

Supplementary material

EpICC: A Bayesian Neural Network model with uncertainty correction for a more accurate classification of cancer

Prasoon Joshi, Riddhiman Dhar

This PDF file contains

Figures S1 to S7

Tables S1 to S4

Figure S1

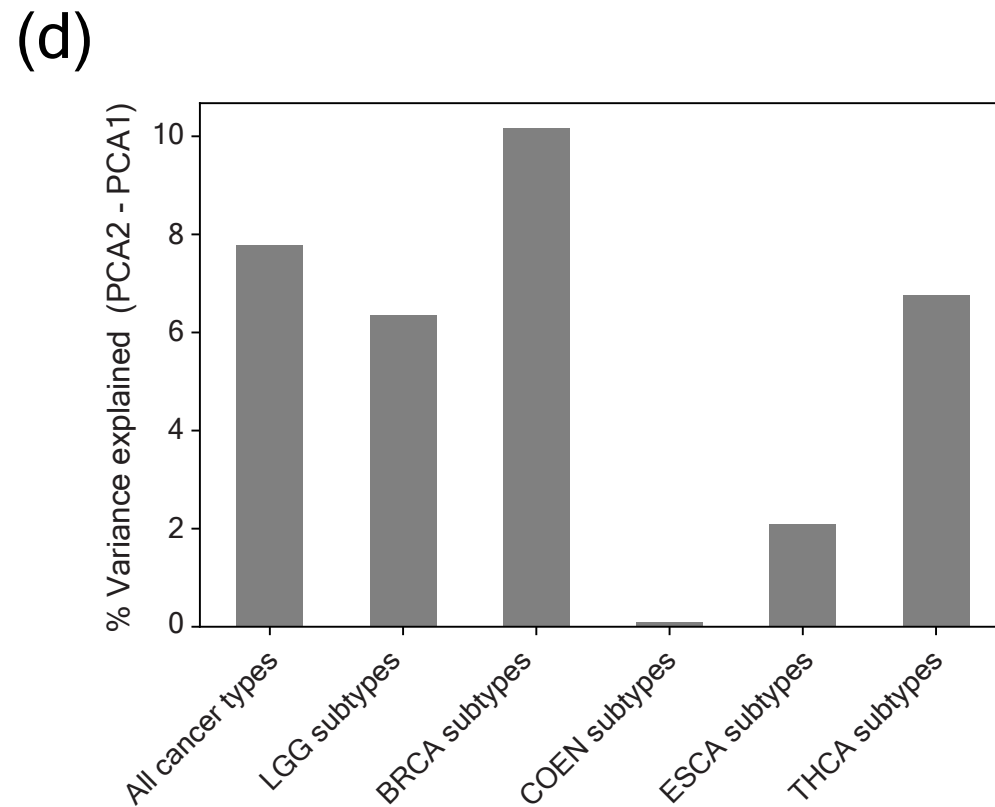
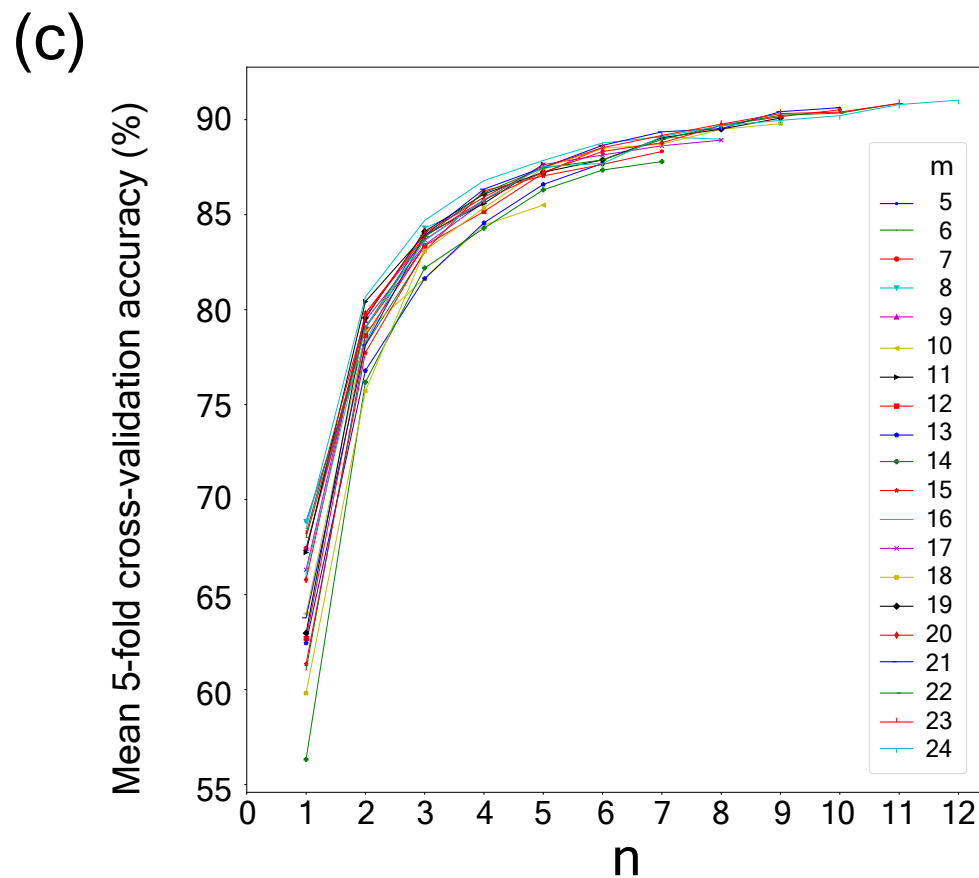
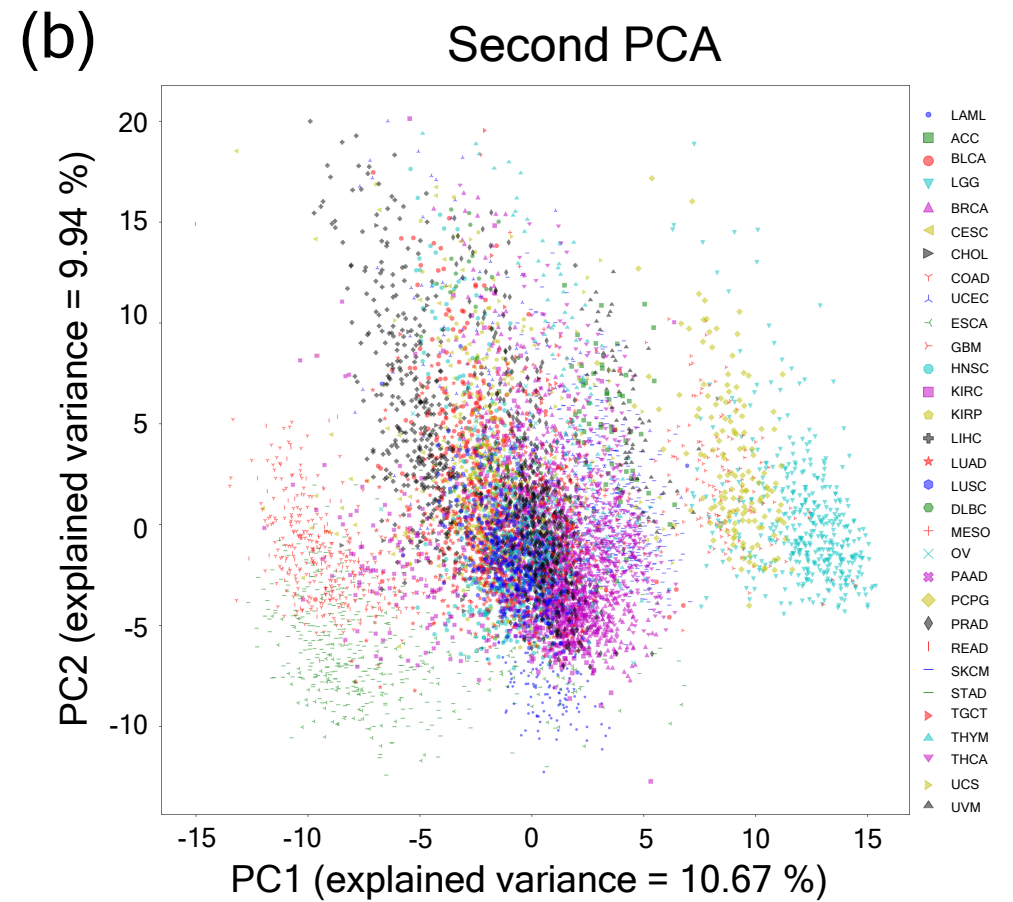
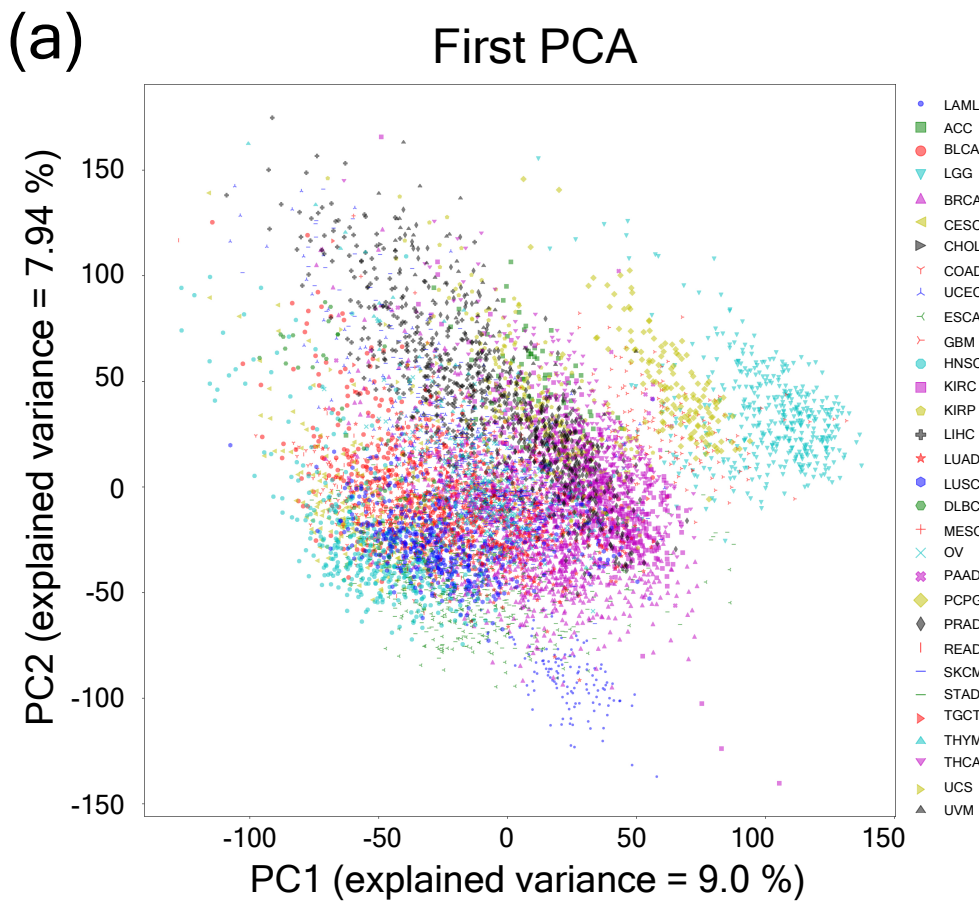


