

SUPPLEMENTAL FIGURES

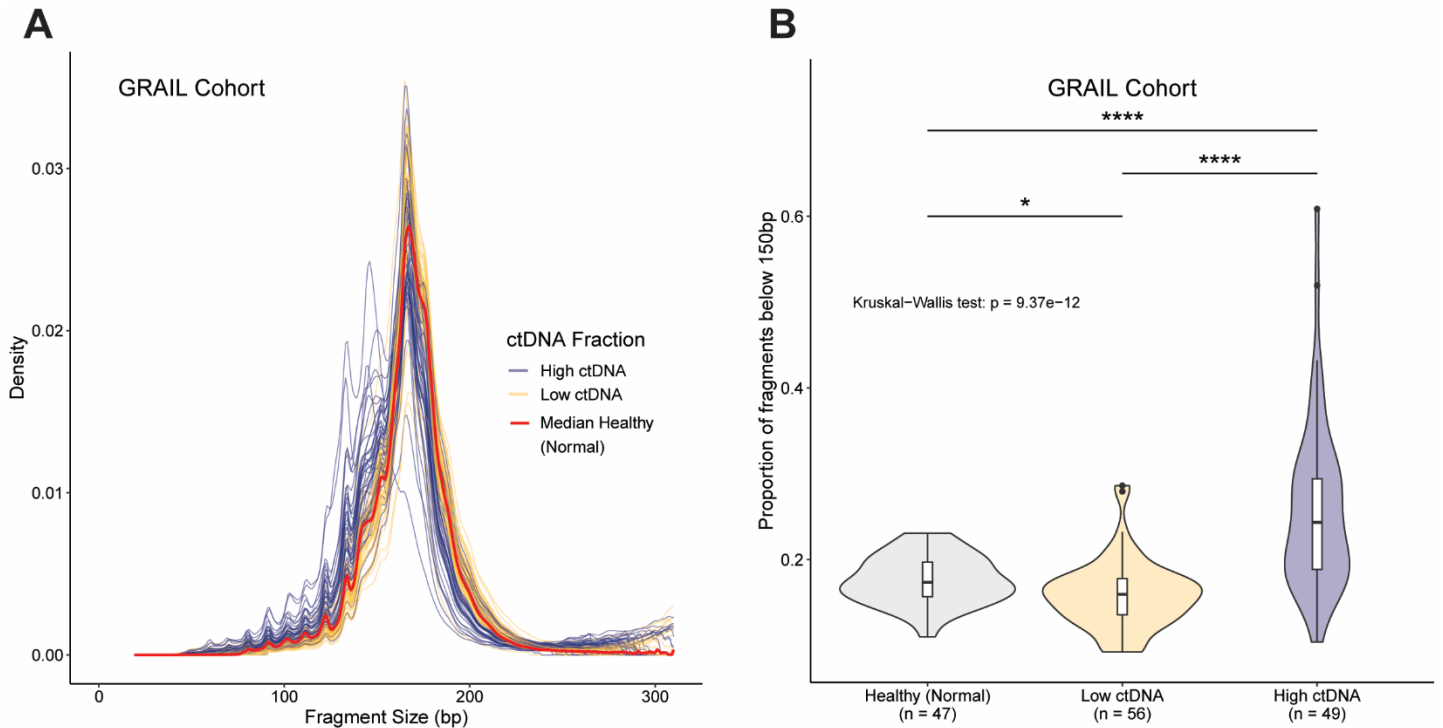


Figure S1: Fragment distribution of GRAIL cohort samples stratified by ctDNA fraction.

(A) Distribution of cfDNA fragments from individual samples colored by low ctDNA (< 10% ctDNA fraction) or high ctDNA (ctDNA fraction \geq 10%). Red line represents the median of all normal healthy samples. (B) Proportion of fragments below 150 bp in healthy, low ctDNA, and high ctDNA samples. A Kruskal-Wallis test was performed to compare all three categories, and a Wilcoxon rank sum test was performed for individual comparisons (* $p < 0.05$; **** $p < 0.0001$)

