

1 **Dysregulated Alternative Splicing Landscape Identifies Intron Retention**  
2 **as a Hallmark and Spliceosome as a Therapeutic Vulnerability in**  
3 **Aggressive Prostate Cancer**

4 Dingxiao Zhang<sup>1,†\*</sup>, Qiang Hu<sup>2†</sup>, Yibing Ji<sup>1</sup>, Hsueh-Ping Chao<sup>3</sup>, Amanda Tracz<sup>1</sup>, Jason Kirk<sup>1</sup>, Silvia  
5 Buonamici<sup>4</sup>, Ping Zhu<sup>4</sup>, Jianmin Wang<sup>2</sup>, Song Liu<sup>2\*</sup>, Dean G. Tang<sup>1,3\*</sup>

6

7 <sup>1</sup>Department of Pharmacology and Therapeutics, Roswell Park Comprehensive Cancer Center, Elm  
8 and Carlton Streets, Buffalo, NY 14263, USA

9 <sup>2</sup>Department of Biostatistics and Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo,  
10 14263 New York, USA

11 <sup>3</sup>Department of Epigenetics and Molecular Carcinogenesis, University of Texas M.D Anderson Cancer  
12 Center, Smithville, 78957 Texas, USA

13 <sup>4</sup>H3 Biomedicine, Inc., 300 Technology Square, Cambridge, MA 02139, USA

14

15

16 **Key words:** prostate cancer, alternative splicing, intron retention, spliceosome, aggressiveness,  
17 stemness

18

19 **Running title:** Linking splicing dysregulation to PCa aggressiveness

20

21

22 †These authors contributed equally to this work.

23

24 **\*Correspondence:** [zhangdingxiao1980@yahoo.com](mailto:zhangdingxiao1980@yahoo.com) or [Song.Liu@Roswellpark.org](mailto:Song.Liu@Roswellpark.org) or

25 [Dean.Tang@Roswellpark.org](mailto:Dean.Tang@Roswellpark.org).

26

27 The authors declare no potential conflicts of interest.

28

29

30

35 **ABSTRACT (150 words)**

36

37 Dysregulation of mRNA alternative splicing (AS) has been implicated in development and progression  
38 of hematological malignancies. Here we describe the first comprehensive AS landscape in the  
39 spectrum of human prostate cancer (PCa) development, progression and therapy resistance. We find  
40 that the severity of splicing dysregulation correlates with disease progression and establish intron  
41 retention (IR) as a hallmark of PCa stemness and aggressiveness. Systematic interrogation of 274  
42 splicing-regulatory genes (SRGs) uncovers prevalent SRG mutations associated with, mainly, copy  
43 number variations leading to mis-expression of ~68% of SRGs during PCa evolution. Consequently,  
44 we identify many SRGs as prognostic markers associated with splicing disruption and patient  
45 outcome. Interestingly, androgen receptor (AR) controls a splicing program distinct from its  
46 transcriptional regulation. The spliceosome modulator, E7107, reverses cancer aggressiveness and  
47 abolishes the growth of castration-resistant PCa (CRPC) models. Altogether, we establish aberrant  
48 AS landscape caused by dysregulated SRGs as a novel therapeutic vulnerability for CRPC.

49

50

51 **Statement of significance (49 words)**

52 We present the first comprehensive AS landscape during PCa evolution and link genomic and  
53 transcriptional alterations in SRGs to global splicing dysregulation. AR regulates splicing in pri-PCa  
54 and CRPC distinct from its transcriptional regulation. Intron retention is a hallmark for and  
55 spliceosome represents a therapeutic vulnerability in aggressive PCa.

56

57

## 58 **INTRODUCTION**

59

60 Prostate cancer (PCa) still causes a significant mortality among men world-wide (1). The prostate is  
61 an exocrine gland containing mainly androgen receptor negative (AR<sup>-</sup>) basal and AR<sup>+</sup> luminal  
62 epithelial cells, together with rare neuroendocrine (NE) cells (2,3). PCa predominantly displays a  
63 luminal phenotype and histologically presents as adenocarcinomas (Ad) largely devoid of basal cells  
64 (4). Most primary PCa (pri-PCa) are diagnosed as low to intermediate grade (i.e., Gleason grade  $\leq 7$ ),  
65 relatively indolent, and treated by radical prostatectomy and/or radiation with a good prognosis.  
66 Locally advanced (Gleason grade 9/10) and metastatic PCa are generally treated with androgen  
67 deprivation therapy (ADT) using LHRH agonists/antagonists, which block testicular androgen  
68 synthesis. Tumors that have failed this first-line therapy are termed castration-resistant PCa (CRPC)  
69 and further treated with anti-androgens such as enzalutamide (Enza) that interfere with the functions  
70 of AR. Enza extends CRPC patients' lives by ~5 months but tumors inevitably become refractory to  
71 Enza. While the majority of CRPC and Enza-resistant tumors histologically present as  
72 adenocarcinoma (i.e., CRPC-Ad), a significant fraction (up to 25%) of them evolve to an aggressive,  
73 AR-indifferent disease with NE features called CRPC-NE (5). In general, all CRPC are relatively  
74 undifferentiated and, molecularly, basal/stem-like (6,7), highlighting lineage plasticity in facilitating  
75 treatment resistance and progression (8). Most metastatic CRPC (mCRPC), including both CRPC-Ad  
76 and CRPC-NE subtypes, remains lethal mainly due to our incomplete understanding of mechanisms  
77 underpinning CRPC emergence, maintenance and progression.

78

79 Dysregulation in pre-mRNA alternative splicing (AS) is emerging as a 'hallmark' of cancer (9,10).  
80 Nearly all multi-exon human genes undergo AS, a tightly regulated process that dramatically expands  
81 diversity of the transcriptome and proteome encoded by the genome (11). As an essential process for  
82 removing non-coding introns and ligating flanking exons to produce mature mRNA in eukaryotic cells,  
83 AS is performed by a dynamic and flexible macromolecular machine, the spliceosome. In addition to

84 the core subunits that constitute five small nuclear ribonucleoprotein particles (snRNPs: U1, U2, U4,  
85 U5 and U6), the spliceosome contains many other auxiliary splicing-regulatory proteins (SRPs)  
86 including families of the serine- and arginine-rich (SR) proteins and heterogeneous nuclear  
87 ribonucleoproteins (hnRNP), and many other proteins that do not belong to these two families but  
88 have a role in modulating splicing (12). In this study, we refer to the genes encoding core subunits  
89 and the SRPs, broadly, as splicing-regulatory genes (SRGs) ([Supplementary Table 1](#)). Aberrant AS is  
90 prevalent in human cancers (13) and many cancer-specific splicing events contribute to disease  
91 development and progression (9,12). Since the initial discovery, via DNA sequencing, of frequent  
92 point mutations in the core spliceosome subunits in myelodysplastic syndromes and, later, in  
93 hematological malignancies (9), splicing dysregulation has been appreciated as a major contributor to  
94 cancer phenotypes. In parallel, therapeutic targeting of mis-splicing by small molecules presents a  
95 new approach for treating hematological malignancies bearing core subunit mutations (10,12) and  
96 solid tumors driven by MYC (14). Nevertheless, despite increasing elucidation of the global and  
97 cancer-associated splicing features by recent RNA sequencing (RNA-seq) analyses of primary tumors  
98 and normal tissues (13), the underlying molecular mechanisms as well as functional and clinical  
99 relevance of splicing misregulation in cancer, especially in solid tumors, remain largely undefined.

100

101 Importantly, many fewer recurrent mutations in the core spliceosome genes have been detected to  
102 date in solid tumors (15,16), suggesting a fundamental and mechanistic difference in splicing  
103 misregulation in hematological versus (vs.) solid cancers. Recently, global analyses of aberrant AS  
104 landscape across many human cancer types, including PCa, have been reported using RNA-seq data  
105 in TCGA (13,17-19), but these studies generally overlooked PCa and only analyzed pri-PCa, leaving  
106 behind life-threatening mCRPC. Furthermore, regarding the functional consequence of splicing  
107 dysregulation in PCa, previous studies have mainly focused on a few well-known genes typified by AR  
108 and CD44 (20), and the potential biological impact and clinical relevance of global splicing  
109 abnormalities in PCa remains unclear. Here, we specifically focus on PCa and provide, to our

110 knowledge, the first comprehensive characterization of the AS landscape during disease development  
111 and progression and upon treatment failure. We report that intron retention (IR) represents the most  
112 salient and consistent feature across the spectrum of PCa entities and positively correlates with PCa  
113 stemness and aggressiveness. We also systematically analyze the dysregulated SRGs and correlate  
114 altered SRGs with aberrant AS patterns in PCa, and examine the deregulated pathways affected by  
115 aberrant splicing events. Finally, we demonstrate that splicing misregulation can be explored  
116 therapeutically for treating CRPC.

117

118

119

## 120 **RESULTS**

121

### 122 **Increased AS Events Accompany PCa Development, Progression and Therapy Resistance**

123 To determine the global dysregulation of AS in PCa development and progression, we employed two  
124 AS mapping algorithms, rMATS (21) and SUPPA (22), to annotate RNA-seq datasets encompassing  
125 pri-PCa and normal (N) prostate tissues (23), CRPC-Ad (24,25), CRPC-NE (5,26), and advanced PCa  
126 treated with hormonal therapy (26,27) (Fig. 1; Supplementary Fig. S1A). We defined ‘progression’  
127 generally as stages beyond pri-PCa and as disease entities that were more aggressive in a  
128 comparative manner. Five main AS patterns, including alternative 3’ splice sites (A3), alternative 5’  
129 splice sites (A5), mutually exclusive exons (MX), exon skipping (SE) and IR, were examined  
130 (Supplementary Fig. S1B). Splicing events with a cutoff of  $\Delta\text{PSI}>0.1$  and  $\text{FDR}<0.1$  (for rMATS) or  
131  $p<0.05$  (for SUPPA) were considered statistically significant (see Methods).

132 Comparative analyses of either bulk or paired tumors and normal tissues indicated that pri-PCa  
133 possessed more AS events (~1.9 fold by rMATS; ~1.7 fold by SUPPA) with preferential increase in  
134 A3, A5 and IR (Fig. 1A and 1B; Supplementary Table S2). PCa post ADT (Fig. 1C) or subjected to  
135 neo-adjuvant hormone therapy (NHT; Fig. 1D) also displayed increased differentially spliced events  
136 (DSEs), suggesting a treatment-induced reshaping of global AS pattern that might have contributed to  
137 therapy resistance. Strikingly, mCRPC that had failed ADT and/or antiandrogens exhibited an  
138 exponential increase in DSEs, with noticeable increase in A3, A5, SE and IR (Fig. 1E and 1F). Within  
139 CRPC, CRPC-NE harbored a distinct splicing landscape relative to CRPC-Ad, although a notably  
140 smaller number of DSEs were observed than that in the CRPC-Ad vs. pri-PCa comparison (1712 vs.  
141 18318; Fig. 1G). Interestingly, when comparing pri-PCa vs. pri-PCa with NE differentiation post NHT  
142 (26), we identified 9364 DSEs (Fig. 1H). Mapping with SUPPA revealed overall similar dysregulated  
143 AS patterns and progressive increase in DSEs in the spectrum of PCa development, therapy  
144 resistance, and progression (Supplementary Fig. S1C and S1D) albeit SUPPA by nature tended to  
145 detect more splicing events (see Methods). These results, taken together, suggest that PCa

146 development is accompanied by increased AS events and that castration resistance and, in particular,  
147 metastasis, are characterized by further significant increases in AS events.

148 Since lineage plasticity facilitates therapeutic resistance and tumor progression (6,28), we  
149 determined the human prostate epithelial lineage-specific AS patterns as basal cells represent the  
150 main pool of prostate stem cells (SCs) and molecularly resemble aggressive PCa subtypes (7).  
151 Results revealed distinct AS profiles for prostatic basal vs. luminal cells, with more IR found in basal  
152 cells (Fig. 1I; Supplementary Fig. S1C). To determine whether basal-specific splicing profile also  
153 resembles that in aggressive PCa, we performed comparative gene set enrichment analysis (GSEA)  
154 and found that PCa with aggressive phenotypes (mCRPC and CRPC-NE) generally possessed a  
155 global basal-like AS profile (Supplementary Fig. S1E). Experimentally, silencing of tumor suppressors  
156 (TS) TP53 and RB1 in LNCaP/AR cells enables a lineage switch from AR<sup>+</sup> luminal cells to AR<sup>-</sup> basal-  
157 like cells (28). Consistently, a large number of DSEs were observed in LNCaP/AR cells with  
158 RB1/TP53 knockdown (Fig. 1J; Supplementary Fig. S1C), suggesting that plasticity driven by loss of  
159 RB1/TP53 is accompanied by a global shift in the AS landscape. Remarkably, GSEA indicated that  
160 the AS signatures of LNCaP/AR cells deficient in RB1/TP53 were significantly enriched in mCRPC  
161 compared to pri-PCa (Supplementary Fig. S1F). These results suggest that inherent lineage  
162 differences in normal prostate epithelial cells and induced lineage plasticity in PCa cells are also  
163 accompanied by dysregulated AS patterns that correlate with increased aggressiveness.

164

### 165 **AS Dysregulation Impacts PCa Biology**

166 We explored the potential impact of AS dysregulation on PCa biology (Supplementary Fig. S2 and  
167 S3). By overlapping the splicing-affected genes (SAGs) and differentially expressed genes (DEGs),  
168 we observed only 9~20% of 'overlapped' genes (Supplementary Fig. S2A), suggesting that the  
169 majority of AS events minimally changes the bulk gene expression but may functionally tune  
170 transcriptomes (29). Gene ontology (GO) analysis (<http://metascape.org>) indicated that SAGs were  
171 enriched in many cancer-associated functional categories with both convergence (e.g., splicing, cell

172 cycle and proliferation, cytoskeleton) and specificity identified at each PCa stage ([Supplementary Fig.](#)  
173 [S2B-D](#)). For instance, GO terms linked to ‘muscle and ion transport’, ‘lipid metabolism’, and ‘cell  
174 polarity’ were pri-PCa specific ([Supplementary Fig. S2B](#)) whereas GO terms ‘DNA damage’,  
175 ‘immunity’, and ‘nuclear pore’ were enriched in CRPC ([Supplementary Fig. S2C](#)), consistent with  
176 recent reports (30). Interestingly, and as expected, GO terms ‘SCs and development’ and ‘neuron and  
177 cell projection’ were greatly enriched in CRPC-NE ([Supplementary Fig. S2D](#)).

