

Table 1: TCGA cancer datasets and corresponding cancer types.

Dataset	Cancer type	Slides Analysed	Protein Alterations
BLCA	Bladder cancer	457	4
BRCA	Breast cancer	719	3
COAD	Colorectal cancer	459	4
GBM	Brain cancer	860	3
HNSC	Head & neck squamous cell carcinoma	472	2
LGG	brain tumor	844	2
LUAD	Lung adenocarcinoma	571	4
LUSC	Lung squamous cell carcinoma	512	2
SKCM	Skin cancer	475	5
THCA	Thyroid cancer	517	2
UCEC	Uterine cancer	566	4

Table 2: Genes and protein variants prediction scores.

dataset	Gene	Variant	Mean AUC (%)	Mean P value
BLCA	PIK3CA	All	64.61	4.39E-06
		p.E545K	70.95	1.50E-05
		p.E542K	49.31	0.9255
	TP53	All	70.33	6.20E-14
		p.R280T	76.04	0.0002
BRCA	PIK3CA	p.R248Q	55.29	0.486
		All	55.28	0.0236
	PIK3CA	p.E545K	68.63	5.01E-05
		p.E542K	61.83	0.0337
		p.H1047R	50.04	0.9913
COAD	PIK3CA	All	57.59	0.0129
		p.E545K	63.34	0.0096
	TP53	All	67.67	5.77E-11
		p.R175H	59.36	0.1088
HNSC	CDKN2A	All	53.89	0.22153
		p.R80*	72.60	0.0003
HNSC	PIK3CA	All	45.97	0.2612
		p.E542K	40.47	0.1953
LUSC	PIK3CA	All	54.21	0.2488
		p.E545K	52.42	0.721
	TP53	All	57.25	0.0554
p.R158L		58.34	0.2291	
UCEC	PIK3CA	All	57.97	0.001
		p.R88Q	60.45	0.0294
		p.E542K	72.29	0.00095
		p.G118D	67.15	0.0234
		p.E545K	50.36	0.963