

Gene mutation and expression data from 10k patients and 26 cancers types for pan-cancer analysis

BioMuta and BioXpress – EDRN Portal

Raja Mazumder mazumder@gwu.edu

Assoc. Prof. Biochemistry and Molecular Medicine
Director, The McCormick Genomic & Proteomic Center
GWU



Outline

I
N
F
R
A
S
T
R
U
C
T
U
R
E

High-Performance Integrated Virtual Environment (HIVE) Tools and Applications for Big Data Analysis

Vahan Simonyan^{1,*} and Raja Mazumder^{2,*} Genes 2014, 5, 957-981

A framework for organizing cancer-related variations from existing databases, publications and NGS data using a High-performance Integrated Virtual Environment (HIVE) Database, Vol. 2014, Article ID bau022

Tsung-Jung Wu¹, Amirhossein Shamsaddini¹, Yang Pan¹, Krista Smith¹, Daniel J. Crichton², Vahan Simonyan³ and Raja Mazumder^{1,4,*}

BioXpress: an integrated RNA-seq-derived gene expression database for pan-cancer analysis

Quan Wan¹, Hayley Dingerdissen¹, Yu Fan¹, Naila Gulzar¹, Yang Pan¹, Tsung-Jung Wu¹, Cheng Yan¹, Haichen Zhang¹ and Raja Mazumder^{1,2,*}

Human germline and pan-cancer variomes and their distinct functional profiles

Yang Pan^{1,†}, Konstantinos Karagiannis^{1,†}, Haichen Zhang¹, Hayley Dingerdissen¹, Amirhossein Shamsaddini¹, Quan Wan¹, Vahan Simonyan² and Raja Mazumder^{1,3,*}

Nucleic Acids Research, 2014, Vol. 42, No. 18

Knowledge discovery

Cancer Disease Ontology

Page 1 of 31

Manuscripts submitted to Database

In press

1
2
3
4
5
6
7
8
9
10
11
12
13

Generating a focused view of Disease Ontology cancer terms for pan-cancer data integration and analysis.

Tsung-Jung Wu^{1#}, Lynn M. Schriml^{2#}, Qing-Rong Chen³, Maureen Colbert⁴, Daniel J. Crichton⁴, Richard Finney³, Ying Hu³, Warren A. Kibbe³, Heather Kincaid⁴, Daoud Meerzaman³, Elvira Mitraka², Yang Pan¹, Krista M. Smith¹, Sudhir Srivastava⁵, Sari Ward⁶, Cheng Yan¹ and Raja Mazumder^{*1,7}

#patient samples in BioMuta (curated cancer mutation database)

CancerDiseaseOntologySLIM	TCGA	ICGC
DOID:3953 / adrenal gland cancer	92	–
DOID:8618 / oral cavity cancer	–	50
DOID:5041 / esophageal cancer	–	183
DOID:4362 / cervical cancer	198	–
DOID:4159 / skin cancer	370	323
DOID:363 / uterine cancer	305	–
DOID:3571 / liver cancer	202	511
DOID:3070 / malignant glioma	291	268
DOID:263 / kidney cancer	685	560
DOID:2531 / hematologic cancer	197	394
DOID:2394 / ovarian cancer	318	181
DOID:219 / colon cancer	217	216
DOID:1993 / rectum cancer	81	185
DOID:184 / bone cancer	–	66
DOID:1793 / pancreatic cancer	147	504
DOID:1781 / thyroid cancer	404	411
DOID:1612 / breast cancer	977	1071
DOID:1324 / lung cancer	178	289
DOID:1319 / brain cancer	289	527
DOID:11934 / head and neck cancer	508	–
DOID:1192 / peripheral nervous system neoplasm	–	41
DOID:11054 / urinary bladder cancer	130	233
DOID:10534 / stomach cancer	289	298
DOID:10283 / prostate cancer	261	275

#patient samples in BioXpress (curated cancer expression database)

CancerDiseaseOntologySLIM	TCGA	ICGC
DOID:3459_Breast Invasive Carcinoma	113	-
DOID:4467_Kidney Renal Clear Cell Carcinoma	72	-
DOID:3963_Thyroid Carcinoma	60	486
DOID:3910_Lung Adenocarcinoma	58	-
DOID:2526_Prostate Adenocarcinoma	51	-
DOID:684_Liver Hepatocellular Carcinoma	50	-
DOID:3907_Lung Squamous Cell Carcinoma	50	-
DOID:1749_Head And Neck Squamous Cell Carcinoma	41	397
DOID:3717_Stomach Adenocarcinoma	31	-
DOID:4465_Kidney Renal Papillary Cell Carcinoma	30	159
DOID:234_Colon Adenocarcinoma	26	434
DOID:4471_Kidney Chromophobe	25	-
DOID:6975_Bladder Urothelial Carcinoma	19	-
DOID:2871_Uterine Corpus Endometrial Carcinoma	17	-
DOID:1107_Esophageal Carcinoma	10	-
DOID:1996_Rectum Adenocarcinoma	6	150
DOID:1793_Pancreatic Cancer	-	147
DOID:11054_Bladder Cancer	-	204
DOID:0060058_Lymphoma	-	50
DOID:2394_Ovarian Cancer	-	336
DOID:1324_Lung Cancer	-	871
DOID:1240_Leukemia	-	173
DOID:4467_Renal Clear Cell Carcinoma	-	510
DOID:1380_Endometrial Cancer	-	484
DOID:1612_Breast Cancer	-	995
DOID:10283_Prostate Cancer	-	197
DOID:3571_Liver Cancer	-	141
DOID:1319_Brain Cancer [Brca]	-	457
DOID:8923_Skin Melanoma [SKCM]	-	66
DOID:3744_Cervical Squamous Cell Carcinoma [CESC]	-	123
DOID:10165_Endocrine Pancreas Cancer [PAEN]	-	17

EDRN portal

You are here: [Home](#) / [Biomarkers](#) / CA125

CA125

Basics | [Organs](#) | [Studies](#) | [Publications](#) | [Resources](#)

Aliases:

This biomarker is also known as:
CA125 ovarian cancer antigen, Mucin 16, MUC16, CA-125, Mucin-16, mucin 16, cell surface associated, Ovarian carcinoma antigen CA125, FLJ14303, MUC-16, Ovarian cancer-related tumor marker CA125,

[View in BioMuta](#)

ATTRIBUTES

QA State:	Accepted
Type:	Protein
Short Name:	
HGNC Name:	MUC16

Description...

MUC16 (CA125) is a h expressed on epithelial c cells. MUC16 is anchored domain and is released i cleavage. It is thought to against particles and infectious agents at mucosal surfaces.

Datasets

- [PLCO Phase III Dataset](#)
- [Pre-PLCO Phase II Dataset](#)

Navigation Menu:

- Network Consulting Team
- Informatics
- Collaborative Groups
- Secure Site
- Public, Patients, Advocates
- Funding Opportunities
- Sites
- Member Directory
- Committees
- Biomarker Informatics Standards
- Division of Cancer