178 We further evaluated the potential functional consequences of AS dysregulation on PCa  
179 transcriptome by identifying transcript-level expression profiles using an isoform-specific alignment  
180 algorithm (31). As shown in [Supplementary Fig. S3A](#), PCa at different stages exhibited distinct splice  
181 isoform signatures. For instance, the widely studied ARv7 was slightly upregulated in pri-PCa and,  
182 together with several other AR variants, was dramatically overexpressed in CRPC-Ad but not in  
183 CRPC-NE (due to loss of AR expression in NE tumors) ([Supplementary Fig. S3B](#)). CD44, a cancer  
184 stem cell (CSC) marker, plays versatile roles in metastasis with CD44-standard (CD44s) suppressing  
185 and CD44 variants (CD44v) promoting cancer cell colonization (32). Consistently, we observed a shift  
186 from no change in pri-PCa to a specific dysregulation of CD44 isoforms in mCRPC, with CD44s being  
187 downregulated in both CRPC-Ad and CRPC-NE and CD44v upregulated in CRPC-Ad  
188 ([Supplementary Fig. S3C](#)). The splicing program driving CRPC-NE emergence is scantily explored.  
189 Recently, an SE event leading to unique upregulation of a MEAF6 isoform containing exon 6 (i.e.,  
190 MEAF6-204), but not the bulk mRNA, was reported in CRPC-NE (33). We observed similar results  
191 ([Supplementary Fig. S3D](#)), thus validating our analytic approaches. These analyses indicate that  
192 splicing abnormalities impact PCa biology by regulating cancer-related pathways, at least partially, via  
193 switching the isoform expression of key relevant genes.

194

### 195 **Increased IR as a Consistent Hallmark of PCa Progression, Stemness and Aggressiveness**

196 Notably, we consistently observed increased IR across the spectrum of PCa development,  
197 progression and therapy resistance whereas the SE represented the most abundant splicing type



198 (Fig. 1A-1J; Supplementary Table S2). We focused our subsequent studies on IR for it is the least  
199 studied AS type (10,34). We observed a >18 fold increase in IR in pri-PCa relative to normal tissues  
200 (Fig. 1K, i), consistent with a previous report that IR is common across multiple primary cancers (34).  
201 PCa progression is tightly associated with ADT failure and subsequent cellular plasticity towards  
202 stemness (35-37). In six different contexts, we consistently observed a preferential upregulation of IR  
203 in association with therapy-resistant, aggressive, and metastatic PCa (Fig. 1K, ii). Similar IR  
204 upregulation was observed in prostate tumors and epithelial cells displaying low vs. high AR activity  
205 (Fig. 1K, iii; see below). Interestingly, increased IR was also found in CSC-enriched PSA<sup>-lo</sup> cell  
206 population isolated from LAPC9 xenografts (38), basal-like LNCaP cells depleted of TP53 and RB1  
207 (28), and LNCaP-CRPC cells that survived long-term Enza treatment (>8 weeks (39) (Fig. 1K, iv). Of  
208 note, SUPPA produced similar results (data not shown). These analyses link the upregulated IR with  
209 PCa stemness. We reanalyzed 3 recently published datasets that examined differentiation of different  
210 SC systems and also observed a positive correlation between IR and normal stemness  
211 (Supplementary Fig. S4A-C). Hence, in genetically matched hESCs – fibroblasts – iPS – fibroblasts  
212 system (40), ESCs lost IR during fibroblast differentiation while fibroblasts regained IR when they  
213 were reprogrammed to iPS cells (Supplementary Fig. S4A), in line with the earlier report (41). During  
214 spermatogenesis, spermatocytes displayed higher levels of IR than differentiated spermatids  
215 (Supplementary Fig. S4B), and these IR events were enriched in genes associated with gamete  
216 function (42). Finally, IR was found to be prevalent in stem-like, resting CD4<sup>+</sup> T cells vs. functionally  
217 activated (differentiated) counterparts (Supplementary Fig. S4C), as reported previously (43).

218 We investigated the splicing ‘code’ of IR (44) in attempt to understand the molecular basis of  
219 preferential IR in aggressive PCa. Retained introns in normal tissues generally have weak 5’ and 3’  
220 splice sites (44). Surprisingly, the splice site strength analysis did not reveal weak 5’ or 3’ sites in  
221 retained introns in pri-PCa, CRPC-Ad and CRPC-NE – in fact, CRPC-Ad showed stronger splice sites  
222 than pri-PCa (Supplementary Fig. S4D). Sequence feature analysis indicated that, compared with  
223 constitutive introns, IR in pri-PCa preferred introns with less GC content (GC%) and longer sequence

224 length whereas CRPC-Ad specifically retained introns that were generally shorter without difference in  
225 GC% ([Supplementary Fig. S4E](#)). No feature variation was observed in the retained introns in CRPC-  
226 NE vs. CRPC-Ad ([Supplementary Fig. S4E](#)). Together, these results suggest that the prevalence of IR  
227 in PCa is not associated with weak 5' and 3' splice sites and may largely be *trans*-regulated.

228 To determine the potential *trans* factors (i.e., SRGs) that may preferentially regulate IR, we  
229 performed motif search for 95 RNA binding proteins (RBP) with known consensus motifs (45-47) on  
230 differentially splicing introns compared with constitutive introns. Based on an RBP-binding score for  
231 each factor, we chose top 20 genes for further analysis. As shown in [Fig. 1L](#), we identified a few  
232 genes that may preferentially regulate IR for each specific comparison. Nonetheless, the majority of  
233 RBPs were shared by introns regardless of the IR status, suggesting that the spliceosome functions  
234 as a group rather than that one particular factor preferentially regulates one AS type. In support, we  
235 decoded the AS events associated with gene expression abundance by fractionating a cohort into two  
236 extremes ([Fig. 1M](#)). As expected, although the expression of ELAVL1 in pri-PCa and RBM38 in  
237 CRPC-Ad cohorts, respectively, both dramatically impacted IR, other splicing types were affected as  
238 well ([Fig. 1M](#)). Interestingly, ELAVL1 was not dysregulated in pri-PCa vs. normal tissues (FC=1.1).  
239 The discrepancy between a potential IR-inhibiting function of ELAVL1 and a marked increase in IR  
240 implied an involvement of other SRGs in preferential (or balanced) regulation of IR in pri-PCa. On the  
241 other hand, an IR-inhibiting function of RBM38 was consistent with its downregulation (FC=2.3) and  
242 an increase in IR in CRPC-Ad vs. pri-PCa ([Fig. 1M](#)). A TS role has been reported for RBM38 (48).

243 Subsequently, we interrogated potential biological impact of the upregulated IR on PCa biology.  
244 IR in normal conditions usually causes nonsense-mediated RNA decay (NMD) to downregulate gene  
245 expression (10,49). We compared the bulk RNA levels of IR-affected genes using two different  
246 mathematical methods and found that, surprisingly, these genes generally exhibited higher expression  
247 than their constitutively spliced counterparts ([Fig. 1N](#); [Supplementary Fig. S4F](#)). To further strengthen  
248 our finding, we overlapped the IR-affected genes with a high-confidence set of human NMD targets  
249 (50) and found that only ~10% of genes in all groups were potentially targeted by NMD, although the

250 genes with upregulated IR in CRPC tended to have slightly higher percentage ( $\chi^2$  test, [Fig. 1O](#)).  
251 These results indicate that IR in PCa minimally causes NMD-mediated downregulation and these IR-  
252 bearing genes are thus likely functional. In support, GO analysis of IR-affected genes revealed that, in  
253 addition to commonly observed category of ‘splicing and RNA metabolism’, several distinct categories  
254 were enriched in aggressive PCa ([Supplementary Fig. S4G-I](#)). For example, GO terms ‘stress  
255 response’, ‘DNA repair’ and ‘cancer-related signaling’ (e.g., ERBB, NOTCH, WNT) were unique to  
256 CRPC-Ad ([Supplementary Fig. S4H](#)) whereas ‘hormone transport’ and ‘SC & development’ were  
257 strongly associated with androgen-insensitive and CSC-enriched CRPC-NE ([Supplementary Fig. S4I](#)).

258

### 259 **AR Regulates a Splicing Program Distinct from the AR-Regulated Transcriptome**

260 AR is obligatory for pri-PCa growth and continues to be expressed and functionally important in CRPC  
261 (51). ADT promotes a stem-like phenotype in PCa (52) and relapsed tumors often exhibit enhanced  
262 SC properties (8,38,53). We set out to determine whether AR may drive splicing dysregulation seen in  
263 PCa development and progression. We first established an AR activity score based on the Z-scores  
264 calculated from the expression of 20 experimentally validated AR targets (23). The TCGA cohort  
265 bearing ‘uninterrupted’ intrinsic AR heterogeneity (23) and CRPC-Ad cohort bearing ‘twisted’ AR  
266 activity by treatments (24) were then fractionated into high and low AR activity groups, followed by  
267 splicing analyses. Not surprisingly, primary tumors with low vs. high AR activities displayed a  
268 significant difference in AS landscape ([Fig. 2A](#)), and this difference was amplified in CRPC ([Fig. 2B](#)),  
269 implicating AR signaling in modulating global AS. Of note, we observed no association between AR  
270 genomic alterations and its potential splicing-modulating activity ([Fig. 2C](#)) since AR is rarely altered in  
271 pri-PCa but frequently amplified in mCRPC (24). To assess the impact of AR-associated splicing on  
272 AR-regulated gene expression, we compared the SAGs with DEGs identified in the AR-low vs. AR-  
273 high comparisons. Surprisingly, only 2% of SAGs overlapped with the DEGs in pri-PCa, although this  
274 overlap was increased to 23.2% in CRPC-Ad (probably due to a much-enlarged repertoire of AR-  
275 regulated molecular events in CRPC) ([Fig. 2D](#)). Thus, AR activity-associated AS events exerted a

276 limited impact on the AR transcriptional targets, leading us to hypothesize that AR regulates a splicing  
277 program distinct from its transcriptional regulation. In support, when we extended the comparison to  
278 three other well-defined AR-target gene sets ((54), and two in this study, see below), we observed  
279 generally <4% overlaps across all comparisons (Fig. 2E).

280 To experimentally validate our hypothesis, we treated AR<sup>+</sup> LNCaP cultures with various  
281 regimens to modulate AR activity (Fig. 2F). Cells cultured in regular fetal bovine serum (FBS)-  
282 containing medium represent an androgen-dependent (AD, or androgen-sensitive) state. Cells grown  
283 for 4 days in medium containing charcoal/dextran stripped serum (CDSS) or treated with Enza (i.e.,  
284 MDV3100, 10  $\mu$ M) were considered androgen-independent (AI). We also utilized siRNA to genetically  
285 silence AR. Finally, cells primed with CDSS for 3 days were treated with 10 nM dihydrotestosterone  
286 (DHT) for 8 h to restore AR signaling. Deep RNA-seq was performed in biological duplicates on  
287 abovementioned LNCaP cultures (Fig. 2F; Supplementary Fig. S5A). Principal component analysis  
288 (PCA) indicated that samples were properly clustered (Fig. 2G) and AR signaling was effectively  
289 modulated as intended, evidenced by expression levels of AR and PSA and by GSEA of AR gene  
290 signature (Supplementary Fig. S5B) and by quantitative reverse transcription PCR (qRT-PCR)  
291 validation (Fig. 2H). Pairwise comparisons uncovered significant differences in DSEs in cells  
292 exhibiting high vs. low AR activities (Fig. 2I; Supplementary Fig. S5C). Also, reanalysis of a recent  
293 RNA-seq dataset (GSE71797) (55) confirmed that in response to R1881 (24~48 h), activated AR  
294 signaling reshaped the AS landscape in 3 AR<sup>+</sup> PCa cell models (i.e., LNCaP, VCaP, 22Rv1)  
295 (Supplementary Fig. S5D). Similarly, by categorizing the DEGs identified in cells depleted of AR (siAR  
296 vs. siNC) or treated with DHT as AR-target sets, we found that, strikingly, these two sets, together  
297 with a previously reported AR signature (54), minimally overlapped with the SAGs (<5%) defined in all  
298 different contexts (Fig. 2J; Supplementary Fig. S5E). Collectively, we conclude that AR regulates a set  
299 of AS-bearing genes distinct from its transcriptional targets, with or without the presence of androgen.

300 We also investigated whether AR might specifically regulate IR, as tumors and basal cells with  
301 low canonical AR activity were associated with increased level of IR (Fig. 1K). Surprisingly, our work

302 in LNCaP system revealed that a decrease in AR activity resulted in no increase, but a decrease, in  
303 IR while stimulation of AR-mediated transcription failed to appreciably repress IR ([Supplementary Fig.  
304 S5F](#)). To further test this experimentally, we utilized a quantitative reporter system (56) in which a  
305 132-nucleotide chimeric  $\beta$ -globin/immunoglobulin intron was inserted into the firefly luciferase gene  
306 ([Supplementary Fig. S5G](#)). Dual luciferase assays ([Supplementary Fig. S5G](#)) indicated that consistent  
307 with previous reports (56), splicing conferred an advantage to gene expression in that equal amounts  
308 of transfected plasmids generated higher signals from intron-containing than intronless luciferases  
309 ([Supplementary Fig. S5H, left](#)). However, luciferases with or without intron generated a similar pattern  
310 of signal changes across conditions with dampened or enhanced AR signaling ([Supplementary Fig.  
311 S5H, right](#)), suggesting that AR does not specifically regulate IR in AR<sup>+</sup> PCa cells.