Figure S2

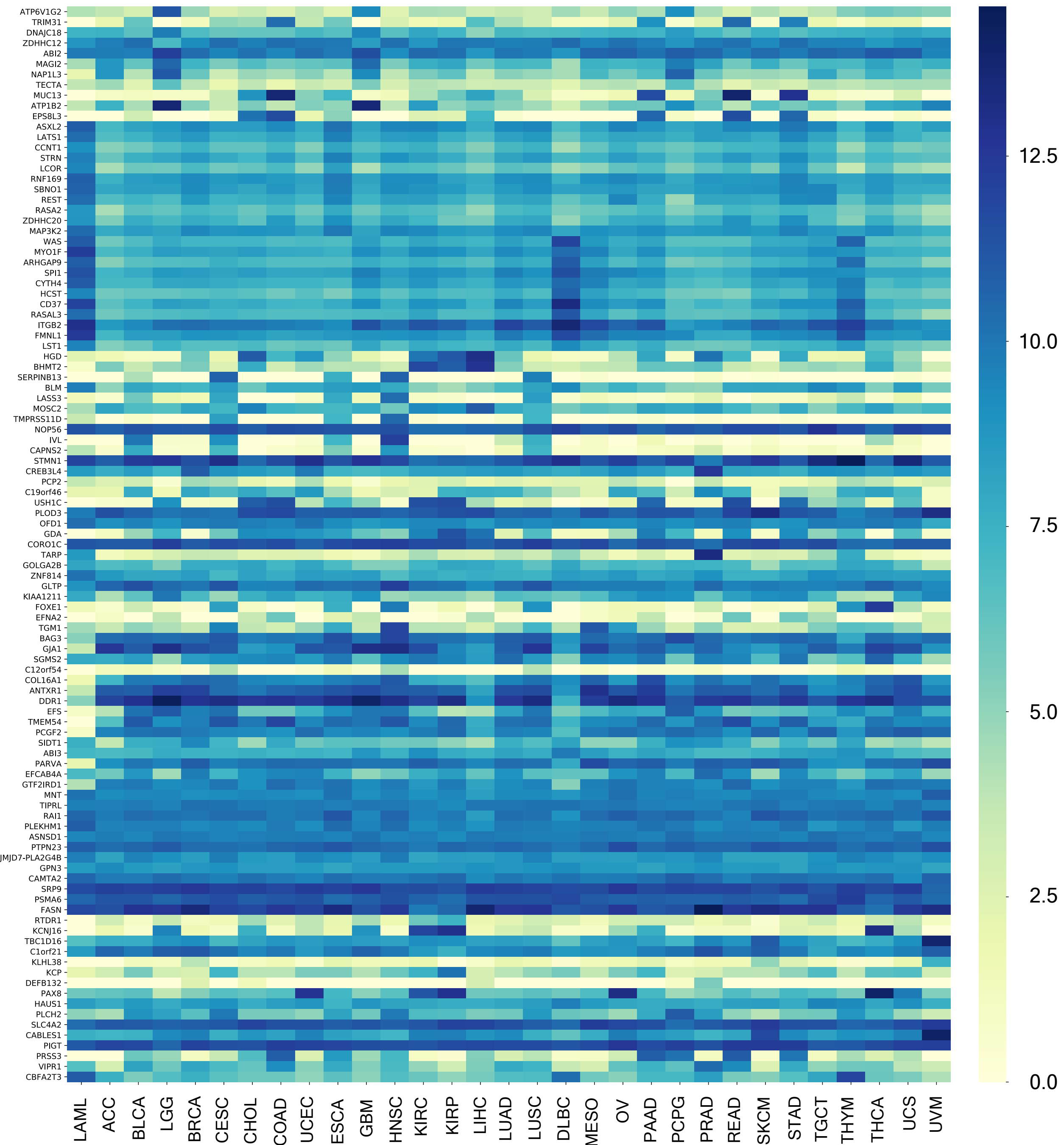
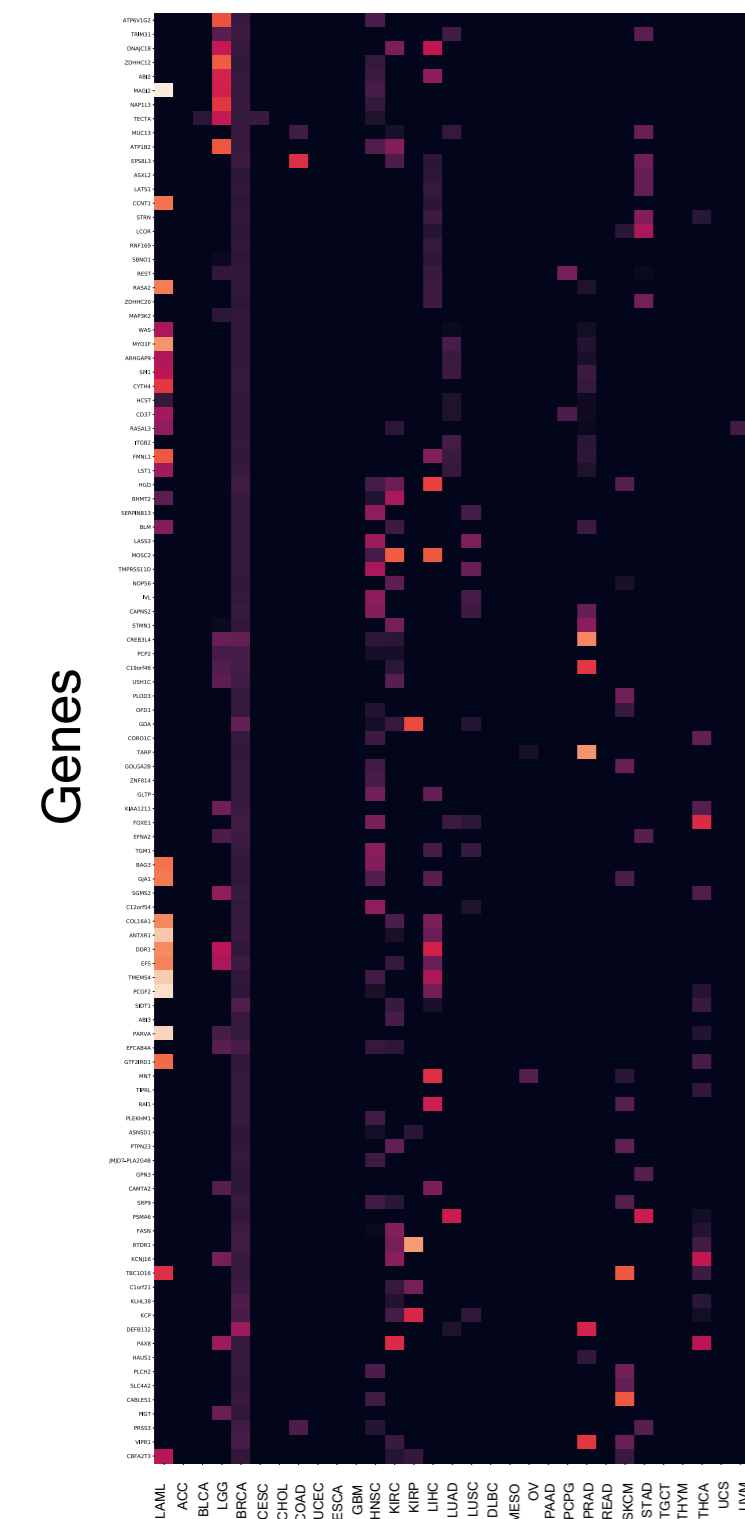