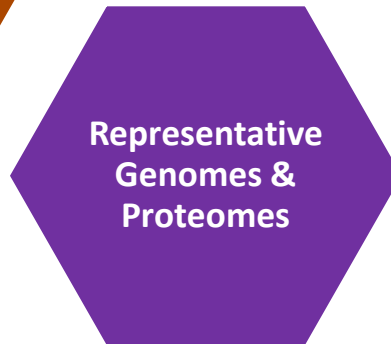
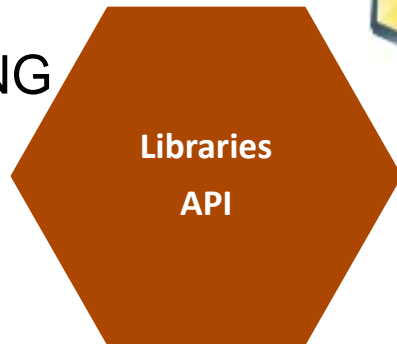
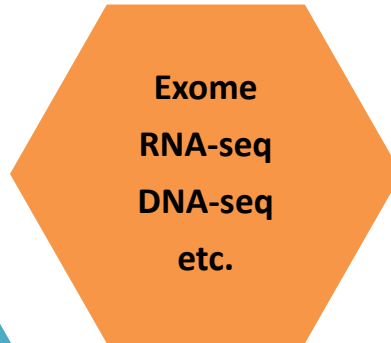
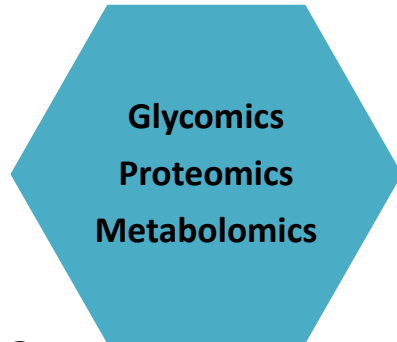
Mutations

The <MUC16> protein has <26> mutation sites from <7> cancer types. This data has been integrated from COSMIC, IntOGen, TCGA, ICGC, ClinVar, CSR and <4> publications. <5> patient samples with <7> mutations in <MUC16> has NGS and associated metadata available for reanalysis.

[View in BioMuta.](#)

HIVE tools + Community tools

GOAL: FDA
COMPLIANT
INTEGRATED
NGS & PROTEOMIC
DATA
STORAGE, SHARING
AND ANALYSIS
PLATFORM



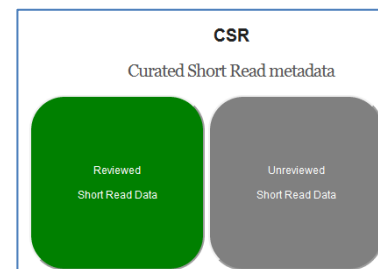
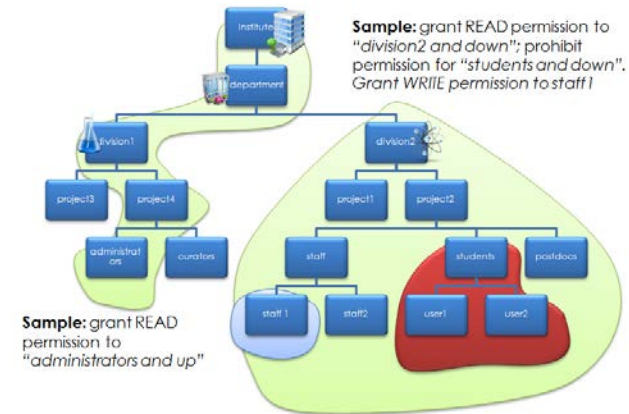
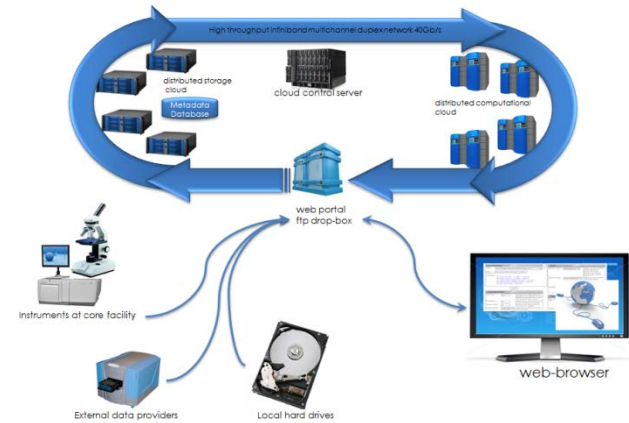
2 portals: Public portal
(GW) & a FDA only portal

References + Standards in collaboration with community



HIVE key features and focus

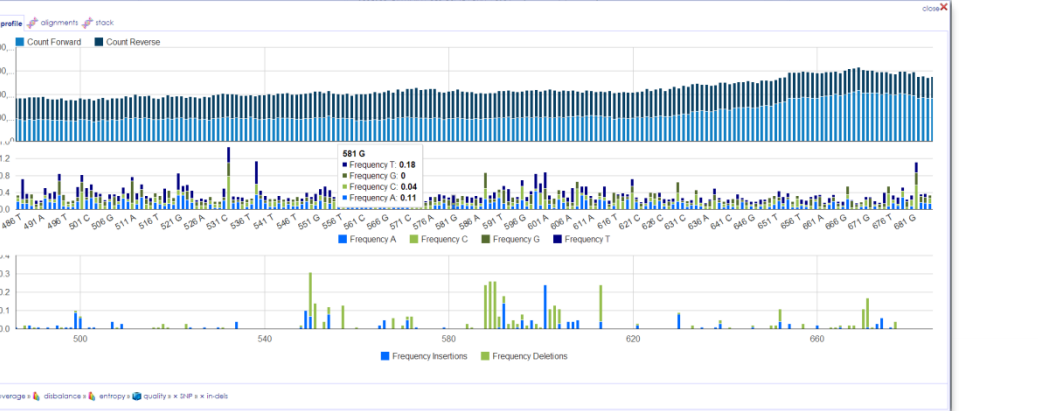
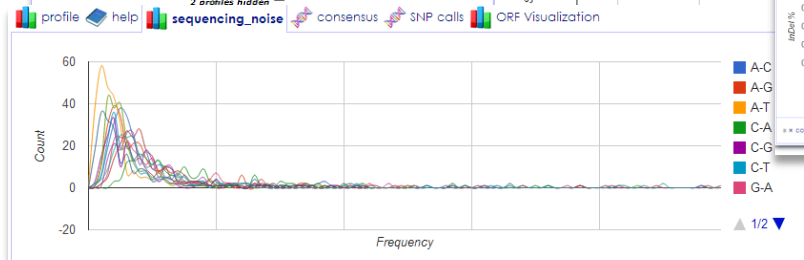
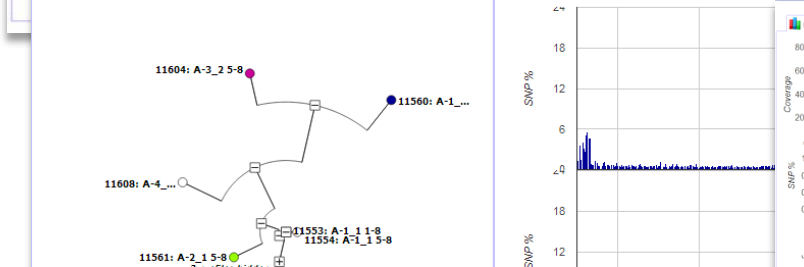
- Distributed storage and computing
- Security features (FDA/industry/patient data in mind).
Authorized to Operate (ATO) in Regulatory Environment
- Granular sharing capabilities
- Biocuration of content
- Compute speed



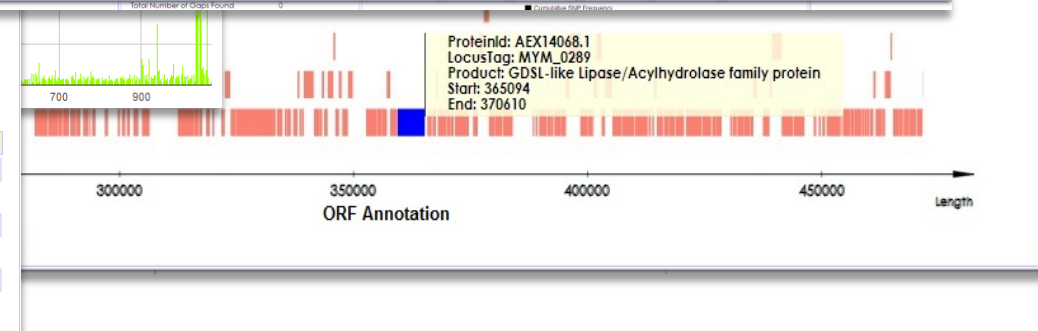
BioMuta v2.0
Older version - BioMuta v1.0

