellular uptake of selected steroids. PLoS One 2019;14:e0224081.
53. Gao K, Jin X, Tang Y, Ma J, Peng J, Yu L, et al. Tumor suppressor SPOP mediates the proteasomal degradation of progesterone receptors (PRs) in breast cancer cells. Am J Cancer Res 2015;5:3210–20.
54. Andersson S, Berman DM, Jenkins EP, Russell DW. Deletion of steroid 5 alpha-reductase 2 gene in male pseudohermaphroditism. Nature 1991;354:159–61.
55. Carter H, Marty R, Hofree M, Gross AM, Jensen J, Fisch KM, et al. Interaction landscape of inherited polymorphisms with somatic events in cancer. Cancer Discov 2017;7:410–23.
56. Houlahan KE, Khan A, Greenwald NF, Vivas CS, West RB, Angelo M, et al. Germline-mediated immunoediting sculpts breast cancer subtypes and metastatic proclivity. Science 2024;384:eadh8697.
57. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. Nat Genet 2018; 50:928–36.
58. Hua JT, Ahmed M, Guo H, Zhang Y, Chen S, Soares F, et al. Risk SNP-mediated promoter-enhancer switching drives prostate cancer through lncRNA PCAT19. Cell 2018;174:564–75.e18.
59. Sharifi N, Azad AA, Patel M, Hearn JWD, Wozniak M, Zohren F, et al. HSD3B1 genotype and outcomes in metastatic hormone-sensitive prostate cancer with androgen deprivation therapy and enzalutamide: ARCHES. Cell Rep Med 2024;5:101644.
60. Sharifi N, Diaz R, Lin HM, Roberts E, Horvath LG, Martin A, et al. Survival of men with metastatic hormone-sensitive prostate cancer and adrenal-permissive HSD3B1 inheritance. J Clin Invest 2024;134: e183583.

61. FitzGerald LM, Agalliu I, Johnson K, Miller MA, Kwon EM, Hurtado-Coll A, et al. Association of TMPRSS2-ERG gene fusion with clinical characteristics and outcomes: results from a population-based study of prostate cancer. BMC cancer 2008;8:230.
62. Kohaar I, Li Q, Chen Y, Ravindranath L, Young D, Ali A, et al. Association of germline genetic variants with TMPRSS2-ERG fusion status in prostate cancer. Oncotarget 2020;11:1321–33.
63. Luedeke M, Linnert CM, Hofer MD, Surowy HM, Rinckleb AE, Hoegel J, et al. Predisposition for TMPRSS2-ERG fusion in prostate cancer by variants in DNA repair genes. Cancer Epidemiol Biomarkers Prev 2009;18:3030–5.
64. Hofer MD, Kuefer R, Maier C, Herkommer K, Perner S, Demichelis F, et al. Genome-wide linkage analysis of TMPRSS2-ERG fusion in familial prostate cancer. Cancer Res 2009;69:640–6.
65. Penney KL, Pettersson A, Shui IM, Graff RE, Kraft P, Lis RT, et al. Association of prostate cancer risk variants with TMPRSS2:ERG status: evidence for distinct molecular subtypes. Cancer Epidemiol Biomarkers Prev 2016;25:745–9.
66. Taylor RA, Fraser M, Livingstone J, Espiritu SMG, Thorne H, Huang V, et al. Germline BRCA2 mutations drive prostate cancers with distinct evolutionary trajectories. Nat Commun 2017;8:13671.
67. Giannareas N, Zhang Q, Yang X, Na R, Tian Y, Yang Y, et al. Extensive germline-somatic interplay contributes to prostate cancer progression through HNF1B co-option of TMPRSS2-ERG. Nat Commun 2022;13:7320.
68. Schoenfelder S, Fraser P. Long-range enhancer-promoter contacts in gene expression control. Nat Rev Genet 2019;20:437–55.
69. Bower G, Kvon EZ. Genetic factors mediating long-range enhancer-promoter communication in mammalian development. Curr Opin Genet Dev 2025;90:102282.
70. Friman ET, Flyamer IM, Marenduzzo D, Boyle S, Bickmore WA. Ultra-long-range interactions between active regulatory elements. Genome Res 2023;33:1269–83.