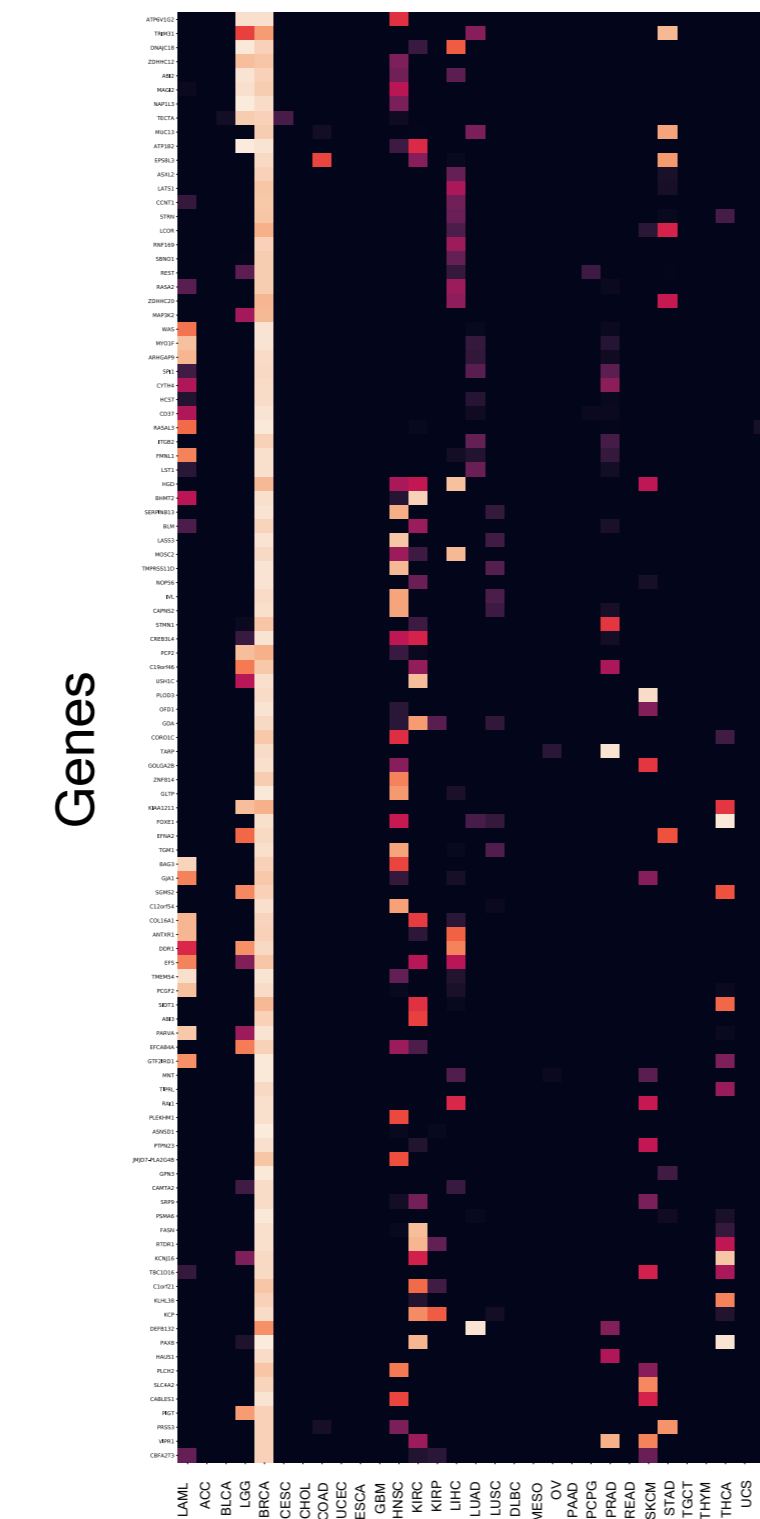
Figure S3

(a) Precision



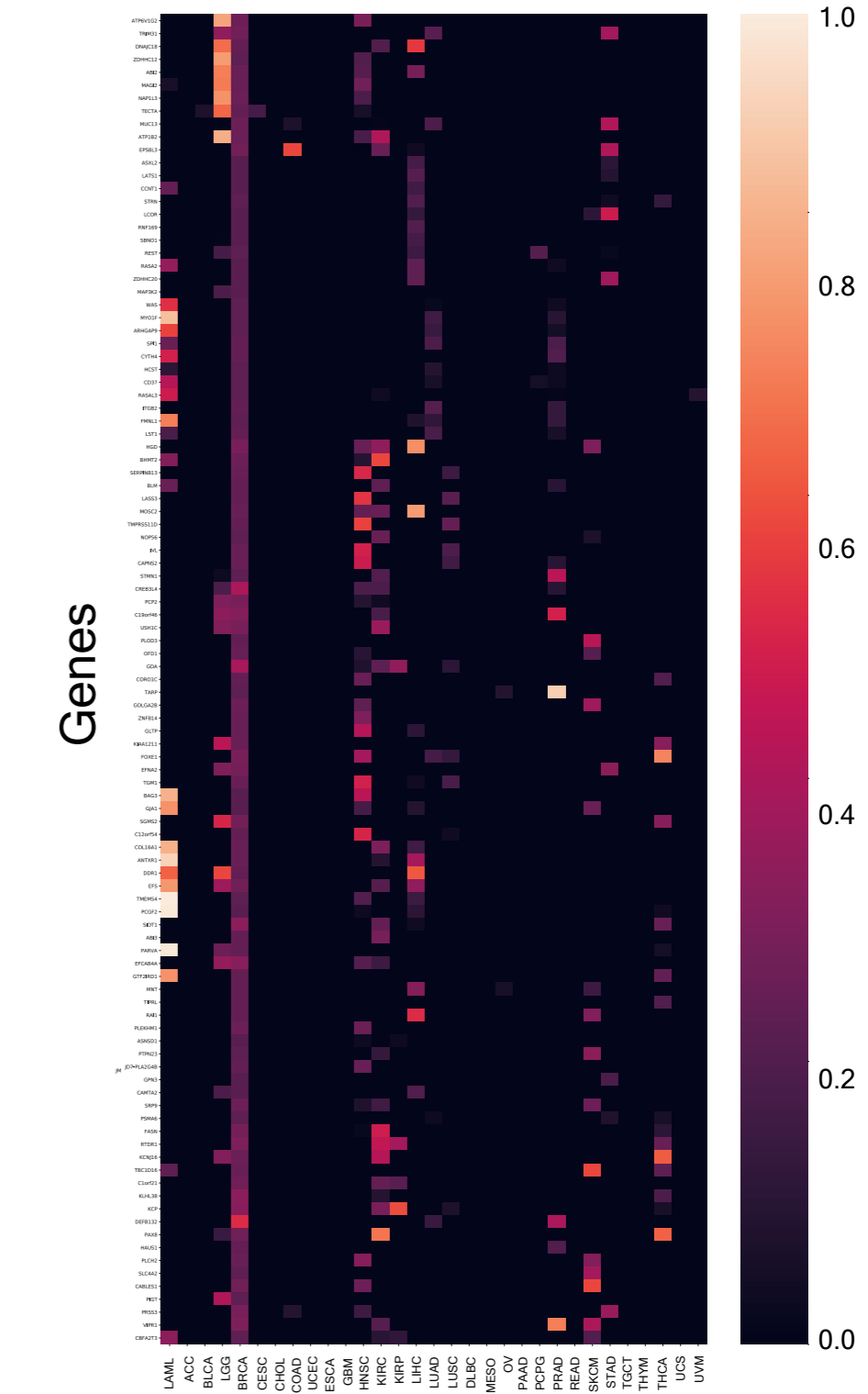
Cancer types

(b) Recall



Cancer types

(c) F1 score



Cancer types

Figure S4

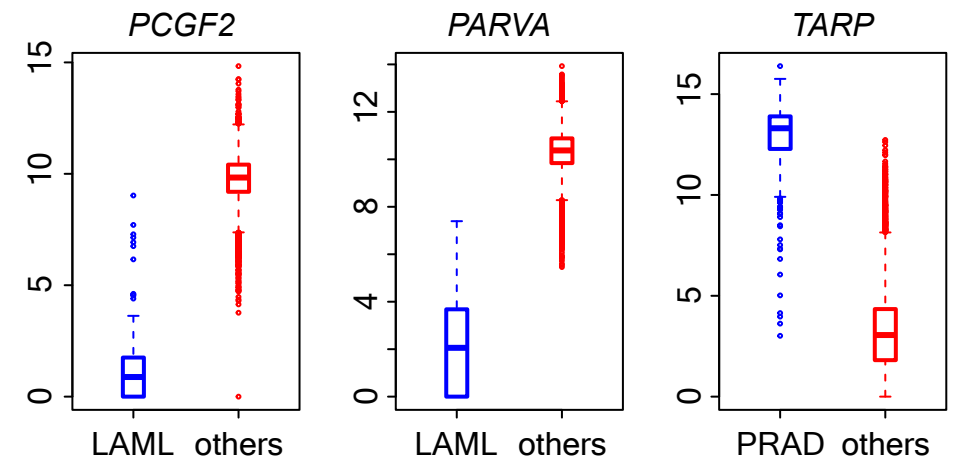
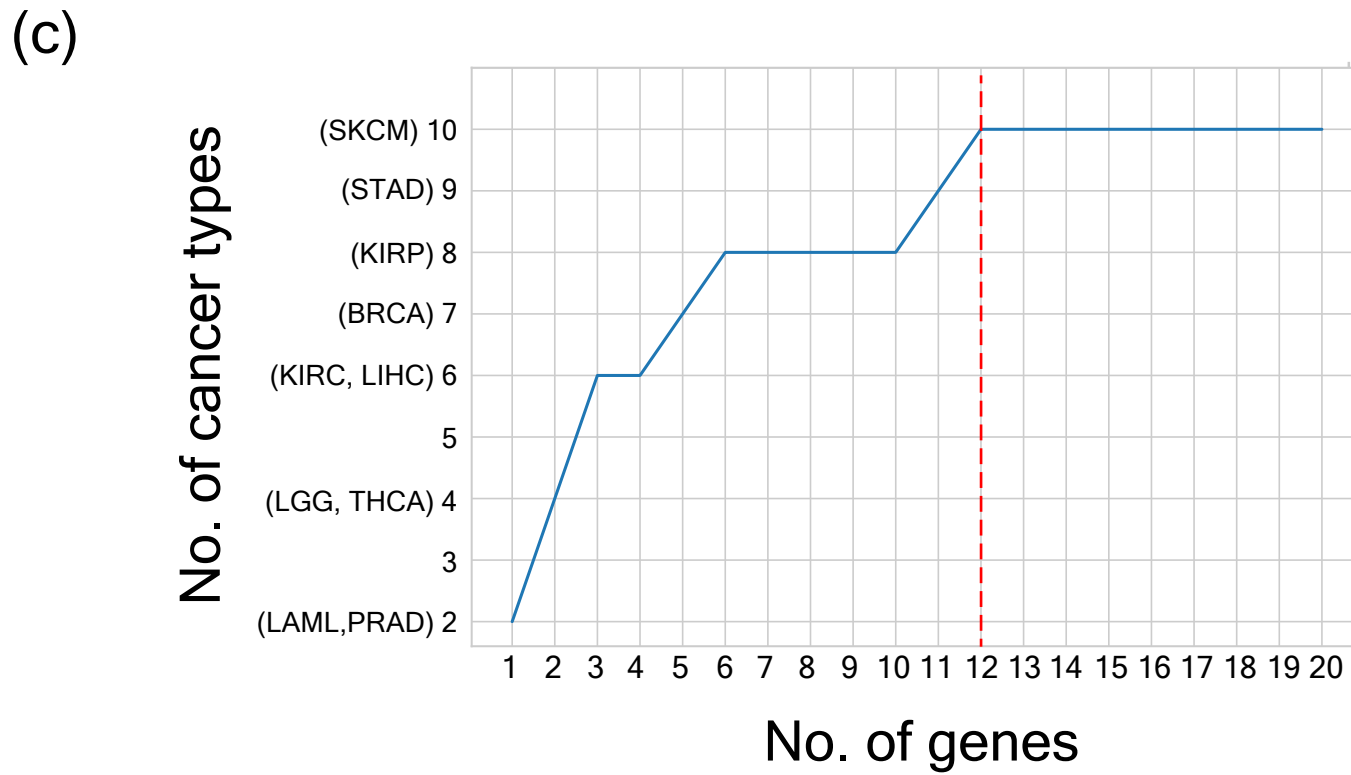
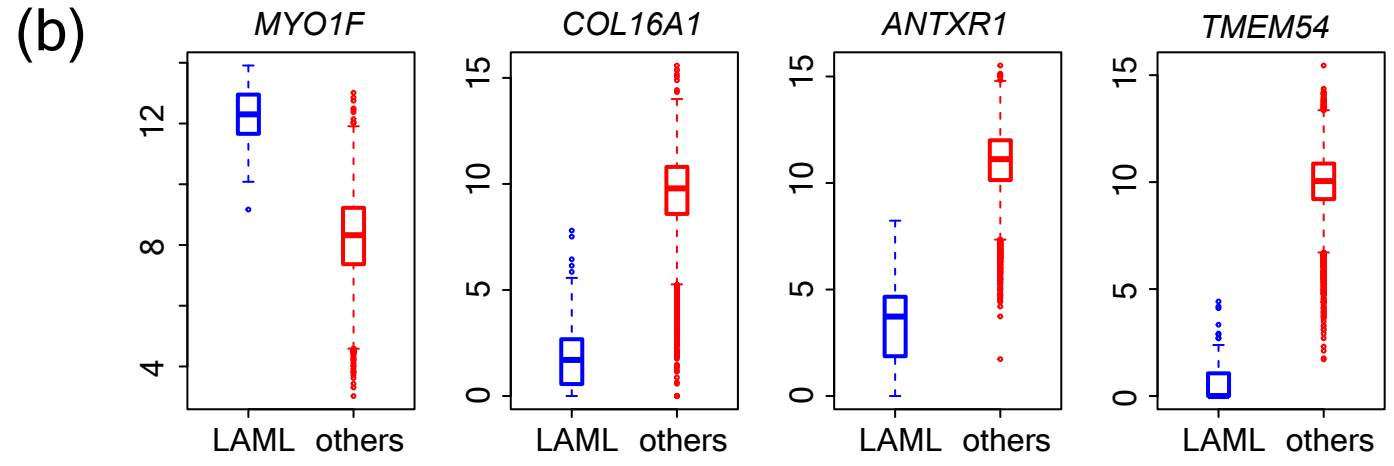
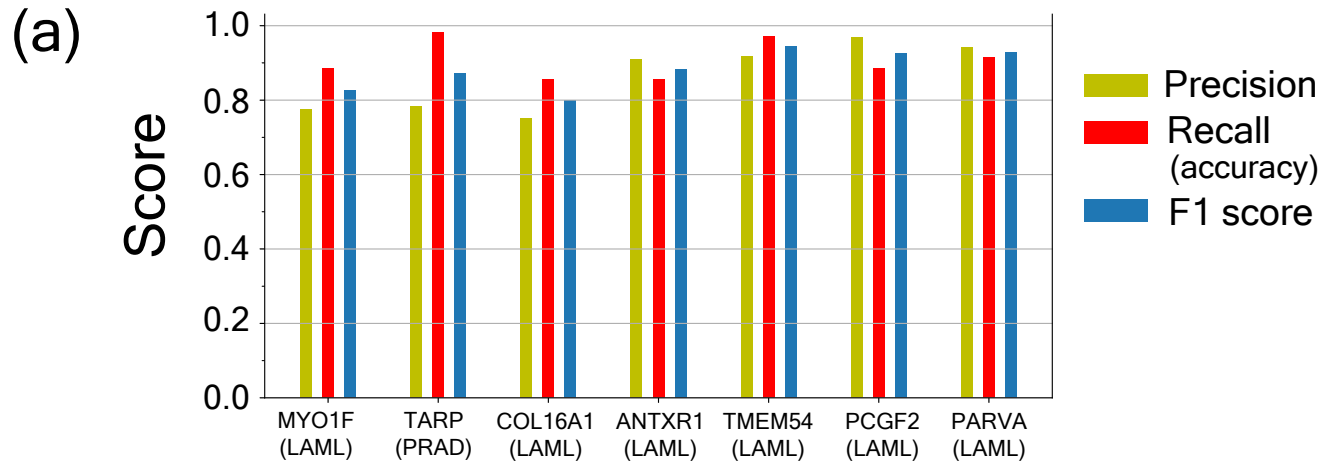