BioXpress v1.0

HIVE: visualizations



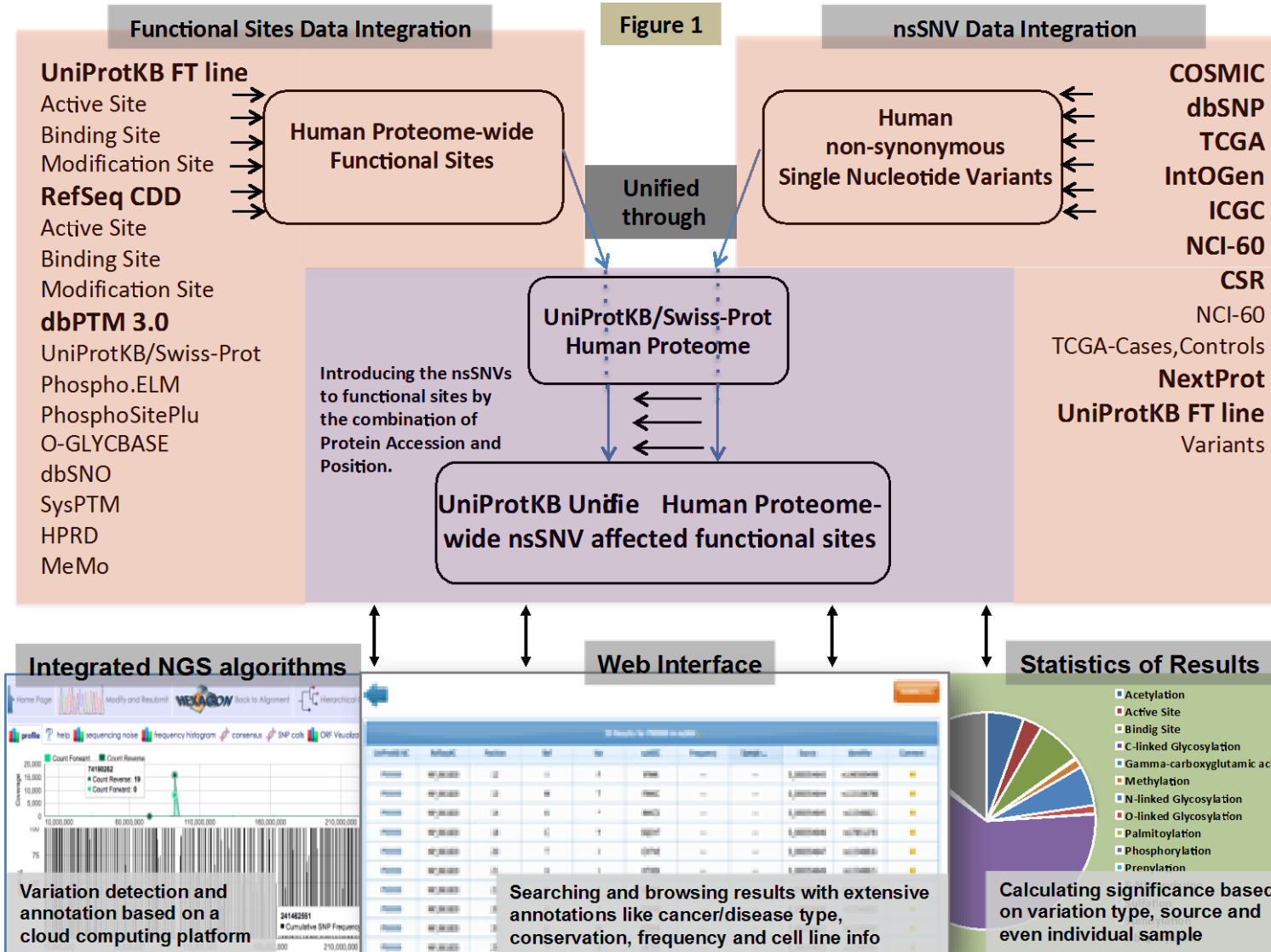
Threshold	A-C	A-G	A-T	C-A	C-G	C-T	G-A	G-C	G-T	T-A	T-C	T-G
50 %	0.0007	0.0007	0.0003	0.0012	0.0006	0.0006	0.0008	0.0005	0.0008	0.0004	0.0007	0.0007
75 %	0.0012	0.0012	0.0005	0.0018	0.0008	0.0010	0.0013	0.0007	0.0011	0.0008	0.0012	0.0011
85 %	0.0017	0.0015	0.0009	0.0028	0.0011	0.0014	0.0019	0.0009	0.0015	0.0012	0.0014	0.0015
90 %	0.0023	0.0020	0.0011	0.0033	0.0013	0.0016	0.0021	0.0011	0.0020	0.0018	0.0018	0.0020
95 %	0.0037	0.0028	0.0017	0.0043	0.0027	0.0023	0.0038	0.0014	0.0033	0.0020	0.0022	0.0030
99 %	0.0066	0.0083	0.0041	0.0070	0.0039	0.0053	0.0068	0.0031	0.0077	0.0057	0.0042	0.0090



BioMuta

- BioMuta integrates all cancer related non-synonymous mutations.
- More than 2.6 million high quality nsSNVs have been integrated into database (largest curated dataset that we know of)
- Disease Ontology nomenclature applied by BioMuta provides more accurate disease description allowing pan-cancer analysis
- 381 DO terms and 54 DO slim terms are included in BioMuta

Workflow



Incomplete variation information

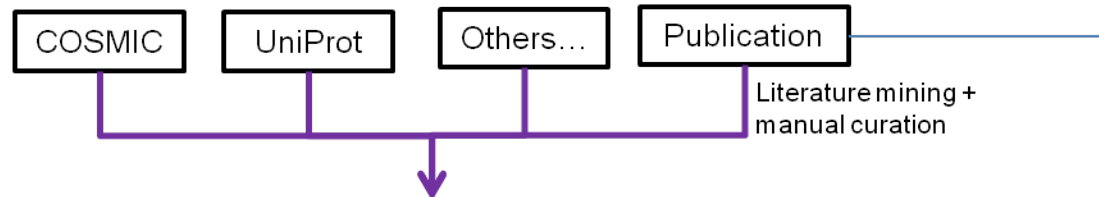
Gene/Protein accession/Gene name	Genomic coordinates	Variation Gene/Protein (position)	Cancer definition	PMID	source
NM_130800.2 O00255 MEN1	64575133-64575133 (chr 11)	C A (1193); G V (230)	Lung, upper right lung, mucous cell, carcinoma	---	COSMIC
---- P40637 TP53	chr17:7579866	---; Q239L	Sporadic cancer	14660012	UniProt
NM_77692.4 ---- TP53	Chr17(7757534)	----; ----	Cancer	1791428	Manual
NM_533167.1 O20147 ---	----	2133(T G); G703P	Pancreas	31229574	IntOGen

DO and DO slim

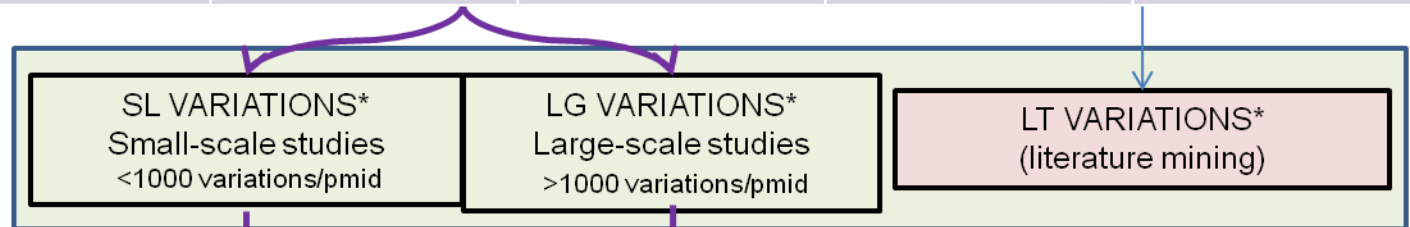
Source	Original cancer term	DOID / DO term	DO_slim
IntOGen	Pancreas	DOID:1793 / pancreatic cancer	DOID:1793 / pancreatic cancer
TCGA	Pancreatic adenocarcinoma [PAAD]	DOID:4074 / pancreas adenocarcinoma	
COSMIC	pancreas,NS,carcinoma, acinar_carcinoma	DOID:5742 / pancreatic acinar cell adenocarcinoma	
UniProt	Pancreatic cancer	DOID:1793 / pancreatic cancer	