71. Benabdallah NS, Williamson I, Illingworth RS, Kane L, Boyle S, Sengupta D, et al. Decreased enhancer-promoter proximity accompanying enhancer activation. Mol Cell 2019;76:473–84.e7.
72. Gupta RM, Hadaya J, Trehan A, Zekavat SM, Roselli C, Klarin D, et al. A genetic variant associated with five vascular diseases is a distal regulator of endothelin-1 gene expression. Cell 2017;170:522–33.e15.
73. Karhadkar SS, Bova GS, Abdallah N, Dhara S, Gardner D, Maitra A, et al. Hedgehog signalling in prostate regeneration, neoplasia and metastasis. Nature 2004;431:707–12.
74. Sanchez P, Hernández AM, Stecca B, Kahler AJ, DeGueme AM, Barrett A, et al. Inhibition of prostate cancer proliferation by interference with SONIC HEDGEHOG-GLI1 signaling. Proc Natl Acad Sci U S A 2004;101:12561–6.
75. Kaushal JB, Raut P, Halder S, Alsafwani ZW, Parte S, Sharma G, et al. Oncogenic potential of truncated-Gli3 via the Gsk3β/Gli3/AR-V7 axis in castration-resistant prostate cancer. Oncogene 2025;44:1007–23.
76. Levina E, Chen M, Carkner R, Shtutman M, Buttyan R. Paracrine Hedgehog increases the steroidogenic potential of prostate stromal cells in a Gli-dependent manner. Prostate 2012;72:817–24.
77. Bernasocchi T, El Tekle G, Bolis M, Mutti A, Vallerga A, Brandt LP, et al. Dual functions of SPOP and ERG dictate androgen therapy responses in prostate cancer. Nat Commun 2021;12:734.
78. Hou Z, Huang S, Mei Z, Chen L, Guo J, Gao Y, et al. Inhibiting 3βHSD1 to eliminate the oncogenic effects of progesterone in prostate cancer. Cell Rep Med 2022;3:100561.
79. Eeles RA, Olama AA, Benlloch S, Saunders EJ, Leongamornlert DA, Tymrakiewicz M, et al. Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. Nat Genet 2013;45:385–91.e1–2.
80. Yamaguchi TN, Houlahan KE, Zhu H, Kurganovs N, Livingstone J, Fox NS, et al. The germline and somatic origins of prostate cancer heterogeneity. Cancer Discov 2025;15:988–1017.
81. Bae S, Park J, Kim J-S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics 2014;30:1473–5.
82. Miyaoka Y, Chan AH, Judge LM, Yoo J, Huang M, Nguyen TD, et al. Isolation of single-base genome-edited human iPS cells without antibiotic selection. Nat Methods 2014;11:291–3.
83. Nelson PS, Clegg N, Arnold H, Ferguson C, Bonham M, White J, et al. The program of androgen-responsive genes in neoplastic prostate epithelium. Proc Natl Acad Sci U S A 2002;99:11890–5.
84. Shaffer PL, Jivan A, Dollins DE, Claessens F, Gewirth DT. Structural basis of androgen receptor binding to selective androgen response elements. Proc Natl Acad Sci U S A 2004;101:4758–63.
85. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 2016;32:3047–8.
86. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–20.
87. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
88. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–40.
89. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation (Camb) 2021;2:100141.
90. Chakravarty D, Sboner A, Nair SS, Giannopoulou E, Li R, Hennig S, et al. The oestrogen receptor alpha-regulated lncRNA NEAT1 is a critical modulator of prostate cancer. Nat Commun 2014;5:5383.
91. Habegger L, Sboner A, Gianoulis TA, Rozowsky J, Agarwal A, Snyder M, et al. RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries. Bioinformatics 2011;27:281–3.
92. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25:2078–9.
93. Alaimo A, Genovesi S, Annesi N, De Felice D, Subedi S, Macchia A, et al. Sterile inflammation via TRPM8 RNA-dependent TLR3-NF-kB/IRF3 activation promotes antitumor immunity in prostate cancer. EMBO J 2024;43:780–805.
94. Genovesi S, Moro R, Vignoli B, De Felice D, Canossa M, Montironi R, et al. Trpm8 expression in human and mouse castration resistant prostate adenocarcinoma paves the way for the preclinical development of TRPM8-based targeted therapies. Biomolecules 2022;12:193.