Figure S5

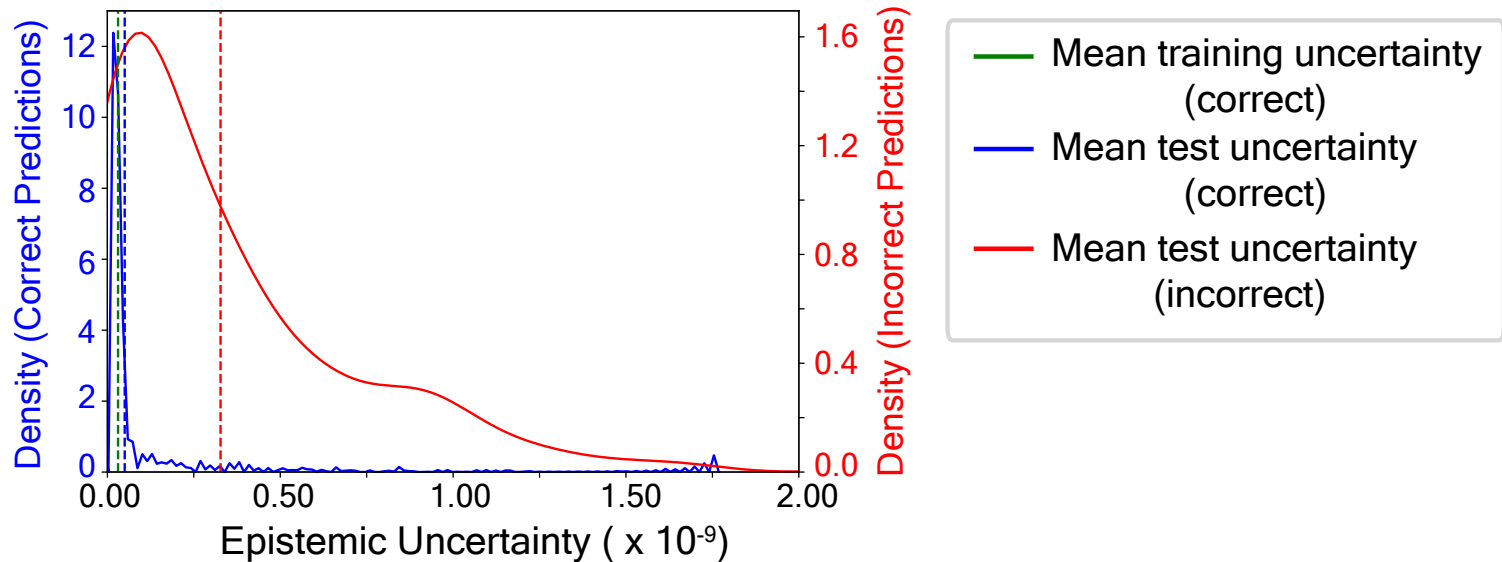
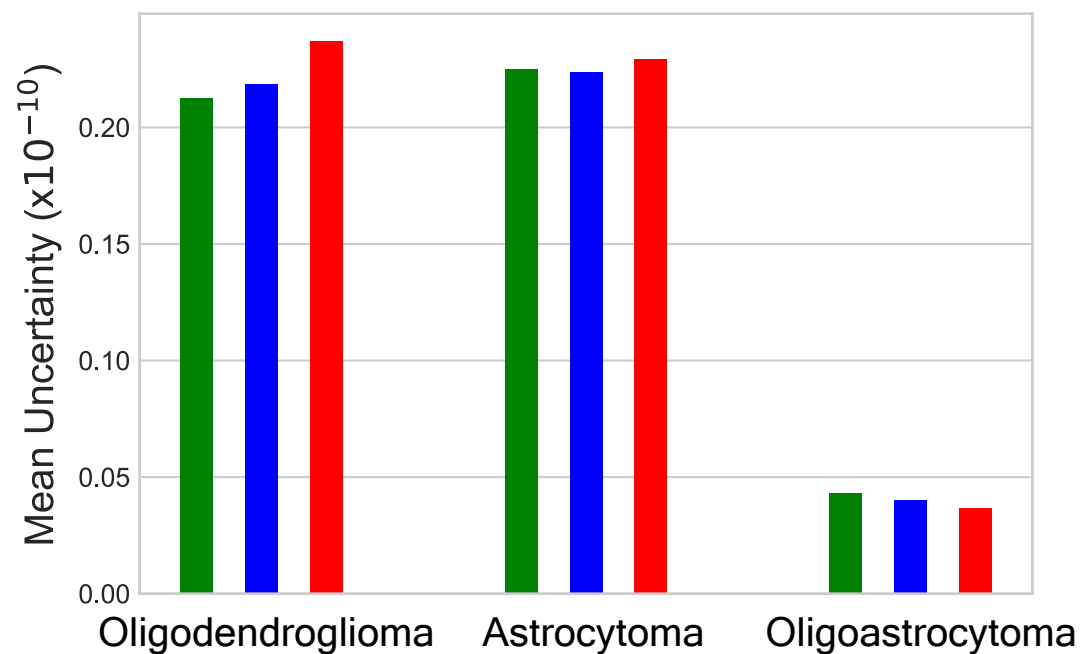
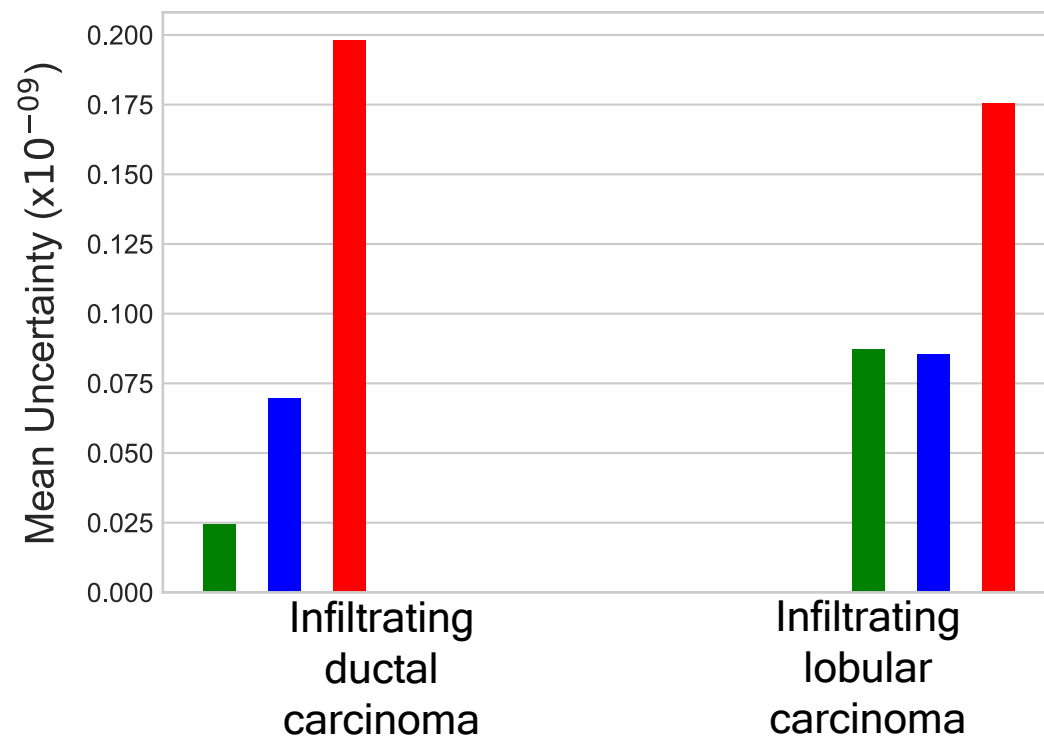


Figure S6

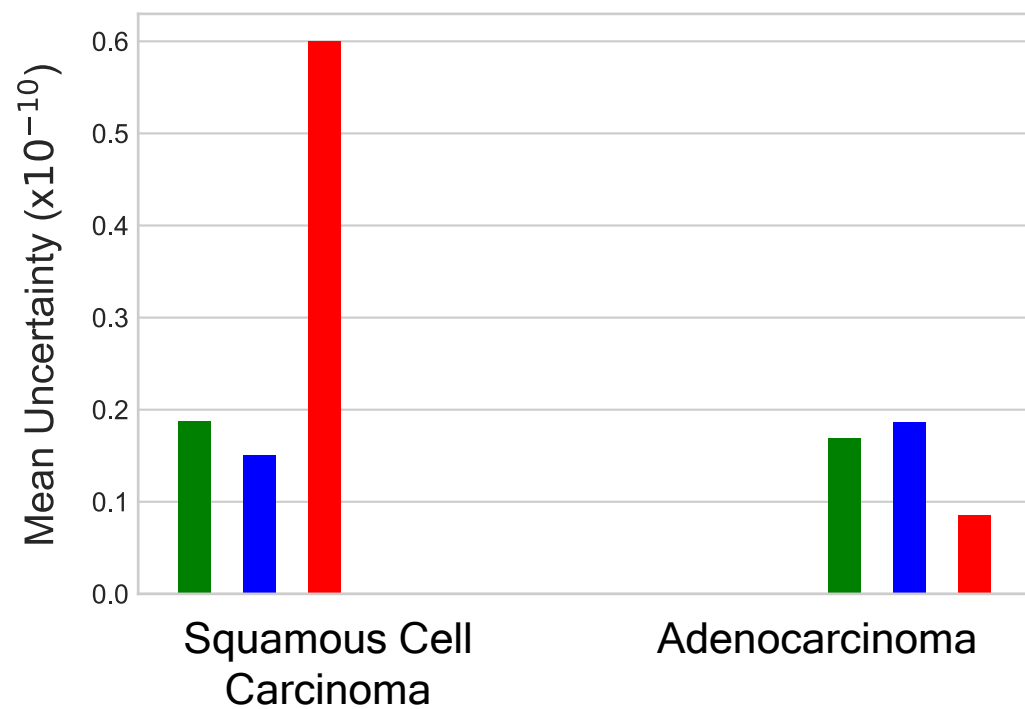
LGG subtypes



BRCA subtypes



ESCA subtypes



THCA subtypes

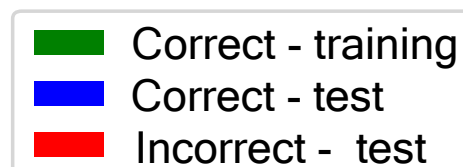
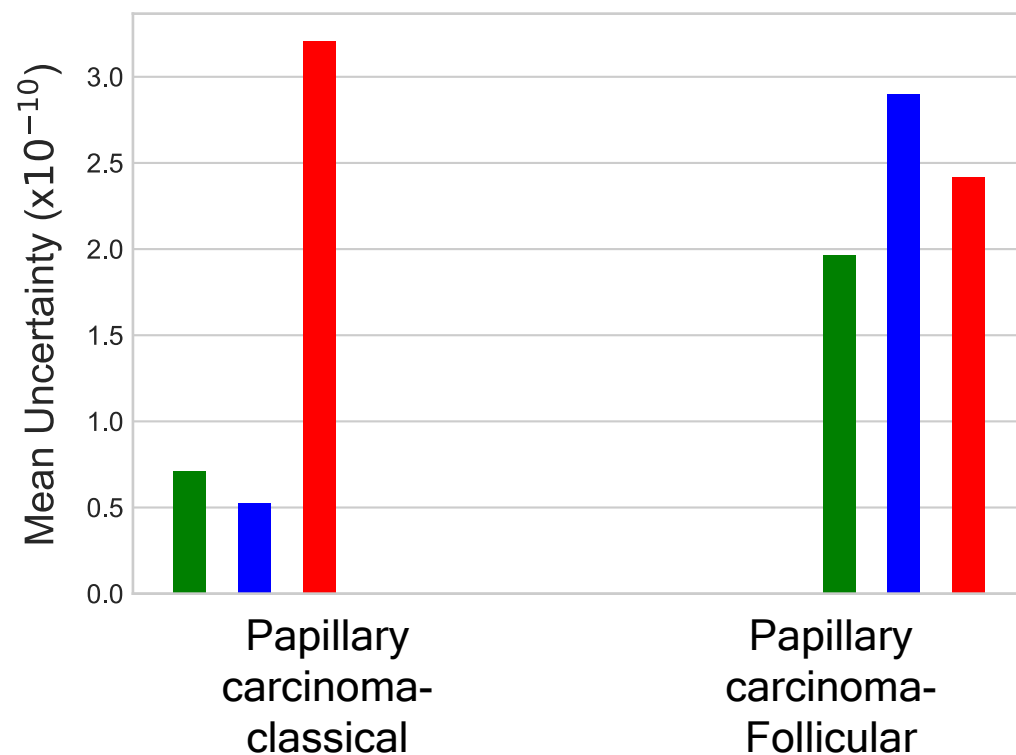
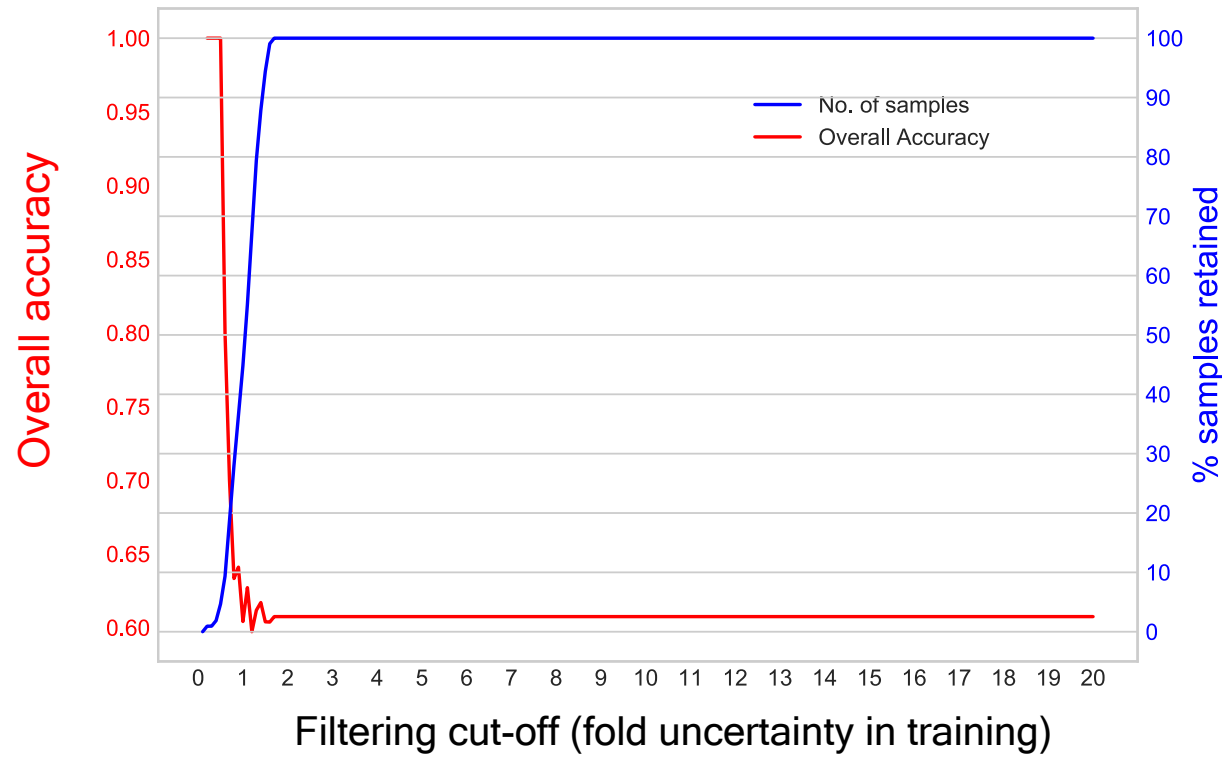
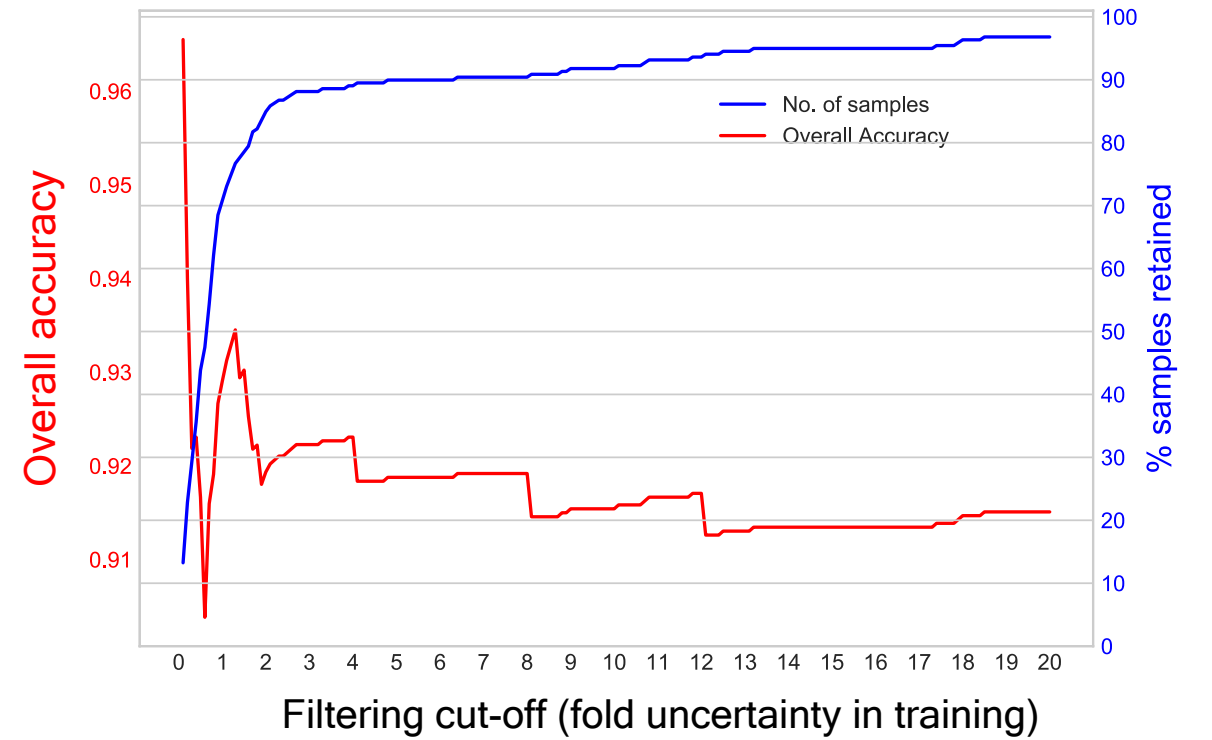