312

### 313 **Distinct Genomic Alterations in SRGs Impact AS and Associate with PCa Aggressiveness**

314 Recent genomic sequencing efforts have revealed the global mutational landscapes of PCa during  
315 development and progression (5,23,24,51,57-61), almost all of which focused their initial analysis on  
316 known PCa-related genes and pathways (e.g., AR, PTEN/PI3K, TP53, RB1, DNA repair, ETS fusion)  
317 whereas alterations in SRGs were overlooked due to a low mutation frequency at individual gene  
318 level. Moreover, point mutations in spliceosome core genes have been recognized as a key driver in  
319 hematological cancers (12). We explored the molecular mechanisms underpinning the AS  
320 dysregulation in PCa by compiling and curating a catalog of 274 SRGs ([Supplementary Table S1](#)) and  
321 systematically surveying their mutational landscape ([Fig. 3; Supplementary Fig. S6-S8](#)). We  
322 interrogated 11 available large-scale clinical datasets in cBioportal (62) and excluded 3 from further  
323 analysis due to limited information available ([Supplementary Table S1](#)). The remaining 8 were  
324 categorized as pri-PCa and CRPC datasets. [Fig. 3A](#) and [3B](#) showed the mutational landscape of top  
325 15 altered SRGs in representative pri-PCa and CRPC datasets (also see [Supplementary Fig. S6 and  
326 S7](#)), respectively. Based on this global mutational landscape, several interesting patterns emerged.  
327 **First**, genomic deletions of SRGs in pri-PCa and amplifications of SRGs in CRPC represented the

328 most prevalent forms of alterations, among others (Fig. 3C and 3D). **Second**, the frequently deleted  
329 and amplified genes often co-occurred with the deletion of TS genes and amplification of oncogenes,  
330 respectively (also see Supplementary Fig. S8A). For example, ENOX1, WBP4, HNRNPA1L2 and RB1  
331 were co-localized and co-deleted on Chr13q ( $p=5.16E-42$ ; Supplementary Table S1). On the other  
332 hand, KHDRBS3, PABPC1, ESRP1, PUF60 were co-amplified with MYC on 8q ( $p\leq 1.50E-15$ ). **Third**,  
333 most of the SRGs mutated at low frequency, as only 20 (7.3%) and 29 (10.58%) of the 274 SRGs  
334 were mutated at a rate of  $\geq 5\%$  in TCGA-PCa and SU2C-CRPC cohorts, respectively. Consequently,  
335 the mutation burden in sum is predominantly contributed by the top 20 altered genes (Fig. 3E;  
336 Supplementary Table S1). **Fourth**, chromosomal distribution of mutated SRGs ( $\geq 5\%$ ) revealed that,  
337 except the top altered genes, the majority of SRGs were localized outside the previously reported  
338 hotspots (23,61) (Supplementary Fig. S8B), in line with their low mutation rates. In aggregate, our  
339 data indicate that, albeit a low alteration frequency at individual gene level, SRGs, collectively,  
340 represent a frequently mutated pathway in PCa, as  $\sim 31-68\%$  and  $87-94\%$  of patients with pri-PCa and  
341 CRPC, respectively, harbor at least one mutation of one SRG (Supplementary Table S1).

342 Evolutionarily, deletion and amplification of selective SRGs might represent early and late  
343 events, respectively, in PCa pathogenesis (Fig. 3C). In support, group analysis of top altered SRGs  
344 showed that deletion of SRGs did not, whereas amplification of SRGs did, associate with increased  
345 Gleason grade (not shown), highlighting a potential survival advantage of clones harboring SRG  
346 amplifications over deletions during PCa progression. This notion is further supported by a recent  
347 study showing that focal genomic amplifications represent a rapid adaptation to selection pressure  
348 and a driving force in metastatic CRPC (63). We also observed an overall increase (e.g., 8% to 14%  
349 for KHDRBS3, 7% to 17% for ESRP1) and decrease (e.g., 18% to 7% for ENOX1, 16% to 6% for  
350 WBP4) in the frequencies of amplified and deleted genes, respectively, in CRPC vs. pri-PCa (Fig. 3A  
351 and 3B). Interestingly, SRG deletions and amplifications seemed to be mutually exclusive (Fig. 3F).

352 We reasoned that copy number variation (CNVs) in SRGs might lead to their differential mRNA  
353 expression, which in turn might be tied to splicing misregulation in PCa. Indeed, gene expression



354 analysis for top altered SRGs in both pri-PCa and CRPC indicated that deletion and amplification  
355 generally correlated with loss and gain of mRNA expression, respectively ([Supplementary Fig. S9A-](#)  
356 [B](#)). Oncomine Concept analysis revealed 72 and 74 dysregulated SRGs ( $p < 0.05$ ) in pri-PCa and  
357 CRPC, respectively ([Supplementary Fig. S9C](#)). In RNA-seq datasets, 33, 89 and 45 SRGs were  
358 significantly deregulated in pri-PCa (vs. normal tissues), CRPC-Ad (vs. pri-PCa) and CRPC-NEPC  
359 (vs. CRPC-Ad), respectively ([Supplementary Fig. S9D](#)). Furthermore, an RNA-seq examining the  
360 response of advanced PCa to ADT (27) revealed 19 DEGs, and, of interest, an exclusive  
361 overexpression of 7 genes was identified in basal vs. luminal cells ([Supplementary Fig. S9D](#)). Notably,  
362 many of the top amplified and deleted SRGs were also found to be, correspondingly, overexpressed  
363 or downregulated in PCa at the population level ([Supplementary Fig. S9E](#)). An integrated summary  
364 ([Fig. 4; Supplementary Table S3](#)) revealed that, in total, 186 out of 274 (67.9%) SRGs were mis-  
365 expressed at different stages of PCa, with more dysregulated SRGs found in CRPC, implicating a  
366 potential dependency of aggressive PCa on spliceosome activity.

367 To further explore the clinical relevance of SRGs, we assessed the prognostic values of  
368 dysregulated SRGs in patient's outcome. We systematically surveyed the 186 misregulated SRGs in  
369 7 Oncomine datasets containing patient survival information and identified two types of 'prognostic'  
370 SRGs: unfavorable genes whose higher expression correlated with poor patient survival and favorable  
371 genes whose higher expression correlated with better patient survival ([Fig. 5A; Supplementary Table](#)  
372 [S4](#)). In general, we observed a consistency between overexpressed SRGs and unfavorable  
373 prognostic genes, but not for down-regulated genes and favorable prognostic genes ([Supplementary](#)  
374 [Table S4](#)). Interestingly, although different datasets revealed varying numbers of prognostic genes  
375 ([Supplementary Table S4](#)), we identified more SRG genes classified as unfavorable genes ([Fig. 5B](#)).  
376 Together with the mutational landscape ([Fig. 3](#)) and deregulated expression patterns of SRGs ([Fig. 4](#))  
377 that cooperatively indicated a potential dependency of CRPC on spliceosome activity, this would  
378 strongly suggest that SRGs, mostly, play oncogenic roles in PCa progression. Importantly, most of the  
379 identified prognostic genes have not previously been linked to PCa patient survival. Towards a better

380 use of these prognostic SRGs in heterogeneous PCa, we established two gene signatures based on  
381 the consistency of the survival results seen in the 7 datasets, corresponding to unfavorable signature  
382 (13 genes, SRSF1, KHDRBS3, ESRP1, HNRNPH1, U2SURP, LSM5, TIA1, CHERP, HNRNPR,  
383 HNRNPH2, HNRNPH3, HNRNPAB and KHDRBS1, with each showing consistent unfavorable  
384 prognosis in  $\geq 3$  datasets) and favorable signature (13 genes, MFAP1, SF3A2, GPATCH1, XAB2,  
385 CELF2, SF3A1, SAP18, SRP54, PPIL2, SF1, MATR3, ELAVL4 and CDK10 with each showing  
386 consistent favorable prognosis in  $\geq 2$  datasets). We found that patients whose cancer gene expression  
387 enriched for the unfavorable or favorable signatures had a worse or a better survival outcome,  
388 respectively (Fig. 5C), suggesting a utility of SRGs as prognostic biomarkers. To further study  
389 underlying link between prognostic SRGs and splicing dysregulation, we investigated impact of  
390 unfavorable signature on disease aggressiveness and splicing in TCGA cohort. As expected, the  
391 unfavorable signature score positively and negatively associated with the tumor grade and disease  
392 recurrence, respectively (Fig. 5D and 5E). Importantly, primary tumors expressing highly or lowly the  
393 unfavorable signature exhibited distinct splicing landscapes, with total DSEs (1.73 fold) and IR (18.91  
394 fold) being specifically upregulated in the high group (Fig. 5F).

395

### 396 **CPPC Cells Are Sensitive to Pharmacological Spliceosome Inhibition In vitro**

397 Based on the observations that mCRPC possess frequent amplifications in (Fig. 3) and deregulation  
398 of SRGs (Fig. 4), that higher expression of SRGs predicts worse outcome (Fig. 5), and that  
399 modulation of AR activity reshapes PCa-associated AS landscape (Fig. 2), we hypothesized that  
400 *spliceosome may represent a preferential CRPC dependency*, thus offering a therapeutic opportunity.  
401 To test this, we first analyzed the mutational profiles of SRGs in 7 PCa cell lines with increasing  
402 aggressiveness and found that the AR<sup>+</sup> and relatively indolent PCa cells tended to have more SRG  
403 deletions whereas AR<sup>-</sup> and aggressive PCa cells showed more SRG amplifications (Supplementary  
404 Fig. S10A). In particular, LNCaP and PC3 cells resembled pri-PCa and CRPC, respectively, with  
405 respect to SRG mutation profiles (Supplementary Fig. S10A). We retrieved two large-scale RNAi



406 screening data (Novartis Project Drive (64) and Broad Project Achilles (65)) and performed GSEA on  
407 ranked lists of essential genes. We observed that aggressive AR<sup>-</sup> PCa cell lines exhibited a  
408 preferential enrichment on two splicing pathway signatures ([Supplementary Fig. S10B](#)). By contrast,  
409 AR signaling and MYC signatures were enriched in AR<sup>+</sup> LNCaP and 22RV1 vs. AR<sup>-</sup> DU145 cells,  
410 respectively ([Supplementary Fig. S10C](#)). These analyses support the postulate that AR<sup>-</sup>, androgen-  
411 independent PCa cells may be particularly dependent on the spliceosome activity.

412 We subsequently tested this postulate using spliceosome inhibitors. Several microbial products,  
413 including Pladienolide B and its derivative E7107 have been shown to bind and inhibit the SF3B1  
414 complex and manifest anti-cancer activities (12,20). The E7107 compound represented the first-in-  
415 class spliceosome inhibitor that underwent phase I clinical trial (66). We found that PCa cells exhibited  
416 preferential sensitivity to E7107 relative to non-tumorigenic prostate epithelial cells RWPE1, with PC3  
417 being more sensitive than LNCaP cells ([Fig. 6A](#); [Supplementary Fig. S10D](#)). Experiments with  
418 Pladienolide B confirmed PC3 as the most sensitive line ([Fig. 6B](#)). While a long-term E7107 treatment  
419 (6~7 days) induced massive cell death ([Fig. 6A](#); [Supplementary Fig. S10D](#)), shorter (<3 days)  
420 treatments generally elicited limited apoptosis but instead arrested PCa cells at the G2/M phase of the  
421 cell cycle ([Fig. 6C](#)). Treatment of PCa cells with E7107 for 20~48 h also inhibited cell migration and  
422 invasion, as measured by both Boyden chamber ([Fig. 6D](#); [Supplementary Fig. S10E](#)) and scratch-  
423 wound ([Supplementary Fig. S10F](#)) assays. Importantly, treatment of PCa cells with 5 nM E7107 for 6  
424 h dramatically reshaped the splicing pattern of the selected genes ([Supplementary Fig. S10G](#); also  
425 see below), suggesting an on-target effect of the drug.

426

#### 427 **E7107 Molecularly ‘Reverses’ PCa Cell Aggressiveness by Inhibiting Spliceosome Activity**

428 To uncover the mechanisms of action of E7107 in PCa, we treated LNCaP and PC3 cells with the  
429 drug for 6 h followed by deep RNA-seq ([Fig. 6E](#)). No gross defects were observed in cell growth  
430 ([Supplementary Fig. S11A](#)), but, as expected, E7107 dramatically inhibited the AS globally in both cell  
431 types ([Fig. 6E](#)) with SE being the major splicing type affected ([Fig. 6F](#)). Sashimi plot visualization of

432 the sequencing data and RT-PCR analysis validated splicing analysis (Fig. 6G; Supplementary Fig.  
433 S11B). We performed GO analysis on genes showing down-regulated splicing events. Analysis of the  
434 top 1000 genes with significant SE events inhibited by E7107 in PC3 cells revealed that many GO  
435 terms associated with cancer-promoting functions, e.g., cell cycle and proliferation, DNA repair,  
436 splicing, and cancer pathways, were markedly enriched (Supplementary Fig. S11C; Supplementary  
437 Table S5), suggesting that E7107 inhibits splicing of a subset of PCa-associated genes important for  
438 survival. At the gene expression level, E7107 reshaped the transcriptomes and exhibited a slightly  
439 suppressive effect, especially in LNCaP cells, on transcription (Fig. 6E; Supplementary Table S6).  
440 qRT-PCR analysis validated DEGs identified in RNA-seq (Supplementary Fig. S11D).

