IRE1 activity regulates tumour and microenvironment cell lineage states while stratifying localised and metastatic prostate cancer.

Doultsinos D.[1,✉], Tomljanovic I.[2], Pilalis E.[3], Abusamra S.[1], Parmentier R.[4], Bridges I.E.[1], Figiel S.[1], Zekri Y.[7,8], Hester J[1]., Leach D.[5], Bevan C.[5], Lamb A.D.[9], Zwart W.[7,8], Chatziioannou A.[2,6], Le Magnen C.[4], Mills I.G.[1,✉]

Affiliations

[1] Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK;

[2]Department of Urology, Erasmus MC Cancer Institute, Erasmus University Medical Center, Wytemaweg 80, 3015 CN, Rotterdam, the Netherlands

[3]e-NIOS PC, Kallithea-Athens, Greece;

[4]Institute of Medical Genetics and Pathology, University Hospital Basel, Switzerland; Department of Urology, University Hospital Basel, Basel, Switzerland; Department of Biomedicine, University of Basel, University Hospital Basel, Basel, Switzerland

[5]Imperial College London, London, UK;

[6]Center of Systems Biology, Biomedical Research Foundation of the Academy of Athens, Athens, Greece.

[7]Division of Oncogenomics, The Netherlands Cancer Insitute, 1066 CX Amsterdam, The Netherlands

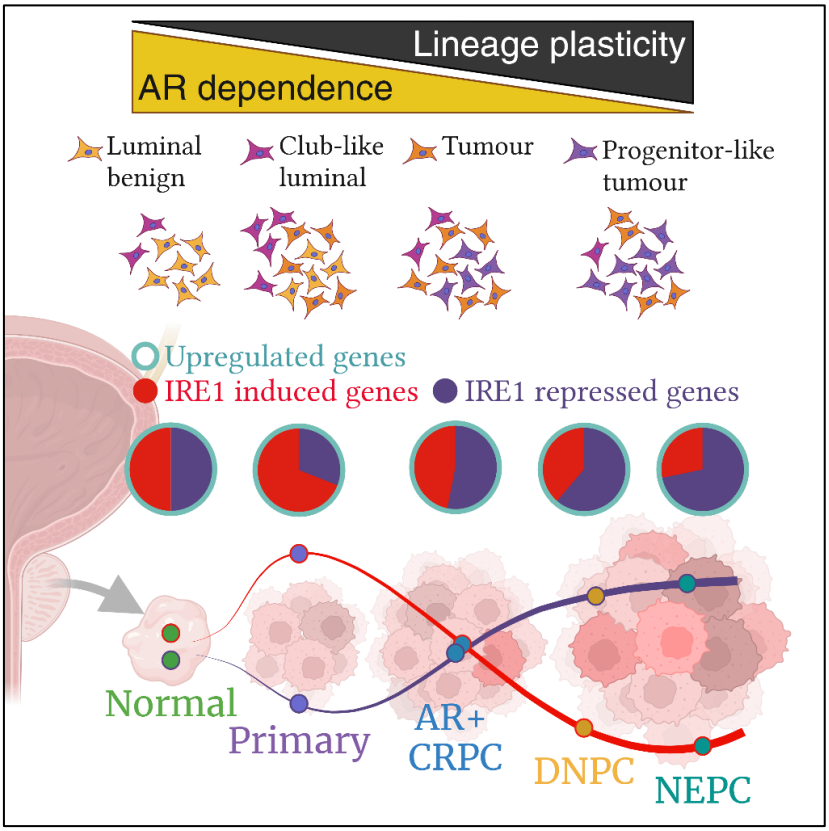
[8]Oncode Institute, The Netherlands

[9]Centre for Cancer Evolution, Barts Cancer Institute, Queen Mary University of London, UK.

✉ Corresponding authors

Abstract

Prostate cancer (PCa) is an androgen receptor (AR) driven, high-incidence disease significantly contributing to cancer mortality. PCa is in need of better risk stratification at diagnosis and treatment outcomes in patients at high risk of metastasis. The unfolded protein response (UPR) is an AR-dependent process. However, the impact of the UPR transducer IRE1 on AR-dependent biology and treatment resistance has not been defined. We use diverse pre-clinical models of stress response to describe IRE1 activity impact on multiple disease stages and demonstrate its involvement with poor prognosis (RB1 loss), and cell lineage determination (club phenotypes). Integrating clinical transcriptomic datasets, we chart IRE1 activity throughout PCa evolution by developing a PCa-specific, IRE1 activity gene set (IRE1_18) reflecting both tumoral and micro-environmental niches. IRE1_18 can determine tumoral identity, inform androgen deprivation treatment suitability, prognosticate localised and metastatic disease independently from AR activity, and guide IRE1 modulation as a novel combination therapeutic.



Graphical Abstract created using Biorender.com.

40 ## Introduction

41 Prostate cancer (PCa) is a heterogenous, multifocal, high-incidence disease that affects one in eight
42 men globally. PCa evolution depends on transcriptional events and epigenetic modifications, forming
43 an adaptive response to intrinsic/extrinsic stresses, culminating in lineage reprogramming and
44 treatment resistance[1]. PCa results in ~360000 deaths annually[2] due to metastatic progression,
45 supported by treatment/stress-resistant cancer cells which experience high proliferative rates and
46 androgen receptor (AR)-driven high metabolic demand.

47 These events require high protein turnover, leading to misfolded protein accumulation, causing
48 Endoplasmic Reticulum (ER) stress. The unfolded protein response (UPR) resolves this stress through
49 three main transducers, Inositol Requiring Enzyme alpha (IRE1α), protein kinase RNA-like endoplasmic
50 reticulum kinase (PERK) and activating transcription factor 6α (ATF6). IRE1α, hereafter referred to as
51 IRE1 and encoded by the gene *ERN1*, is the most evolutionarily conserved transducer of the UPR and
52 signals through two distinct downstream functional outputs[3]. Firstly, through the unconventional
53 splicing of the transcript encoding XBP1, a transcription factor, resulting in a non-canonical splice
54 variant lacking a 60-nucleotide sequence which encodes the active protein (XBP1s). Secondly, through
55 the targeted degradation/decay (RIDD) of a subset of mRNA (e.g., *DGAT2*[4], *BLOC1S1*[5]) and miRNA
56 transcripts to ease transcriptional load[6]. Both functions require IRE1 endoribonuclease activity and can
57 be thought of as a continuum of early (XBP1s) and late (RIDD) overlapping responses that are designed
58 to adapt to stress.

59 Using short (up to 24 hour) pre-clinical model treatments we have previously shown that the IRE1-XBP1
60 axis contains AR-induced genes, notably *ERN1* itself[7]. Targeting IRE1 in androgen-dependent pre-
61 clinical models can restrain tumorigenesis[8]. However, it remains unclear which IRE1-dependent genes
62 or biological processes underpin these responses and how IRE1 activity may change in response to
63 sustained stress, reflecting PCa evolution. Most studies so far have focused on the role of IRE1 as an

64 activator of XBP1 impacting tumour development[9] largely ignoring IRE1-RIDD impact on RNA stability
65 which could be restraining the emergence of cancers that are more difficult to treat.

66 In this study we used bulk, single cell (SC) and spatial (ST) RNA sequencing data from two-dimensional
67 (cell lines) and three-dimensional models (patient derived explants (PDEs)), as well as patient biopsy
68 samples to describe how IRE1 activity changes throughout PCa evolution. Crucially, we found that
69 adaptation to IRE1 loss correlated with cell identities usually driven by prognostically poor genomic
70 alterations such as RB1 loss. We also observed that IRE1-deficient cells develop a phenotype
71 resembling club cells, epithelial cells that interact with immunosuppressive microenvironmental cells.
72 We produced a novel IRE1 activity gene set, IRE1_18, representing significant ERN1 transcriptomic
73 perturbations in PCa-specific models. IRE1_18 is a proxy of multiple signalling pathways linked to PCa
74 progression and can chart PCa evolution from early stage to treatment resistant disease while
75 prognosticating large clinical cohorts. Current clinical practices impact therapeutic effectiveness on
76 PCa mortality[10] and therefore, using IRE1 activity to identify tumours with treatment resistance traits
77 early, could be transformative of patient care. IRE1 activity could be used to stratify patients for therapy
78 intensification, surveillance or suitability of IRE1/UPR modulation as novel combination therapeutics
79 alongside conventional treatments.

80

81 Results

82 ***Changes in AR activity correlate with significant changes in IRE1 activity in PCa cell lines and***
83 ***clinical datasets.***

84 To capture steady state adaptations to AR-induced physiological stresses we cultured LNCaP cells in
85 charcoal stripped serum media (CSM) or full serum media (FM) for 72 hours or treated them with the
86 synthetic androgen R1881 or EtOH for 12 hours. Whole transcriptome sequencing of these conditions
87 followed by gene set enrichment analysis (GSEA) revealed biological pathways negatively enriched in
88 CSM vs FM while positively enriched in the R1881 vs EtOH conditions including the androgen response,

89 fatty acid metabolism, MTORC1 signalling, oxidative phosphorylation and cholesterol homeostasis
90 (Figure 1A, B, Supplementary tables S1, S2)). Canonical AR targets reflecting AR activity were
91 significantly upregulated in the R1881 treatment whilst significantly downregulated in the CSM
92 condition (Supplementary Figure S1A, Supplementary tables S3, S4).

93 Some Hallmarks were only significant upon androgen treatment or androgen deprivation. These
94 included a positive enrichment of protein secretion and negative enrichment of the Unfolded Protein
95 Response (UPR) in R1881 vs EtOH and CSM vs FM conditions respectively (Figure 1A, B). The UPR
96 Hallmark is not IRE1 specific containing target genes of all UPR transducers (ATF4, ATF6 and XBP1s)
97 while RIDD may not be represented within it. Consequently, we used a curated list of XBP1, and RIDD
98 target genes to correlate alterations in AR signalling with IRE1 activity. Whereas XBP1 target genes were
99 largely androgen-induced, RIDD targets were not consistently androgen dependent (Figure 1C,
100 Supplementary figure S1A).

101 UPR responses are tissue dependent and therefore, we determined how AR and IRE1 target gene
102 expression changed over 72-hour CSM culture of LNCaP compared to FM. A decrease in IRE1 activity
103 (*XBP1s* downregulated, RIDD targets *DGAT2*, *BLOC1S1* upregulated) was seen as early as 8 hours,
104 mirroring full length AR levels and activity (Figure 1D). Similar profiles were observed at the protein level
105 (Supplementary Figure S1B). Simultaneously we observed that levels of *ATF4* and *eIF2α* either did not
106 significantly or consistently change (Supplementary Figure S1C).

107 Hallmarks including E2F targets, G2M checkpoint and epithelial mesenchymal transition EMT (Figure
108 1A) were partially androgen dependent. Therefore, we hypothesised that they may be regulated by other
109 AR-dependent global mechanisms such as the IRE1 axis. We expanded our pre-clinical model
110 repertoire and assessed canonical AR and IRE1 target gene expression in patient derived explants
111 (PDEs)[11] treated with, or deprived of, dihydrotestosterone (DHT) for 12 hours. AR response markers
112 *KLK3* and *CAMKK2* were significantly upregulated upon DHT treatment. *ERN1* itself and *DGAT2*,

113 *BLOC1S1* were not significantly altered but XBP1s was significantly upregulated (Figure 1E,
114 Supplementary figure S1D).

115 We then tested if this AR-IRE1 activity correlation could be replicated within a clinical cohort
116 representing AR positive and AR negative tumours. Regulon analyses of 48 castration resistant prostate
117 cancer (CRPC) samples split into CRPC-Adenocarcinoma (n=35, AR positive) versus CRPC-
118 neuroendocrine prostate cancer (NEPC, n=13, AR negative)[12] showed that AR and XBP1 regulons were
119 both less active in NEPC samples while reprogramming markers pertaining to stem-like lineages such
120 as SOX2 and POU3F2 were increasingly active (Figure 1F, G). As such, transcription factor (TF) activity
121 correlates with IRE1 activity and PCa lineage.