Figure S7

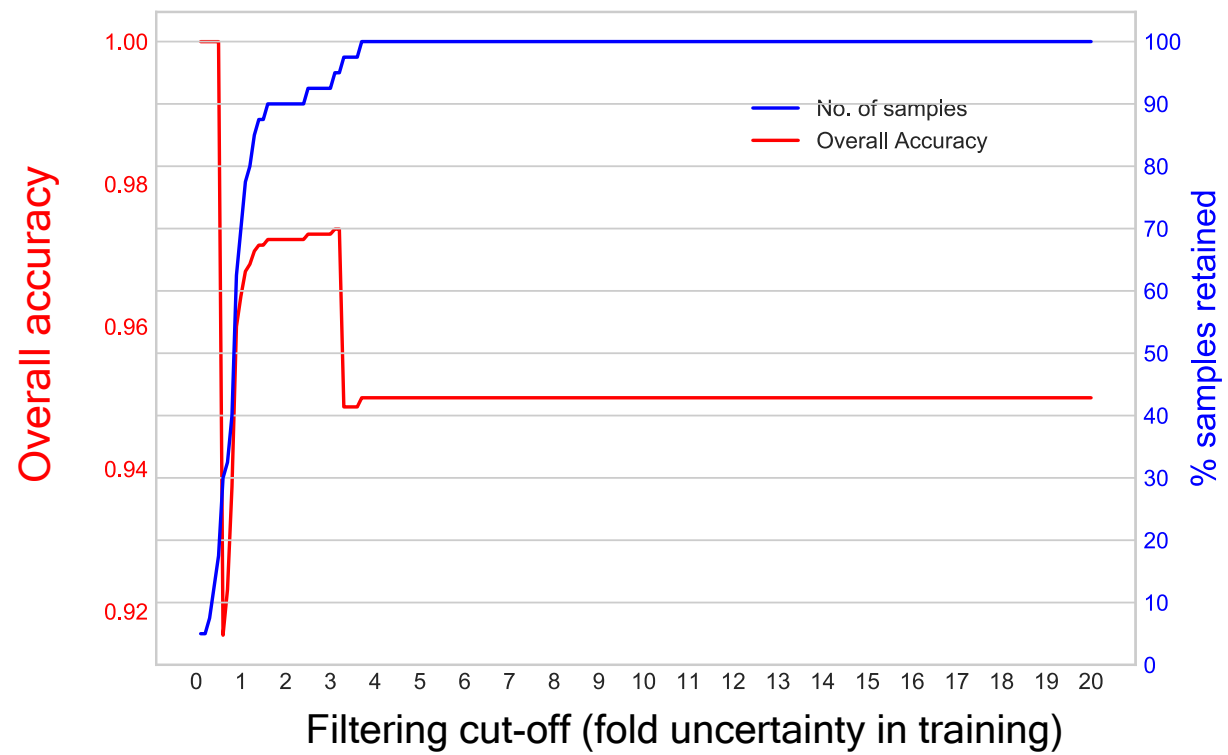
LGG Subtypes



BRCA Subtypes



ESCA Subtypes



THCA Subtypes

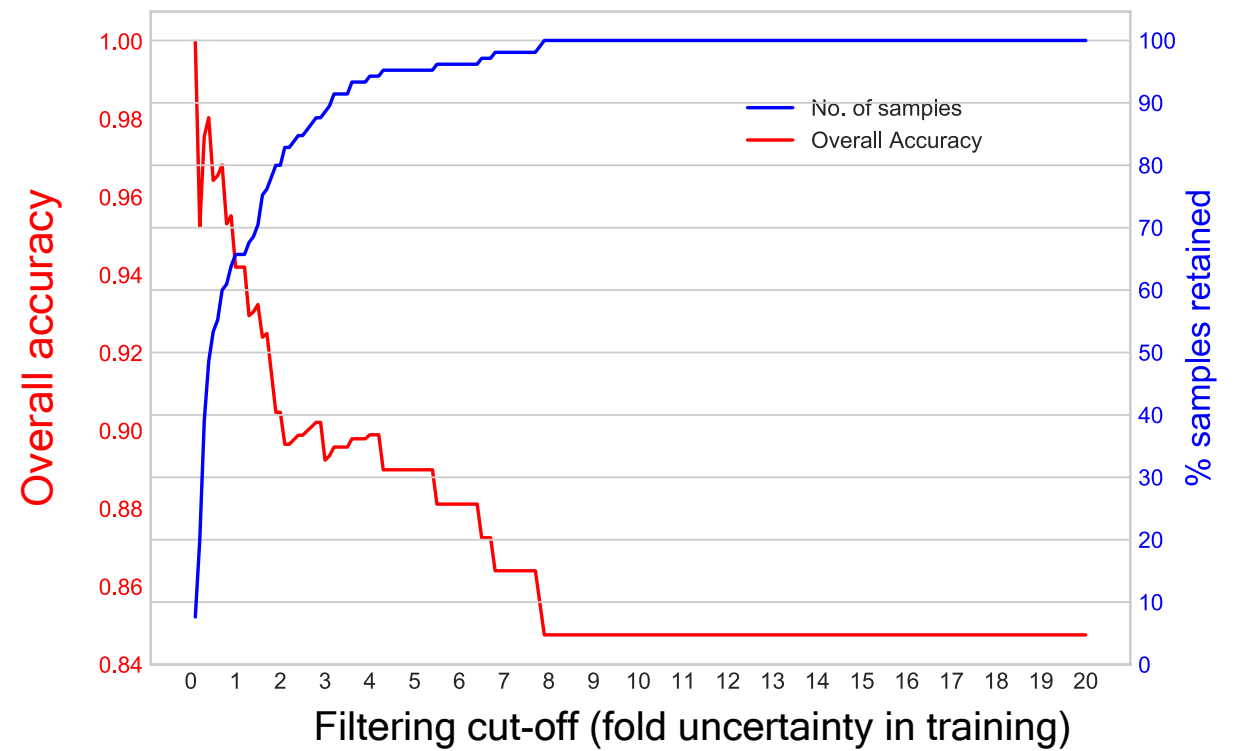


Table S1 - Details of 103 feature genes

S. No.	Gene	Description	Function	Reported Oncogenic? (Y/N)	Reported tumor suppressor? (Y/N)	Biomarker/Other association with cancer? (Y/N)
1	ABI2	Abl interactor 2	Actin polymerization/depolymerization, cell migration, cytoskeleton organization	N	Y (Dai and Pendergast, 1995)	N
2	ABI3	ABI family member 3	Regulation of actin polymerization, cell migration	N	Y (Latini <i>et al.</i> , 2011; Moraes <i>et al.</i> , 2016)	N
3	ANTXR1	ANTXR cell adhesion molecule 1	Role in cell attachment and migration, interacts with extracellular matrix proteins and with the actin cytoskeleton.	N	N	Y (Chen <i>et al.</i> , 2013; Geng <i>et al.</i> , 2019)
4	ARHGAP9	Rho GTPase activating protein 9	GTPase activator for the Rho-type GTPases	N	Y (Piao <i>et al.</i> , 2020)	Y (Piao <i>et al.</i> , 2020)
5	ASNSD1	Asparagine synthetase domain containing 1	Asparagine synthetic pathway	N	N	N
6	ASXL2	ASXL transcriptional regulator 2	Epigenetic regulators that bind various histone-modifying enzymes and involved in the assembly of transcription factors. Mutations in this gene are associated with cancer in several tissues (breast, bladder, pancreas, ovary, prostate, and blood)	Putative (Park <i>et al.</i> , 2016; Cui <i>et al.</i> , 2020)	Y (Daou <i>et al.</i> , 2015; Micol <i>et al.</i> , 2017)	Y (Park <i>et al.</i> , 2016; Cui <i>et al.</i> , 2020)
7	ATP1B2	ATPase Na ⁺ /K ⁺ Transporting Subunit Beta 2	Non-catalytic component of Na ⁺ -K ⁺ -ATPase	Putative (Li <i>et al.</i> , 2019)	Putative (van den Boome <i>et al.</i> , 2006; Demircan <i>et al.</i> , 2009)	N
8	ATP6V1G2	ATPase H ⁺ Transporting V1 Subunit G2	Component of vacuolar ATPase (V-ATPase) that is responsible for acidification of intracellular compartment	N	N	Y (Sennoune <i>et al.</i> , 2004; Xu <i>et al.</i> , 2013; Terrasi <i>et al.</i> , 2019)
9	BAG3	BAG Cochaperone 3	Co-chaperone for HSP70 and HSC70 chaperone proteins, has anti-apoptotic activity	Y (Chiappetta <i>et al.</i> , 2007; Shi <i>et al.</i> , 2018; Li <i>et al.</i> , 2018)	Y (Kong <i>et al.</i> , 2016)	N
10	BHMT2	Betaine--Homocysteine S-Methyltransferase 2	Regulation of homocysteine metabolism, converts homocysteine to methionine using S-methylmethionine (SMM) as a methyl donor	N	N	Putative (Dufresne <i>et al.</i> , 2019)
11	BLM	Bloom Syndrome RecQ Like Helicase	TP-dependent DNA helicase that unwinds single- and double-stranded DNA, participates in DNA replication and repair	N	Y (Payne and Hickson, 2009; de Voer <i>et al.</i> , 2015)	Y (Arora <i>et al.</i> , 2015)
12	C12orf54	Chromosome 12 Open Reading Frame 54	-	N	N	N
13	C19orf46	SYNE4 (Spectrin Repeat Containing Nuclear Envelope Family Member 4)	component of the LINC (Linker of Nucleoskeleton and Cytoskeleton) complex, involved in the connection between the nuclear lamina and the cytoskeleton	N	N	Y (Xu <i>et al.</i> , 2020)
14	C1orf21	Chromosome 1 Open Reading Frame 21	Cell Proliferation-Inducing Gene 13	N	N	Y (Hannemann <i>et al.</i> , 2006; Lee <i>et al.</i> , 2007; Yoshihara <i>et al.</i> , 2015)
15	CABLES1	CDK5 And ABL1 Enzyme Substrate 1	Cyclin-dependent kinase binding protein. Enhances cyclin-dependent kinase tyrosine phosphorylation by nonreceptor tyrosine kinases, such as that of CDK5 by activated ABL1	N	Y (Armason <i>et al.</i> , 2013; Shi <i>et al.</i> , 2015)	Putative (Huang <i>et al.</i> , 2017)
16	CAMTA2	Calmodulin-Binding Transcription Activator 2	Calmodulin-binding transcription activator protein	Putative (Luan <i>et al.</i> , 2019)	N	N
17	CAPNS2	Calcium-Dependent Protease Small Subunit 2	Calcium-regulated non-lysosomal thiol-protease	N	N	N
18	CBFA2T3	Core-Binding Factor, Runt Domain, Alpha Subunit 2; Translocated To, 3	Transcriptional corepressor, facilitates transcriptional repression via association with DNA-binding transcription factors and recruitment of other corepressors and histone-modifying enzymes	Y (Masetti <i>et al.</i> , 2019)	Y (Kochetkova <i>et al.</i> , 2002)	N
19	CCNT1	Cyclin T1	Regulatory subunit of the cyclin-dependent kinase pair (CDK9/cyclin-T1) complex	Y (Moiola <i>et al.</i> , 2010)	N	N
20	CD37	CD37 Molecule	Leukocyte antigen CD37	N	Y (de Winde <i>et al.</i> , 2016; Schaper and van Spruiel, 2018)	Y (Zhang <i>et al.</i> , 2020)
21	COL16A1	Collagen, Type XVI, Alpha 1	Involved in cell attachment and cell spreading	Y (Senner, 2008; Bauer, 2011; Ratzinger, 2011)	N	Y (Yoshihara <i>et al.</i> , 2015)
22	CORO1C	Coronin, Actin-Binding Protein, 1C	Role in directed cell migration by regulation of RAC1	Y (Castagnino <i>et al.</i> , 2018; Cheng <i>et al.</i> , 2019)	N	Y (Wu <i>et al.</i> , 2010; Cheng <i>et al.</i> , 2019; Wang <i>et al.</i> , 2020)
23	CREB3L4	Cyclic AMP-Responsive Element-Binding Protein 3-Like Protein 4	Transcriptional activator that may play a role in the unfolded protein response	Y (Chen <i>et al.</i> , 2010; Kim <i>et al.</i> , 2017)	N	Y (Kim <i>et al.</i> , 2017)
24	CYTH4	Cytohesin 4	Promotes the activation of ARF factors	N	N	N
25	DDR1	Discoidin Domain Receptor Tyrosine Kinase 1	Tyrosine kinase that functions as cell surface receptor for fibrillar collagen and regulates cell attachment to the extracellular matrix	Y (Valiathan <i>et al.</i> , 2012; Hu <i>et al.</i> , 2014)	Y (Takai <i>et al.</i> , 2018)	Y (Song <i>et al.</i> , 2016)
26	DEFB132	Defensin Beta 132	Cysteine-rich cationic polypeptides that are important in the immunologic response to invading microorganisms	N	N	N
27	DNAJC18	DnaJ Heat Shock Protein Family (Hsp40) Member C18	DNAJC18 (DnaJ Heat Shock Protein Family (Hsp40) Member C18)	N	N	Y (Guey <i>et al.</i> , 2010; Pohl <i>et al.</i> , 2019)
28	EFCAB4A	EF-hand calcium binding domain 4A	calcium ion binding	N	N	Y (Milani <i>et al.</i> , 2010)