441 We also performed GO analysis of DEGs upregulated after E7107 treatment. In AR+p53<sup>+</sup> LNCaP  
442 cells, four main categories of related GO terms were identified (Supplementary Fig. S11E) with  
443 ‘Splicing’ being the most significant one, consistent with a recent report (67). AR and its target  
444 expression and AR signaling were not significantly affected by E7107 in LNCaP cells (Supplementary  
445 Fig. S11F). Interestingly, p53 was activated, along with several other TS genes including RBM4 (68)  
446 and MIR34A (69) (Supplementary Fig. S11E and S11G). Consistently, GO terms ‘cell cycle arrest’  
447 and ‘differentiation’ were enriched (Supplementary Fig. S11E). We have previously shown that the  
448 LNCaP gene expression profile resembles that in pri-PCa (7). GSEA of gene signatures specific to  
449 normal prostate tissues vs. pri-PCa revealed that the ‘normal’, but not the ‘tumor’, gene signature, was  
450 significantly enriched in E7107-treated LNCaP cells (Supplementary Fig. S11H), suggesting a  
451 reversion of LNCaP transcriptome from PCa-like to normal-like. Similarly, pathway analysis in AR-p53<sup>-</sup>  
452 PC3 cells identified both convergent (e.g., splicing, differentiation, cell cycle arrest and proliferation  
453 inhibition) and unique (i.e., steroid hormone and muscle development) GO categories, when  
454 compared to the analysis in LNCaP cells (Fig. 6H). Enrichment of ‘differentiation’ and ‘steroid  
455 hormone’ categories in PC3 cells prompted us to examine the androgen/AR signaling. Strikingly,  
456 transcript levels of AR itself and many typical AR targets were upregulated in PC3 cells treated with  
457 E7107, leading to a dramatic enrichment of AR pathway (Fig. 6I). Furthermore, a LNCaP gene

458 signature was highly enriched in E7107-treated PC3 cells (Fig. 6J). Experimentally, E7107 treatment  
459 increased cell size in both PC3 and DU145 cells (Fig. 6K), indicating morphological differentiation.  
460 Moreover, although p53 was not activated in PC3 cells due to its genetic loss, several other TS genes  
461 (e.g., ALOX15, NKX3-1, RBM4, MIR34A) were upregulated (Fig. 6L). These data, together, suggest a  
462 reversal, molecularly and phenotypically, of aggressive PCa cells (PC3) to a more indolent, “LNCaP-  
463 like” cell state upon spliceosome inhibition by E7107.

464

### 465 **Therapeutic Targeting of CRPC *in vivo* via Inhibition of Spliceosome Activity**

466 To evaluate the activity of E7107 against CRPC *in vivo*, we treated 3 distinct castration-resistant (AI)  
467 PCa xenograft models, i.e., the AR<sup>+hi</sup> LNCaP-AI (36), AR<sup>-lo</sup> LAPC9-AI (36) and AR<sup>-</sup> PC3, with E7107  
468 or vehicle (Fig. 7A). The LNCaP-AI and LAPC9-AI models were established by serially passaging the  
469 corresponding parent AD tumor cells in castrated immunodeficient mice (36). The LNCaP-AI was  
470 initially responsive to Enza but quickly became Enza-resistant whereas LAPC9-AI was refractory to  
471 Enza *de novo* (36). Treatment of Enza-refractory LAPC9-AI tumors with either one cycle (i.e., tail vein  
472 injection for 5 consecutive days) or two cycles (with 1 week of ‘drug holiday’ between the 2 cycles)  
473 effectively inhibited tumor growth (Fig. 7B and 7C; Supplementary Fig. S12A and S12B, left). Similarly,  
474 treatment of mice bearing LNCaP-AI with two cycles of E7107 (Fig. 7D; Supplementary Fig. S12C, left)  
475 and PC3 xenografts with one cycle of E7107 (Fig. 7E; Supplementary Fig. S12D, left) also inhibited  
476 tumor growth. Although a certain degree of toxicity of E7107 was observed, treated mice returned to  
477 the range of normal body weight within a week after cessation of treatment (Supplementary Fig.  
478 S12A-S12D, right). The endpoint tumors frequently displayed a more ‘differentiated’ morphology  
479 manifested by an enrichment of enlarged and polynucleated cells (Supplementary Fig. S12E).

480 To determine whether the tumor-inhibitory effects of E7107 are associated with spliceosome  
481 inhibition, we performed RNA-seq analysis in LAPC9-AI and PC3 tumors 4 h after the fifth injection of  
482 E7107 (Fig. 7F; see Methods). Consistent with the *in vitro* data (Fig. 6E), E7107 suppressed the AS  
483 globally in both AR<sup>-lo</sup> CRPC models (Fig. 7G), evidenced by decreases in A3, A5, and SE (Fig. 7H).

484 Sashimi plot visualization and RT-PCR were performed to validate our splicing analysis (Fig. 7I;  
485 [Supplementary Fig. S12F](#)). GO analysis of the top 1000 genes bearing down-regulated SE events in  
486 LAPC9-AI model upon E7107 treatment revealed an enrichment of several cancer-promoting  
487 functional categories including ‘cell cycle and proliferation’, ‘DNA repair’, ‘splicing’, and ‘cancer  
488 pathways’ ([Supplementary Fig. S12G](#)). Analysis of the gene expression changes after E7107  
489 treatment revealed 3299 and 2289 DEGs, respectively, in LAPC9-AI and PC3 systems without  
490 obvious bias on transcription (Fig. 7G). qRT-PCR analysis in tumor samples confirmed the differential  
491 expression of selected genes ([Supplementary Fig. S13A](#)).

492 GO analysis of the genes upregulated in E7107-treated LAPC9-AI revealed a broad spectrum of  
493 functional categories linked to inhibition of cell proliferation and promotion of normal (prostate)  
494 developmental, differentiation, inflammation, and TS pathways, among others ([Supplementary Fig.  
495 S13B](#)). Since the LAPC9-AI has the AR<sup>-lo</sup> phenotype (36), transcription of AR signaling, as expected,  
496 remained unaltered ([Supplementary Fig. S13C](#)). Notably, gene signatures specific to pri-PCa and  
497 CRPC were significantly enriched in E7107- vs. vehicle-treated LAPC9-AI tumors, respectively  
498 ([Supplementary Fig. S13D](#)), again suggesting that spliceosome inhibition by E7107 reverses the gene  
499 expression pattern of LAPC9-AI from CRPC-Ad-like (more aggressive) to pri-PCa-like (less  
500 aggressive). We have recently shown that LAPC9-AI molecularly resembles CRPC-Ad (36). In the  
501 PC3 model, we observed an increase in the expression of genes involved in muscle development,  
502 inflammation, immune cell infiltration, and androgen response, among others, after E7107 treatment  
503 ([Supplementary Fig. S13E](#)). Interestingly, despite the upregulated category of ‘androgen response’,  
504 AR signaling and many targets remained inactivated ([Supplementary Fig. S13F](#)). Compared with in  
505 vitro data showing that E7107 strongly boosted the AR signaling (Fig. 6I), this discrepancy could be  
506 explained by an in vivo environment lacking androgen in castrated hosts such that the ‘E7107  
507 reprogramed’ AR<sup>+</sup> PC3 cells may not survive. Endpoint PC3 tumors tended to be less aggressive in  
508 terms of molecular signatures ([Supplementary Fig. S13G-I](#)). For instance, E7107 treatment  
509 significantly inhibited pathways associated with cancer metastasis and stemness ([Supplementary Fig.](#)

510 [S13G](#)), decreased the expression of a PC3-cell signature ([Supplementary Fig. S13H](#)), and reverted  
511 the gene expression pattern from CRPC-NE like to CRPC-Ad like ([Supplementary Fig. S13I](#)). We  
512 have previously demonstrated that PC3 cells molecularly resemble the CRPC-NE (7,36).

513

514 **DISCUSSION**

515

516 Studies of AR variants, ARv7 in particular, have implicated splicing dysregulation in PCa resistance to  
517 ADT/Enza (20). Recently, splicing factor HNRNPL was identified as a dependency for LNCaP cells  
518 (70) and SFPQ (i.e., PSF) was reported to promote AR splicing and CRPC cell survival (71). An  
519 examination of race-specific AS changes in PCa in African American (AA) vs. European American  
520 (EA) men discovered an AA-enriched PIK3CD isoform that promotes tumor aggressiveness and drug  
521 resistance (72). Globally, splicing dysregulation in pri-PCa vs. normal tissues has been observed  
522 (13,19). While these studies implicate splicing dysregulation in PCa pathogenesis, the global AS  
523 landscape unraveling the dynamic evolution of PCa has not been reported and the impact of aberrant  
524 AS alterations on PCa progression, therapy resistance, and patient outcome remains unclear.

525 Here, we provide the first comprehensively annotated splicing map in PCa using clinical and  
526 experimental RNA-seq data covering the entire spectrum of PCa development and progression.  
527 Aberrantly spliced genes specific to different PCa stages are both convergently and uniquely enriched  
528 in diverse GO terms and pathways linked to many key cellular processes important for cell survival,  
529 which establishes aberrant splicing as a distinct mechanism (vs. gene expression regulation) driving  
530 PCa progression and therapy resistance. In particular, we observe increasing severity of AS  
531 dysregulation and identify IR as a hallmark of stemness and aggressiveness during PCa progression  
532 and therapy resistance. Recently, widespread IR, associated with somatic single-nucleotide variations  
533 in six cancer types (excluding PCa), has been observed to be more enriched in TS genes leading to  
534 their loss of expression (73). Surprisingly, we did not observe a similar trend in PCa. Rather, our data  
535 reveals that IR generally enhances gene expression and thus likely functions in PCa biology,  
536 suggesting different roles of IR in distinct cancer types. Particularly, we find that IR in PCa impacts  
537 genes involved in stemness and cancer-promoting functions ([Supplementary Fig. S4G-I](#)), and that AR  
538 regulates a splicing program, but not IR specifically, distinct to its transcriptional regulation,  
539 suggesting IR as a PCa-regulating mechanism independent of AR axis. In fact, we have observed a

540 general negative association of AR activity with IR level in multiple clinical datasets. Together, our  
541 results may establish IR as a common mechanism of cellular stemness, as supported by studies in  
542 mouse ESCs (49). The IR prevalent in PCa is not associated strongly with *cis*-genomic features, but  
543 seems to be regulated by *trans*-regulatory mechanisms involving the combinatorial effects of multiple  
544 SRGs (Fig. 1G and 1H). In support, candidate RBPs modulate not only the IR but also other splicing  
545 types as well (Fig. 1H). Alternatively, besides altered spliceosome activity, IR might also be modulated  
546 by other molecular alterations. For example, loss-of-function mutations in SETD2 (a H3K36  
547 methyltransferase) and subsequent loss of H3K36 trimethylation at target exons are associated with  
548 increased IR in renal cancers (74). Our work expands the view of molecular complexity underlying  
549 and justifies further exploration on the role of IR in PCa etiology and progression.

550       There are many ways by which RNA splicing can be dysregulated in cancer relative to normal  
551 cells. Previously, recurrent point mutations in core spliceosome genes (e.g., SF3B1, U2AF1, SRSF2,  
552 ZRSR2) have been identified as the driving mechanism underpinning splicing dysregulation in  
553 hematological cancers (12). Our genomic analyses of SRGs reveal copy-number variations (CNVs) as  
554 the main driver of AS alterations in PCa, which generally alter the expression of affected SRGs and  
555 illustrate cancer type-specific differences in mechanisms of splicing dysregulation. Remarkably, the  
556 majority of the top altered SRGs are located in regions containing either TS genes or oncogenes  
557 (Supplementary Fig. S8A), all of which have not been highlighted in previous large-scale DNA  
558 sequencing studies. This raises an interesting question of whether these alterations are just  
559 passenger mutations or they causally contribute to PCa pathogenesis. While direct functional  
560 evidence implicating them in PCa biology awaits experimental validation, the involvement of these  
561 genes in other types of cancer has been reported (9,10,20). Particularly, splicing dysregulation has  
562 been recently proposed as a 'driver' of transformation independently of oncogenic processes (17).  
563 Intriguingly, CNVs of SRGs exerts, comparatively, a dramatic impact on global splicing landscape in  
564 CRPC (data not shown), suggesting an enhanced dependency of CRPC on aberrant spliceosome  
565 activity. Therefore, these mutated SRGs may bear some of the 'driver' properties, and it would be



566 interesting to dissect, in future studies, whether deletion or amplification of CNV-associated SRGs  
567 with or without collateral alterations in RB1 or MYC loci, or vice versa, could change cancer  
568 phenotypes. Another potential mechanism that may cause splicing abnormality is the mutations in  
569 splice sites (12). However, mutations in splice sites constitute the minority of all somatic mutations (as  
570 low as ~0.6%) in PCa (75); consequently, we reason that deregulation of SRGs is the main  
571 mechanism underpinning splicing abnormalities. In support, the majority of SRGs are mis-expressed  
572 in various stages of PCa, consistent with studies showing that altered expression of SRGs, even in  
573 the absence of mutations, promotes oncogenesis (10). Of clinical significance, our study has identified  
574 many SRGs that can be linked, individually or in combination, to clinical features of advanced PCa,  
575 indicating a 'biomarker' value. Almost all of these identified prognostic SRGs and DSEs have not  
576 previously been implicated in PCa, thus warranting further investigation. Notably, the unfavorable  
577 SRG signature that we developed herein predicts PCa progression and correlates with poor patient  
578 survival, associates with much 'twisted' splicing landscape, and establishes splicing misregulation as  
579 a promoter of PCa aggressiveness.

580 Multiple lines of evidence reveal a preferential dependency of CRPC on aberrant spliceosome  
581 activity. *First*, the number of DSEs increases exponentially along the spectrum of cancer progression  
582 ([Supplementary Fig. S1D](#)), linking the severity of splicing dysregulation to PCa aggressiveness.  
583 *Second*, amplifications of SRGs are predominantly observed, and CNVs of SRGs mainly impact  
584 global splicing in CRPC. *Third*, more SRGs are dysregulated in CRPC, highlighting a potentially  
585 critical role of SRG misexpression in driving CRPC evolution. *Fourth*, the majority of altered SRGs are  
586 predictive of worse patient outcome and the unfavorable SRG signature associates with high tumor  
587 grade and more prominent disruption in the splicing landscape. *Fifth*, chemical castration and Enza,  
588 both of which target AR signaling, reshape the splicing landscape in PCa cells, and the distorted  
589 splicing landscape likely contributes to subsequent treatment failure and disease progression ([Fig. 2](#)),  
590 as documented in other cancer types (12). *Finally*, E7107, the spliceosome modulator, effectively  
591 inhibits the growth of multiple experimental CRPC models in vivo ([Fig. 6 and 7](#)) regardless of the AR



592 status. In this study, we did not explore the combination of E7107 and Enza because all CRPC  
593 models we utilized are already Enza-resistant. A phase-I study of E7107 in patients with advanced  
594 solid tumors (PCa excluded) was terminated due to side effects (66) and we also observed certain  
595 toxicities of E7101 in animals (this study), suggesting the need to define intricate treatment window  
596 and doses for E7101. Our results may point to a new strategy of administering E7107, or other  
597 splicing inhibitors, as we show that, interestingly, E7107 promotes PCa cell differentiation and  
598 reprograms PCa cells from an ‘androgen-insensitive’ to an ‘androgen-sensitive’ state. We thus  
599 envision a potential treatment regimen in which CRPC is first subject to a short-term splicing inhibition  
600 (to avoid toxicity and also to reprogram aggressive PCa cells) followed by Enza treatment. Ongoing  
601 studies are exploring the value of this sequential treatment protocol. Overall, our findings suggest that  
602 there may be a therapeutic window for spliceosome modulators in the treatment of CRPC. Future  
603 studies that aim to determine the origins and consequences of aberrant splicing in aggressive PCa  
604 could enhance our understanding of disease pathogenesis and aid novel drug development.