122

123

124

125

126

127

128

129

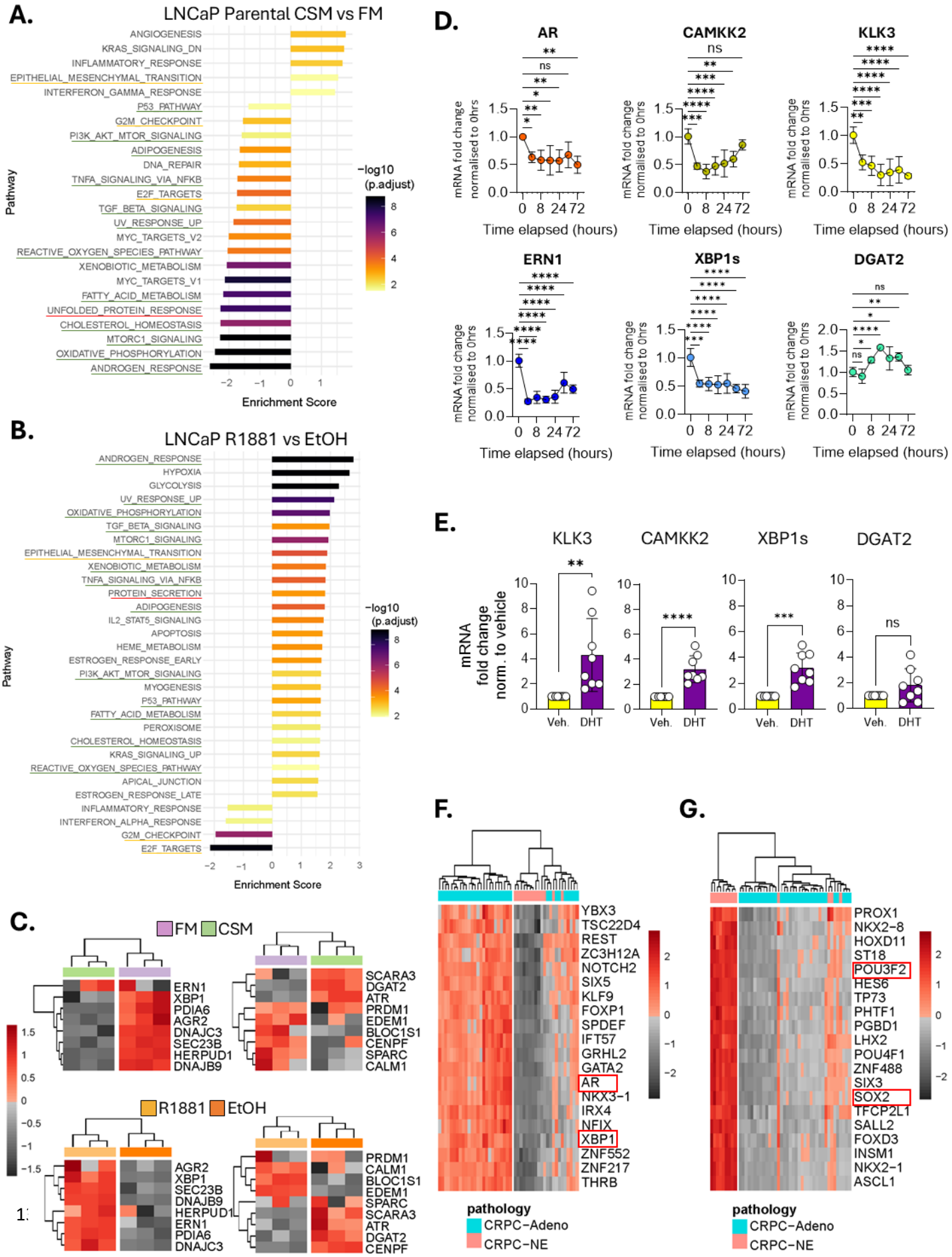
130

131

132

133

Figure 1. ***Changes in AR activity correlate with significant changes in IRE1 activity in pre-clinical models and clinical datasets.***



135 ***Figure 1. Changes in AR activity correlate with significant changes in IRE1 activity in pre-clinical***
136 ***models and clinical datasets.*** Barplots displaying GSEA normalized enrichment scores (NES) for significantly
137 enriched MSigDB Hallmark gene sets in: LNCaP cells grown in charcoal stripped media (CSM) for 72 hours
138 compared to LNCaP cells grown in full media (FM) for 72 hours (A), and LNCaP cells treated with R1881 for 12
139 hours compared to LNCaP cells treated with ethanol (EtOH) for 12 hours (B). Bar length indicates the normalized
140 enrichment score and bar colour signifies the statistical significance. Hallmarks activated in B) but suppressed in
141 A) are underlined in green. Hallmarks activated or suppressed in both A) and B) are underlined in orange. Hallmarks
142 representing protein homeostasis are underlined in red. C) Expression heatmaps of selected "IRE1 induced" (left)
143 and "IRE1 repressed" (right) genes according to the literature in LNCaP cells in CSM (green) vs FM (purple), (top);
144 and R1881 (yellow) vs EtOH (orange) (bottom). Colour scale represents expression z-score from black (lowest)
145 to red (highest). D) mRNA fold change of androgen receptor activity markers (full length AR, CAMKK2, KLK3) and
146 IRE1 activity markers (ERN1, XBP1s, DGAT2) in LNCaP parental cells grown over 72 hours in CSM, normalised to
147 untreated FM control. Statistical analysis was performed using an ordinary one-way ANOVA with Dunnett's
148 multiple comparisons test and single pooled variance. Four stars signify adjusted p value of <0.0001. E) mRNA
149 fold change of AR and IRE1 activity markers KLK3, CAMKK2, XBP1s and DGAT2 in patient derived explants (PDEs)
150 treated with dihydrotestosterone (DHT), normalised to PDEs treated with vehicle control. Statistical analysis was
151 performed using a T-test per condition. Each star signifies an order of magnitude of significance <0.05. F, G)
152 Heatmaps representing regulons significantly altered in neuroendocrine (NEPC, pink) castration resistant (CRPC)
153 tumours compared to adenocarcinoma CRPC (light blue). Regulon levels are represented as single-sample
154 enrichment scores with individual regulons denoted by the name of their corresponding transcription factor (TF).
155 Significantly activated and suppressed regulons are shown, with AR, XBP1 (AR, IRE1 activity), SOX2 and POU3F2
156 (reprogramming) in red boxes. Colour scale represents regulon enrichment z-scores from lowest in greyscale to
157 highest in red-scale.

158

159

160

161

162 ***Knocking out IRE1 induces changes in lineage characteristics of AR dependent PCa cells.***

163 Having established that IRE1 activity changes with PCa progression alongside the AR we investigated
164 the adaptation of AR positive cells (LNCaP) to stable IRE1 ablation, using an IRE1 CRISPR-knockout
165 (LNCaP ERN1$^{-/-}$)[13], lacking IRE1 at steady state or chemical ER stress (tunicamycin/thapsigargin)
166 (Supplementary Figure S2A, B). Bulk RNAseq confirmed multiple canonical "IRE1 induced" and "IRE1
167 repressed" genes were downregulated and upregulated respectively in ERN1$^{-/-}$ cells (Supplementary
168 figure S2C).

169 Any observed biology in the ERN1$^{-/-}$ cells could be attributed to the cell line generation selection
170 pressure. To definitively link physiological observations to the absence of IRE1 activity we stably
171 expressed an IRE1 mutant in the ERN1$^{-/-}$ cells carrying a truncation mutation at tryptophan 780[14] thus
172 lacking the effector ribonuclease domain. These new ERN1$^{-/-}$ Q780* cells were IRE1 activity deficient
173 at steady state and ER stress conditions (Supplementary Figures S2D-G).

174 We then documented the growth of LNCaP, LNCaP ERN1$^{-/-}$ and LNCaP ERN1$^{-/-}$ Q780* in FM or CSM.
175 ERN1$^{-/-}$ and ERN1$^{-/-}$ Q780* doubling time was 25% and 32% lower than parental LNCaP in FM,
176 respectively. In CSM, ERN1$^{-/-}$ and ERN1$^{-/-}$ Q780* doubling time was increased by 50% compared to FM
177 whilst parental cells stopped growing (Figure 2A) entirely. To document the biological functions
178 underlying this phenotype we performed pathway analysis on whole transcriptome data from these
179 conditions. E2F targets, G2M checkpoint, and the EMT were significantly enriched in ERN1$^{-/-}$ samples
180 (Figure 2B, Supplementary Table S5).

181 Consequently, androgen deprivation-induced growth arrest and DNA repair abrogation on PCa cells
182 might be a consequence of both the repression of XBP1 activity and the stabilization of RIDD targets.
183 We thus compared significantly altered Hallmarks induced by androgen deprivation with those altered
184 by knocking out IRE1. ERN1$^{-/-}$ cells were still sensitive to CSM as E2F targets, G2M checkpoint and AR
185 Response were negatively enriched compared to ERN1$^{-/-}$ cells grown in FM (Figure 2C, Supplementary
186 Table S6). However, these Hallmarks were not affected by androgen deprivation to the same degree as

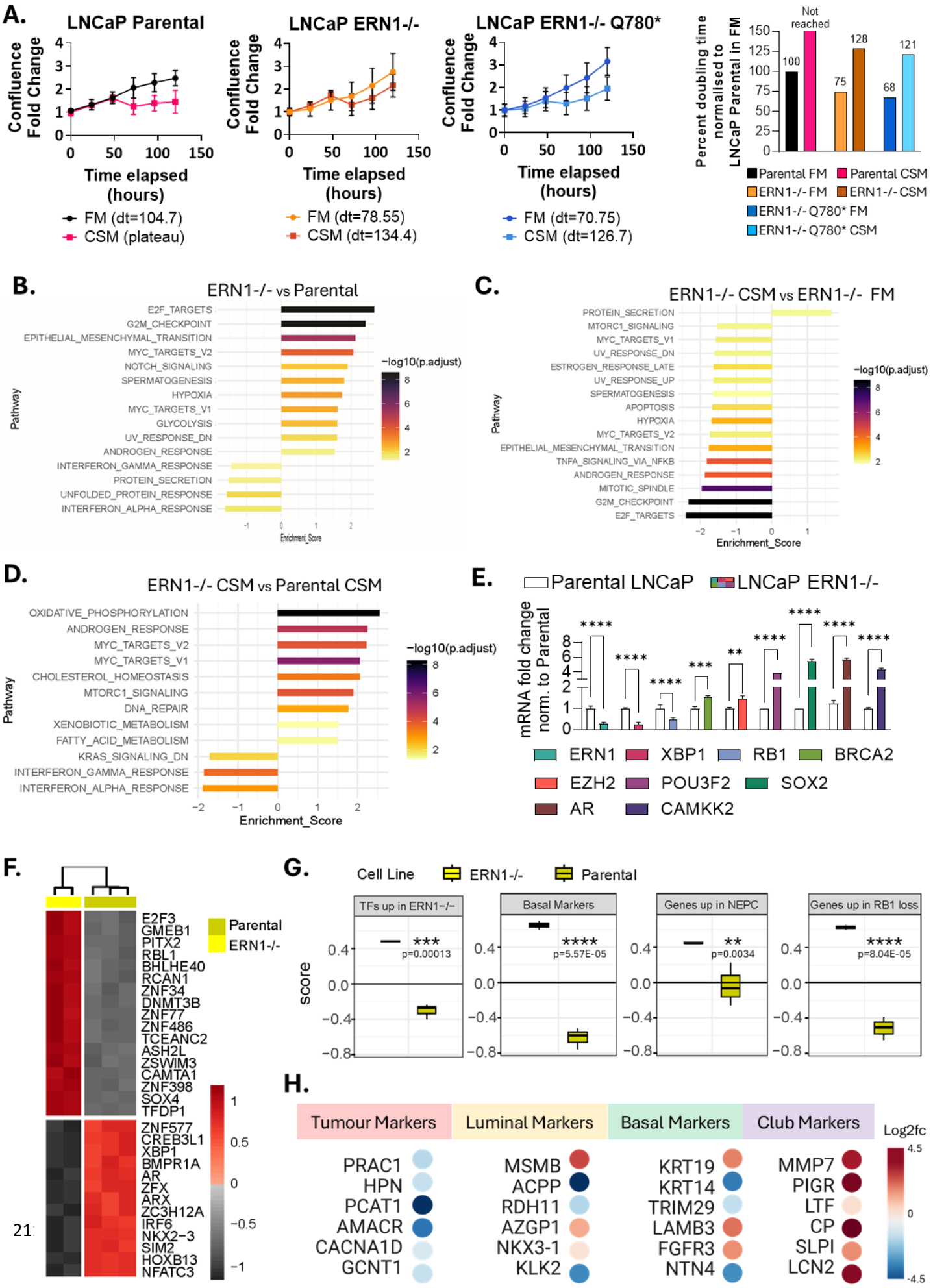
187 DNA repair, AR response and MTORC1 were upregulated in ERN1$^{-/-}$ compared to parental cells grown
188 in CSM (Figure 2D, Supplementary Table S7).

189 We observed that low *XBP1* correlated with high activity of reprogramming TFs (*SOX2, POU3F2*) in
190 clinical samples (Figure 1F, G) and that IRE1 loss leads to cell cycle de-regulation in PCa cell lines
191 (Figure 2A-D). Consequently, we measured mRNA of lineage reprogramming[15,16] and cell cycle
192 regulation[17] genes with known clinical association in PCa. *RB1* was significantly downregulated whilst
193 *BRCA2, SOX2 EZH2* and *POU3F2* mRNA was significantly upregulated in ERN1$^{-/-}$ cells in FM (Figure
194 2E). Therefore, adaptation to IRE1 activity loss changed the capacity of LNCaP cells to respond to
195 androgen deprivation whilst retaining their AR axis.

196 We then mapped TF repertoires (regulons) significantly affected by *ERN1* knockout. Stress response
197 (CREB3L1), inflammatory signalling (NFATC3, IRF6), cell cycling (E2F3) and
198 embryogenesis/stemness/differentiation regulons (SOX4, PITX2, HOXB13) were significant. RBL1 was
199 positively enriched (Figure 2F, Supplementary table S8) which has been shown to regulate cell cycle in
200 response to RB1 loss[18].