29	EFNA2	Ephrin A2	Cell surface GPI-bound ligand for Eph receptors, a family of receptor tyrosine kinases crucial for migration, repulsion and adhesion during neuronal, vascular and epithelial development	N	N	N
30	EFS	Embryonal Fyn-Associated Substrate	Central coordinating role for tyrosine-kinase-based signaling related to cell adhesion	N	Y (Sertkaya et al., 2015)	Y (Neumann et al., 2011)
31	EPS8L3	Epidermal Growth Factor Receptor Pathway Substrate 8-Related	Epidermal growth factor receptor pathway substrate 8 (EPS8), a substrate for the epidermal growth factor receptor.	Y (Zeng et al., 2018; Li et al., 2019; Chen et al., 2020; Xuan et al., 2020)	N	Y (Li et al., 2019)
32	FASN	Fatty Acid Synthase	Catalyze the synthesis of palmitate from acetyl-CoA and malonyl-CoA, in the presence of NADPH, into long-chain saturated fatty acids	Y (Patel et al., 2015; Corominas-Faja et al., 2017)	N	Y (Visca et al., 2004; Cruz et al., 2014; Ricklefs et al., 2020)
33	FMNL1	Formin Like 1	Formin-related proteins have been implicated in morphogenesis	Putative (Chen et al., 2018)	N	Y (Chen et al., 2018; Higa et al., 2019)
34	FOXE1	Forkhead Box E1	Forkhead family of transcription factors	Putative (Ma et al., 2019)	Y (Dai et al., 2020)	Y (Kallel et al., 2010; Papadia et al., 2014; Sugimachi et al., 2016)
35	GDA	Guanine Deaminase	Enzyme responsible for the hydrolytic deamination of guanine	N	N	N
36	GJA1	Gap Junction Protein Alpha 1	Gap junction protein - a gap junction consists of a cluster of pairs of transmembrane channels, the connexons, through which low MW materials diffuse	N	Y (Talhouk et al., 2013; Teleki et al., 2014; Naser Al Deen et al., 2019)	Putative (Aasen et al., 2019)
37	GLTP	Glycolipid Transfer Protein	Involved in intermembrane transfer of glycolipids	N	Y (Mishra et al., 2019)	N
38	GOLGA2B	Golgin A2 Pseudogene 5	Pseudogene	N	N	N
39	GPN3	GPN-Loop GTPase 3	Small GTPase required for proper localization of RNA polymerase II	Putative (Lara-Chacón et al., 2019)	N	N
40	GTF2IRD1	GTF2I Repeat Domain Containing 1	Transcription regulator involved in cell-cycle progression and skeletal muscle differentiation	Y (Huo et al., 2016; Nambara et al., 2020)	N	N
41	HAUS1	HAUS Augmin Like Complex Subunit 1	Mitotic spindle assembly, maintenance of centrosome integrity and completion of cytokinesis as part of the HAUS augmin-like complex	N	N	Y (Campregher et al., 2016)
42	HCST	Hematopoietic Cell Signal Transducer	Transmembrane adapter protein forms an activation receptor KLRK1-HCST; KLRK1-HCST receptor plays a role in immune surveillance against tumors and is required for cytolysis of tumors cells	N	N	N
43	HGD	Homogentisate 1,2-Dioxygenase	Involved in the catabolism of the amino acids tyrosine and phenylalanine	N	N	N
44	ITGB2	Integrin Subunit Beta 2	participate in cell adhesion as well as cell-surface mediated signalling	N	N	Y (Laskowska et al., 2016; Peng et al., 2020)
45	IVL	Involucrin	Part of the insoluble cornified cell envelope (CE) of stratified squamous epithelia	N	N	Y (Walts et al., 1985)
46	JMJD7-PLA2G4B	Jumonji Domain Containing 7-Phospholipase A2, Group IVB (Cytosolic) Read-Through	-	Y (Cheng et al., 2017)	N	N
47	KCNJ16	Potassium Channel, Inwardly Rectifying Subfamily J Member 16	An integral membrane protein and inward-rectifier type potassium channel	N	N	Y (Adam et al., 2010; Jiang et al., 2017)
48	KCP	Kielin/Chordin-Like Protein	Enhances bone morphogenetic protein (BMP) signaling	N	N	N
49	KIAA1211	-	Uncharacterized	Y (Liu et al., 2019)	N	Y (Zhou et al., 2019)
50	KLHL38	Kelch Like Family Member 38	-	N	N	Y (Schleifer et al., 2018; Kuang et al., 2020)
51	LASS3	Ceramide Synthase 3	Regulate sphingolipid synthesis by catalyzing the formation of ceramides from sphingoid base and acyl-coA substrates	N	N	Y (Dany et al., 2015; Moro et al., 2018; Brachtendorf et al., 2019)
52	LATS1	Large Tumor Suppressor Kinase 1	Serine/threonine-protein kinase	N	Y (Xia et al., 2002; Yang et al., 2004; Hao et al., 2008)	Y (Ji et al., 2012; Furth et al., 2018)
53	LCOR	Ligand Dependent Nuclear Receptor Corepressor	Repressor of ligand-dependent transcription activation	N	N	Y (Jalaguier et al., 2017; Triki et al., 2017; Vogelsang et al., 2020)
54	LST1	Leukocyte Specific Transcript 1	A membrane protein that can inhibit the proliferation of lymphocytes, possible role in immunomodulation	N	N	Y (Hsu et al., 2019)
55	MAGI2	Membrane Associated Guanylate Kinase, WW And PDZ Domain	Act as scaffold molecule at synaptic junctions by assembling neurotransmitter receptors and cell adhesion proteins	Putative (Li et al., 2020)	Y (Thomas et al., 2002; Hu et al., 2007; Berger et al., 2011; Ma et al., 2016)	Y (Mahdian et al., 2014; Valle et al., 2020)

56	MAP3K2	Mitogen-Activated Protein Kinase Kinase 2	Component of protein kinase signal transduction cascade. Mediates activation of the NF-kappa-B, AP1 and DDIT3 transcriptional regulators	Y (Mazur et al., 2014; Huang et al., 2016; Wang et al., 2018; Shi et al., 2020)	N	N
57	MNT	MAX Network Transcriptional Repressor	Binds DNA as a heterodimer with MAX and represses transcription.	Y (Link et al., 2012; Link and Hurlin, 2015)	N	N
58	MOSC2	Mitochondrial Amidoxime Reducing Component 2 (MARC2)	Catalyzes the reduction of N-oxygenated molecules, acting as a counterpart of cytochrome P450 and flavin-containing monoxygenases in metabolic cycles	Y (Wu et al., 2020)	N	Y (Wu et al., 2020)
59	MUC13	Mucin 13, Cell Surface Associated	Epithelial and hemopoietic transmembrane mucin that may play a role in cell signaling.	Y (Sheng et al., 2017; Khan et al., 2017; Kumari et al., 2018; Sheng et al., 2019; Tiemin et al., 2020)	N	Y (Khan et al., 2018; Filippou et al., 2018)
60	MYO1F	Unconventional myosin type 1F	Use energy from ATP hydrolysis to generate force on actin filaments, works in intracellular environment	Y (Diquigiovanni et al., 2018)	N	Y (Taki et al., 2005)
61	NAP1L3	Nucleosome Assembly Protein 1 Like 3	A member of the nucleosome assembly protein (NAP) family	N	N	Y (Ross et al., 2015)
62	NOP56	NOP56 Ribonucleoprotein	Involved in 60S ribosomal subunit biogenesis. Core component of snoRNP particles	Y (Cowling et al., 2014)	N	Y (Gong et al., 2017; Toth et al., 2019)
63	OFD1	OFD1 Centriole And Centriolar Satellite Protein	Component of the centrioles controlling mother and daughter centrioles length	N	N	N
64	PARVA	Parvin Alpha	Plays a role in sarcomere organization and in smooth muscle cell contraction	Y (Fukuda et al., 2013; Huang et al., 2015; Velazquez-Torres et al., 2018)	N	N
65	PAX8	Paired Box 8	Involved in thyroid follicular cell development and expression of thyroid-specific genes	Y (Kroll et al., 2000; Allison, 2019)	N	Y (Nonaka et al., 2008; Wang et al., 2013; Xiang and Kong, 2013; Wang et al., 2015)
66	PCGF2	Polycomb Group Ring Finger 2	Transcriptional repressor, May play a role in control of cell proliferation and/or neural cell development.	N	Y (Guo et al., 2007; Lee et al., 2014)	Y (Zakrzewska et al., 2011)
67	PCP2	Purkinje Cell Protein 2	May function as a cell-type specific modulator for G protein-mediated cell signaling	N	Putative (Yan et al., 2006)	N
68	PIGT	Phosphatidylinositol Glycan Anchor Biosynthesis Class T	Involved in glycosylphosphatidylinositol (GPI)-anchor biosynthesis	Y (Wu et al., 2002; Nagpal et al., 2008)	N	Y (Wu et al., 2002; Nagpal et al., 2008)
69	PLCH2	Phospholipase C Eta 2	Member of the PLC-eta family of the phosphoinositide-specific phospholipase C (PLC) superfamily of enzymes	N	N	N
70	PLEKHM1	Pleckstrin Homology And RUN Domain Containing M1	Essential for bone resorption, and may play a critical role in vesicular transport in the osteoclast	N	N	N
71	PLOD3	Procollagen-Lysine,2-Oxoglutarate 5-Dioxygenase 3	Membrane-bound homodimeric enzyme that catalyzes the hydroxylation of lysyl residues in collagen-like peptide	Y (Baek et al., 2018; Tsai et al., 2018; Baek et al., 2019; Xie et al., 2020)	N	Y (Shen et al., 2018)
72	PRSS3	Serine Protease 3	Member of the trypsin family of serine proteases.	Y (Hockla et al., 2012; Qian et al., 2017; Wang et al., 2019)	Putative (Marsit et al., 2005)	Y (Qian et al., 2017; Wang et al., 2019)
73	PSMA6	Proteasome 20S Subunit Alpha 6	Component of the 20S core proteasome complex involved in the proteolytic degradation of intracellular proteins	Y (Kakumu et al., 2017; Bakke et al., 2019)	N	N
74	PTPN23	Protein Tyrosine Phosphatase Non-Receptor Type 23	may be involved in the regulation of small nuclear ribonucleo protein assembly and pre-mRNA splicing	N	Y (Lin et al., 2011; Tanaka et al., 2013; Li et al., 2015; Li et al., 2019)	Y (Lin et al., 2011; Chen et al., 2020)
75	RAI1	Retinoic Acid Induced 1	Transcriptional regulator of the circadian clock components	N	N	N
76	RASA2	RAS P21 Protein Activator 2	Inhibitory regulator of the Ras-cyclic AMP pathway.	N	Y (Arafah et al., 2015; Arafah et al., 2019)	N
77	RASAL3	RAS Protein Activator Like 3	Functions as a Ras GTPase-activating protein	N	Putative (Mishra et al., 2018; Mishra et al., 2019)	N
78	REST	RE1 Silencing Transcription Factor	Transcriptional repressor which binds neuron-restrictive silencer element (NRSE) and represses neuronal gene transcription in non-neuronal cells	Y (Lawinger et al., 2000; Fuller et al., 2005; Huang et al., 2012; Negrini et al., 2013)	Y (Westbrook et al., 2005; Wagoner et al., 2010; Huang et al., 2012; Negrini et al., 2013)	N
79	RNF169	Ring Finger Protein 169	Probable E3 ubiquitin-protein ligase that acts as a negative regulator of double-strand breaks (DSBs) repair	N	N	Y (An et al., 2017; An et al., 2018)
80	RTDR1	Rhabdoid tumor deletion region 1	-	N	N	Y (Zhou et al., 2000)
81	SBNO1	Strawberry Notch Homolog 1	-	N	N	N
82	SERPINB13	Serpin Family B Member 13	A member of the serpin family of serine protease inhibitors, putative role in the proliferation/differentiation of keratinocytes	N	Y (Nakashima et al., 2000; Martin et al., 2008)	Y (Kim et al., 2008; Sahu and Grandis, 2011)
83	SGMS2	Sphingomyelin Synthase 2	Synthesize the sphingolipid, sphingomyelin, through transfer of the phosphatidyl head group, phosphatidylcholine, on to the primary hydroxyl of ceramide	Y (Zheng et al., 2019)	N	N