- DO provides accurate disease description for all cancer term
- DO slim group several terms of DO
- DO slim is easy for later analysis

BioMuta workflow



Gene/Protein accession/Gene name	Genomic coordinates	Variation Gene/Protein (position)	Disease Ontology term/ ID Polyphen	PMID	Source / scale
NM_130800.2 O00255 MEN1	64575133-64575133 (chr 11)	1193 C A 230 G V	lung squamous cell carcinoma / 3907 probably damage	2383612	COSMIC / LG

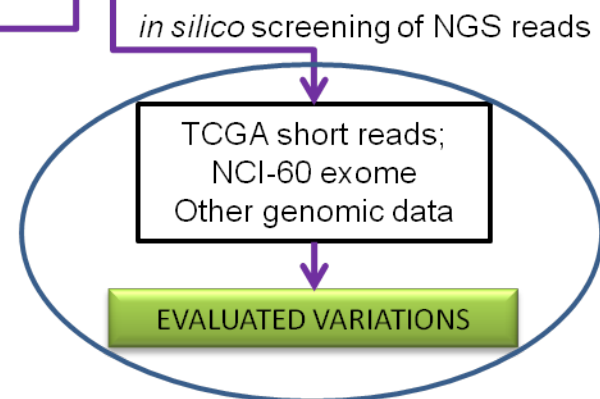


→ Phase I: Ongoing

→ Phase II: Future plans

*Variations associated with multiple cancer types can be viewed

EDRN Portal





BioMuta SNV table

UCSC Genome Bioinformatics



Source

DOWNLOAD

1450 Results for MUC16 in BioMuta v2.0

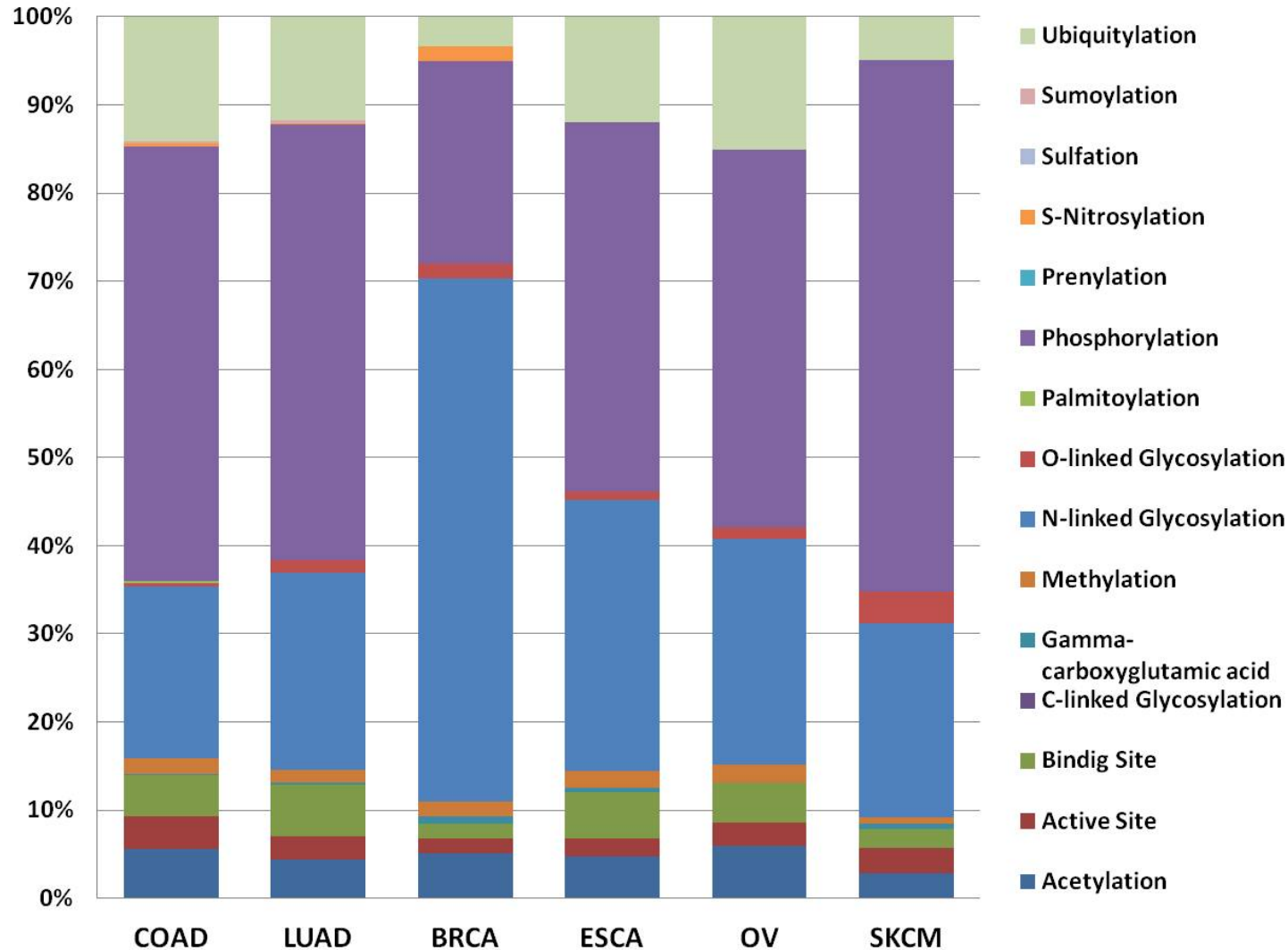
UniProtKB A	Gene	Accession	SNV position	Pos(N)	Ref(N)	Var(N)	Pos(A)	Ref(A)	Var(A)	Polyphen Pred.	PMID	Disease Ontology Name	Source	Status	Comment
Q8WXI7	MUC16	NM_02...	chr19:9091786-90...	233	G	T	10	S	*	-	20111...	DOID:1324 / lung cancer	IntOGen	LG	
Q8WXI7	MUC16	NM_02...	chr19:9091771-90...	248	C	T	15	R	H	benign	20393...	DOID2871 / endometri...	ICGC	LG	
Q8WXI7	MUC16	NM_02...	chr19:9091756-90...	263	C	T	20	G	E	benign	20393...	DOID3907 / lung squa...	ICGC	LG	
Q8WXI7	MUC16	NM_02...	chr19:9091636-90...	383	G	A	60	P	L	benign	20393...	DOID3907 / lung squa...	ICGC	LG	
Q8WXI7	MUC16	NM_02...	chr19:9091636-90...	383	G	A	60	P	L	benign	20111...	DOID:1324 / lung cancer	IntOGen	LG	
Q8WXI7	MUC16	NM_02...	chr19:9091606-90...	413	G	A	70	S	L	benign	20393...	DOID2871 / endometri...	ICGC	LG	
Q8WXI7	MUC16	NM_02...	chr19:9091579-90...	440	A	G	79	L	S	possibly dama...	24467...	DOID:3459 / breast car...	CSR-T...	LG	
Q8WXI7	MUC16	NM_02...	chr19:9091578-90...	441	C	A	79	L	F	possibly dama...	20393...	DOID1996 / rectum ad...	ICGC	LG	
Q8WXI7	MUC16	NM_02...	chr19:9091544-90...	475	A	G	91	S	P	benign	20393...	DOID3459 / breast car...	ICGC	LG	
Q8WXI7	MUC16	NM_02...	chr19:9091544-90...	475	A	G	91	S	P	benign	20111...	DOID:1612 / breast ca...	IntOGen	LG	