605

## 606 **ACKNOWLEDGMENTS**

607 We acknowledge the support of several Shared Resources at the Roswell Park Comprehensive  
608 Cancer Center (RPCCC), including Histology, Flow and Image Cytometry and Genomics Resources.  
609 This project was supported, in part, by grants from the U.S NIH (R01CA155693, R01CA237027 and  
610 R01CA240290) and Department of Defense (W81XWH-14-1-0575 and W81XWH-16- 1-0575), and  
611 from the CPRIT (RP120380) (all to D.G.T.) and by RPCCC and NCI center grant P30CA016056. S.L.  
612 and J.W. were supported, in part, by the NIH grant U24CA232979. D.Z. was supported, in part, by the  
613 NIH/NCI grant R21CA218635 and Huazhong Agricultural University (HZAU) Startup Fund for  
614 Advanced Talents. We apologize to the colleagues whose work was not cited due to space constraint.

615

616

## 617 **AUTHOR CONTRIBUTIONS**

618 D.Z. and D.T. conceived and designed the study, interpreted data, and wrote the manuscript;  
619 D.Z. and Y.J. performed most experiments;  
620 D.Z., Q.H., H.C., Y.L., J.W. and S.L. conducted Bioinformatic analysis;  
621 A.T. and J.K. provided assistance in animal experiments;  
622 S.B and P.Z. provided reagents and expertise for the spliceosome modulation experiments;  
623 All authors read and approved the manuscript.

624

625

## 626 **DECLARATION OF INTERESTS**

627 The authors declare no competing financial interests.

628 **Figure Legend**

629 **Figure 1. Splicing landscape identifies IR as a hallmark of PCa stemness and progression.**

- 630 (A-J) The landscape of AS during PCa development and progression. Two related datasets are  
631 interrogated for each PCa stage. Shown are splicing patterns and the number of DSEs decoded  
632 by rMATS.  
633 (K) Changes of IR and SE across the 14 comparisons detected by rMATS. “Baseline” refers to 1 as  
634 data presented as fold change in a comparative manner.  
635 (L) RBP motif analysis of retained introns specific to the PCa stages indicated. A total of 95 RBPs are  
636 examined and shown are the top 20 genes ranked by a binding score that takes into account  
637 both binding frequency and binding strength for each RBP.  
638 (M) DSEs associated with high or low expression level of ELAVL1 and RBM38 in pri-PCa and CRPC-  
639 Ad, respectively.  
640 (N) Pairwise comparison of expression of the genes showing significant IR events during PCa  
641 progression. Expression variability is quantified for each gene as a Z-score relative to the mean  
642 expression in normal prostate samples. Genes exhibiting both up- and down-regulated IR events  
643 are removed, and the resultant gene number is indicated. Significance was calculated by a paired  
644 Student’s *t*-test.  
645 (O) Overlap of significant IR-bearing genes with a high-confidence set of 1000 human NMD targets.  
646 Significance was calculated by a  $\chi^2$  test.  
647 DSEs, differentially spliced events; RBP, RNA binding protein; NMD, nonsense-mediated mRNA  
648 decay.

649  
650  
651 **Figure 2. AR activity impacts AS landscape distinctively from its regulation of transcription.**

- 652 (A-B) DSEs associated with high and low AR activity (cutoff, Z-score >7 or <-7) in pri-PCa (A) and  
653 CRPC-Ad (B), respectively. AR activity (see Methods) was used to fractionate patients cohorts  
654 followed by splicing analysis by rMATS.  
655 (C) Genomic alterations do not contribute to the diversity of AR activities across PCa populations.  
656 Shown are frequency and AR mutation types observed in TCGA and CRPC cohorts. AR activities  
657 of samples grouped as in A were displayed.  
658 (D-E) Overlap between SAGs and DEGs (D) and between SAGs and three sets of AR-regulated  
659 genes (E) in indicated contexts. The number in parentheses denotes a percentage of overlapped  
660 genes proportioned to all SAGs. Circles are not drawn to scale.  
661 (F-H) Experimental design (F), principal component analysis (PCA) showing proper clustering of  
662 samples (G), and qPCR validation of intended modulations of AR signaling in LNCaP cells (H).  
663 (I) The AR-regulated AS program in PCa cells. Shown are the DSEs associated with high (up arrow)  
664 or low (down arrow) AR activity in LNCaP cells detected by rMATS.  
665 (J) Overlap between SAGs and three sets of AR-regulated genes in indicated contexts. The number  
666 in parentheses denotes a percentage of overlapped genes proportioned to all SAGs. Circles are  
667 not drawn to scale.  
668 DSEs, differentially spliced events; DEGs, differentially expressed genes; SAGs, splicing-affected  
669 genes.

670  
671  
672 **Figure 3. The mutational landscape of SRGs in human PCa.**

- 673 (A-C) A comprehensive survey for genomic alterations in 274 SRGs in available clinical cohorts in  
674 cBioportal. The top 15 mutated SRGs are shown in the representative pri-PCa (A) and metastatic  
675 CRPC-Ad (B) cohorts. Frequently deleted RB1 and PTEN (colored in blue) and amplified MYC  
676 and AR (colored in red) are included as reference genes. Each bar represents the alteration  
677 status of an individual gene for a single patient and the percentage of alterations for each gene in

678 the indicated cohort is provided. Shown in C are bar graphs summarizing the cumulative genomic  
679 alterations of SRGs in the largest and representative pri-PCa (TCGA) and CRPC (SU2C)  
680 cohorts.  
681 **(D-E)** Bar plots illustrating the cumulative aberration frequencies of all 274 SRGs combined (D) and  
682 the top 20 mutated SRGs (10 most amplified and 10 most deleted) (E) across all cohorts, with  
683 numbers above and within the bars representing the total frequency and a frequency of  
684 amplification or deletion of indicated genes, respectively.  
685 **(F)** Integrated mutational landscape of top 20 mutated SRGs in PCa showing mutual exclusivity, in  
686 large part, between deletions and amplifications of SRGs.  
687 See [Table S1](#) for detail.

688  
689

#### 690 **Figure 4. Dysregulation of SRGs in PCa.**

691 Integrated heatmap of differentially expressed SRGs identified in Oncomine ( $p < 0.05$ ) and RNA-seq  
692 (fold-change (FC)  $\geq 1.5$  and FDR  $< 0.1$ ). In Oncomine (O), the medium-rank of  $< 2500$ ,  $< 4000$ , and  
693  $> 4000$  for a gene denotes high, moderate and low level of expression, respectively. For visualization,  
694 DEGs revealed by RNA-seq (R) data are categorized into three groups according to FC differences  
695 (FC  $\geq 3$ ,  $\geq 2$  and  $\geq 1.5$ ). Based on pairwise comparisons, the stages of PCa are defined as tumor  
696 development (pri-PCa vs. normal tissues), ADT treatment response (ADT-after vs. -before), CRPC  
697 progression (CRPC-Ad vs. pri-PCa) and plasticity (CRPC-NE vs. CRPC-Ad).

698  
699

#### 700 **Figure 5. SRGs are prognostic and associated with splicing dysregulation.**

701 **(A)** Examples of Kaplan-Meier plots for unfavorable and favorable genes associated with patient  
702 overall survival in 7 different cohorts.  
703 **(B)** Numbers of genes showing favorable and unfavorable prognostic effects in 5 distinct cohorts.  
704 Patient numbers for each cohort are shown in parentheses.  
705 **(C)** Meta-analysis showing higher level of unfavorable signature and lower level of favorable signature  
706 correlating with reduced overall patient survival, respectively. Data were based on the Setlur and  
707 Glinsky studies.  
708 **(D-E)** Unfavorable signature is associated with higher Gleason score (D) and higher level of  
709 unfavorable signature positively correlates with disease recurrence in TCGA cohort (E).  
710 **(F)** DSEs associated with high or low expression level of unfavorable signature in primary PCa cohort.  
711 The p value was calculated using Student's *t*-test (D) and log-rank test (A, C, E) \* $p < 0.05$ .  
712 See [Table S4](#) for details.

713  
714

#### 715 **Figure 6. CRPC cells are sensitive to pharmacologic modulation of spliceosome activity.**

716 **(A-B)** Cell proliferation (MTT; left) and colony formation (right) assays in indicated cell lines treated  
717 with varying concentration of E7107 (A) or Pladienolide B (B) *in vitro*.  
718 **(C-D)** Cell cycle analysis (C) and migration and invasion assays (D) in indicated PCa cells treated with  
719 varying concentrations of E7107. Results shown were representative of 2–3 repeat experiments.  
720 For D, data represent mean  $\pm$  S.D. from cell number counting of 5–6 random high magnification  
721 (X20) areas. The P value was calculated using Student's *t*-test. \* $P < 0.05$  and \*\* $P < 0.01$ .  
722 **(E)** Effect of 10nM E7107 on PCa transcriptome *in vitro*. Shown are schematic of RNA-seq  
723 experiments (top) and the number of total DSEs (bottom left) and DEGs (bottom right) identified  
724 upon E7107 treatment in indicated PCa cells.  
725 **(F)** AS pattern showing that E7107 reshapes the splicing landscape of PCa cells indicated.  
726 **(G)** Sashimi plots visualization and RT-PCR validation of IR in DDIT3 gene after an acute E7107 (10  
727 nM, 6 h) treatment. The event  $\Delta$ PSI values calculated by rMATS were provided in parentheses.

- 728 (H) GO analysis of genes upregulated at bulk RNA level in PC3 cells after E7107 treatment (10nM, 6  
729 h).
- 730 (I-J) GSEA showing enrichment of AR signaling related signatures (I) in E7107-treated PC3 cells,  
731 indicating that E7107 reprograms the AR<sup>-</sup> PC3 cells back into relatively AR<sup>+</sup> LNCaP-like cells. In  
732 support, a LNCaP gene signature (defined as top 300 genes solely expressed or overexpressed in  
733 LNCaP compared with PC3) was significantly enriched in PC3 cells treated with E7107 (J).
- 734 (K) Representative FACS plots of PC3 and DU145 cells treated with E7107 (5 nM) for 3 days showing  
735 an increase in cell size.
- 736 (L) Upregulation of tumor suppressors (ALOX15, KNX3-1, RBM4 and MIR34A) in PC3 cells after  
737 E7107 treatment.

738  
739

740 **Figure 7. Therapeutic targeting of CRPC cells *in vivo*.**

- 741 (A) Schematic of *in vivo* E7107 treatment.
- 742 (B-E) Inhibitory effect of E7107 on the growth of indicated Enza-resistant CRPC models *in vivo*.  
743 Shown are growth curve (left), endpoint tumor image (middle) and tumor weight (right) of  
744 LAPC9-AI (B and C), LNCaP-AI (D) and PC3 (E) models treated with vehicle or E7107.
- 745 (F-G) Effect of E7107 on CRPC transcriptome *in vivo*. Shown are schematic of RNA-seq experiment  
746 (F) and the ratio of total DSEs (G; left) and total DEGs (G; right) identified upon E7107 treatment  
747 in indicated CRPC models.
- 748 (H) AS pattern showing that E7107 reshapes the splicing landscape of CRPC xenografts *in vivo*.
- 749 (I) Sashimi plots and RT-PCR validation of IR in DDIT3 gene after E7107 treatment *in vivo*. The event  
750  $\Delta$ PSI values calculated by rMATS were provided in parentheses. E, E7107; V, vehicle.