201 Gene signatures associated with RB1 loss, NEPC and basal phenotypes (Supplementary Table S9)
202 segregated ERN1$^{-/-}$ from parental cells, in a similar manner to the regulons (Figure 2G). Thus, knocking
203 out IRE1 in cells that are AR-dependent, and retain some luminal epithelial character, could change
204 their lineage characteristics. Comparing luminal, basal and club cells within a PCa SC dataset[19] we
205 observed that TFs with regulons activated upon IRE1-loss were significantly upregulated in basal and
206 club cells compared to luminal epithelial cells (Supplementary Figure 3A). We then documented the
207 enrichment of genes upregulated in tumoral, luminal, basal and club cells[20] (Supplementary Table S10)
208 in the ERN1$^{-/-}$ transcriptome. ERN1$^{-/-}$ cells were positively enriched for club[21–23] while negatively
209 enriched for tumoral and luminal cell markers, showing decreases of *AMACR*, ACPP and NTN4 but
210 increases in *MMP7, PIGR, MSMB* and *KRT19* (Figure 2H, Supplementary Figure S3B).

Figure 2. ***Knocking out IRE1 induces lineage changes in AR dependent PCa cells.***



212 ***Figure 2. Knocking out IRE1 induces lineage changes in AR-dependent PCa cells.*** A) Growth curves of LNCaP
213 parental, LNCaP ERN1-/- and LNCaP ERN1-/- Q780* cells grown over 120 hours in FM or CSM. Confluence fold
214 change was normalised to Incucyte readings taken at the 0-hour time point of each condition. Exponential
215 (Malthusian) growth curve equations were used to estimate the doubling time of each cell line in FM or CSM.
216 GSEA normalized enrichment scores (NES) for significantly enriched MSigDB Hallmark gene sets in LNCaP ERN1-
217 /- cells compared to LNCaP parental cells grown in FM (B), LNCaP ERN1-/- cells grown in CSM compared to
218 LNCaP ERN1-/- cells grown in FM (C), and LNCaP ERN1-/- cells compared to LNCaP parental cells grown in CSM
219 (D). Bar length indicates the normalized enrichment score (NES) and bar colour signifies the statistical
220 significance. E) mRNA fold change of ERN1, XBP1, cell cycle markers RB1, BRCA2, and lineage reprogramming
221 markers EZH2, POU3F2, SOX2 in LNCaP ERN1-/- normalized to parental LNCaP. Statistical analysis was
222 performed using an ordinary two-way ANOVA and Sidak's multiple comparisons test with single pooled variance.
223 Heatmap representing regulons significantly altered in LNCaP ERN1-/- (IKO, light blue) compared to parental
224 LNCaP (pink). Top 30 statistically significant regulons shown ($p<0.001$). Regulon levels are represented as single-
225 sample enrichment scores with individual regulons denoted by the name of their corresponding transcription
226 factor (TF). Colour scale represents regulon enrichment z-scores (from lowest in greyscale to highest in red-
227 scale). G) GSVA single-sample gene set enrichment scores enrichment scores in LNCaP ERN1-/- cells (IKO,
228 yellow) compared to LNCaP parental cells (mustard) shown from left to right for gene sets including: TFs with
229 regulons activated in ERN1-/- cells, genes upregulated in the RB1-loss signature from Ertel et al., a basal cell
230 signature from Aggarwal et al, and the genes upregulated in the NEPC signature from Beltran et al. Statistical
231 testing was performed using limma. H) Bulk RNAseq log2 fold change expression levels in $ERN1^{-/-}$ cells compared
232 to parental, of genes widely reported to be associated with tumour, luminal epithelial, basal epithelial and club
233 epithelial cells. Downregulated genes in blue, upregulated genes in red. Created using Biorender.com. Each
234 experiment was performed at least in duplicate. Four stars signify adjusted p value of <0.0001.

235

236

237

238 ***IRE1 perturbation affects cell cycle and lineage reprogramming markers and responses to***
239 ***physiological androgen deprivation therapy (ADT).***

240 Adaptation to IRE1 loss leads to an epithelial club-like lineage shift histologically associated with
241 inflammatory atrophy[21], that can result from AR blockade[23]. Low IRE1 activity correlates with ADT
242 resistant clinical presentations (Figure 1) and IRE1 deficient cells express high RBL1 and low RB1 RNA
243 levels. Lineage plasticity is facilitated by knocking RB1 out, or knocking RB1 out alongside TP53 in
244 LNCaP cells [24,25] which are reported to be resistant to replication stress, have high proliferative rates
245 and increased DNA repair capacity[26]. Intriguingly, Hallmarks affected by RB1 loss or RB1 and TP53 loss
246 were similar to IRE1 loss (Figures 3A, 3B, Supplementary Tables S11, S12) showing upregulation in E2F,
247 G2M checkpoint and MYC targets.

248 We then tested whether IRE1 activity was altered in RB1 deficient cells. *ERN1, BRCA2*, *RB1*, *EZH2* and
249 *POU3F2* were significantly downregulated in RB1$^{-/-}$ cells (Figure 3C). RB1$^{-/-}$/TP53$^{-/-}$ cells had
250 significantly lower levels of both *ERN1*, *XBP1s, BRCA2* and *RB1* (Figure 3D) while higher levels of *EZH2*,
251 *POU3F2* and *SOX2* (Figure 3E) compared to parental.

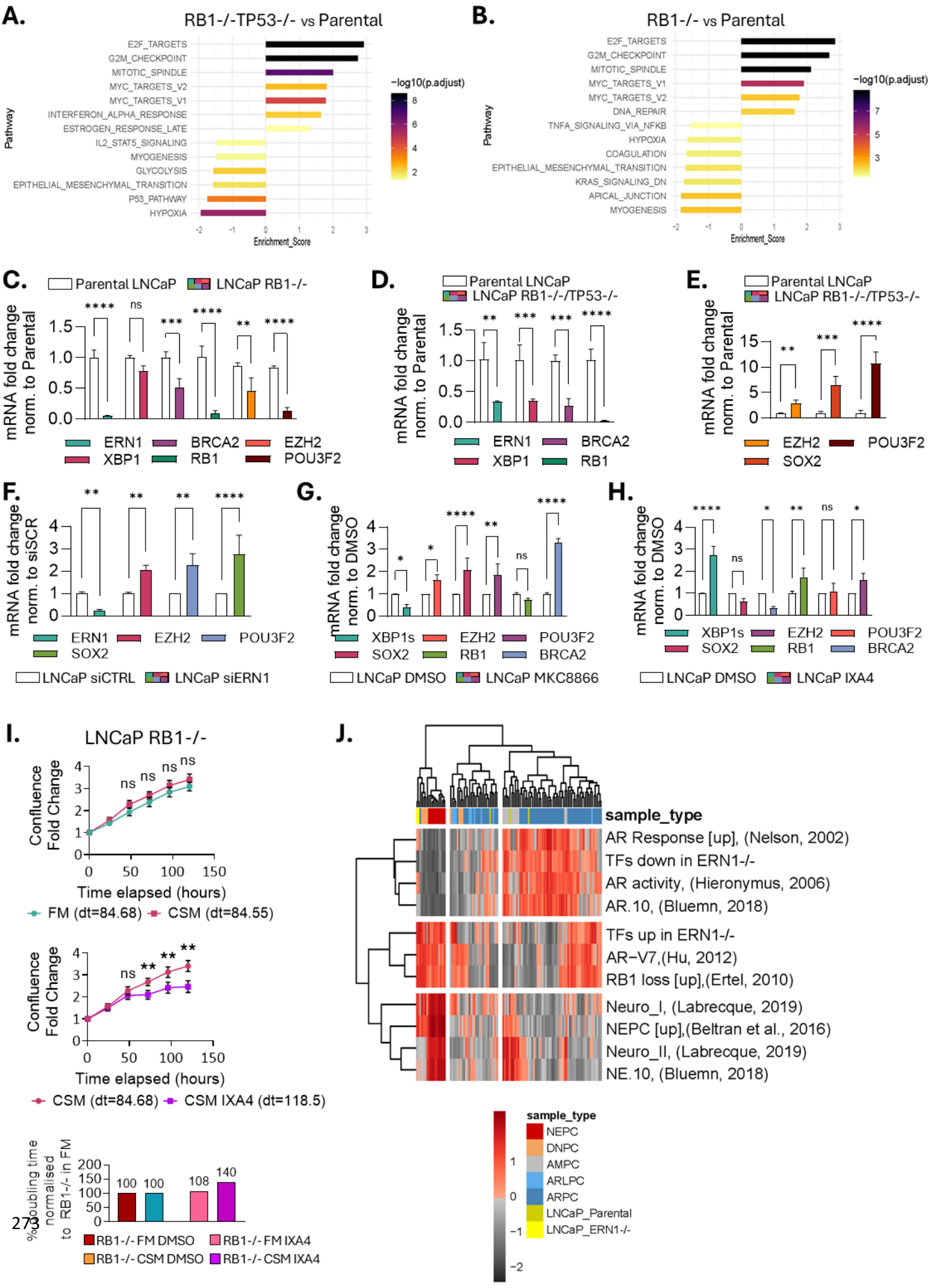
252 *ERN1* siRNA silencing significantly upregulated *POU3F2*, *EZH2*, *SOX2* (Figure 3F). Treatment with the
253 salicylaldehyde IRE1 inhibitor MKC8866[27] caused significant upregulation of *EZH2*, *POU3F2*, *SOX2* and
254 *BRCA2* (Figure 3G) while treatment with the IRE1 activator IXA4[28], caused significant upregulation of
255 *RB1* and downregulation of *BRCA2* with reprogramming markers showing little change (Figure 3H).

256 We correlated resistance to ADT with IRE1 activity loss. Therefore, we tested the effect of
257 pharmacological IRE1 activity upregulation on the ability of treatment resistant cells to grow in CSM.
258 CSM did not affect growth of either RB1$^{-/-}$ or RB1$^{-/-}$/TP53$^{-/-}$ LNCaP cells. However, addition of subtoxic
259 doses of the IRE1 activator IXA4, slowed their growth in CSM. RB1-/- doubling time increased by 40%
260 after 24 hours whilst RB1$^{-/-}$/TP53$^{-/-}$ only changed by 10% after 72-hours compared to FM and DMSO
261 (Figure 3I, Supplementary Figure S4A).

262 Then, in a metastatic, castration resistant PCa (mCRPC, 98 tumours from 55 men)[29] cohort, we
263 investigated whether the patterns of enrichment of gene signatures comprising TFs with regulons
264 significantly activated or suppressed in ERN1$^{-/-}$ (Figure 2F) could differentiate AR positive from AR
265 negative mCRPC alongside several AR and NEPC gene signatures. The set of TFs with regulons
266 suppressed in ERN1-/- - clustered with AR activity signatures whilst the TFs with regulons activated in
267 ERN1-/- clustered with RB1-loss and ARV-7 activity signatures[30,31], distinguishing AR naïve from AR
268 agnostic or NEPC/DNPC PCa. Overall, ERN1-/- cells clustered with NEPC and double-negative for
269 NEPC and AR PCa (DNPC) as well as some AR-low (ARLPC) mCRPC samples (Figure 3J).
270 Consequently, IRE1 transcriptomic repertoires may correlate with AR driven and agnostic phenotypes
271 by playing a key role in lineage plasticity, cell cycle and ultimately treatment response.

272

Figure 3. ***IRE1 activity impacts AR responses in RB1 loss.***



274 ***Figure 3. IRE1 activity impacts AR responses in RB1 loss.*** GSEA normalized enrichment scores (NES) for
275 significantly enriched MSigDB Hallmark (H) gene sets in: LNCaP RB1$^{-/-}$TP53$^{-/-}$ cells compared to LNCaP parental
276 cells grown in FM (A), and LNCaP RB1-/- cells grown in CSM compared to LNCaP parental cells grown in FM (B).
277 Bar length indicates the normalized enrichment score (NES) and bar colour signifies the statistical significance.
278 C, D, E) mRNA fold change of *ERN1, XBP1, RB1, BRCA2, EZH2, POU3F2, SOX2* in C) LNCaP RB1-/- normalised
279 to parental LNCaP and D, E) LNCaP RB1$^{-/-}$TP53$^{-/-}$ cells normalised to parental LNCaP. F) mRNA fold change of
280 *ERN1, EZH2, POU3F2, SOX2* in LNCaP parental cells treated with siRNA against ERN1 normalised to scramble
281 control. mRNA fold change of *XBP1s, EZH2, POU3F2, SOX2, RB1, BRCA2* in LNCaP parental cells treated with
282 G) MKC8866 or H) IXA4 normalised to DMSO. Statistical analysis was performed using an ordinary two-way
283 ANOVA and Sidak's multiple comparisons test with single pooled variance. I) Growth curves of LNCaP RB1$^{-/-}$ cells
284 grown over 120 hours in FM (green), CSM (red) or in CSM in the presence of subtoxic (25µM) doses of IXA4
285 (purple). Confluence fold change was normalised to Incucyte readings taken at the 0-hour time point of each
286 condition. T-tests were used for individual time point comparisons between FM and CSM; CSM and CSM plus
287 IXA4. Exponential (Malthusian) growth curve equations were used to estimate the doubling time of each condition.
288 J) Heatmap of GSVA enrichment scores for 11 selected PCa gene signatures shown across samples from the
289 Labrecque patient cohort (N=98) together with LNCaP parental and ERN1-/- cell line samples (N=5). Colour
290 scale represents GSVA enrichment z- scores (from lowest in greyscale to highest in red-scale). Each experiment
291 was performed at least in duplicate. Four stars signify adjusted p value of <0.0001.