Swiss-Prot

RefSeq



Effect of mutations on PTMs

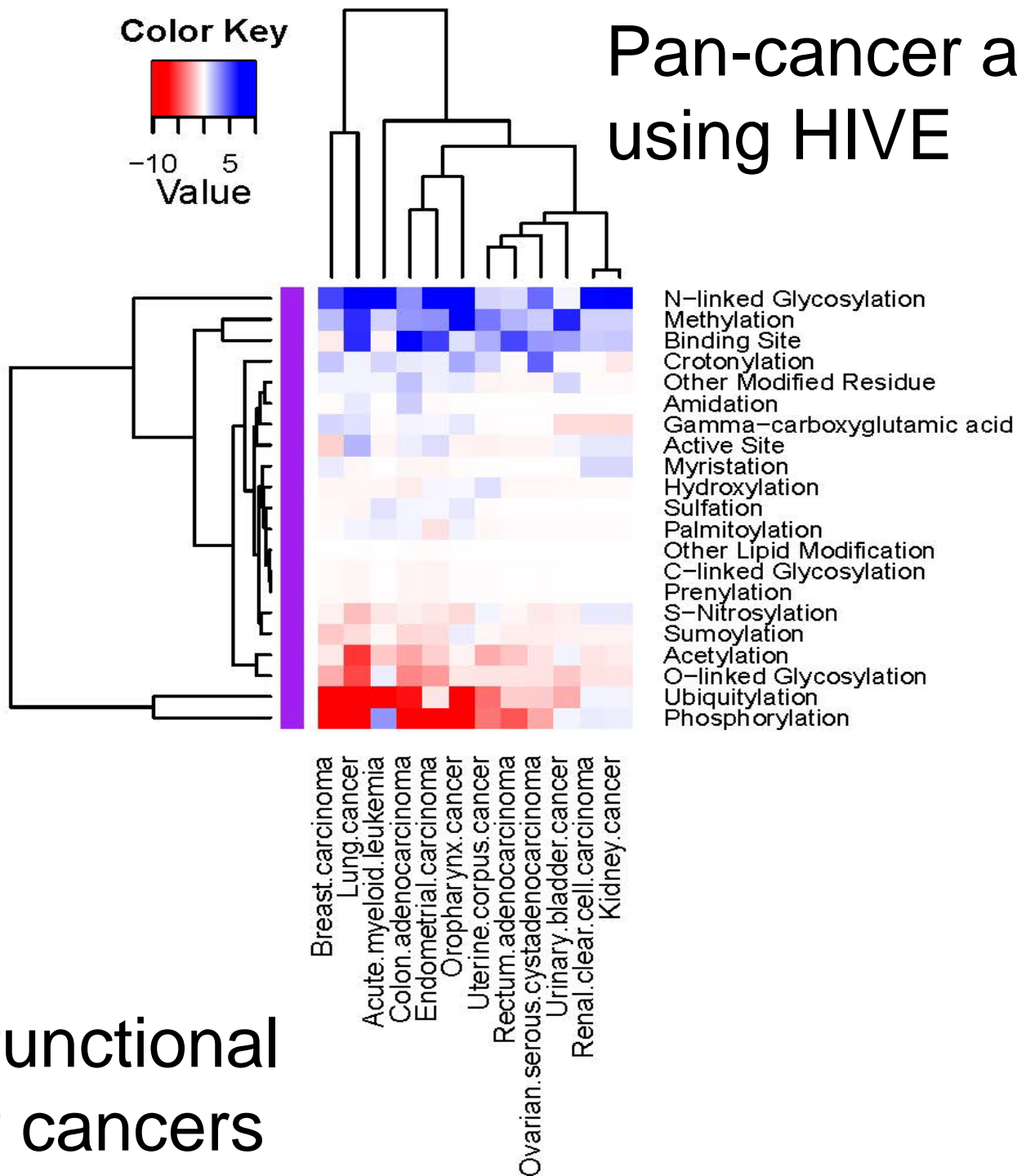


PTM: post-translational modification

Variant type based comparison of nsSNV's impact on different types of functional sites

	Total	+/- (Somatic) ¹	Not in dbSNP (Somatic)	+/- (Germline) ²	In dbSNP (Germline)
Acetylation	2.30E-30	-	4.39E-17	-	3.82E-15
Active Site	1.83E-04	+	1.56E-14	-	8.53E-05
Amidation	1.12E-01	+	2.28E-01	+	2.05E-01
Binding Site	4.81E-32	+	2.26E-110	-	7.86E-20
C-linked Glycosylation	1.68E-02	-	2.44E-01	-	1.09E-02
Crotonylation	2.89E-04	+	8.62E-07	-	2.69E-01
Gamma-carboxyglutamic acid	3.46E-08	+	1.80E-07	+	1.66E-02
Hydroxylation	4.42E-01	-	4.85E-01	-	4.90E-01
Methylation	2.24E-25	+	2.86E-28	+	4.84E-03
Myristation	4.85E-01	+	3.47E-01	-	2.26E-01
N-linked Glycosylation	4.53E-273	+	2.09E-118	+	3.95E-163
O-linked Glycosylation	8.62E-24	-	9.79E-22	-	7.98E-06
Other Lipid Modification	2.21E-01	-	5.03E-01	-	3.03E-01
Other Modified Residue	6.75E-04	+	5.24E-04	+	1.67E-01
Palmitoylation	6.88E-02	-	1.67E-01	-	1.73E-01
Phosphorylation	1.28E-110	-	6.85E-66	-	7.78E-47
Prenylation	4.94E-03	-	1.22E-02	-	1.64E-01
S-Nitrosylation	7.20E-05	-	2.24E-06	-	2.91E-01
Sulfation	5.15E-01	+	3.33E-01	-	3.84E-01
Sumoylation	7.41E-08	-	2.25E-03	-	8.66914E-07
Ubiquitylation	1.18E-136	-	1.57E-71	-	1.10915E-61