751  
752  
753

754  
755

## REFERENCES

- 756 1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA: a cancer journal for clinicians* **2019**  
757 doi 10.3322/caac.21551.
- 758 2. Zhang D, Jeter C, Gong S, Tracz A, Lu Y, Shen J, *et al.* Histone 2B-GFP Label-Retaining  
759 Prostate Luminal Cells Possess Progenitor Cell Properties and Are Intrinsically Resistant to  
760 Castration. *Stem Cell Reports* **2018**;10(1):228-42 doi 10.1016/j.stemcr.2017.11.016.
- 761 3. Zhang D, Zhao S, Li X, Kirk JS, Tang DG. Prostate Luminal Progenitor Cells in Development  
762 and Cancer. *Trends Cancer* **2018**;4(11):769-83 doi 10.1016/j.trecan.2018.09.003.
- 763 4. Zhang D, Tang DG. "Splice" a way towards neuroendocrine prostate cancer. *EBioMedicine*  
764 **2018**;35:12-3 doi 10.1016/j.ebiom.2018.08.037.
- 765 5. Beltran H, Prandi D, Mosquera JM, Benelli M, Puca L, Cyrta J, *et al.* Divergent clonal evolution  
766 of castration-resistant neuroendocrine prostate cancer. *Nat Med* **2016**;22(3):298-305 doi  
767 10.1038/nm.4045.
- 768 6. Ku SY, Rosario S, Wang Y, Mu P, Seshadri M, Goodrich ZW, *et al.* Rb1 and Trp53 cooperate  
769 to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance.  
770 *Science* **2017**;355(6320):78-83 doi 10.1126/science.aah4199.
- 771 7. Zhang D, Park D, Zhong Y, Lu Y, Rycak K, Gong S, *et al.* Stem cell and neurogenic gene-  
772 expression profiles link prostate basal cells to aggressive prostate cancer. *Nature*  
773 *communications* **2016**;7:10798 doi 10.1038/ncomms10798.
- 774 8. Tang DG. Understanding cancer stem cell heterogeneity and plasticity. *Cell research*  
775 **2012**;22(3):457-72 doi 10.1038/cr.2012.13.
- 776 9. Sveen A, Kilpinen S, Ruusulehto A, Lothe RA, Skotheim RI. Aberrant RNA splicing in cancer;  
777 expression changes and driver mutations of splicing factor genes. *Oncogene*  
778 **2016**;35(19):2413-27 doi 10.1038/onc.2015.318.
- 779 10. Zhang J, Manley JL. Misregulation of pre-mRNA alternative splicing in cancer. *Cancer*  
780 *discovery* **2013**;3(11):1228-37 doi 10.1158/2159-8290.CD-13-0253.
- 781 11. Park E, Pan Z, Zhang Z, Lin L, Xing Y. The Expanding Landscape of Alternative Splicing  
782 Variation in Human Populations. *Am J Hum Genet* **2018**;102(1):11-26 doi  
783 10.1016/j.ajhg.2017.11.002.
- 784 12. Lee SC, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. *Nat Med* **2016**;22(9):976-  
785 86 doi 10.1038/nm.4165.
- 786 13. Kahles A, Lehmann KV, Toussaint NC, Huser M, Stark SG, Sachsenberg T, *et al.*  
787 Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer*  
788 *cell* **2018**;34(2):211-24 e6 doi 10.1016/j.ccell.2018.07.001.
- 789 14. Hsu TY, Simon LM, Neill NJ, Marcotte R, Sayad A, Bland CS, *et al.* The spliceosome is a  
790 therapeutic vulnerability in MYC-driven cancer. *Nature* **2015**;525(7569):384-8 doi  
791 10.1038/nature14985.
- 792 15. Sebestyen E, Singh B, Minana B, Pages A, Mateo F, Pujana MA, *et al.* Large-scale analysis of  
793 genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing  
794 networks. *Genome research* **2016**;26(6):732-44 doi 10.1101/gr.199935.115.
- 795 16. Seiler M, Peng S, Agrawal AA, Palacino J, Teng T, Zhu P, *et al.* Somatic Mutational  
796 Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer  
797 Types. *Cell reports* **2018**;23(1):282-96 e4 doi 10.1016/j.celrep.2018.01.088.
- 798 17. Climente-Gonzalez H, Porta-Pardo E, Godzik A, Eyraas E. The Functional Impact of Alternative  
799 Splicing in Cancer. *Cell reports* **2017**;20(9):2215-26 doi 10.1016/j.celrep.2017.08.012.
- 800 18. Li Y, Sahni N, Pancsa R, McGrail DJ, Xu J, Hua X, *et al.* Revealing the Determinants of  
801 Widespread Alternative Splicing Perturbation in Cancer. *Cell reports* **2017**;21(3):798-812 doi  
802 10.1016/j.celrep.2017.09.071.



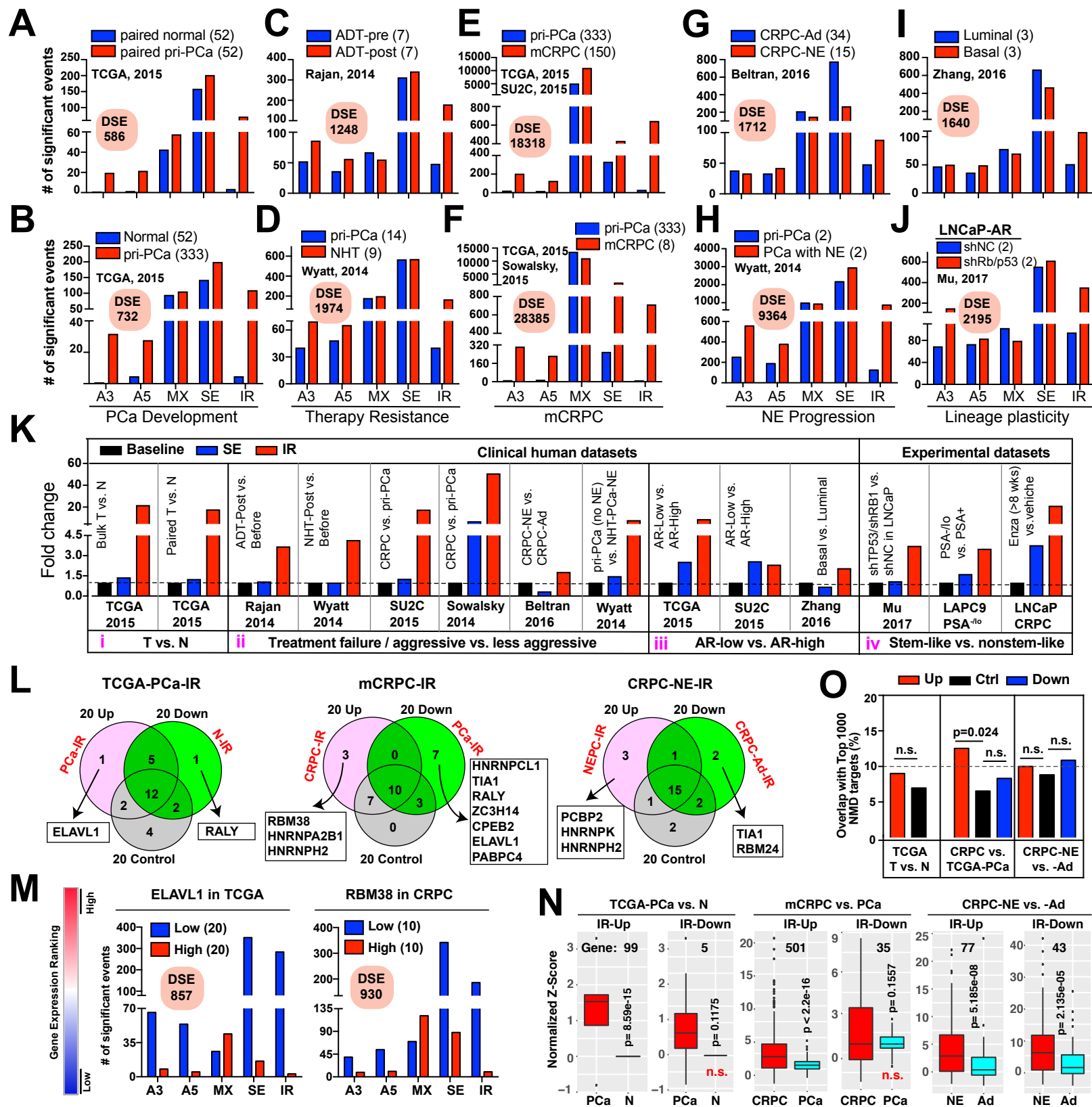
- 803 19. Ryan M, Wong WC, Brown R, Akbani R, Su X, Broom B, *et al.* TCGASpliceSeq a compendium  
804 of alternative mRNA splicing in cancer. *Nucleic Acids Res* **2016**;44(D1):D1018-22 doi  
805 10.1093/nar/gkv1288.
- 806 20. Paschalis A, Sharp A, Welti JC, Neeb A, Raj GV, Luo J, *et al.* Alternative splicing in prostate  
807 cancer. *Nat Rev Clin Oncol* **2018**;15(11):663-75 doi 10.1038/s41571-018-0085-0.
- 808 21. Shen S, Park JW, Lu ZX, Lin L, Henry MD, Wu YN, *et al.* rMATS: robust and flexible detection  
809 of differential alternative splicing from replicate RNA-Seq data. *Proceedings of the National*  
810 *Academy of Sciences of the United States of America* **2014**;111(51):E5593-601 doi  
811 10.1073/pnas.1419161111.
- 812 22. Alamancos GP, Pages A, Trincado JL, Bellora N, Eyraas E. Leveraging transcript quantification  
813 for fast computation of alternative splicing profiles. *RNA* **2015**;21(9):1521-31 doi  
814 10.1261/rna.051557.115.
- 815 23. Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate Cancer. *Cell*  
816 **2015**;163(4):1011-25 doi 10.1016/j.cell.2015.10.025.
- 817 24. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, *et al.* Integrative  
818 clinical genomics of advanced prostate cancer. *Cell* **2015**;161(5):1215-28 doi  
819 10.1016/j.cell.2015.05.001.
- 820 25. Sowalsky AG, Xia Z, Wang L, Zhao H, Chen S, Bublely GJ, *et al.* Whole transcriptome  
821 sequencing reveals extensive unspliced mRNA in metastatic castration-resistant prostate  
822 cancer. *Molecular cancer research : MCR* **2015**;13(1):98-106 doi 10.1158/1541-7786.MCR-14-  
823 0273.
- 824 26. Wyatt AW, Mo F, Wang K, McConeghy B, Brahmhatt S, Jong L, *et al.* Heterogeneity in the  
825 inter-tumor transcriptome of high risk prostate cancer. *Genome biology* **2014**;15(8):426 doi  
826 10.1186/s13059-014-0426-y.
- 827 27. Rajan P, Sudbery IM, Villasevil ME, Mui E, Fleming J, Davis M, *et al.* Next-generation  
828 sequencing of advanced prostate cancer treated with androgen-deprivation therapy. *European*  
829 *urology* **2014**;66(1):32-9 doi 10.1016/j.eururo.2013.08.011.
- 830 28. Mu P, Zhang Z, Benelli M, Karthaus WR, Hoover E, Chen CC, *et al.* SOX2 promotes lineage  
831 plasticity and antiandrogen resistance in TP53- and RB1-deficient prostate cancer. *Science*  
832 **2017**;355(6320):84-8 doi 10.1126/science.aah4307.
- 833 29. Zhang Z, Pal S, Bi Y, Tchou J, Davuluri RV. Isoform level expression profiles provide better  
834 cancer signatures than gene level expression profiles. *Genome Med* **2013**;5(4):33 doi  
835 10.1186/gm437.
- 836 30. Rodriguez-Bravo V, Pippa R, Song WM, Carceles-Cordon M, Dominguez-Andres A, Fujiwara  
837 N, *et al.* Nuclear Pores Promote Lethal Prostate Cancer by Increasing POM121-Driven E2F1,  
838 MYC, and AR Nuclear Import. *Cell* **2018**;174(5):1200-15 e20 doi 10.1016/j.cell.2018.07.015.
- 839 31. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware  
840 quantification of transcript expression. *Nature methods* **2017**;14(4):417-9 doi  
841 10.1038/nmeth.4197.
- 842 32. Yae T, Tsuchihashi K, Ishimoto T, Motohara T, Yoshikawa M, Yoshida GJ, *et al.* Alternative  
843 splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell.  
844 *Nature communications* **2012**;3:883 doi 10.1038/ncomms1892.
- 845 33. Lee AR, Li Y, Xie N, Gleave ME, Cox ME, Collins CC, *et al.* Alternative RNA splicing of the  
846 MEAF6 gene facilitates neuroendocrine prostate cancer progression. *Oncotarget*  
847 **2017**;8(17):27966-75 doi 10.18632/oncotarget.15854.
- 848 34. Dvinge H, Bradley RK. Widespread intron retention diversifies most cancer transcriptomes.  
849 *Genome Med* **2015**;7(1):45 doi 10.1186/s13073-015-0168-9.
- 850 35. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, *et al.* An embryonic stem  
851 cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nature*  
852 *genetics* **2008**;40(5):499-507 doi 10.1038/ng.127 [pii]10.1038/ng.127.

- 853 36. Li Q, Deng Q, Chao HP, Liu X, Lu Y, Lin K, *et al.* Linking prostate cancer cell AR heterogeneity  
854 to distinct castration and enzalutamide responses. *Nature communications* **2018**;9(1):3600 doi  
855 10.1038/s41467-018-06067-7.
- 856 37. Zhang D, Tang DG, Rycaj K. Cancer stem cells: Regulation programs, immunological  
857 properties and immunotherapy. *Semin Cancer Biol* **2018** doi  
858 10.1016/j.semcancer.2018.05.001.
- 859 38. Qin J, Liu X, Laffin B, Chen X, Choy G, Jeter CR, *et al.* The PSA(-/lo) prostate cancer cell  
860 population harbors self-renewing long-term tumor-propagating cells that resist castration. *Cell*  
861 *stem cell* **2012**;10(5):556-69 doi 10.1016/j.stem.2012.03.009.
- 862 39. Rycaj K, Cho EJ, Liu X, Chao HP, Liu B, Li Q, *et al.* Longitudinal tracking of subpopulation  
863 dynamics and molecular changes during LNCaP cell castration and identification of inhibitors  
864 that could target the PSA-/lo castration-resistant cells. *Oncotarget* **2016**;7(12):14220-40 doi  
865 10.18632/oncotarget.7303.
- 866 40. Choi J, Lee S, Mallard W, Clement K, Tagliazucchi GM, Lim H, *et al.* A comparison of  
867 genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nat*  
868 *Biotechnol* **2015**;33(11):1173-81 doi 10.1038/nbt.3388.
- 869 41. Middleton R, Gao D, Thomas A, Singh B, Au A, Wong JJ, *et al.* IRFinder: assessing the impact  
870 of intron retention on mammalian gene expression. *Genome biology* **2017**;18(1):51 doi  
871 10.1186/s13059-017-1184-4.
- 872 42. Naro C, Jolly A, Di Persio S, Bielli P, Setterblad N, Alberdi AJ, *et al.* An Orchestrated Intron  
873 Retention Program in Meiosis Controls Timely Usage of Transcripts during Germ Cell  
874 Differentiation. *Developmental cell* **2017**;41(1):82-93 e4 doi 10.1016/j.devcel.2017.03.003.
- 875 43. Ni T, Yang W, Han M, Zhang Y, Shen T, Nie H, *et al.* Global intron retention mediated gene  
876 regulation during CD4+ T cell activation. *Nucleic Acids Res* **2016**;44(14):6817-29 doi  
877 10.1093/nar/gkw591.
- 878 44. Braunschweig U, Barbosa-Morais NL, Pan Q, Nachman EN, Alipanahi B, Gonatopoulos-  
879 Pournatzis T, *et al.* Widespread intron retention in mammals functionally tunes transcriptomes.  
880 *Genome research* **2014**;24(11):1774-86 doi 10.1101/gr.177790.114.
- 881 45. Lu ZX, Huang Q, Park JW, Shen S, Lin L, Tokheim CJ, *et al.* Transcriptome-wide landscape of  
882 pre-mRNA alternative splicing associated with metastatic colonization. *Molecular cancer*  
883 *research : MCR* **2015**;13(2):305-18 doi 10.1158/1541-7786.MCR-14-0366.
- 884 46. Paz I, Kosti I, Ares M, Jr., Cline M, Mandel-Gutfreund Y. RBPmap: a web server for mapping  
885 binding sites of RNA-binding proteins. *Nucleic Acids Res* **2014**;42(Web Server issue):W361-7  
886 doi 10.1093/nar/gku406.
- 887 47. Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, Li X, *et al.* A compendium of RNA-  
888 binding motifs for decoding gene regulation. *Nature* **2013**;499(7457):172-7 doi  
889 10.1038/nature12311.
- 890 48. Ye J, Liang R, Bai T, Lin Y, Mai R, Wei M, *et al.* RBM38 plays a tumor-suppressor role via  
891 stabilizing the p53-mdm2 loop function in hepatocellular carcinoma. *J Exp Clin Cancer Res*  
892 **2018**;37(1):212 doi 10.1186/s13046-018-0852-x.
- 893 49. Boutz PL, Bhutkar A, Sharp PA. Detained introns are a novel, widespread class of post-  
894 transcriptionally spliced introns. *Genes & development* **2015**;29(1):63-80 doi  
895 10.1101/gad.247361.114.
- 896 50. Colombo M, Karousis ED, Bourquin J, Bruggmann R, Muhlemann O. Transcriptome-wide  
897 identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6-  
898 and SMG7-mediated degradation pathways. *RNA* **2017**;23(2):189-201 doi  
899 10.1261/rna.059055.116.
- 900 51. Kumar A, Coleman I, Morrissey C, Zhang X, True LD, Gulati R, *et al.* Substantial  
901 interindividual and limited intraindividual genomic diversity among tumors from men with  
902 metastatic prostate cancer. *Nat Med* **2016**;22(4):369-78 doi 10.1038/nm.4053.