292

293

294

295

296

297

298

299 ***Enzalutamide treatment induces a UPR downregulation which may act as a negative feedback***
300 ***loop that leads to treatment resistance to ADT in PCa.***

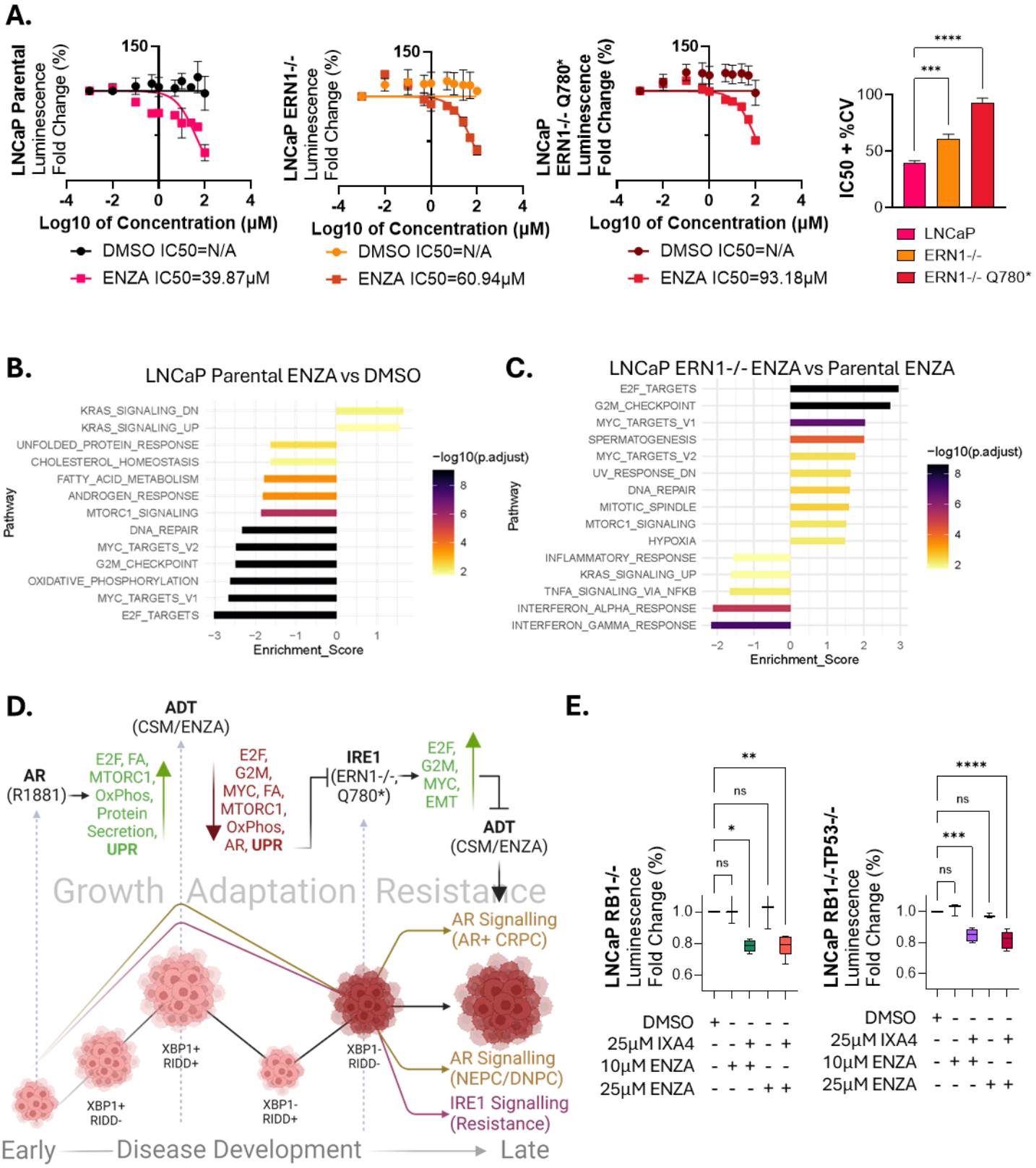
301 Adaptation to IRE1 activity loss may push AR responsive cells towards an ADT resistant phenotype,
302 marking the starting point of IRE1 activity saturation and the shift from an active phenotype to a passive,
303 adapted phenotype. ADT could catalyse this key event in the shift from early stage to late-stage disease
304 progression. Consequently, we investigated the effect of increasing concentrations of Enzalutamide
305 (ENZA) on LNCaP, LNCaP ERN1$^{-/-}$ and LNCaP ERN1$^{-/-}$Q780* cells over 12 hours. Both IRE1-deficient
306 cell lines had significantly higher ENZA IC50s (Figure 4A, Supplementary Figure S4B).

307 ENZA treatment of parental LNCaP, induced suppression of AR response, DNA repair, MTORC1, E2F
308 targets and the UPR (Figure 4B, Supplementary Table S13). Comparing ERN1-/- to parental LNCaP
309 treated with ENZA, we observed positive significant enrichment for E2F targets, G2M checkpoint and
310 DNA repair (Figure 4C, Supplementary Table S14), corroborating our CSM results (Figures 1,2). In
311 addition, we compared ERN1$^{-/-}$ responses to ENZA, to publicly available RB1$^{-/-}$/TP53$^{-/-}$ LNCaP ENZA
312 response data observing strong similarities between the two. Only inflammatory, hypoxia and MTOR
313 signalling significantly differentiating the responses of these two cell lines to ENZA (Supplementary
314 Figure S4C, Supplementary Tables S15, S16).

315 Thus, IRE1 activity suppression may play a pivotal role in acquired treatment resistance (Figure 4D)
316 through lineage adaptation to ADT. We then tested whether IRE1 activation could sensitise RB1
317 deficient cells to ENZA. Whereas the IRE1 activator IXA4 had no effect on viability when used alone
318 (not shown), it sensitised both ENZA resistant (Supplementary Figure S4D, E) RB1$^{-/-}$ and RB1$^{-/-}$TP53$^{-/-}$
319 cells to physiologically relevant doses of ENZA over a 12-hour treatment (Figure 4E).

320 Consequently, treatment resistant PCa pre-clinical models have low IRE1 activity recapitulating clinical
321 observations but a short-term increase in IRE1 activity enhances the response to AR antagonism.

Figure 4. ***Enzalutamide treatment downregulates UPR responses which contribute to acquired ADT resistance.***



324 ***Figure 4. Enzalutamide treatment downregulates UPR responses which contribute to acquired ADT***
325 ***resistance.*** A) Enzalutamide (ENZA) concentration gradient measuring cytotoxicity through ATP MTT assays in
326 LNCaP parental, LNCaP ERN1$^{-/-}$ and LNCAP ERN1$^{-/-}$ Q780* cells treated for 12 hours with increasing
327 concentrations of ENZA. Logarithmic concentration vs normalised response curves were fit to calculate the IC50
328 of ENZA for each cell line. Bar chart represents ENZA IC50 comparison between the cell lines. Maximum variance
329 used to calculate significance by ordinary one-way ANOVA with multiple comparisons. B) GSEA normalized
330 enrichment scores (NES) of LNCaP parental cells treated with 10µM ENZA for 12 hours compared to DMSO or
331 C) of LNCaP ERN1$^{-/-}$ cells treated with 10µM ENZA for 12 hours compared to parental LNCaP cells treated with
332 the same conditions. Bar length indicates the normalized enrichment score (NES) and bar colour signifies the
333 statistical significance. D) Schematic summarising observations of figures 1-4. During AR driven tumour growth
334 represented by the R1881 dataset, multiple key mechanisms are upregulated including the UPR. In the eventuality
335 of ADT represented by the CSM and ENZA datasets, these processes are reversed/ downregulated. However,
336 removing IRE1 activity in isolation leads to an increase in processes that directly contradict the effects of ADT
337 leading to partial treatment resistance. Created using Biorender.com. E) Cell titer glo assay, fluorescence fold
338 change of treatment resistant LNCaP RB1$^{-/-}$ and LNCaP RB1$^{-/-}$TP53$^{-/-}$ cells treated with 50µM and 100µM ENZA
339 in the presence or absence of 25µM IXA4. Statistical analysis was performed using an ordinary one-way ANOVA
340 with Dunnett's multiple comparisons test and single pooled variance. Each experiment was performed at least in
341 duplicate. Four stars signify adjusted p value of <0.0001.

342

343

344

345

346

347

348

349

350 ***Integrative analysis of sequencing datasets gives rise to a novel IRE1 activity gene set (IRE1_18)***
351 ***that represents multiple biological functions that drive PCa dynamic disease progression.***

352 We identified PCa-specific, IRE1-responsive genes, labelling them "IRE1 activity induced" (significantly
353 downregulated in ERN1$^{-/-}$) or "IRE1 activity repressed" (significantly upregulated in ERN1$^{-/-}$)
354 (Supplementary Table S17) by integrating results from multiple RNAseq datasets (41 genes, Figure 5A,
355 Supplementary Figure S5A). This gene set represents key PCa progression biological functions
356 comprising cell adhesion, motility, cell cycle, DNA repair, and protein trafficking, while correlating with
357 changes in cell cycle, EMT and metabolism PCa-associated genes (Supplementary Figure S5B, S5C
358 Supplementary Tables S18, S19).

359 We then investigated whether this approach yielded direct IRE1 activity targets within IRE1_41. XBP1
360 ChIP-seq analysis of LNCaP cells treated with R1881 for 24 hours showed that multiple "IRE1 induced"
361 genes displayed XBP1 peaks directly in their promoter (Supplementary figure S6A). RNA silencing of
362 IRE1 or XBP1 downregulated "IRE1 induced" and upregulated "IRE1 repressed" genes (Figure 5B). The
363 same effect was achieved by MKC8866 (IRE1 RNAse inhibitor) treatment (Figure 5C) while the opposite
364 with IXA4 (Figure 5D) or thapsigargin (Figure 5E) treatment. To ensure that these effects were not cell-
365 line specific, the same analyses were also performed in C4-2B cells and similar effects were observed
366 (Supplementary Figure S6B-F).

367 We then assessed AR/IRE1 activity canonical genes in a compendium of integrated RNA-seq datasets
368 comprising over 1000 clinical tissue samples ranging from benign to NEPC (Prostate Cancer Atlas[32]).
369 We also applied signatures of basal and luminal epithelial cell lineages to these data[33]. The basal
370 signature expression mirrored AR repression while AR induction signature mirrored both the luminal
371 signature and IRE1 induction (Supplementary Figure S7A). We then investigated the new IRE1 activity
372 gene set in these samples, documenting the pseudo-time disease progression trajectory attributed to
373 the "IRE1 induced" and "IRE1 repressed" components (Supplementary Figure S7B). A particular

374 subgroup of genes (*MYL9, THBS1, TRANK1, TGM2, RND3 and CHI3L2)* was consistently
375 downregulated compared to normal tissue (Figure 5F).

376 The IRE1 activity gene set was derived from significant differential expression upon IRE1 perturbation
377 in pre-clinical cancer models. However, some of these genes under IRE1 control could be enriched in
378 other cell types. ST[34] data expression maps encompassing benign, stromal and localised grade group
379 (GG)4 histological annotations, showed that signal was concentrated exclusively in the benign/stromal
380 region; either ubiquitously (*RND3, MYL9*) or in very specific benign sub-compartments (*TGM2*). Some
381 genes were tumour-region specific with the majority of "IRE1 induced" genes (e.g., SNHG8) enriched
382 in GG4 (Figure 5G, Supplementary Figure S7C) while also significantly upregulated in tumoral compared
383 to normal epithelial cells in two separate SC datasets of primary tumour and matched benign samples
384 [19,35] (Supplementary Figure S7D). Recapitulating multicellularity, IRE1 activity could distinguish
385 between SC luminal, basal and club cells.

386 Using these ST and SC sequencing filters, we selected refined IRE1 activity gene sets that identify
387 tumoral and peri-tumoral regions. We applied further redundancy filters (Supplementary Figure S8A)
388 based on expression patterns across clinical designations in the Prostate Cancer Atlas, which
389 segregated AR+ from AR- tumours (*VSTM2L, TRPM8*) or displayed increasing gradient expression
390 (*HELLS, XRCC2, BRCA2*) and decreasing gradient expression patterns (*SNHG8, EEF2*). We thus
391 produced shorter filtered gene sets that were applied to segregating mCRPC (Labrecque) or localised
392 (PRAD TCGA) cohorts (Supplementary Figure S8B, S8C). The final set of 18 genes (labelled IRE1_18)
393 still represented diverse biological functions and cell types (Supplementary Figure S8D, Supplementary
394 Table S20). Both "IRE1 induced" and "IRE1 repressed" parts of IRE1_18 differentiated AR-positive and
395 AR-negative tumours from normal tissue (Supplementary Figure S8E, Supplementary Tables S21, S22).