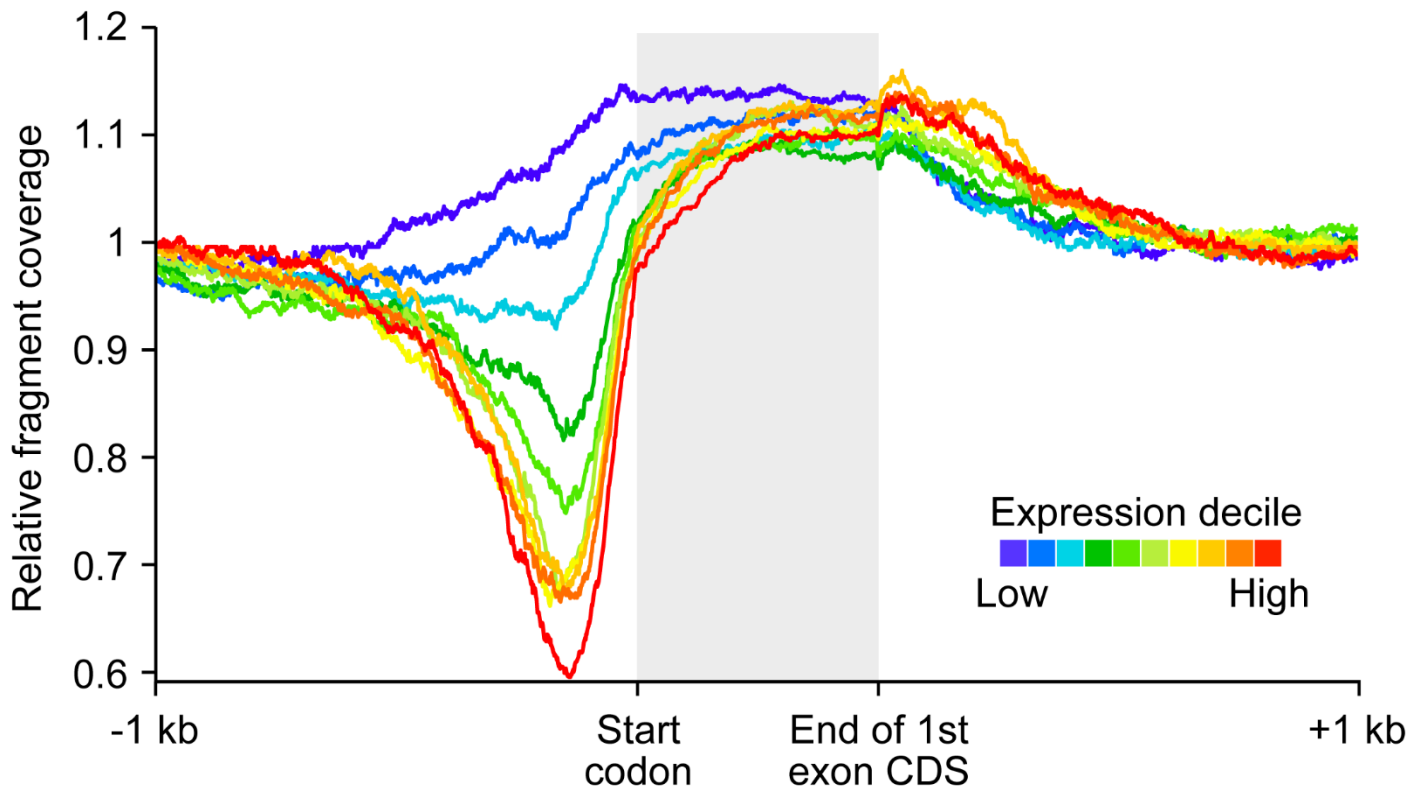


Figure S2: Relative fragment coverage in first coding exon by gene expression decile.

Average plasma cell-free DNA fragment coverage near the exon 1 coding sequence (CDS) of 11748 genes annotated in MANE version 0.93, calculated across 41 whole genome sequenced ctDNA-positive samples from the NCT02125357 trial (Herberts et al. Nature 2022). Genes were separated into ten quantile groups based on their average expression in prostate cancer tissue samples. Fragment coverage is normalized relative to 1kb distant flanks. Only multi-exon genes with a CDS containing exon 1 were included in the analysis. Gene orientation and exon 1 CDS length were normalized between the genes for visualization. One kilobase of upstream and downstream flanking region is also shown (without normalization).