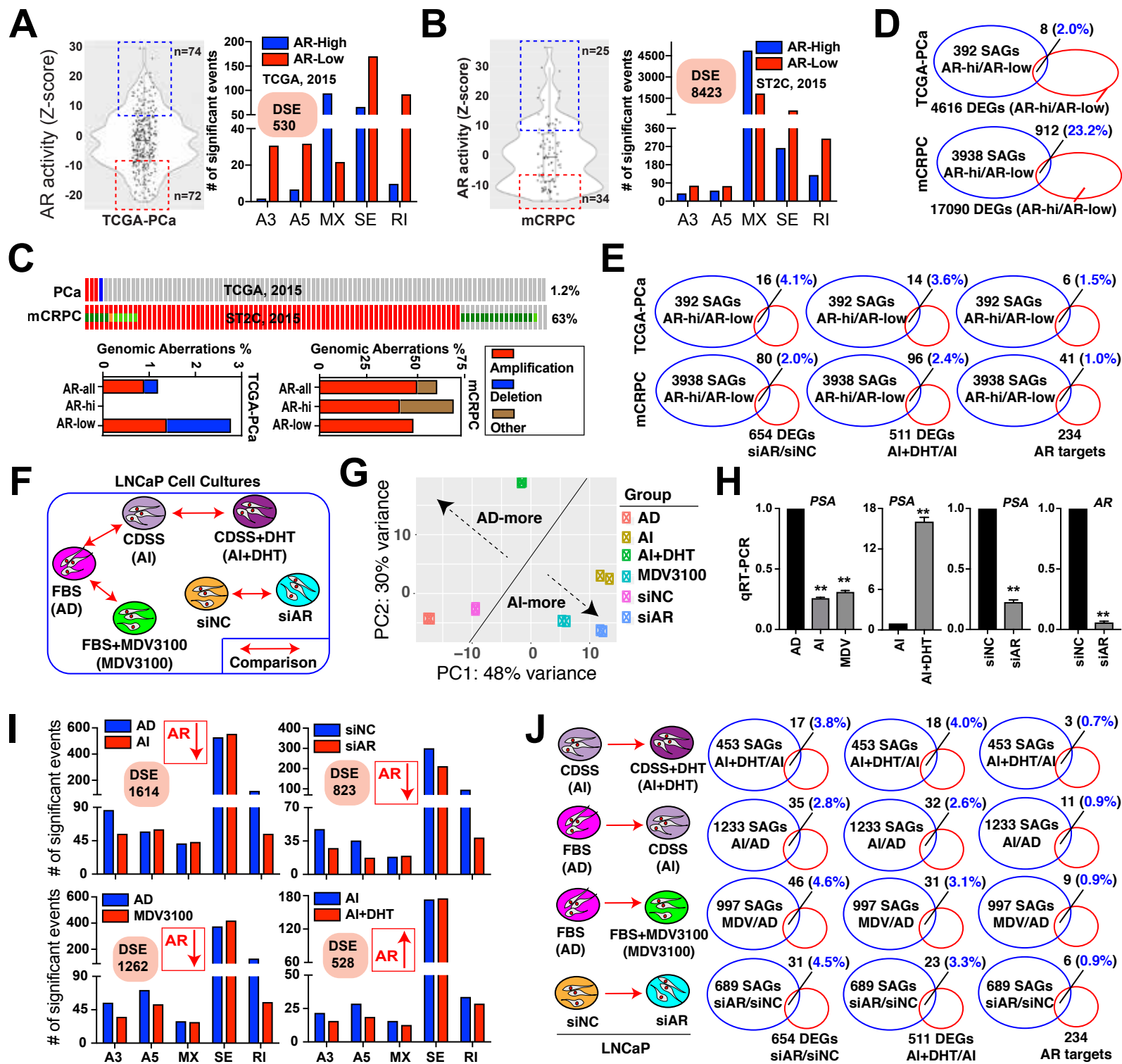
- 903 52. Schroeder A, Herrmann A, Cherryholmes G, Kowolik C, Buettner R, Pal S, *et al.* Loss of  
904 androgen receptor expression promotes a stem-like cell phenotype in prostate cancer through  
905 STAT3 signaling. *Cancer research* **2014**;74(4):1227-37 doi 10.1158/0008-5472.CAN-13-0594.  
906 53. Liu X, Chen X, Rycaj K, Chao HP, Deng Q, Jeter C, *et al.* Systematic dissection of phenotypic,  
907 functional, and tumorigenic heterogeneity of human prostate cancer cells. *Oncotarget*  
908 **2015**;6(27):23959-86.  
909 54. Asangani IA, Dommeti VL, Wang X, Malik R, Cieslik M, Yang R, *et al.* Therapeutic targeting of  
910 BET bromodomain proteins in castration-resistant prostate cancer. *Nature*  
911 **2014**;510(7504):278-82 doi 10.1038/nature13229.  
912 55. Olsen JR, Azeem W, Hellem MR, Marvyin K, Hua Y, Qu Y, *et al.* Context dependent regulatory  
913 patterns of the androgen receptor and androgen receptor target genes. *BMC Cancer*  
914 **2016**;16:377 doi 10.1186/s12885-016-2453-4.  
915 56. Younis I, Berg M, Kaida D, Dittmar K, Wang C, Dreyfuss G. Rapid-response splicing reporter  
916 screens identify differential regulators of constitutive and alternative splicing. *Mol Cell Biol*  
917 **2010**;30(7):1718-28 doi 10.1128/MCB.01301-09.  
918 57. Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, *et al.* Punctuated  
919 evolution of prostate cancer genomes. *Cell* **2013**;153(3):666-77 doi  
920 10.1016/j.cell.2013.03.021.  
921 58. Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, *et al.* Exome  
922 sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer.  
923 *Nature genetics* **2012**;44(6):685-9 doi 10.1038/ng.2279.  
924 59. Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, *et al.* The mutational  
925 landscape of lethal castration-resistant prostate cancer. *Nature* **2012**;487(7406):239-43 doi  
926 nature11125 [pii]10.1038/nature11125.  
927 60. Hieronymus H, Schultz N, Gopalan A, Carver BS, Chang MT, Xiao Y, *et al.* Copy number  
928 alteration burden predicts prostate cancer relapse. *Proceedings of the National Academy of*  
929 *Sciences of the United States of America* **2014**;111(30):11139-44 doi  
930 10.1073/pnas.1411446111.  
931 61. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, *et al.* Integrative genomic  
932 profiling of human prostate cancer. *Cancer cell* **2010**;18(1):11-22 doi  
933 10.1016/j.ccr.2010.05.026.  
934 62. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, *et al.* Integrative analysis of  
935 complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **2013**;6(269):p11  
936 doi 10.1126/scisignal.2004088.  
937 63. Ulz P, Belic J, Graf R, Auer M, Lafer I, Fischereder K, *et al.* Whole-genome plasma  
938 sequencing reveals focal amplifications as a driving force in metastatic prostate cancer. *Nature*  
939 *communications* **2016**;7:12008 doi 10.1038/ncomms12008.  
940 64. McDonald ER, 3rd, de Weck A, Schlabach MR, Billy E, Mavrakis KJ, Hoffman GR, *et al.*  
941 Project DRIVE: A Compendium of Cancer Dependencies and Synthetic Lethal Relationships  
942 Uncovered by Large-Scale, Deep RNAi Screening. *Cell* **2017**;170(3):577-92 e10 doi  
943 10.1016/j.cell.2017.07.005.  
944 65. Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, *et al.* Defining a  
945 Cancer Dependency Map. *Cell* **2017**;170(3):564-76 e16 doi 10.1016/j.cell.2017.06.010.  
946 66. Eskens FA, Ramos FJ, Burger H, O'Brien JP, Piera A, de Jonge MJ, *et al.* Phase I  
947 pharmacokinetic and pharmacodynamic study of the first-in-class spliceosome inhibitor E7107  
948 in patients with advanced solid tumors. *Clin Cancer Res* **2013**;19(22):6296-304 doi  
949 10.1158/1078-0432.CCR-13-0485.  
950 67. Pellagatti A, Armstrong RN, Steeples V, Sharma E, Repapi E, Singh S, *et al.* Impact of  
951 spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and  
952 clinical associations. *Blood* **2018**;132(12):1225-40 doi 10.1182/blood-2018-04-843771.