Pan-cancer analysis using HIVE

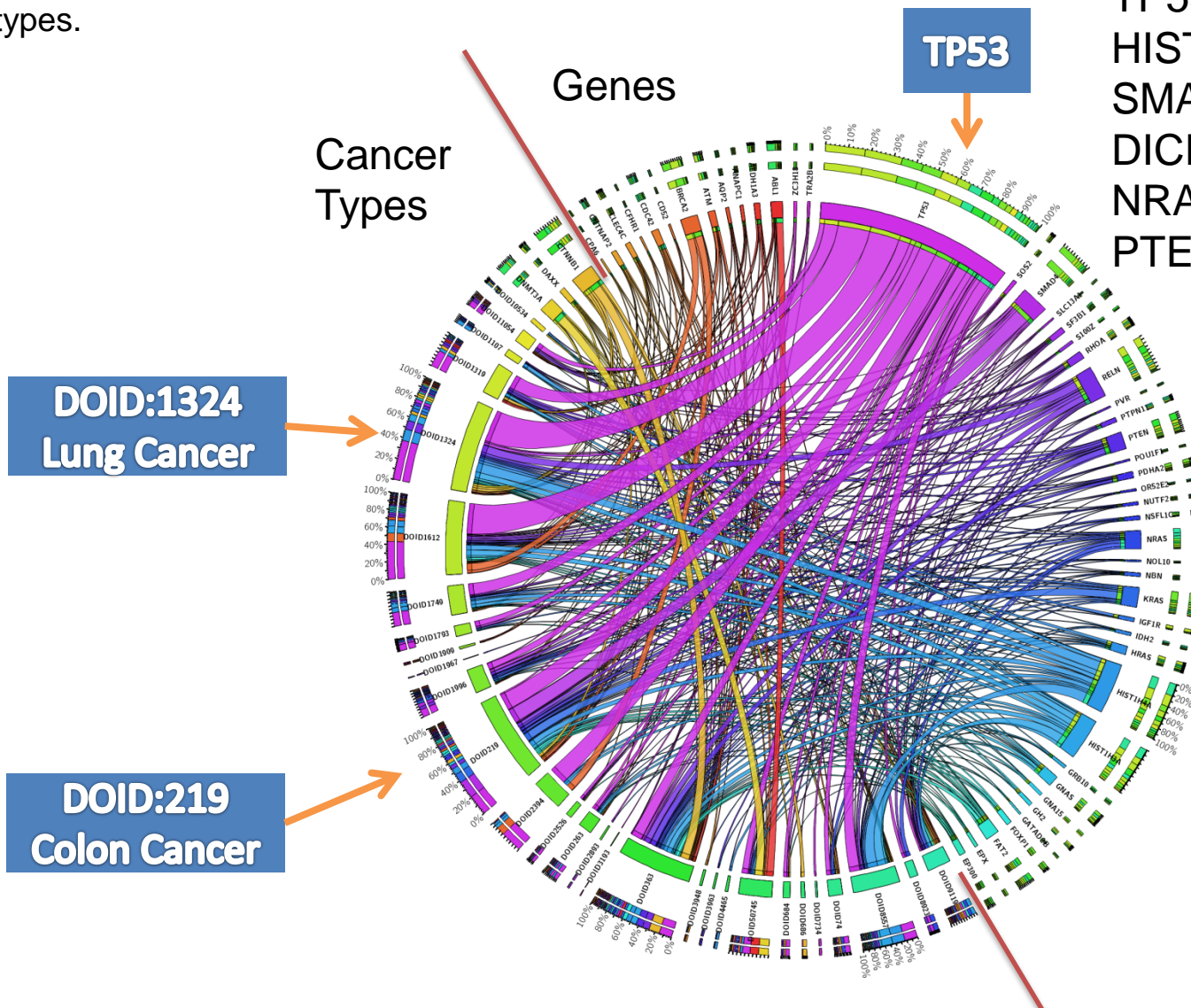


Creating functional profiles of cancers

Results

990 cancer-associated mutations from 51 genes containing mutations that are across 3 or more cancer types.

Top 10 out of 51 key genes:
TP53, HIST1H4A, HIST1H3A, RELN, SMAD4, CTNN81, DICER1, KRAS, NRAS, BRCA2 and PTEN



Priority targets
13 genes
106 mutations

HGNCGeneName	ProteinName	#byPosition	#byMutation
PPP6C PPP6	Serine/threonine-protein phosphatase 6 catalytic subunit	1	1
SF3B1 SAP155	Splicing factor 3B subunit 1	1	1
IDH1 PICD	Isocitrate dehydrogenase [NADP] cytoplasmic	1	4
NRAS HRAS1	GTPase Nras	1	3
HRAS HRAS1	GTPase Hras	1	2
KRAS KRAS2 RASK2	GTPase KRas	1	1
RAF1 RAF	RAF proto-oncogene serine/threonine-protein kinase	1	1
TP53 P53	Cellular tumor antigen p53	12	19
ITGA2B GP2B ITGAB	Integrin alpha-IIb	1	1
C1S	Complement C1s subcomponent	1	1
THRA EAR7 ERBA1 NR1A1 THRA1 THRA2	Thyroid hormone receptor alpha	1	1
PABPC1 PAB1 PABP1 PABPC2	Polyadenylate-binding protein 1	1	1
CFTR ABCC7	Cystic fibrosis transmembrane conductance regulator	1	1
RXRA NR2B1	Retinoic acid receptor RXR-alpha	1	1
MST1 D3F15S2 DNF15S2 HGFL	Hepatocyte growth factor-like protein	1	1

CTNNB1 CTNNB OK/SW-c1.35 PR02286

IDUA

IDH2

NLRP12 NALP12 PYPAF7 RNO

PTEN MMAC1 TEP1

GNAS GNAS1 GSP

Our criteria (≥ 2 DOs, ≥ 5 TCGA patientIDs, ≥ 1 Loss of Functional sites):

-they can be counted by # of distinct position on protein reference. e.g. TP53 position 31.

-they can be counted by # of distinct mutations on protein, e.g. TP53 position 31 A->S, TP53 position 31 A->H.

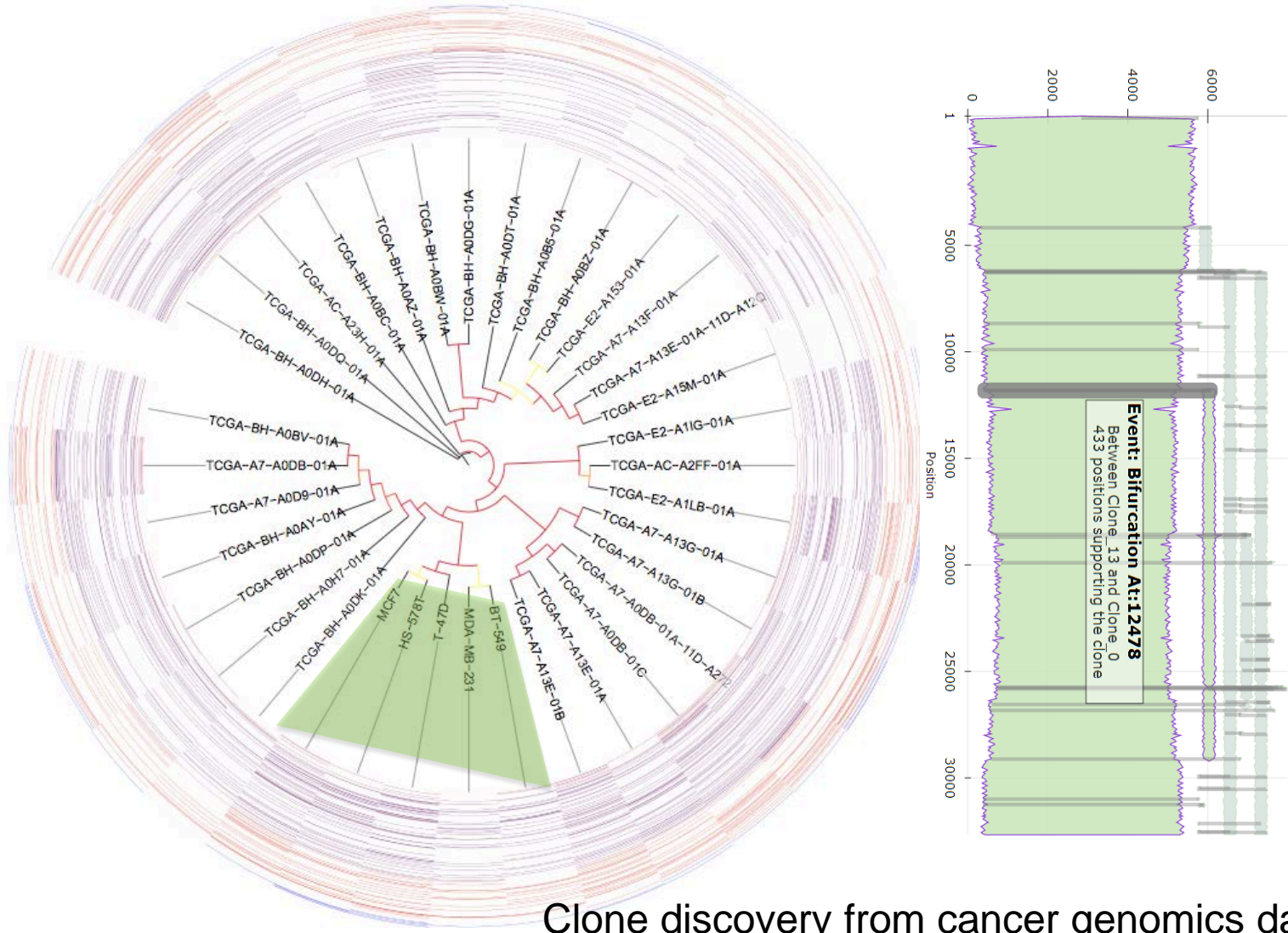
HIST1H3A H3FA; HIST1H3B H3FL; HIST1H3C H3FC; HIST1H3D H3FB; HIST1H3E H3FD; HIST1H3F H3FI; HIST1H3G H3FH; HIST1H3H H3FK; HIST1H3I H3FF; HIST1H3J H3FJ