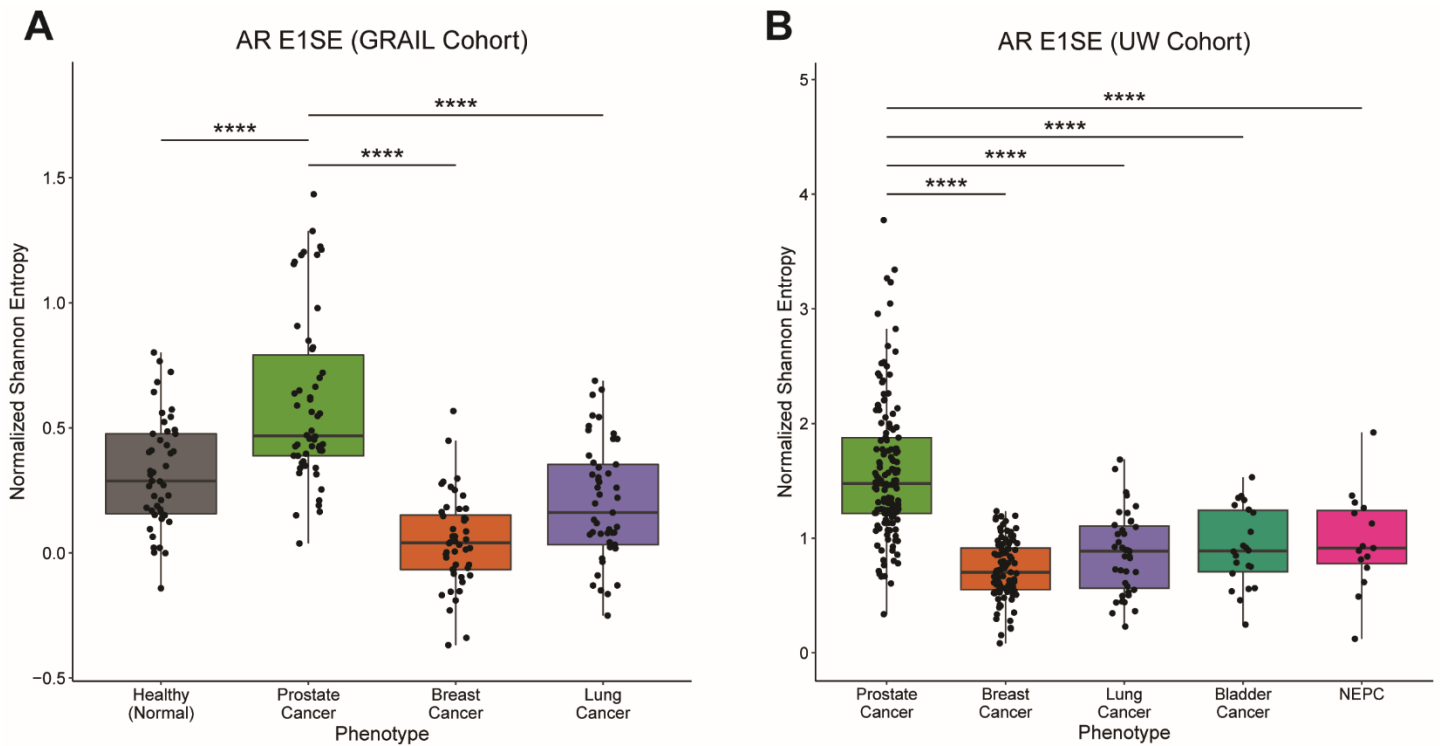


Figure S3: Exon 1 Shannon entropy of the AR by cancer type.

Normalized Shannon entropy was calculated for the first coding exon of the androgen receptor gene (AR) for all samples in the GRAIL cohort (A) and UW cohort (B). AR E1SE displays significantly higher normalized Shannon entropy in prostate cancer samples compared to other cancer types and healthy normal samples. Two-sided Student's t-test was used for significance testing (**** $p < 0.0001$).