84	SIDT1	SID1 Transmembrane Family Member 1	Role in RNA transport	N	Y (Wang et al., 2019)	Y (Wang et al., 2019)
85	SLC4A2	Solute Carrier Family 4 Member 2	Plasma membrane anion exchange protein	N	N	Y (Yang et al., 2008; Wang et al., 2013)
86	SPI1	Spi-1 Proto-Oncogene	Binds to PU-box, transcriptional activator specifically involved in the differentiation/activation of macrophages or B-cells.	Y (Moreau-Gachelin et al., 1988; Nguyen et al., 1990; Ray et al., 1990)	Y (Özdemir et al., 2018)	N
87	SRP9	Signal Recognition Particle 9	Crucial role in targeting secretory proteins to the rough endoplasmic reticulum membrane	N	N	Y (Rho et al., 2008)
88	STMN1	Stathmin 1	Microtubule destabilizer protein with an important role in cell cycle progression, cell proliferation, migration and survival.	Y (Chen et al., 2013; Zhang et al., 2020)	N	Y (Nemunaitis et al., 2012; Nie et al., 2015; Liu et al., 2017; Verdanet et al., 2017; Shimizu et al., 2019; Li et al., 2019)
89	STRN	Striatin	Calmodulin-binding protein which may function as scaffolding or signaling protein	N	N	Y (Kusano et al., 2016; Bastos et al., 2018; Du et al., 2020)
90	TARP	TCR Gamma Alternate Reading Frame Protein	-	Putative (Wolfgang et al., 2000; Wolfgang et al., 2001)	N	Y (Fritzsche et al., 2010)
91	TBC1D16	TBC1 Domain Family Member 16	May act as a GTPase-activating protein for Rab family protein	Putative (Akavia et al., 2010; Vizoso et al., 2015; Rodger et al., 2019)	N	Y (Wouters et al., 2017; Yang et al., 2018)
92	TECTA	Tectorin Alpha	One of the major non-collagenous components of the tectorial membrane	N	N	Y (Hernández et al., 2007)
93	TGM1	Transglutaminase 1	Catalyzes the cross-linking of proteins and the conjugation of polyamines to proteins	N	N	Y (Martinet et al., 2003; Huang et al., 2017)
94	TIPRL	TOR Signaling Pathway Regulator	May be an allosteric regulator of serine/threonine-protein phosphatase 2A (PP2A)	N	Y (Luan et al., 2020)	Y (Song et al., 2012; Jeon et al., 2019; Jun et al., 2019; Xu et al., 2020)
95	TMEM54	Transmembrane Protein 54	-	N	N	N
96	TMPRSS11D	Transmembrane Serine Protease 11D	Trypsin-like serine protease released from the submucosal serous glands onto mucous membrane	Y (Cao et al., 2017; Yan et al., 2019)	N	Y (Cao et al., 2017; Zhao et al., 2019)
97	TRIM31	Tripartite Motif Containing 31	Functions as an E3 ubiquitin-protein ligase	Y (Sugiura and Miyamoto, 2008; Hatakeyama, 2011; Guo et al., 2018; Yu et al., 2018; Li et al., 2018; Zhou et al., 2019)	Y (Li et al., 2014)	Y (Yu et al., 2018)
98	USH1C	USH1 Protein Network Component Harmonin	Scaffold protein that functions in the assembly of Usher protein complexes	N	N	N
99	VIPR1	Vasoactive Intestinal Peptide Receptor 1	A receptor for vasoactive intestinal peptide, a small neuropeptide	N	Y (Mlakar et al., 2010; Zhao et al., 2019b)	Y (Yin et al., 2016; Sun et al., 2020)
100	WAS	WASP Actin Nucleation Promoting Factor	Effector protein for Rho-type GTPases that regulates actin filament reorganization	N	Y (Menotti et al., 2019)	Y (Ishihara et al., 2013)
101	ZDHHC12	Zinc Finger DHHC-Type Palmitoyltransferase 12	protein-cysteine S-palmitoyltransferase activity and palmitoyltransferase activity	N	N	Y (Chen et al., 2020; Meng et al., 2020)
102	ZDHHC20	Zinc Finger DHHC-Type Palmitoyltransferase 20	Catalyzes palmitoylation of Cys residues on target proteins	Y (Ducker et al., 2004; Draper and Smith, 2010; Jung et al., 2017)	N	N
103	ZNF814	zinc finger protein 814	Nucleic acid binding	N	N	Y (Karasaki et al., 2016; Ma et al., 2017)

Table S2 - Abbreviations used for the 31 different cancer types used in our analysis

Abbreviation	Cancer Type
LAML	acute myeloid leukaemia
ACC	adrenocortical carcinoma
BLCA	bladder urothelial carcinoma
LGG	brain lower grade glioma
BRCA	breast invasive carcinoma
CESC	cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	cholangiocarcinoma
COAD	colon adenocarcinoma
UCEC	corpus endometrioid carcinoma
ESCA	oesophageal carcinoma
GBM	glioblastoma multiforme
HNSC	head & neck squamous cell carcinoma
KIRC	kidney renal clear cell carcinoma
KIRP	kidney renal papillary cell carcinoma
LIHC	liver hepatocellular carcinoma
LUAD	lung adenocarcinoma
LUSC	lung squamous cell carcinoma
DLBC	lymphoid neoplasm diffuse large B-cell lymphoma
MESO	Mesothelioma
OV	ovarian serous cystadenocarcinoma
PAAD	pancreatic adenocarcinoma
PCPG	pheochromocytoma and paraganglioma
PRAD	prostate adenocarcinoma
READ	rectum adenocarcinoma
SKCM	skin cutaneous melanoma
STAD	stomach adenocarcinoma
TGCT	testicular germ cell tumour
THYM	thymoma
THCA	thyroid carcinoma
UCS	uterine carcinosarcoma
UVM	uveal melanoma

Table S3 - Number of training and testing samples for each cancer type in our analysis

Cancer type	No. of Samples	
	Train	Test
LAML	138	35
ACC	63	16
BLCA	340	86
LGG	424	106
BRCA	974	244
CESC	246	62
CHOL	36	9
COAD	263	66
UCEC	160	41
ESCA	156	40
GBM	137	35
HNSC	452	114
KIRC	484	122
KIRP	258	65
LIHC	338	85
LUAD	460	116
LUSC	442	111
DLBC	38	10
MESO	69	18
OV	246	62
PAAD	146	37
PCPG	149	38
PRAD	440	110
READ	84	21
SKCM	379	95
STAD	360	90
TGCT	124	32
THYM	97	25
THCA	457	115
UCS	45	12
UVM	64	16

Table S4 - List of genes for each cancer type that showed combined precision and recall values greater than 0.75

Cancer Type	Genes
LAML	<i>TMEM54</i>
PRAD	<i>TARP</i>
LGG	<i>ATP1B2, ATPV1G2</i>
THCA	<i>FOXE1, PAX8</i>
KIRC	<i>PAX8, BHMT2, FASN</i>
LIHC	<i>MOSC2, HGD, DDR1</i>
BRCA	<i>DEFB132, CREB3L4, GDA, SIDT1, KLHL38</i>
KIRP	<i>KCP, RTDR1, GDA, C1orf21, CBFA2T3, ASNSD1</i>
STAD	<i>LCOR, EPS8L3, MUC13, TRIM31, ZDHHC20, PRSS3, EFNA2, GPN3, ASXL2, LATS1, PSMA6</i>
SKCM	<i>CABLES1, TBC1D16, PLOD3, VIPR1, SLC4A2, GOLGA2B, PTPN23, PLCH2, RAI1, HGD, SRP9, GJA1</i>

Talhouk RS, Fares MB, Rahme GJ, et al. Context dependent reversion of tumor phenotype by connexin-43 expression in MDA-MB231 cells and MCF-7 cells: role of β -catenin/connexin43 association. *Exp Cell Res.* 2013;319(20):3065-3080. doi:10.1016/j.yexcr.2013.10.002

Tanaka K, Kondo K, Kitajima K, Muraoka M, Nozawa A, Hara T. Tumor-suppressive function of protein-tyrosine phosphatase non-receptor type 23 in testicular germ cell tumors is lost upon overexpression of miR142-3p microRNA. *J Biol Chem.* 2013;288(33):23990-23999. doi:10.1074/jbc.M113.478891

Teleki I, Szasz AM, Maros ME, et al. Correlations of differentially expressed gap junction connexins Cx26, Cx30, Cx32, Cx43 and Cx46 with breast cancer progression and prognosis. *PLoS One.* 2014;9(11):e112541. Published 2014 Nov 10. doi:10.1371/journal.pone.0112541

Terrasi A, Bertolini I, Martelli C, et al. Specific V-ATPase expression sub-classes IDHW1 lower-grade gliomas and impacts glioma growth in vivo. *EBioMedicine.* 2019;41:214-224. doi:10.1016/j.ebiom.2019.01.052

Thomas M, Laura R, Hepner K, et al. Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation. *Oncogene.* 2002;21(33):5088-5096. doi:10.1038/sj.onc.1205668

Tiemin P, Fanzheng M, Peng X, et al. MUC13 promotes intrahepatic cholangiocarcinoma progression via EGFR/P13K/AKT pathways. *J Hepatol.* 2020;72(4):761-773. doi:10.1016/j.jhep.2019.11.021

Toth R, Schiffmann H, Hube-Maggs C, et al. Random forest-based modelling to detect biomarkers for prostate cancer progression. *Clin Epigenetics.* 2019;11(1):148. Published 2019 Oct 22. doi:10.1186/s13148-019-0736-8

Triki M, Ben Ayed-Guerfali D, Sagueim L, et al. RPI140 and LCoR expression in gastrointestinal cancers. *Oncotarget.* 2017;8(67):111161-111175. Published 2017 Nov 25. doi:10.18632/oncotarget.22686

Tsai CK, Huang LC, Tsaui WC, Huang SM, Lee JT, Huang DY. Overexpression of PLOD3 promotes tumor progression and poor prognosis in gliomas. *Oncotarget.* 2018;9(21):15705-15720. Published 2018 Feb 28. doi:10.18632/oncotarget.24594

Valiathan RR, Marco M, Leitingger B, Kleer CG, Fridman R. Discoidin domain receptor tyrosine kinases: new players in cancer progression. *Cancer Metastasis Rev.* 2012;31(1-2):295-321. doi:10.1007/s10555-012-9346-2

Valls BL, Rodriguez-Torres S, Kuhn E, et al. HIST1H2BB and MAGI2 Methylation and Somatic Mutations as Precision Medicine Biomarkers for Diagnosis and Prognosis of High-grade Serous Ovarian Cancer [published online ahead of print, 2020 Jun 24]. *Cancer Prev Res (Phila).* 2020;10.1158/1940-6207.CAPR-19-0412. doi:10.1158/1940-6207.CAPR-19-0412

van den Boom J, Wolter M, Blaschke B, Knobbe CB, Reifemberger G. Identification of novel genes associated with astrocytoma progression using suppression subtractive hybridization and real-time reverse transcription-polymerase chain reaction. *Int J Cancer.* 2006;119(10):2330-2338. doi:10.1002/ijc.22108

Velasquez-Torres G, Shoshan E, Ivan C, et al. A-to-I miR-378a-3p editing can prevent melanoma progression via regulation of PARVA expression. *Nat Commun.* 2018;9(1):461. Published 2018 Jan 31. doi:10.1038/s41467-018-02851-7

Verdanet E, Dereure O, René C, et al. Diagnostic value of STMN1, LMO2, HGAL, AID expression and lp36 chromosomal abnormalities in primary cutaneous B cell lymphomas. *Histopathology.* 2017;71(4):648-660. doi:10.1111/his.13279

Visca P, Sebastiani V, Botti C, et al. Fatty acid synthase (FAS) is a marker of increased risk of recurrence in lung carcinoma. *Anticancer Res.* 2004;24(6):4169-4173.