MEF2A MEF2	Myocyte-specific enhancer factor 2A	1	1
PAK2	Serine/threonine-protein kinase PAK 2	3	3
SMAD4 DPC4 MADH4	Mothers against decapentaplegic homolog 4	1	2
TDG	G/T mismatch-specific thymine DNA glycosylase	1	1
GNAS GNAS1	Guanine nucleotide-binding protein G(s) subunit alpha isoforms Xlas	1	2
ERICH6B FAM194B	Glutamate-rich protein 6B	1	1
DND1 RBMS4	Dead end protein homolog 1	1	1
ZNF563	Zinc finger protein 563	1	1
PRG4 MSF SZP	Proteoglycan 4	1	1
ANAPC1 TSG24	Anaphase-promoting complex subunit 1	1	1
BCLAF1 BTF KIAA0164	Bcl-2-associated transcription factor 1 (Btf)	1	1
SEC63 SEC63L	Translocation protein SEC63 homolog	1	1
JAG2	Protein jagged-2 (Jagged2) (hJ2)	1	1

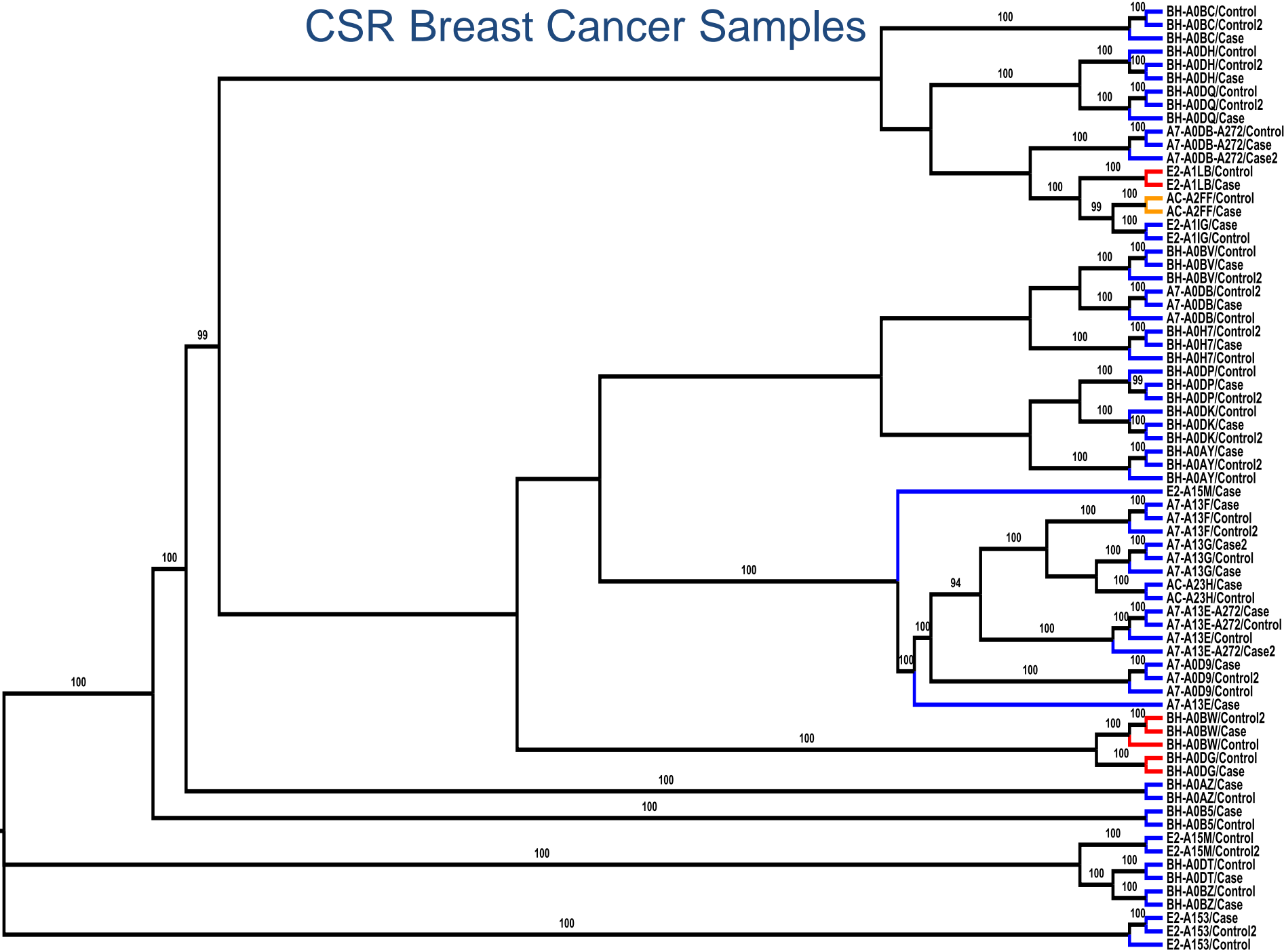
Whole genome single-nucleotide variation profile-based phylogenetic tree building methods for analysis of viral, bacterial and human genomes

William J. Faison ^{a,1}, Alexandre Rostovtsev ^{b,1}, Eduardo Castro-Nallar ^c, Keith A. Crandall ^c, Konstantin Chumakov ^b, Vahan Simonyan ^b, Raja Mazumder ^{a,d,*}