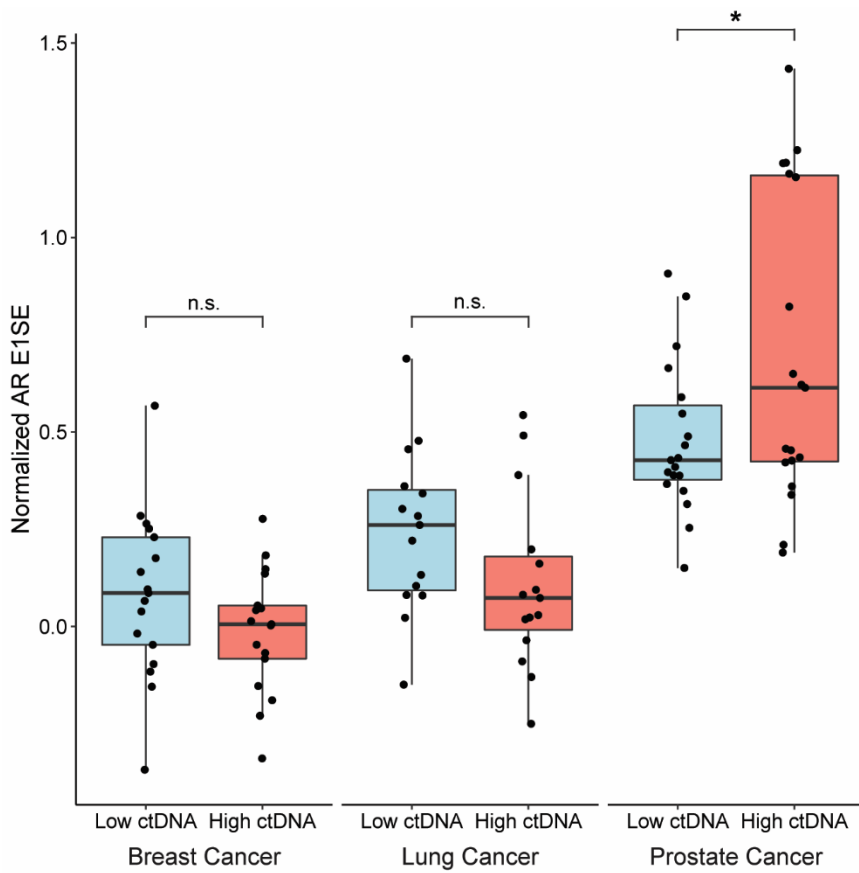


Figure S4: Normalized AR Shannon entropy stratified by ctDNA fraction

Within each cancer type, samples were stratified into low and high ctDNA fraction using the median ctDNA fraction as the cutoff. Normalized Shannon entropy at the first coding exon of AR was calculated and plotted by cancer type and ctDNA level. High ctDNA fraction prostate cancer samples were found to have significantly higher AR E1SE compared to low ctDNA fraction prostate cancer samples only. Two-sided Student's t-test was used for significance testing (* $p < 0.05$; n.s. – non significant)

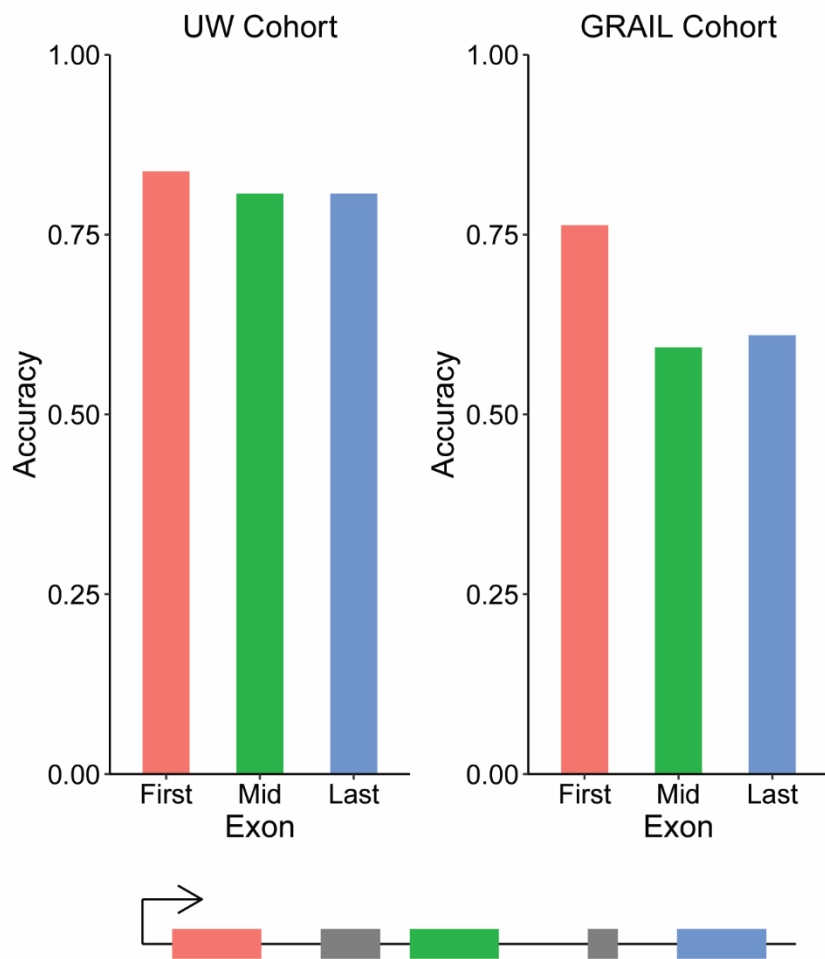


Figure S5: Model performance using alternative exons.