396 We next assessed how IRE1_18 clustered metastatic (Labrecque) and localised (TCGA-PRAD) patient
397 cohorts compared to AR signatures. The "IRE1 induced" part of IRE1_18 was significantly reduced in
398 NEPC/DNPC mCRPC samples compared to AR positive mCRPC samples showing similarity to a 10-

399 gene AR induced signature[36]. Within the AR positive mCRPC groups, “IRE1 repressed” subset of
400 IRE1_18 most negatively enriched in the amphicrine group whilst most positively enriched in the AR-
401 low group (Figure 5H). In the TCGA-PRAD, the “IRE1 induced” gene set was significantly decreased in
402 primary disease compared to benign tissue. The exact opposite was observed in the “IRE1 repressed”
403 gene set (Figure 5I). Encouragingly, both parts of IRE1_18 segregated AR-positive from AR-negative
404 tumours and primary tumours from benign tissue in a manner comparable to AR activity signatures..

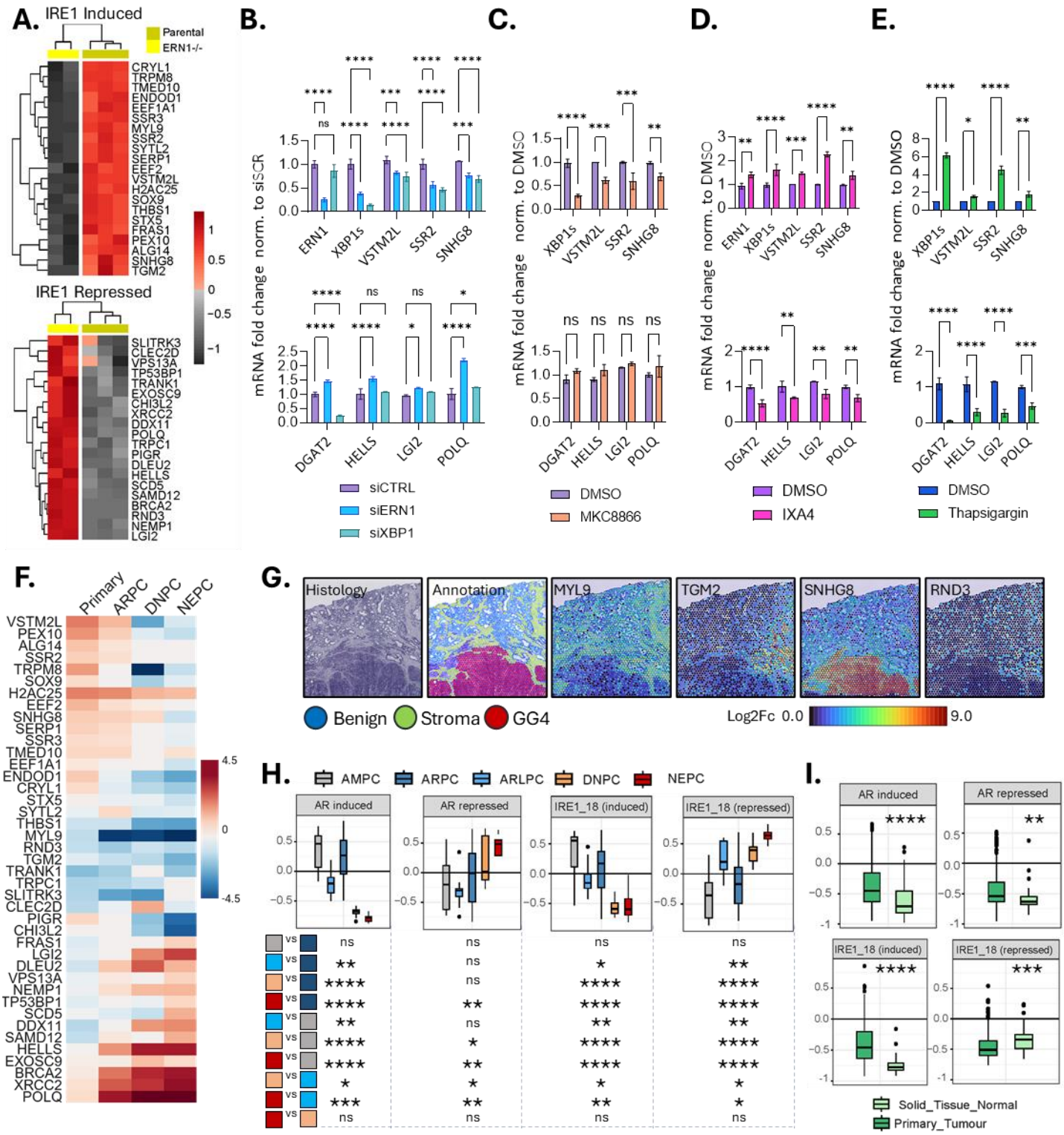
405 Since IRE1_18 represents different cell types and biological functions it should have specific prognostic
406 power in different clinical settings. Indeed, splitting IRE1_18 in two subgroups comprising both
407 “induced” and “repressed” genes significantly prognosticated both mCRPC (Figure 5I) and localised
408 disease (Figure 5J) for overall and progression free survival respectively. In the PCF/SU2C cohort,
409 patients with significantly differentially expressed mRNA of IRE1_18 had worse probability of overall
410 survival compared to patients without (Supplementary Table S23). In the TCGA/PRAD cohort, patients
411 with significantly differentially expressed mRNA of IRE1_18 showed significantly worse probability of
412 progression free survival, compared to patients without (Supplementary Table S24).

413

414

415

Figure 5. ***Multiple sequencing datasets give rise to a novel IRE1 activity gene set that represents multiple key biological functions and dynamic disease progression.***



417 **Figure 5. Multiple sequencing datasets give rise to a novel IRE1 activity gene set that represents key PCa**
418 **biological functions and dynamic disease progression.** A) Gene expression heatmaps of genes in the IRE1
419 activity signature defined as "IRE1 induced" (top) when downregulated in LNCaP ERN1$^{-/-}$ cells (yellow) compared
420 to parental (mustard) and as "IRE1 repressed" (bottom) when upregulated in LNCaP ERN1$^{-/-}$ cells compared to
421 parental. mRNA fold change of *ERN1, XBP1s, DGAT2, VSTM2L, SSR2, SNHG8, HELLS, LGI2* in LNCaP parental
422 cells treated with B) siRNA against ERN1 or XBP1 for 48 hours normalised to scramble control, C) 5µM MKC8866
423 for 5 hours normalised to DMSO, D) 25µM IXA4 for 5 hours normalised to DMSO and E) 100nM Thapsigargin for
424 5 hours normalised to DMSO. Statistical analysis was performed using an ordinary two-way ANOVA and Sidak's
425 multiple comparisons test with single pooled variance. F) Heatmap of the IRE1 activity gene set expression fold
426 change in primary, AR positive CRPC, NEPC and DNPC patient cohorts compared to normal tissue (Prostate
427 Cancer Atlas). Fold changes calculated using "Normal" as control in the DESeq2 Prostate Cancer Atlas function.
428 Downregulated genes in blue, upregulated in red. G) Histological annotations of 6.5 mm^2 tissue slide where benign
429 (blue), stroma (green) and grade group 4 (GG4) cribiform tumour is represented (red). Loupe browser generated
430 distribution maps of individual genes (*MYL9, TGM2, RND3, SNHG8*) across these histological designations
431 showing either generalised or specific localisation to benign or tumoral tissue. Box plots of GSVA enrichment
432 scores for AR activity and IRE1_18 gene sets, benchmarked in H) mCRPC (Labrecque) and I) localised disease
433 (TCGA PRAD) datasets. Patient populations represented include AR positive (dark blue), AR low (light blue),
434 amphicrine (grey), NEPC (red), double negative (DNPC, orange), normal tissue (light green) and primary tumour
435 (dark green). Statistical testing was performed using limma (lmFit function). Shortened versions of the IRE1
436 activity signature ((*VPS13A, EEF2, TRPM8, POLQ, TGM2, THBS1, LGI2, STX5, TRPC1*), (*VSTM2L, HELLS, CHI3L2,*
437 *SNHG8, DDX11, SLITRK3, CRYL1, NEMP1, MYL9*)) tested for prognostication power showing significance in I)
438 probability of overall survival and J) progression free survival in an mCRPC (PCF/SU2C) and primary disease
439 (PRAD TCGA) cohort respectively. Altered expression defined as J) significant log-transformed mRNA z-scores
440 compared to expression distribution of all samples (RNA Seq FPKM) and K) significant z-scores of tumour
441 samples compared to the expression distribution of all log-transformed mRNA expression of adjacent normal
442 samples in the cohort. Statistical analysis was performed using a Logrank (Mantel-Cox test). Each experiment
443 was performed at least in duplicate. Four stars signify adjusted p value of <0.0001. Heatmap in F) was created
444 using Biorender.com.

445

446 ***Novel IRE1 activity gene set distinguishes between histological grades, between benign and***
447 ***tumour-bearing virtual biopsies and correlates with multiple clinical attributes.***

448 IRE1_18 is a proxy of multiple biological functions, cell types and treatment responses and identifies
449 or prognosticates both localised and metastatic disease. Recently, we described how gene expression
450 profiling of virtual tumour and benign biopsies reveals differences between grade groups and tumour
451 clones not reflected in bulk analyses. Therefore, bulk analyses of whole biopsies or tumour-only areas,
452 as used in clinical practice, may provide inaccurate assessments[37] as tissue type, tumour grade and
453 clonal composition all influence gene expression.

454 We then tested if IRE1_18 could be confounded by heterogeneity. Pseudo bulk analysis of tumour
455 bearing virtual biopsy using IRE1_18 showed transcriptomic changes attributed to both tumoral or peri-
456 tumoral tissue. Upregulation of *CHI3L2*, *TGM2* and *TRPM8*, *SNHG8*, *VSTM2L* (Supplementary Figure
457 S9A, B) reflected peri-tumoral (*TGM2*, *CHI3L2*) or tumoral (*SNHG8*, *TRPM8, VSTM2L*) regions
458 (Supplementary Figure S9C). We proceeded to determine if IRE1_18 could identify high grade tumour,
459 profile a bulk virtual biopsy and place this profile in an exact subgroup of patients within a larger external
460 clinical cohort. IRE1_18 genes separating GG4 from local benign in a 6.5mm[34], showed the same
461 pattern in a virtual biopsy containing both GG2 and GG4. Since this patient solely underwent radical
462 prostatectomy, we expected their virtual biopsy to correspond to primary disease and indeed, the
463 expression profile of IRE1_18 corresponds to the primary disease subgroup of the Prostate Cancer
464 Atlas (Figure 6A).

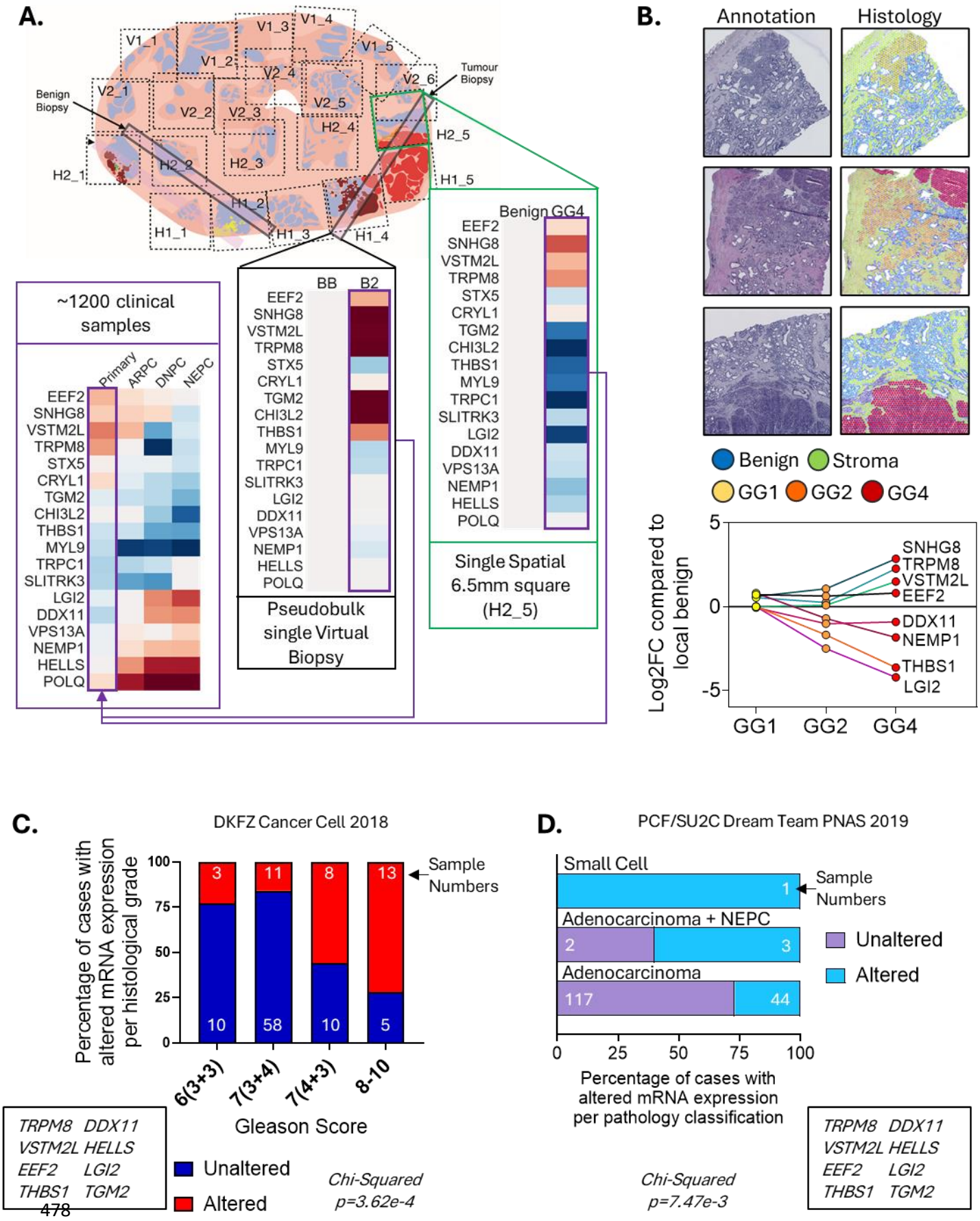
465 Thereafter we correlated IRE1_18 expression with increasing grade group in this radical prostatectomy
466 using GG1, GG2 and GG4 spots to recapitulate the full spectrum of multifocality and heterogeneity. The
467 "IRE1 induced" component (*TRPM8, EEF2, SNHG8, VSTM2L*) increased whilst the "IRE1 repressed"
468 component (*THBS1, DDX11, HELLS, NEMP1, LGI2*) decreased with tumour grade (Figure 6B).