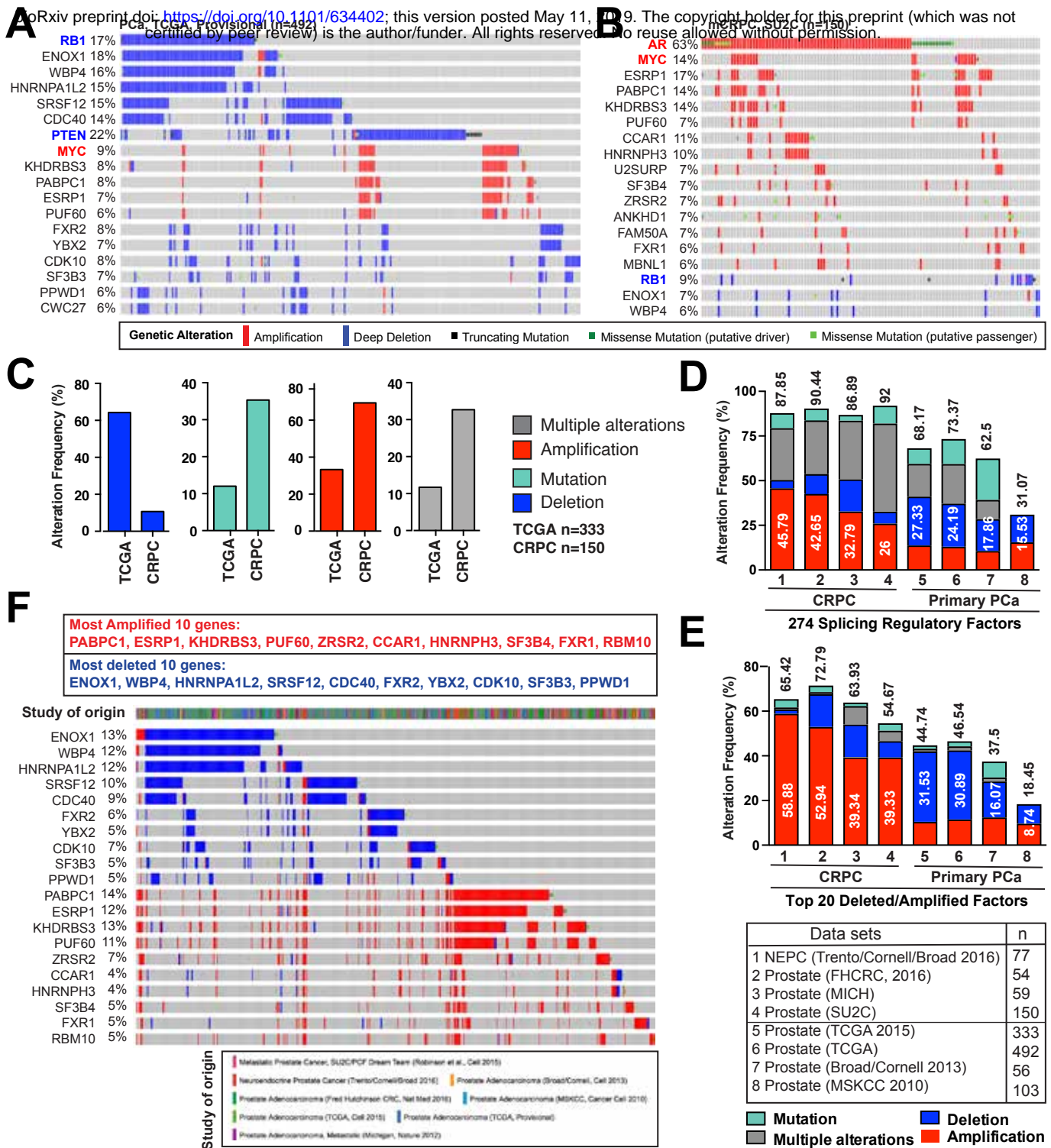


- 953 68. Wang Y, Chen D, Qian H, Tsai YS, Shao S, Liu Q, *et al.* The splicing factor RBM4 controls  
954 apoptosis, proliferation, and migration to suppress tumor progression. *Cancer cell*  
955 **2014**;26(3):374-89 doi 10.1016/j.ccr.2014.07.010.
- 956 69. Liu C, Kelnar K, Liu B, Chen X, Calhoun-Davis T, Li H, *et al.* The microRNA miR-34a inhibits  
957 prostate cancer stem cells and metastasis by directly repressing CD44. *Nat Med*  
958 **2011**;17(2):211-5 doi 10.1038/nm.2284.
- 959 70. Fei T, Chen Y, Xiao T, Li W, Cato L, Zhang P, *et al.* Genome-wide CRISPR screen identifies  
960 HNRNPL as a prostate cancer dependency regulating RNA splicing. *Proceedings of the*  
961 *National Academy of Sciences of the United States of America* **2017**;114(26):E5207-E15 doi  
962 10.1073/pnas.1617467114.
- 963 71. Takayama KI, Suzuki T, Fujimura T, Yamada Y, Takahashi S, Homma Y, *et al.* Dysregulation  
964 of spliceosome gene expression in advanced prostate cancer by RNA-binding protein PSF.  
965 *Proceedings of the National Academy of Sciences of the United States of America*  
966 **2017**;114(39):10461-6 doi 10.1073/pnas.1706076114.
- 967 72. Wang BD, Ceniccola K, Hwang S, Andrawis R, Horvath A, Freedman JA, *et al.* Alternative  
968 splicing promotes tumour aggressiveness and drug resistance in African American prostate  
969 cancer. *Nature communications* **2017**;8:15921 doi 10.1038/ncomms15921.
- 970 73. Jung H, Lee D, Lee J, Park D, Kim YJ, Park WY, *et al.* Intron retention is a widespread  
971 mechanism of tumor-suppressor inactivation. *Nature genetics* **2015**;47(11):1242-8 doi  
972 10.1038/ng.3414.
- 973 74. Simon JM, Hacker KE, Singh D, Brannon AR, Parker JS, Weiser M, *et al.* Variation in  
974 chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with  
975 widespread RNA processing defects. *Genome research* **2014**;24(2):241-50 doi  
976 10.1101/gr.158253.113.
- 977 75. Kumar A, White TA, MacKenzie AP, Clegg N, Lee C, Dumpit RF, *et al.* Exome sequencing  
978 identifies a spectrum of mutation frequencies in advanced and lethal prostate cancers.  
979 *Proceedings of the National Academy of Sciences of the United States of America*  
980 **2011**;108(41):17087-92 doi 10.1073/pnas.1108745108.
- 981



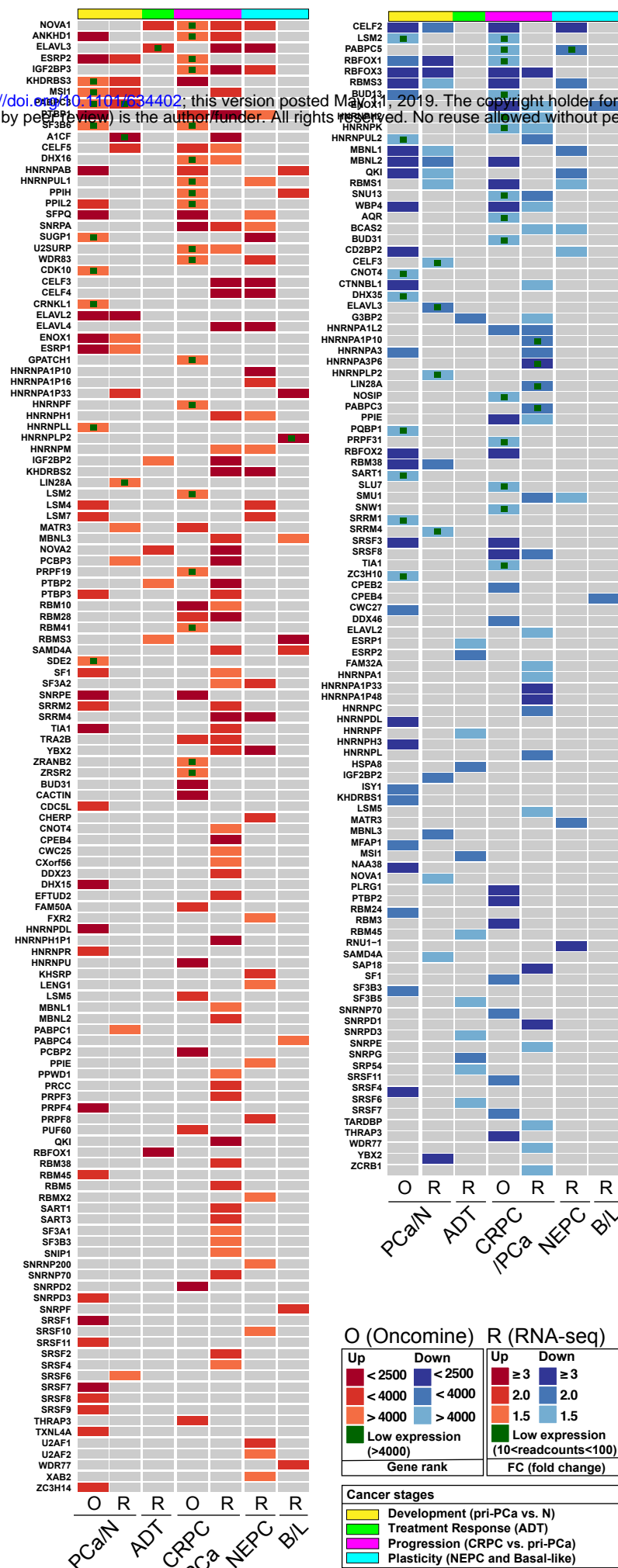
Zhang et al., Fig. 1





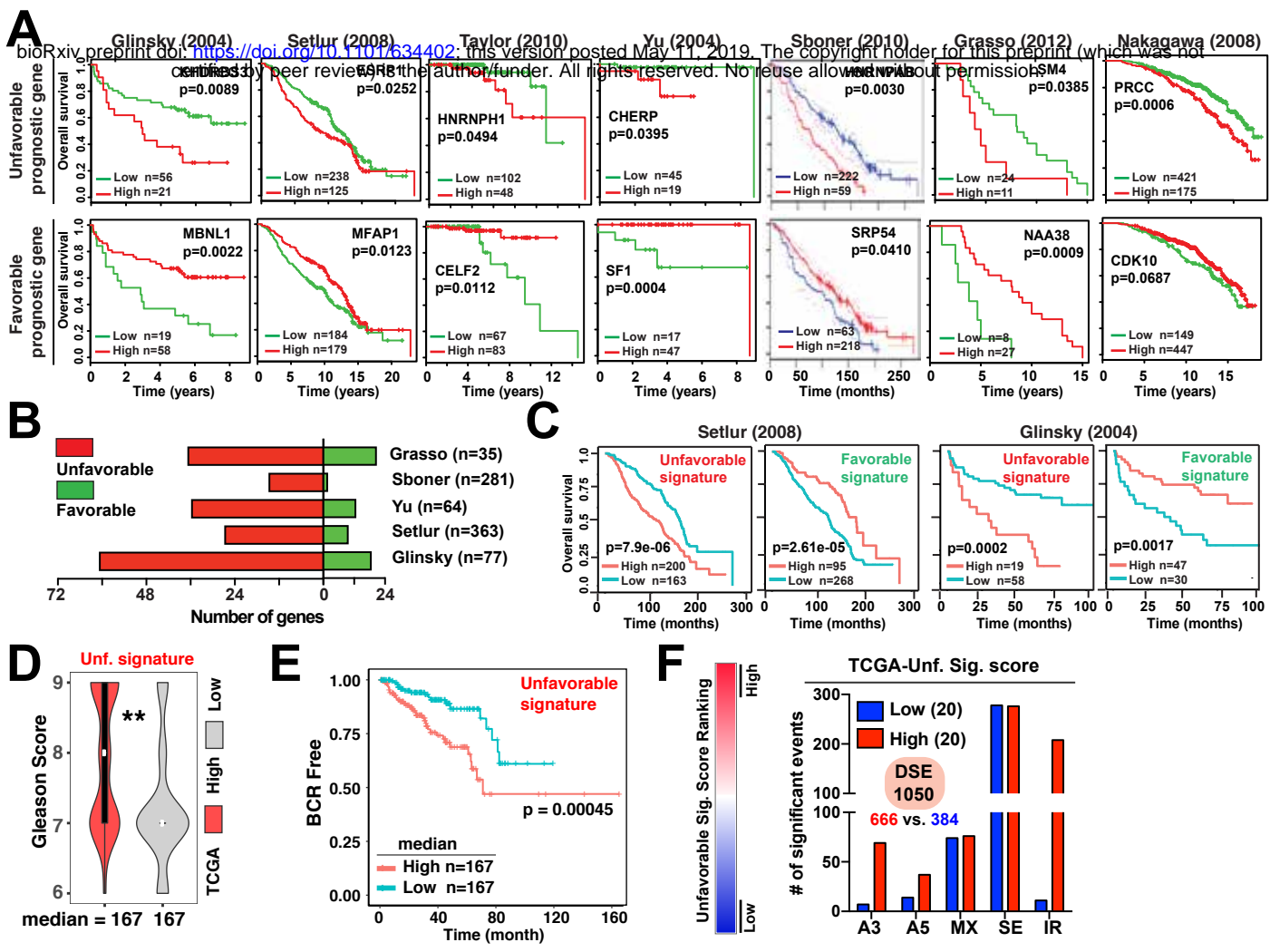
Zhang et al., Figure 3

bioRxiv preprint doi: <https://doi.org/10.1101/334402>; this version posted May 11, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

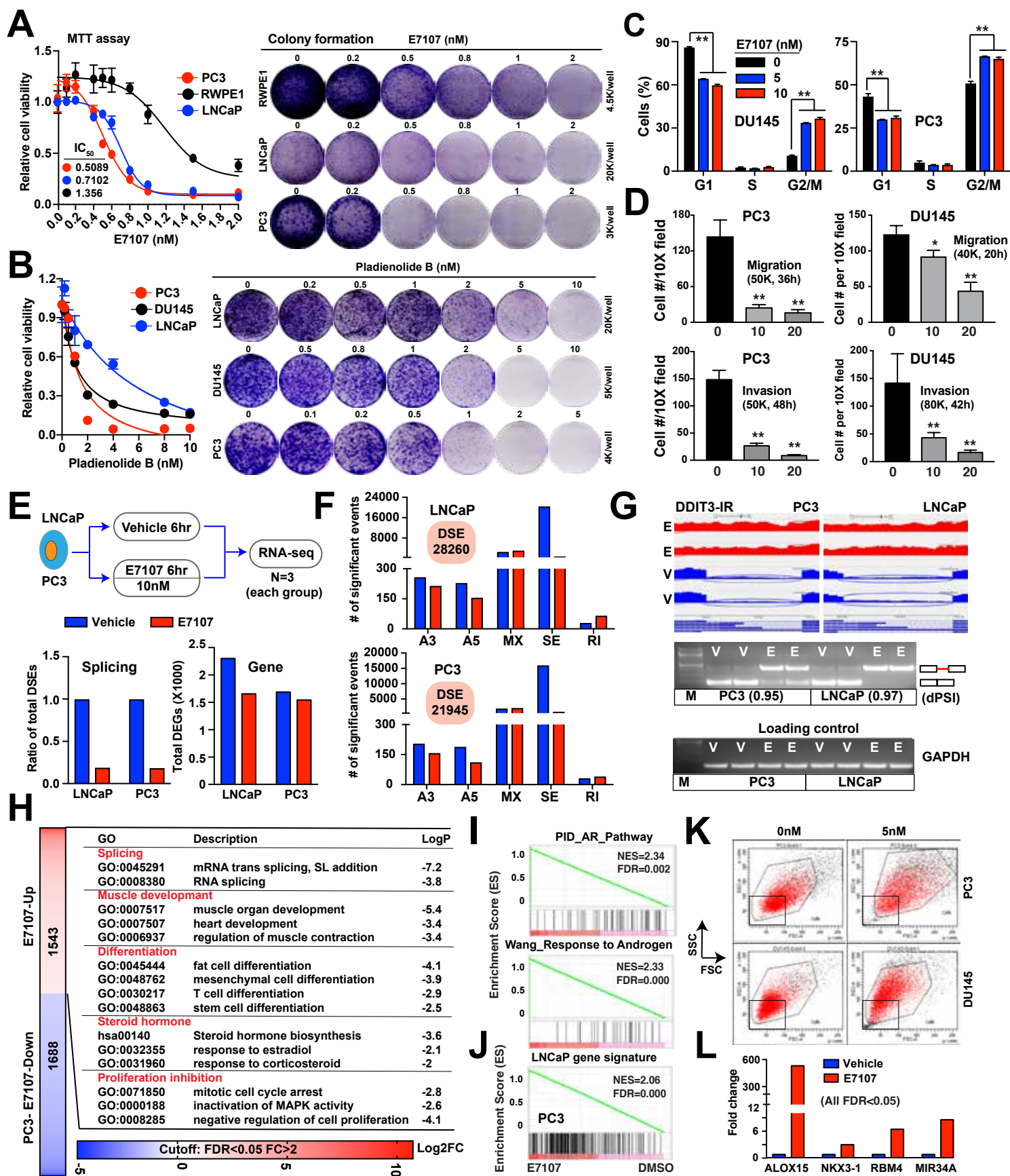


Zhang et al., Fig. 4





Zhang et al., Figure 5



Zhang et al., Fig. 6