Model performance was assessed using Shannon entropies calculated from reads overlapping either the first, middle (mid), or last exons of the genes in each panel (see bottom schematic). For genes with an even number of exons, the exon closest to the TSS of the two middle exons was used. Accuracy was calculated for the UW cohort (left) and the GRAIL cohort (right). In both cohorts, Shannon entropies calculated from the first exon had the highest accuracy.

SUPPLEMENTAL METHODS

Summary of differences between GRAIL and UW cohorts:

Patients:

GRAIL: Patients with metastatic cancer who were progressing on stable doses of treatment. The normal (non-cancer) blood samples were obtained from the San Diego Blood Bank.

UW: Patients with metastatic cancer. While in general, patients who were treatment naïve or progressing were preferred, this also included patients who were responding to treatment. Neuroendocrine prostate cancer and bladder cancer were also included, which were not in the GRAIL dataset. No normal blood samples were included, as this was not allowed on the institutional blood collection protocol.

Sample tubes:

GRAIL: Streck tubes were used

UW: EDTA or CellSave tubes were used

cfDNA extraction:

GRAIL: QIAamp Circulating Nucleic Acid Kit (Qiagen)

UW: QIAamp Circulating Nucleic Acid kit (Qiagen)

Library preparation:

GRAIL: Illumina TruSeq DNA nano protocol with 6mer UMIs (Illumina)

UW: xGen Prism DNA library preparation kit with 8mer UMIs (Integrated DNA Technologies)

Target capture:

GRAIL: Custom 2.1Mb panel with 508 cancer genes using Illumina Nextera Rapid Capture protocol (Illumina)

UW: Custom 2.4Mb panel with 822 cancer genes using the xGen hybridization capture kit (Integrated DNA Technologies)

Sequencing Depth:

GRAIL: average raw cfDNA sequencing depth 71,749X

UW: average raw cfDNA sequencing depth 3,042X

SUPPLEMENTAL TABLES

Table S1: cfDNA fraction of samples with available germline sequencing.

ID	Phenotype	ctDNA fraction
BL_1	Bladder	0.0080
BL_2	Bladder	0.0119
BL_3	Bladder	0.0080
BL_4	Bladder	0.0421
BL_5	Bladder	0.1516
BL_6	Bladder	0.6395
BL_7	Bladder	0.6537
BL_8	Bladder	0.6250
BL_9	Bladder	0.6362
BL_10	Bladder	0.6193
BL_11	Bladder	0.6408
BL_12	Bladder	0.6390
BR_1	Breast	0.0040
BR_2	Breast	0.2788
BR_3	Breast	0.0040
BR_4	Breast	0.0344
BR_5	Breast	0.3264
BR_6	Breast	0.1421
BR_7	Breast	0.0149
BR_8	Breast	0.0769
BR_9	Breast	0.1843
BR_10	Breast	0.0040
BR_11	Breast	0.0208
BR_12	Breast	0.0373
BR_13	Breast	0.0080
BR_14	Breast	0.0070
BR_15	Breast	0.0139
BR_16	Breast	0.0139
BR_17	Breast	0.0469
BR_18	Breast	0.0383
BR_19	Breast	0.0139
BR_20	Breast	0.0070
BR_21	Breast	0.0188
BR_22	Breast	0.0070
BR_23	Breast	0.6395
BR_24	Breast	0.6357
BR_25	Breast	0.6418
BR_26	Breast	0.6343
BR_27	Breast	0.0247
BR_28	Breast	0.0090
BR_29	Breast	0.0325