469 To expand on this correlation, we explored data from 118 samples (DKFZ, Cancer Cell cohort) and
470 observed significant differential expression changes in IRE1_18 genes with increasing Gleason score

471 (Figure 6C). Furthermore, in the PCF/SU2C cohort, significant differential expression of these genes
472 correlated with worsening pathology classification (NEPC/small cell disease) (Figure 6D). Indeed,
473 change in mRNA status of these genes significantly correlated with clinical attributes in multiple
474 datasets including NEPC scoring, BCR status and mutation count (Supplementary Table S25).
475 Therefore, IRE1 activity represented by IRE1_18 showcases the importance of stress response biology
476 in PCa pathophysiology, progression and treatment response.

477

Figure 6. ***IRE1_ 18 may act as a proxy of multiple clinical attributes.***



478

479 ***Figure 6. IRE1 activity may act as proxy of multiple clinical attributes.*** A) Schematic of whole radical
480 prostatectomy axial disc with coded dashed squares signifying corresponding spatial transcriptomic 6.5mm2
481 squares. As previously described three rectangles show virtual biopsy placements corresponding to completely
482 benign or tumour bearing 1.2mm x 18 mm standard transurethral (TRUS) sampling. From right to left, log fold
483 change expression of the IRE1 activity signature genes is represented in heatmaps locally comparing GG4 to
484 benign; comparing a benign virtual biopsy to a tumour bearing virtual biopsy carrying GG2 and GG4; and comparing
485 primary, AR positive CRPC, DNPC and NEPC bulk samples to normal tissue. B) Histological annotations (left)
486 from three separate 6.5 mm^2 squares from the same radical prostatectomy representing multifocality and
487 heterogeneity. Grade groups (GG) 1 (yellow), 2 (orange) and 4 (red) can be seen. On the right, log fold change in
488 expression of "IRE1 induced" (*EEF2, TRPM8, SNHG8, VSTM2L*) and "IRE1 repressed" (*THBS1, LGI2, DDX11,*
489 *HELLS, NEMP1*) genes when comparing increasing GG to each corresponding benign region. Each locally
490 distinguishing comparison involves 331 GG1 spots, 938 GG2 spots, 1169 GG4 spots and 942, 703 and 1078
491 benign spots respectively. C) Correlation of expression pattern of the IRE1 activity gene set to radical
492 prostatectomy Gleason score in the DKFZ Cancer Cell 2018 cohort. Statistical analysis was performed using Chi-
493 squared test. Altered (red) status is defined as differentially significant z-score expression in tumour samples
494 compared to the expression distribution of all log-transformed mRNA expression in adjacent normal samples in
495 the cohort. D) Correlation of expression pattern of the IRE1 activity signature to pathology classification including
496 adenocarcinoma (green), small cell (orange) and adenocarcinoma with NEPC features (purple). Statistical
497 analysis was performed using Chi-squared test. Altered gene expression is defined as differentially significant log-
498 transformed mRNA z-scores compared to mRNA signal from all samples in the cohort. Heatmaps in A) were
499 created using Biorender.com.

500

501

502

503

504

505 ## Discussion

506 We have previously reported that the IRE1-XBP1 axis of the unfolded protein response is regulated by
507 androgen receptor activity and that inhibiting IRE1 restricts tumorigenesis in androgen-dependent
508 models of prostate cancer. Sequencing well-annotated clinical samples has revealed that whilst AR
509 activity is maintained in many cases as PCa progresses to metastasis, AR variants and target genes that
510 support progression change. In addition, NEPC/DNPC PCa subtypes do not seem to rely on AR function
511 or expression. By contrast, the UPR/IRE1 axis is capable of adapting as these disease states emerge.
512 Given that the UPR-IRE1 axis is a feature of a wide range of cell types we have sought to determine how
513 IRE1 activity relates to the emergence of acquired treatment resistance and to a spectrum of PCa sub-
514 types or cell lineages. To assign IRE1 activity to cell lineages and PCa subtypes we derived a PCa-
515 specific gene set (IRE1_18) by initially relying on data derived from PCa cell-lines and pre-clinical
516 models. When refining IRE1_18 we selected genes that could discriminate between PCa sub-types in
517 patient datasets and to prognosticate both localized and metastatic disease. De-convoluting IRE1_18
518 spatially, we determined that in clinical samples these genes are associated with a variety of cell types
519 and biological functions likely to reflect niches in which treatment resistant disease can emerge.

520 Our cell-line/pre-clinical models cannot capture the cellular heterogeneity or biological complexity of
521 these niches, but a potentially significant finding was that knocking out IRE1 in an androgen-dependent
522 PCa cell-line results in the increased expression of lineage markers associated with club-like epithelial
523 progenitor (*MMP7, PIGR, CP, LCN2*), and lineage reprogramming (*SOX2, EZH2, POU3F2*). IRE1 activity
524 may impact cellular differentiation being associated with the acquisition of cancer stem cell
525 characteristics in glioblastoma through XBP1s-driven repression of mir-148, which de-represses
526 stemness TFs such as SOX2[38]. In PCa, *SOX2* has been associated with stemness and treatment
527 resistance arising from knocking out TP53 and RB1 in androgen-dependent pre-clinical models[39].

528 Intriguingly, knocking out IRE1 also leads to transcriptional changes resembling RB1 loss and oncogenic
529 TF coregulation[40] while correlating with AR loss and NEPC genotypes in mCRPC datasets. Importantly,

530 pharmacologically increasing IRE1 activity in RB1 deficient cells sensitised them to physiological and
531 pharmacological AR deprivation.

532 In the lifetime of PCa evolution and treatment, UPR/IRE1 activity has been shown to increase as an
533 early stress-response, adaptive, pro-survival mechanism due to mounting metabolic stress and protein
534 folding demand, parallel to the AR. But what happens when AR activity starts to decrease
535 physiologically, or as a response to therapy? ADT-induced IRE1 signalling ablation may lead to
536 adaptation that pushes a subset of tumours towards an AR+ CRPC or NEPC/DNPC phenotype, thus
537 placing IRE1 activity at the centre of acquired treatment resistance.

538 Our study thus introduces two major novel themes for further investigation. Firstly, in more effectively
539 stratifying patients for ADT and perhaps in the future, modulators of IRE1/UPR activity (ONC201[41],
540 NXP800[42]) by mapping IRE1 responses to AR antagonism (STAMPEDE), to metastatic site (Hartwig
541 Medical Foundation) and to RB1-loss, BRCA2-amplification phenotypes lacking somatic RB1/BRCA2
542 changes (TCGA, SU2C). Secondly, for further developing models of lineage plasticity that can capture
543 transition states between high AR activity, differentiated luminal epithelial cells and low AR activity,
544 club-like epithelial progenitors aiming to predict the emergence of AR-negative basal or stem-like cells.
545 IRE1 activity proxies (e.g., IRE1_18) representing multicellular tumoral niches are not dependent on
546 accurate tumour sampling and can identify peri-tumoral tissue in false negative biopsies or inform risk
547 stratification in histologically benign tissue with pre-tumoral genotypes. Indeed, clinically approved
548 genomic tests (Decipher) contain genes attributed to multiple cell types and thus significantly associate
549 with risk of metastasis and PCa specific mortality[43].

550 The majority of significant work in this area of lineage plasticity has thus far focused on transcription
551 factors (*ACSL1, POU3F2*) and chromatin-modifying enzymes (*EZH2*). Against this backdrop it will be
552 interesting to determine whether changes in IRE1 activity create permissive states in which these more
553 established drivers of lineage plasticity can function more effectively, or whether IRE1's contribution is
554 mechanistically independent.

555 ## Summary

556 In conclusion, we explored PCa dynamic disease progression through the homeostatic mechanism of
557 the UPR. By focusing on the UPR transducer IRE1, we produced an IRE1 activity gene set that can act
558 as a proxy of AR activity, stress response, histological grade, cell type, clinical designation and
559 potentially treatment response. IRE1 activity is commonly thought to be a driver of tumorigenesis. We
560 here, report for the first time that adaptation to sustained stress and suppression of the IRE1 axis may
561 induce lineage characteristics that mimic a club-like pre-cursor to RB1 loss and treatment resistant
562 disease. This has major implications for IRE1 modulation in treatment resistant PCa as a differentiation
563 therapy and for use of IRE1 activity as early identifier of lethal disease.

564

565

566

567

568

569

570

571

572

573

574

575 ## References

576 1. Doultsinos, D. & Mills, I. The role of the androgen receptor as a driver and mitigator of cellular
577 stress. *J Mol Endocrinol* **1**, (2020).

578 2. Rawla, P. Epidemiology of Prostate Cancer. *World J Oncol* **10**, 63–89 (2019).

579 3. Almanza, A. *et al.* Endoplasmic reticulum stress signalling – from basic mechanisms to clinical
580 applications. *FEBS Journal* vol. 286 241–278 Preprint at https://doi.org/10.1111/febs.14608
581 (2019).

582 4. Almanza, A. *et al.* Regulated IRE1α-dependent decay (RIDD)-mediated reprograming of lipid
583 metabolism in cancer. *Nature Communications 2022 13:1* **13**, 1–13 (2022).

584 5. Bright, M. D., Itzhak, D. N., Wardell, C. P., Morgan, G. J. & Davies, F. E. Cleavage of BLOC1S1
585 mRNA by IRE1 Is Sequence Specific, Temporally Separate from XBP1 Splicing, and Dispensable
586 for Cell Viability under Acute Endoplasmic Reticulum Stress. *Mol Cell Biol* **35**, 2186–2202
587 (2015).

588 6. Raymundo, D. P. *et al.* Pharmacological Targeting of IRE1 in Cancer. *Trends Cancer* **6**, 1018–
589 1030 (2020).

590 7. Sharma, N. L. *et al.* The Androgen Receptor Induces a Distinct Transcriptional Program in
591 Castration-Resistant Prostate Cancer in Man. *Cancer Cell* **23**, 35–47 (2013).

592 8. Sheng, X. *et al.* IRE1α-XBP1s pathway promotes prostate cancer by activating c-MYC signaling.
593 *Nat Commun* **10**, 323 (2019).

594 9. Sheng, X. *et al.* Divergent androgen regulation of unfolded protein response pathways drives
595 prostate cancer. *EMBO Mol Med* **7**, 788–801 (2015).

596 10. Hamdy, F. C. *et al.* Fifteen-Year Outcomes after Monitoring, Surgery, or Radiotherapy for
597 Prostate Cancer. *New England Journal of Medicine* **388**, 1547–1558 (2023).

598 11. Centenera, M. M. *et al.* A patient-derived explant (PDE) model of hormone-dependent cancer.
599 *Mol Oncol* **12**, 1608 (2018).

600 12. Conteduca, V. *et al.* Clinical features of neuroendocrine prostate cancer. *Eur J Cancer* **121**, 7–
601 18 (2019).

602 13. Stelloo, S. *et al.* Androgen modulation of XBP1 is functionally driving part of the AR
603 transcriptional program. *Endocr Relat Cancer* **27**, 67–79 (2020).

604 14. Lhomond, S. *et al.* Dual IRE1 RNase functions dictate glioblastoma development. *EMBO Mol*
605 *Med* **10**, (2018).

606 15. Mu, P. *et al.* SOX2 promotes lineage plasticity and antiandrogen resistance in TP53-and RB1-
607 deficient prostate cancer. *Science (1979)* **355**, (2017).

608 16. Venkadakrishnan, V. B. *et al.* Lineage-specific canonical and non-canonical activity of EZH2 in
609 advanced prostate cancer subtypes. *Nature Communications 2024 15:1* **15**, 1–15 (2024).

610 17. Miao, C. *et al.* RB1 loss overrides PARP inhibitor sensitivity driven by RNASEH2B loss in
611 prostate cancer. *Sci Adv* **8**, 9794 (2022).