Vizzoso M, Ferreira HJ, Lopez-Serra P, et al. Epigenetic activation of a cryptic TBC1D16 transcript enhances melanoma progression by targeting EGFR. *Nat Med.* 2015;21(7):741-750. doi:10.1038/nm.3863

Vogelsang TLR, Schmeckel E, Kuhn C, et al. Regulation of LCoR and RPI140 expression in cervical intraepithelial neoplasia and correlation with CIN progression and deafferentiation. *J Cancer Res Clin Oncol.* 2020;146(7):1847-1855. doi:10.1007/s00432-020-03178-x

Wagner MP, Gunsalus KT, Schoenike B, Richardson AL, Friedl A, Roopra A. The transcription factor REST is lost in aggressive breast cancer. *PLoS Genet.* 2010;6(6):e1000979. Published 2010 Jun 10. doi:10.1371/journal.pgen.1000979

Walts AE, Said JW, Siegel MB, Banks-Schlegel S. Involucrin, a marker of squamous and urothelial differentiation. An immunohistochemical study on its distribution in normal and neoplastic tissues. *J Pathol.* 1985;145(4):329-340. doi:10.1002/path.1711450406

Wang C, Yang Y, Yin L, et al. Novel Potential Biomarkers Associated With Epithelial to Mesenchymal Transition and Bladder Cancer Prognosis Identified by Integrated Bioinformatic Analysis. *Front Oncol.* 2020;10:931. Published 2020 Jun 30. doi:10.3389/fonc.2020.00931

Wang F, Hu YL, Feng Y, et al. High-level expression of PRSS3 correlates with metastasis and poor prognosis in patients with gastric cancer. *J Surg Oncol.* 2019;119(8):1108-1121. doi:10.1002/jso.25448

Wang LL, Zhang M, miR-582-5p is a potential prognostic marker in human non-small cell lung cancer and functions as a tumor suppressor by targeting MAP3K2. *Eur Rev Med Pharmacol Sci.* 2018;22(22):7760-7767. doi:10.26355/eurrev_201811_16397

Wang M, Ma H, Pan Y, et al. PAX2 and PAX8 reliably distinguishes ovarian serous tumors from mucinous tumors. *Appl Immunohistochem Mol Morphol.* 2015;23(4):280-287. doi:10.1097/PAL0000000000000065

Wang T, Zhao L, Yang Y, et al. EGR1 is critical for gastrin-dependent upregulation of anion exchanger 2 in gastric cancer cells. *FEBS J.* 2013;280(1):174-183. doi:10.1111/febs.12058

Wang Y, Li H, Ma J, et al. Integrated Bioinformatics Data Analysis Reveals Prognostic Significance Of SIDT1 In Triple-Negative Breast Cancer. *Oncotargets Ther.* 2019;12:8401-8410. Published 2019 Oct 11. doi:10.2147/OTT.S215898

Wang Y, Wang Y, Li J, et al. PAX8: a sensitive and specific marker to identify cancer cells of ovarian origin for patients prior to neoadjuvant chemotherapy. *J Hematol Oncol.* 2013;6:60. Published 2013 Aug 19. doi:10.1186/1756-8722-6-60

Westbrook TF, Martin ES, Schlabach MR, et al. A genetic screen for candidate tumor suppressors identifies REST. *Cell.* 2005;121(6):837-848. doi:10.1016/j.cell.2005.03.033

Wolfgang CD, Essand M, Lee B, Pastan I. T-cell receptor gamma chain alternate reading frame protein (TARF) expression in prostate cancer cells leads to an increased growth rate and induction of caveolins and amphiregulin. *Cancer Res.* 2001;61(22):8122-8126.

Wolfgang CD, Essand M, Vincent JJ, Lee B, Pastan I. TARF: a nuclear protein expressed in prostate and breast cancer cells derived from an alternate reading frame of the T cell receptor gamma chain locus. *Proc Natl Acad Sci U S A.* 2000;97(17):9437-9442. doi:10.1073/pnas.160270597

Wouters J, Vizzoso M, Martinez-Cardus A, et al. Comprehensive DNA methylation study identifies novel progression-related and prognostic markers for cutaneous melanoma. *BMC Med.* 2017;15(1):101. Published 2017 Jun 5. doi:10.1186/s12916-017-0851-3

Wu D, Wang Y, Yang G, et al. A novel mitochondrial amidoxime reducing component 2 is a favorable indicator of cancer and suppresses the progression of hepatocellular carcinoma by regulating the expression of p27 [published online ahead of print, 2020 Aug 18]. *Oncogene.* 2020;10.1038/s41388-020-01417-6. doi:10.1038/s41388-020-01417-6

Wu G, Guo Z, Chatterjee A, et al. Overexpression of glycosylphosphatidylinositol (GPI) transamidase subunit phosphatidylinositol glycan class T and/or GPI anchor attachment 1 induces tumorigenesis and contributes to invasion in human breast cancer. *Cancer Res.* 2006;66(20):9829-9836. doi:10.1158/0008-5472.CAN-06-0506

Wu L, Peng CW, Hou JX, et al. Corin-1C is a novel biomarker for hepatocellular carcinoma invasive progression identified by proteomics analysis and clinical validation. *J Exp Clin Cancer Res.* 2010;29(1):17. Published 2010 Feb 24. doi:10.1186/1756-9966-29-17

Xia H, Qi H, Li Y, et al. LATS1 tumor suppressor regulates G2/M transition and apoptosis. *Oncogene.* 2002;21(8):1233-1241. doi:10.1038/sj.onc.1205174

Xiang L, Kong B. PAX8 is a novel marker for differentiating between various types of tumor, particularly ovarian epithelial carcinomas. *Oncol Lett.* 2013;5(3):735-738. doi:10.3892/ol.2013.1121

Xie D, Li J, Wei S, et al. Knockdown of PLOD3 suppresses the malignant progression of renal cell carcinoma via reducing TWIST1 expression [published online ahead of print, 2020 Jun 22]. *Mol Cell Probes.* 2020;101608. doi:10.1016/j.mcp.2020.101608

Xu K, Mao X, Mehta M, et al. Elucidation of low cancer cells avoid acidosis through comparative transcriptomic data analysis. *PLoS One.* 2013;8(8):e71177. Published 2013 Aug 14. doi:10.1371/journal.pone.0071177

Xu X, Zhu H, Yang M, et al. Knockdown of TOR signaling pathway regulator suppresses cell migration and invasion in non-small cell lung cancer via the regulation of epithelial-to-mesenchymal transition. *Exp Ther Med.* 2020;19(3):1925-1932. doi:10.3892/etm.2019.8358

Xu Y, Wu G, Li J, et al. Screening and Identification of Key Biomarkers for Bladder Cancer: A Study Based on TCGA and GEO Data. *Biomed Res Int.* 2020;2020:8283401. Published 2020 Jan 23. doi:10.1155/2020/8283401

Xuan Z, Zhao L, Li Z, et al. EPRS3L3 promotes hepatocellular carcinoma proliferation and metastasis by modulating EGFR dimerization and internalization. *Am J Cancer Res.* 2020;10(1):60-77. Published 2020 Jan 1.

Yan HX, Yang W, Zhang R, et al. Protein-tyrosine phosphatase PCP-2 inhibits beta-catenin signaling and increases E-cadherin-dependent cell adhesion. *J Biol Chem.* 2006;281(22):15423-15433. doi:10.1074/jbc.M602607200

Yan K, Hu C, Liu C, Chu G, Wang X, Mac S, Li L. Knockdown of TMPPRSS11D inhibits the proliferation, migration and invasion of cervical cancer cells. *RSC Adv.* 2019; 9:21591

Yang X, Yu K, Hao Y, et al. LATS1 tumour suppressor affects cytokinesis by inhibiting LIMK1. *Nat Cell Biol.* 2004;6(7):609-617. doi:10.1038/ncb1140

Yang Y, Wu PP, Wu J, et al. Expression of anion exchanger 2 in human gastric cancer. *Exp Oncol.* 2008;30(1):81-87.

Yang Z, Shen H, He W, et al. Expression of TBC1D16 Is Associated with Favorable Prognosis of Epithelial Ovarian Cancer. *Tohoku J Exp Med.* 2018;245(3):141-148. doi:10.1620/tjem.245.141

Yin F, Shi L, Liu X, et al. Microarray-based identification of genes associated with cancer progression and prognosis in hepatocellular carcinoma. *J Exp Clin Cancer Res.* 2016;35(1):127. Published 2016 Aug 27. doi:10.1186/s13046-016-0403-2

Yoshihara K, Wang Q, Torres-Garcia W, et al. The landscape and therapeutic relevance of cancer-associated transcription fusions. *Oncogene.* 2015;34(37):4845-4854. doi:10.1038/onc.2014.406

Yu C, Chen S, Guo Y, Sun C. Oncogenic TRIM31 confers gemcitabine resistance in pancreatic cancer via activating the NF- κ B signaling pathway. *Theranostics.* 2018;8(12):3224-3236. Published 2018 May 11. doi:10.7150/thno.23259

Zakrzewska M, Zakrzewski K, Gręśner SM, Piaskowski S, Zalewska-Szewczyk B, Liberski PP. Polycyomb genes expression as a predictor of poor clinical outcome in children with medulloblastoma. *Childs Nerv Syst.* 2011;27(1):79-86. doi:10.1007/s00381-010-1260-5

Zeng CX, Tang LY, Xie CY, et al. Overexpression of EPRS3L3 promotes cell proliferation by inhibiting the transactivity of FOXO1 in HCC. *Neoplasma.* 2018;65(5):701-707. doi:10.4149/neo_2018_170725NS03

Zhang Q, Han Q, Zi J, Song C, Ge Z. CD37 high expression as a potential biomarker and association with poor outcome in acute myeloid leukemia. *Biosci Rep.* 2020;40(5):BSR20200008. doi:10.1042/BSR20200008

Zhang R, Gao X, Zuo J, et al. STMN1 upregulation mediates hepatocellular carcinoma and hepatic stellate cell crosstalk to aggravate cancer by triggering the MET pathway. *Cancer Sci.* 2020;111(2):406-417. doi:10.1111/cas.14262

Zhao L, Chi W, Cao H, et al. Screening and clinical significance of tumor markers in head and neck squamous cell carcinoma through bioinformatics analysis. *Mol Med Rep.* 2019;19(1):143-154. doi:10.3892/mmr.2018.9639

Zhao L, Yu Z, Zhao B. Mechanism of VPR1 gene regulating human lung adenocarcinoma H1299 cells. *Med Oncol.* 2019b;36(11):91. Published 2019 Sep 27. doi:10.1007/s12032-019-1312-y

Zheng K, Chen Z, Feng H, et al. Sphingomyelin synthase 2 promotes an aggressive breast cancer phenotype by disrupting the homeostasis of ceramide and sphingomyelin. *Cell Death Dis.* 2019;10(3):157. Published 2019 Feb 15. doi:10.1038/s41419-019-1303-0

Zhou J, Fogelgren B, Wang Z, Roe BA, Biegel JA. Isolation of genes from the rhabdoid tumor deletion region in chromosome band 22q11.2. *Gene.* 2000;241(1):133-141. doi:10.1016/s0378-1119(99)00456-4

Zhou L, Deng ZZ, Li HY, et al. TRIM31 promotes glioma proliferation and invasion through activating NF- κ B pathway. *Oncotargets Ther.* 2019;12:2289-2297. Published 2019 Mar 27. doi:10.2147/OTT.S183625

Zhou M, Fan J, Li Z, et al. Prognostic impact of tumor mutation burden and the mutation in KIAA1211 in small cell lung cancer. *Respir Res.* 2019;20(1):248. Published 2019 Nov 7. doi:10.1186/s12931-019-1205-9