BR_30	Breast	0.0109
BR_31	Breast	0.1326
BR_32	Breast	0.0090
BR_33	Breast	0.0564
BR_34	Breast	0.0109
BR_35	Breast	0.1575
BR_36	Breast	0.0090
BR_37	Breast	0.0060
BR_38	Breast	0.0554
BR_39	Breast	0.0889
BR_40	Breast	0.0421
BR_41	Breast	0.0198
BR_42	Breast	0.0060
BR_43	Breast	0.6092
BR_44	Breast	0.2847
BR_45	Breast	0.0305
BR_46	Breast	0.3871
BR_47	Breast	0.2222
BR_48	Breast	0.5185
BR_49	Breast	0.0188
BR_50	Breast	0.0354
BR_51	Breast	0.0129
BR_52	Breast	0.1843
BR_53	Breast	0.1876
BR_54	Breast	0.0478
BR_55	Breast	0.0658
BR_56	Breast	0.1793
BR_57	Breast	0.0139
BR_58	Breast	0.0440
BR_59	Breast	0.0601
LU_1	Lung	0.0109
LU_2	Lung	0.0119
LU_3	Lung	0.0100
LU_4	Lung	0.0080
LU_5	Lung	0.2293
LU_6	Lung	0.0050
LU_7	Lung	0.0169
LU_8	Lung	0.2159
LU_9	Lung	0.0276
LU_10	Lung	0.6702
LU_11	Lung	0.1456
LU_12	Lung	0.0100
LU_13	Lung	0.0149
LU_14	Lung	0.0159
LU_15	Lung	0.0198
LU_16	Lung	0.5385

NE_1	NEPC	0.3812
PR_1	Prostate	0.2087
PR_2	Prostate	0.0060
PR_3	Prostate	0.0149
PR_4	Prostate	0.3008
PR_5	Prostate	0.0169
PR_6	Prostate	0.0237
PR_7	Prostate	0.0178
PR_8	Prostate	0.0227
PR_9	Prostate	0.0450
PR_10	Prostate	0.5130
PR_11	Prostate	0.0354
PR_12	Prostate	0.0129
PR_13	Prostate	0.0159
PR_14	Prostate	0.0198
PR_15	Prostate	0.0080
PR_16	Prostate	0.0526
PR_17	Prostate	0.0686
PR_18	Prostate	0.4019
PR_19	Prostate	0.4436
PR_20	Prostate	0.4550
PR_21	Prostate	0.3819
PR_22	Prostate	0.4544
PR_23	Prostate	0.5866
PR_24	Prostate	0.1079
PR_25	Prostate	0.0227
PR_26	Prostate	0.1007
PR_27	Prostate	0.5856
PR_28	Prostate	0.2087
PR_29	Prostate	0.0129
PR_30	Prostate	0.6550
PR_31	Prostate	0.0276
PR_32	Prostate	0.1061
PR_33	Prostate	0.3122
PR_34	Prostate	0.3165
PR_35	Prostate	0.0852
PR_36	Prostate	0.1626
PR_37	Prostate	0.3580
PR_38	Prostate	0.2766
PR_39	Prostate	0.0218
PR_40	Prostate	0.6038
PR_41	Prostate	0.3532
PR_42	Prostate	0.1481
PR_43	Prostate	0.5648
PR_44	Prostate	0.1369
PR_45	Prostate	0.1481

PR_46	Prostate	0.0497
PR_47	Prostate	0.6283
PR_48	Prostate	0.5658
PR_49	Prostate	0.3580
PR_50	Prostate	0.0188
PR_51	Prostate	0.0257
PR_52	Prostate	0.2127
PR_53	Prostate	0.0169
PR_54	Prostate	0.5846
PR_55	Prostate	0.2095
PR_56	Prostate	0.6320
PR_57	Prostate	0.1635
PR_58	Prostate	0.1958
PR_59	Prostate	0.6418
PR_60	Prostate	0.0159
PR_61	Prostate	0.2935
PR_62	Prostate	0.0080
PR_63	Prostate	0.0119
PR_64	Prostate	0.0030
PR_65	Prostate	0.0050
PR_66	Prostate	0.0100
PR_67	Prostate	0.5658
PR_68	Prostate	0.7707
PR_69	Prostate	0.3607
PR_70	Prostate	0.0611
PR_71	Prostate	0.0080
PR_72	Prostate	0.2533
PR_73	Prostate	0.1282
PR_74	Prostate	0.0714
PR_75	Prostate	0.0276
PR_76	Prostate	0.0070
PR_77	Prostate	0.0227
PR_78	Prostate	0.4797
PR_79	Prostate	0.5343