612 18. Flores, M. & Goodrich, D. W. Retinoblastoma Protein Paralogs and Tumor Suppression. *Front*
613 *Genet* **13**, 818719 (2022).

614 19. Hirz, T. *et al.* Dissecting the immune suppressive human prostate tumor microenvironment via
615 integrated single-cell and spatial transcriptomic analyses. *Nature Communications 2023 14:1*
616 **14**, 1–20 (2023).

617 20. Kiviaho, A. *et al.* Single cell and spatial transcriptomics highlight the interaction of club-like
618 cells with immunosuppressive myeloid cells in prostate cancer. *Nature Communications 2024*
619 *15:1* **15**, 1–16 (2024).

620 21. Huang, F. W. *et al.* Club-like cells in proliferative inflammatory atrophy of the prostate. *J Pathol*
621 **261**, 85–95 (2023).

622 22. Song, H. *et al.* Single-cell analysis of human primary prostate cancer reveals the heterogeneity
623 of tumor-associated epithelial cell states. *Nature Communications 2022 13:1* **13**, 1–20 (2022).

624 23. Joseph, D. B. *et al.* 5-Alpha reductase inhibitors induce a prostate luminal to club cell
625 transition in human benign prostatic hyperplasia. *J Pathol* **256**, 427–441 (2022).

626 24. Ku, S. Y. *et al.* Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity,
627 metastasis, and antiandrogen resistance. *Science (1979)* **355**, (2017).

628 25. Mu, P. *et al.* SOX2 promotes lineage plasticity and antiandrogen resistance in TP53-and RB1-
629 deficient prostate cancer. *Science (1979)* **355**, 84–88 (2017).

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

632 27. Logue, S. E. *et al.* Inhibition of IRE1 RNase activity modulates the tumor cell secretome and
633 enhances response to chemotherapy. *Nat Commun* **9**, 3267 (2018).

634 28. Grandjean, J. M. D. *et al.* Pharmacologic IRE1/XBP1s activation confers targeted ER
635 proteostasis reprogramming. *Nature Chemical Biology 2020 16:10* **16**, 1052–1061 (2020).

636 29. Labrecque, M. P. *et al.* Molecular profiling stratifies diverse phenotypes of treatment-
637 refractory metastatic castration-resistant prostate cancer. *J Clin Invest* **129**, 4492 (2019).

638 30. Thangavel, C. *et al.* RB loss promotes prostate cancer metastasis. *Cancer Res* **77**, 982–995
639 (2017).

640 31. He, M. X. *et al.* Transcriptional mediators of treatment resistance in lethal prostate cancer.
641 *Nature Medicine 2021 27:3* **27**, 426–433 (2021).

642 32. Bolis, M. *et al.* Dynamic prostate cancer transcriptome analysis delineates the trajectory to
643 disease progression. *Nature Communications 2021 12:1* **12**, 1–15 (2021).

644 33. Aggarwal, R. *et al.* Prognosis Associated With Luminal and Basal Subtypes of Metastatic
645 Prostate Cancer Supplemental content. *JAMA Oncol* **7**, 1644–1652 (2021).

646 34. Erickson, A. *et al.* Spatially resolved clonal copy number alterations in benign and malignant
647 tissue. *Nature 2022 608:7922* **608**, 360–367 (2022).

648 35. Chen, S. *et al.* Single-cell analysis reveals transcriptomic remodellings in distinct cell types that
649 contribute to human prostate cancer progression. *Nat Cell Biol* **23**, 87–98 (2021).

650 36. Bluemn, E. G. *et al.* Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained
651 through FGF Signaling. *Cancer Cell* **32**, 474 (2017).

652 37. Figiel, S. *et al.* Spatial transcriptomic analysis of virtual prostate biopsy reveals confounding
653 effect of tissue heterogeneity on genomic signatures. *Mol Cancer* **22**, 1–5 (2023).

654 38. Doultsinos, D. *et al.* Constitutive UPRER activation sustains tumor cell differentiation. *bioRxiv*
655 (2019) doi:10.1101/594630.

656 39. Venkadakrishnan, V. B., Yamada, Y., Weng, K., Idahor, O. & Beltran, H. Significance of RB loss in
657 unlocking phenotypic plasticity in advanced cancers. *Mol Cancer Res* **21**, 497 (2023).

658 40. Mandigo, A. C. *et al.* Novel Oncogenic Transcription Factor Cooperation in RB-Deficient
659 Cancer. *Cancer Res* **82**, 221–234 (2022).

660 41. Amoroso, F. *et al.* Modulating the unfolded protein response with ONC201 to impact on
661 radiation response in prostate cancer cells. *Scientific Reports 2021 11:1* **11**, 1–16 (2021).

662 42. Welti, J. *et al.* NXP800 activates the unfolded protein response, altering AR and E2F function
663 to impact castration-resistant prostate cancer growth. *Clin Cancer Res* OF1–OF18 (2025)
664 doi:10.1158/1078-0432.CCR-24-2386/750830/AM/NXP800-ACTIVATES-THE-UNFOLDED-
665 PROTEIN-RESPONSE.

666 43. Feng, F. Y. *et al.* Validation of a 22-Gene Genomic Classifier in Patients With Recurrent Prostate
667 Cancer: An Ancillary Study of the NRG/RTOG 9601 Randomized Clinical Trial. *JAMA Oncol* **7**,
668 544 (2021).

669 44. Doultsinos, D. *et al.* Peptidomimetic-based identification of FDA-approved compounds
670 inhibiting IRE1 activity. *FEBS Journal* **288**, 945–960 (2021).

671 45. Abida, W. *et al.* Genomic correlates of clinical outcome in advanced prostate cancer. *Proc Natl*
672 *Acad Sci U S A* **166**, 11428–11436 (2019).

673 46. Gerhauser, C. *et al.* Molecular Evolution of Early-Onset Prostate Cancer Identifies Molecular
674 Risk Markers and Clinical Trajectories. *Cancer Cell* **34**, 996-1011.e8 (2018).

675 47. Taylor, B. S. *et al.* Integrative genomic profiling of human prostate cancer. *Cancer Cell* **18**, 11–
676 22 (2010).

677 48. Miller, H. E. & Bishop, A. J. R. Correlation AnalyzeR: functional predictions from gene co-
678 expression correlations. *BMC Bioinformatics* **22**, (2021).

679 49. Dunning, M. J. *et al.* Mining Human Prostate Cancer Datasets: The “camcAPP” Shiny App.
680 *EBioMedicine* **17**, 5 (2017).

681

682

683

684

Acknowledgments & Contributions

Dr Doultsinos and Prof. Mills designed the study, analysed data and wrote the manuscript. Dr Doultsinos carried out all experiments presented (qPCR, Western Blot, growth assays, viability assays, RNAseq) supported by Ms Abusamra, Ms Bridges and Prof. Hester. Ms Tomljanovic (bulk preclinical and clinical datasets), Dr Pilalis (bulk preclinical datasets), Dr Parmentier (single cell datasets), Dr Figiel (spatial datasets), Dr Zekri (ChIP-seq) and Dr Doultsinos (bulk clinical/preclinical, spatial datasets) carried out data analysis. Dr Leach and Prof. Bevan generated and provided PDEs. Prof. Chatziioannou, Dr Lamb, Dr Le Magnen and Prof. Zwart provided study design input and access to datasets and samples.

Special thanks to:

Dr Ursula Arndt for supporting sequencing experiments at the Granta Park Illumina Accelerator site. Dr Eric Chevet, Dr Tony Avril and Dr Diana Pelizarri (INSERM U1242) for providing IRE1 mutant constructs (Q780*) and MKC8866. Professor Nelson (Fred Hutchinson Cancer Research Centre) for providing LNCaP RB1 knockout and LNCaP RB1/TP53 knockout cell lines. Professor Zwart (NKI) for providing LNCaP ERN1 knockout cell lines. Dr Adam Sharp (ICR), Professor Scott Dehm (University of Minnesota), Dr Danielle Fairbrass (University of Oxford) and Dr Anette Magnussen (University of Oxford) for reading and commenting on the manuscript.

Funding

This study was supported by: a Prostate Cancer Foundation- CRIS Cancer Foundation Young Investigator Award, awarded to Dr Doultsinos; Prof. Mills is supported by the John Black Charitable Foundation; an Illumina Accelerator Sequencing Grant awarded to Dr Doultsinos, Dr Pilalis and Prof Chatziioannou.

Conflicts of interest: AC, EP are co-founders of e-NIOS Applications PC (https://e-nios.com/).

709 # Materials & Methods

710

711 **Cell lines & Culture Conditions**

712 LNCaP and C4-2B cells were purchased from ATCC and cultured in RPMI supplemented with
713 10% foetal bovine serum (FBS) or DMEM supplemented with 10% foetal bovine serum (FBS)
714 respectively. LNCaP knocked out for RB1 or knocked out for both RB1 and TP53 were kindly
715 donated by Professor Peter Nelson (Fred Hutchinson Cancer Centre) and were cultured in RPMI
716 supplemented with 10% foetal bovine serum (FBS). LNCaP knocked out for ERN1 were kindly
717 donated by Professor Wilbert Zwart (The Netherlands Cancer Institute) and were cultured in
718 RPMI supplemented with 10% foetal bovine serum (FBS). No antibiotics were used in the
719 culturing of these cells as they may have the capacity to bind a variety of kinases including IRE1
720 as we previously reported[44]. For androgen deprived media, Charcoal Stripped FBS was used
721 instead of normal FBS to supplement RPMI.

722

723 **Primers and qPCR**

724 RNA was extracted from cell pellets using TRIzol reagent (Fisher Scientific, #12034977)
725 according to manufacturer's instructions. cDNA was reverse transcribed from 2000ng of RNA
726 (measured by Nanodrop) using Maxima Reverse Transcriptase (ThermoFisher Scientific
727 EP0743) and associated reagents (RiboLock-Life Technologles-EO0381; DNTPs-Thermo
728 Scientific-18427013; Random Hexamer-Life Technologies-SO142). qPCR was executed on a
729 QuantStudio5 (Applied Biosystems) using Power SYBR Green PCR master mix (ThermoFisher
730 Scientific) and results were analysed using the Design & Analysis Software 2.8.3 (Applied
731 Biosystems).

Gene	Direction	Sequence	Manufacturer
DGAT2	Forward	GGCTGGTGTTTGACTGGAA	Eurofins Genomics
	Reverse	TCAGCAGGTTGTGTGTCTTC	
ACTIN	Forward	AGAGCTACGAGCTGCCTGAC	
	Reverse	AGCACTGTGTTGGCGTACAG	
XBP1s	Forward	TGCTGAGTCCGCAGCAGGTG	
	Reverse	GCTGGCAGGCTCTGGGGAAG	
ERN1	Forward	GCCACCCTGCAAGAGTATGT	
	Reverse	ATGTTGAGGGAGTGGAGGTG	
POLQ	Forward	CCTGTGGCATCTCCTTGGAGCA	
	Reverse	AATCCCTTGGCTGGTCTCCATC	
KLK3	Forward	CGTGACGTGGATTGGTGC	
	Reverse	ACTGCCCTGCCACGAGAG	
HELLS	Forward	GATTTTGGATCGAATGCTGCCAG	
	Reverse	ATGGACCCATCAAGCCTGCTGA	
AR	Forward	AAGACCTGCCTGATCTGTGG	
	Reverse	CGAAGACGACAAGATGGACA	
CAMKK2	Forward	TGAAGACCAGGCCCGTTTCTACTT	
	Reverse	TGGAAGGTTTGATGTCACGGTGGA	
BRCA2	Forward	AATGTCAGACAAGCTCAAAG	Merck
	Reverse	TCATGTATTTTTCAGGTGGC	

RB1	Forward	GACCCAGAAGCCATTGAAATCT	Eurofins Genomics
	Reverse	GGTGTGCTGGAAAAGGGTCC	
SSR2	Forward	CTTCACCTCGGCAACAATTACT	Sigma Kicqstart
	Reverse	GGGGAGAATCGC CTGTCAAAC	
SOX2	Forward	GAGCTTTGCAGGAAGTTTGC	Eurofins Genomics
	Reverse	GCAAGAAGCCTCTCCTTGAA	
POU3F2	Forward	ACACTGACCGATCTCCACGCAGTA	
	Reverse	GAGGGTGTGGGACCCTAAATATGAC	
EZH2	Forward	AATAAAAGCGATGGCGATTGG	
	Reverse	TCCACTGCCTTCTGAGTCC	
ATF4	Forward	GTTCTCCAGCGACAAGGCTA	
	Reverse	ATCCTGCTTGCTGTTGTTGG	
eIF2α	Forward	GCCTTTCTGTCCTCAGTCAAGC	
	Reverse	CTCTCCATCTGCCTCGCAAGTT	
ATF6	Forward	CTCCGAGATCAGCAGAGGAA	
	Reverse	AATGACTCAGGGATGGTGCT	
LGI2	Forward	TGGAGAACTGCATGGTGCTGGA	
	Reverse	CTTACAGCCCACGATGGACTGA	
VSTM2L	Forward	GATCCAGTGGTGGTATGTACGG	
	Reverse	ACACTTATTTTGGTTGCCTCCTTC	
SNHG8	Forward	AAGTTTACAAGCATGCGCGG	
	Reverse	TCAAACTGACGGTTCTCGGG	

732

733 **Drug Treatments**

734 Cells were seeded in 96, 24 or 6 well plate format depending on the assay (96 well plate for
735 viability, 96 or 24 well plate for growth, 6 well plate for RNA extraction) and drugs were added
736 using the range of concentrations below.

Compound	Manufacturer	Lot number	Concentrations
IXA4	Cayman Chemical	0653659-8	10 - 25 µM
MKC8866	Cayman Chemical	0664887-3	5 µM
Thapsigargin	Sigma	0000120608	100 nM
Tunicamycin	Sigma	079M4004V	500 nM
Enzalutamide	Selleck Chemicals	MDV3100	10 µM

737

738 **Cell Titer Glo sensitivity assays**

739 Cells were seeded into 96-well plates and incubated at 37 °C for 24 hours. They were then
740 treated with drug combinations for a further 12 hours. Percentage survival was measured
741 photometrically using the CellTiter-Glo 2.0 viability assay (Promega) in a microplate reader
742 (FLUOstar OMEGA 415-2812). Survival was normalized to the control treatment.

743 **Growth assays**

744 Cells were seeded into 96-well plates and incubated at 37 °C for 24 hours. Complete media
745 change followed to either full (FM) or charcoal stripped media (CSM) for 120 hours. Cell growth
746 was monitored using Incucyte technology taking confluence measurements every 24 hours.
747 Growth was normalised to the initial measurement at time 0.

748 **Transfection assays and RNA silencing.**

749 For stable expression of IRE1 mutant Q780*, LNCaP ERN1$^{-/-}$ cells were transfected with the
750 Q780* construct[14] using LTX Plus reagent (ThermoFisher Scientific, 15338100) according to
751 the manufacturer's instructions. Successfully transfected cells were selected by cell sorting
752 for GFP expression using a BD FACSAriaIII cell sorter. Thereafter, constant GFP expression was
753 monitored using the ZOE™ Fluorescent Cell Imager. Lipofectamine RNAiMax transfection

754 reagent (Thermofisher Scientific, 13778150) was used for transfecting siRNAs. Cells were
755 incubated at 37 °C for 48 hours before extracting RNA for downstream analysis.

siRNA target	Sequence		Manufacturer
	Sense	Antisense	
XBP1	CCUGCCUACUGGAUGCUUAtt	UAAGCAUCCAGUAGGCAGGaa	Ambion
IRE1	GAAACUUCCUUUUACCAUCtt	GAUGGUAAAAGGAAGUUUCgt	Ambion
Control	Silencer Select Negative Control #1		Ambion

756

757 **Immunoblotting and Antibodies**

758

Antibody	Manufacturer	Catalog number	Incubation dilution
Anti-IRE1 (phospho S724) antibody (rabbit)	Abcam	ab48187	1 in 1000
Anti-XBP1 antibody (rabbit)	Abcam	ab37152	1 in 1000
Anti-CAMKK2 antibody (rabbit)	Sigma-Aldrich	HPA017389	1 in 1000
Anti-alpha tubulin (DM1A) antibody (mouse)	Santa Cruz Animal Health	sc-32293	1 in 5000
Monoclonal beta Actin antibody (mouse)	Abcam	ab6276	1 in 5000
Anti-rabbit IgG, HRP-linked Antibody (goat)	Cell Signaling	#7074	1 in 2000
Anti-mouse IgG, HRP-linked Antibody (horse)	Cell Signaling	#7076	1 in 2000

759

760 **Online Tools/Databases**

761 • cBioportal: RNA sequencing datasets investigated included the Metastatic Prostate
762 Adenocarcinoma (SU2C/PCF Dream Team, PNAS 2019[45]), Prostate Cancer (DKFZ,
763 Cancer Cell 2018[46]), Prostate Adenocarcinoma (MSK, Cancer Cell 2010[47]) and
764 Prostate Adenocarcinoma (TCGA, Firehose Legacy or PanCancer). The comparisons
765 were based on log-transformed mRNA expression z-scores compared to expression
766 distribution of all samples (RNA Seq FPKM) or of tumour samples compared to the
767 expression distribution of all adjacent normal samples in the cohort when available
768 with a z score threshold of +/- 2.0. Altered mRNA was defined as mRNA significantly
769 high or low compared to all or normal samples in each dataset using the z score
770 threshold. Kaplan Meier curves were generated for the Prostate Adenocarcinoma
771 (SU2C/PCF Dream Team, PNAS 2019) and Prostate Adenocarcinoma (TCGA,
772 PanCancer) datasets. Only samples from patients with survival data or progression
773 data were used in the generation of these curves respectively. Samples with any
774 significant alteration in mRNA expression were compared to samples with zero
775 significant changes in the former whilst samples with significant alterations in at least
776 three genes were compared to samples with zero significant changes in the latter.
777
778
779 • Prostate Cancer Atlas[32]: DESeq2 differential expression analysis was carried out
780 foreach of sample groups designated as Primary, AR+ CRPC, NEPC and DNPC
781 compared to the sample group designated as normal tissue. Differentially significant
782 mRNA expression changes (padj</=0.01) were documented to mine for genes of
783 interest representing IRE1 activity. Pseudotime disease progression plots were
784 generated for “IRE1 induced” and “IRE1 repressed” sub-groups of the IRE1 activity
785 gene set using the in-built functionality of the Prostate Cancer Atlas

786 (https://prostatecanceratlas.org/app/home). Further in-built functionality was used to
787 analyse gene set enrichment of our IRE1 activity gene set in the normal, primary, AR+
788 CRPC, NEPC and DNPC samples groups.

789 • Correlation AnalyzeR[48]: The online interface of this tool was used to verify the
790 correlation of the IRE1 activity gene set with other significantly differentially expressed
791 genes in prostate cancer and the biological processes it represents.

792 • CAMCAPP[49]: This tool was used to verify gene by gene correlation of the IRE1 activity
793 gene set with clinical co-variates in three separate clinical cohorts (Stockholm,
794 Cambridge, MSKCC).

795

796 **RNA extraction, library prep and sequencing**

797 RNA was extracted from cell line pellets using the RNeasy Plus Mini kit. Samples were QC'd
798 using Nanodrop and the Qubit Flex High Sensitivity ID kit. Libraries were prepared using the
799 TruSeq Stranded Total RNA with Illumina Ribo-Zero plus rRNA Depletion workflow
800 (https://emea.support.illumina.com/downloads/truseq-stranded-total-rna-with-ribo-zero-
801 plus-rrna-depletion-ref-guide.html) . Index adapters, rRNA depletion kit were acquired in
802 addition to the TruSeq Stranded Total Library Prep Gold kit
803 (https://emea.illumina.com/products/by-type/sequencing-kits/library-prep-kits/truseq-
804 stranded-total-rna.html#tabs-bb5d9c771f-item-01fd0b3d08-order). cDNA library QC was
805 carried out on Tapestation and Libraries were pooled, diluted to ~10 nM, denatured, and further
806 diluted for paired-end sequencing (NovaSeq6000 platform, Illumina, NovaSeq 6000 S2/S4
807 reagent kit v1.5, 300 cycles). A test run prior to full sequencing was executed on a MiSeq
808 system. Bulk paired-end RNA sequencing was performed to obtain at least 20mn reads per
809 sample. Raw data were processed using a standardized computational workflow to obtain
810 gene-level raw counts. FASTQ files were quality-checked using FastQC to assess read quality

811 and adapter contamination. High-quality reads were aligned to the human reference genome
812 (GRCh38, Ensembl release 108) using STAR aligner (Dobin et al., 2013). Gene-level
813 quantification was performed using HTSeq-count (Anders et al., 2015).

814 **RNA Sequencing Data analysis**

815 1. Spatial transcriptomics data were obtained from a previously published dataset derived
816 from radical prostatectomy tissue of a patient diagnosed with multifocal prostate cancer
817 (Nature). Each spatial transcriptomics spot was reviewed by two independent pathologists
818 as part of a consensus approach for pathological annotation.

819 • Virtual Biopsy: The analysis was performed using the same approach as previously
820 described (Molecular Cancer). Gene expression data were extracted from the barcode ID
821 of designed biopsies (LoupeBrowser, v8.0.0) using R (v4.4.0). Libraries were normalised
822 using the spaceranger aggr (v2.0) function, which corrects for batch effects to facilitate the
823 comparison of gene expression data across samples and minimise technical variability. The
824 analysis included seven distinct tissue sections, representing a total of 3721 spatial
825 transcriptomics spots. To maintain data integrity, spots with fewer than 500 Unique
826 Molecular Identifier (UMI) counts were removed. Fold changes were calculated using
827 EdgeR (v3.40.1). as previously described[37].

828 Single Visium 6.5mm squares: The raw sequencing reads were directly processed using
829 Space Ranger and mapped using reference genome GRCh38, release 93 as previously
830 described[34]. Each histological square was treated as a distinct spatial transcriptomic
831 experiment to determine if, in an independent manner, gene sets could be used as
832 distinguishing tools between histological characteristics. Histological annotations (benign,
833 GG1, GG2, GG4, stroma) were compared using either the locally distinguishing or globally
834 distinguishing functions of the Loupe Browser which was also used to generate ST images
835 displayed.

836

837 2. Single Cell

838 PCa tissue data from Chen et al. were accessed via the GEO online portal (GSM4203181) as a
839 raw count matrix. Quality control filtering was applied based on the criteria described in the
840 original publication: cells with fewer than 200 or more than 5,218 detected genes, more than
841 20% mitochondrial content, and fewer than 56 housekeeping genes were discarded. The
842 filtered data were then converted into an SCE object, separated by patient ID, and normalized
843 using a cluster-based normalization factor computed after each sample was first partitioned
844 into biologically relevant clusters using the quickCluster(min.size = 100) function from the
845 scran R package (v1.26.2).

846 Next, to correct for differences in sequencing depth across batches, all SCE objects were
847 rescaled using the MultiBatchNorm(min.mean = 0.1) function from the Batchelor R package
848 (v1.14.1). Epithelial cells were retrieved using the singleR package (v2.0.0), leveraging the
849 BlueprintEncodeData reference accessed via the celldex package (v1.16.0). Using scran
850 (v1.26.2), variance modeling of gene expression was performed for each sample individually
851 with the modelGeneVar() function; results were merged with combineVar(), and the top 1,000
852 most highly variable genes were selected using getTopHVGs() for subsequent analysis. Batch
853 correction was applied using the fastMNN(k = 300, d = 20) function from the Batchelor
854 package.

855 To separate luminal malignant cells from healthy epithelial cell types, clustering was performed
856 on the integrated data using the clusterCells() function from the scran R package (v1.26.2). The
857 clustering was conducted with the Walktrap algorithm, exploring different values of k = 10,
858 which resulted in 14 distinct clusters. All clusters were considered tumor except cluster 7 and
859 14, which were negative of both tumor and luminal markers (e.g AMACR, PCA3, KLK3) and
860 were positive for basal markers such as KRT19, KRT18 and KRT14.

861 PCa tissue data from Hirz et al. were accessed via the GEO online portal (GSE181294) as a raw
862 count matrix. Cells passing quality thresholds and cell type annotation, as defined in the
863 original publication, were retrieved from the annotation table provided by the authors and
864 stored in an SCE object. Similar to the Chen et al. dataset, each sample first underwent cluster-
865 based normalization followed by inter-sample correction to account for sequencing depth
866 differences between batches.

867 3. Bulk

868 The RNA sequencing data (raw counts) of the Labrecque mCRPC patient cohort (accession
869 number GSE126078) and the $RB1^{-/-}/TP53^{-/-}$ cell line dataset (accession number GSE147250)
870 was retrieved from GEO. TCGA data was accessed via the GDC Data Portal
871 (https://portal.gdc.cancer.gov/projects/TCGA-PRAD, TCGA-PRAD.2022-04-27). The Beltran
872 patient cohort FPKM data was obtained from the supplementary material of Beltran et al.,
873 2016.

874 Data visualization and statistical analyses of the bulk RNA sequencing datasets were
875 performed using R statistical software (v4.4.1). Batch effect correction was executed using the
876 sva package (v.3.54.0). Differential gene expression analysis was performed using the DESeq2
877 package (v1.40.2), with subsequent LFC shrinkage using the 'apeglm' estimator. Gene set
878 enrichment analysis (GSEA) was conducted on a subset of the Molecular Signatures Database
879 (MSigDB) v2023.2.Hs using the clusterProfiler package (v4.8.2). Gene Set Variation Analysis
880 (GSVA) was used for single-sample pathway enrichment as implemented in the GSVA package
881 (v1.48.3). Differential expression at the pathway level was assessed using the lmFit function
882 from limma (v3.62.2). For visualization and clustering purposes, the variance stabilizing
883 transformation (VST) was applied using the DESeq2 package. Visualization was performed
884 using ggplot2 (v3.5.0) and heatmaps were generated using the pheatmap package (v1.0.12).
885 The viper R package (v1.4.0) was used for single-sample regulon enrichment and differential

testing of regulon activity levels between sample groups. The regulons of 1695 human transcription factors were previously inferred using ARACNE -AP on an in-house set of mCRPC patient tumors. Unless otherwise specified, the statistical significance threshold across all analyses was set at 0.05 and a Benjamini–Hochberg correction was applied to account for multiple testing.