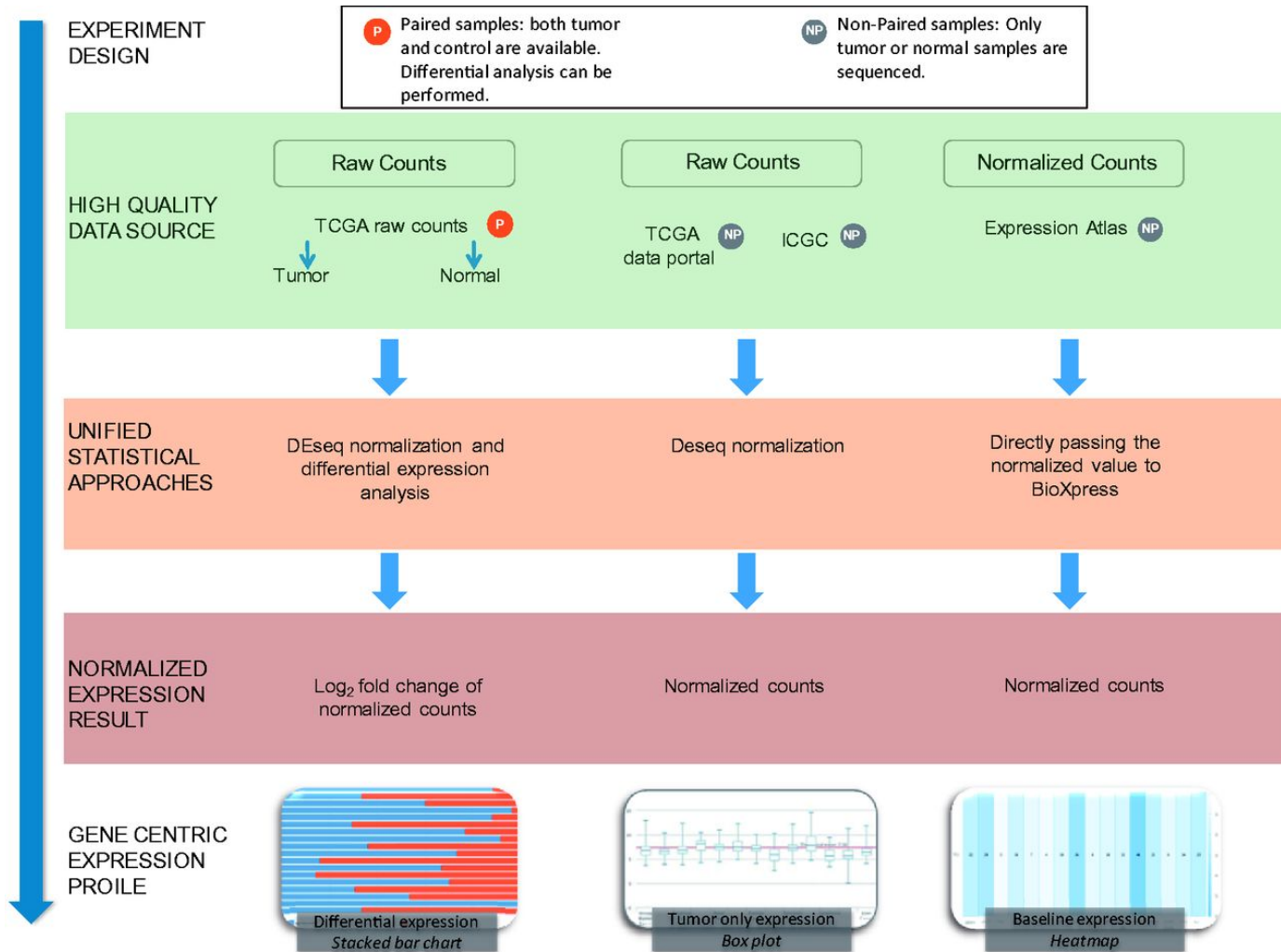
Phylogenetic tree of the whole exome sequencing results using PhyloSNP



CSR Breast Cancer Samples



Flow chart of the workflow used to create BioXpress.




Quan Wan et al. Database 2015;2015:bav019




BioXpress: an integrated RNA-seq-derived gene expression database for pan-cancer analysis

Table 1.
Statistics of data collected in BioXpress

Source	Data type	No. of samples/individuals ^a	Tumor/normal
TCGA	Raw read count	1320/660 ^b	Tumor and normal
ICGC and TCGA	Raw read count	6397/6324	Tumor
Expression Atlas baseline	Normalized count	1/1	Normal
Literature	Published literature	Not applicable (135 publications)	Tumor and normal comparison

 ^aTypically, each patient contains more than one sequencing sample. Therefore, we provide the number of both samples and individuals.

 ^bThe number of patients is collected from TCGA, ICGC and Expression Atlas baseline projects. Some TCGA patient IDs overlap with the ICGC patient IDs.

Snapshot of BioXpress interface.

BioXpress v1.0

BioXpress is a curated gene expression and disease association database where the expression levels are mapped to genes.

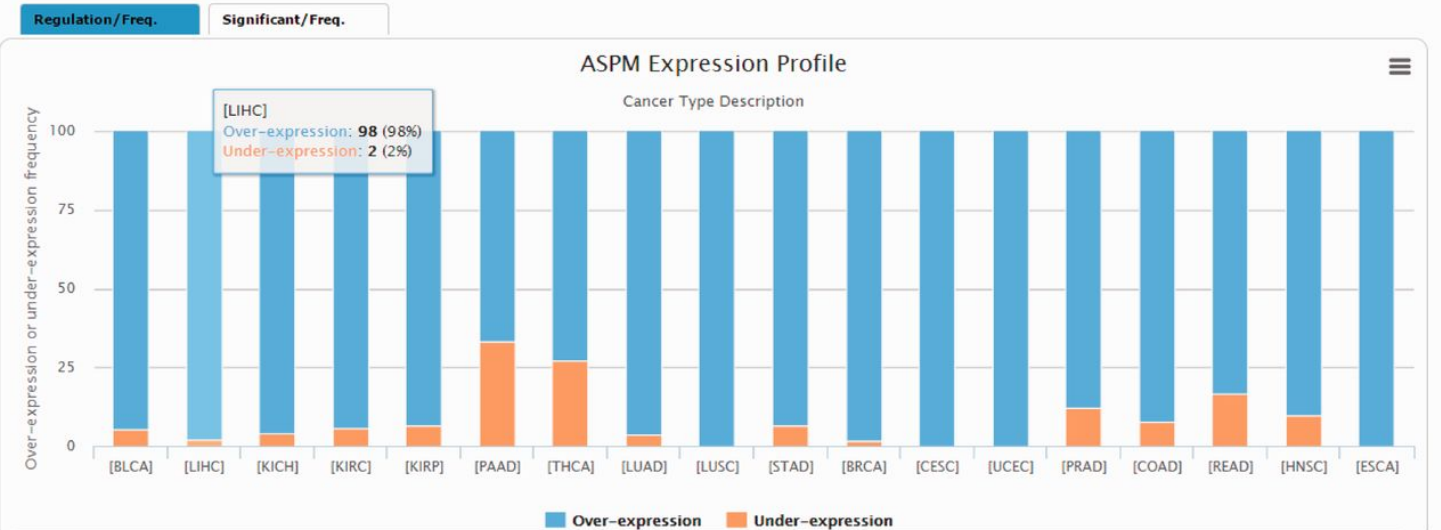


Table Column description Download

Abnormal spindle-like microcephaly-associated protein



919 Results for ASPM in Differential Expression

UniProtKB_AC	RefSeq	Gene	log2FoldChange	p_value	adjusted_p_value	Significant	Expression	Cancer Type	#Patients	Freq(sig%)	Sample ID	Data Source	PMID	Freq(up%)	Freq(Down%)
Q8IZT6	NP_001...	ASPM	NA	NA	NA	Yes	Up	DOI:3908...	NA	NA	NA	Manual	10...	NA	NA
Q8IZT6	NP_001...	ASPM	5.42	3.10E-04	3.63E-02	Yes	Up	DOI:1107...	9	22.22	TCGA-L5-A4OR	RNASeqV1	-	100.0	0.0
Q8IZT6	NP_001...	ASPM	6.2	1.35E-04	4.42E-02	Yes	Up	DOI:3459...	266	33.63	TCGA-BH-A18L	RNASeqV1	-	98.23	1.77
Q8IZT6	NP_001...	ASPM	4.78	3.28E-05	5.70E-03	Yes	Up	DOI:3459...	266	33.63	TCGA-E2-A1L7	RNASeqV1	-	98.23	1.77
Q8IZT6	NP_001...	ASPM	8.09	1.11E-04	1.35E-02	Yes	Up	DOI:6975...	30	26.32	TCGA-BL-A13J	RNASeqV1	-	94.74	5.26

Quan Wan et al. Database 2015;2015:bav019

Table 3.

Top five genes significantly differentially expressed in tumor and normal samples in >50% of the patients

Gene	UniProtKB AC	Protein name	Over-expressed cancer types	Under-expressed cancer types
COL10A1	Q03692	Collagen alpha-1(X) chain	BRCA, STAD, BLCA, COAD, HNSC, LUAD	
COL11A1	P12107	Collagen alpha-1(XI) chain	BRCA, COAD, HNSC, LUAD, LUSC,	
MMP11	P24347	Stromelysin-3	BRCA, BLCA, COAD, HNSC, LUAD	
TMPRSS4	Q9NRS4	Transmembrane protease serine 4	KIRC, LUAD, LUSC, THCA, UCEC	
MMP1	P03956	Interstitial collagenase	COAD, LUAD, LUSC, HNSC	
ADH1B	P00325	Alcohol dehydrogenase 1B		BLCA, THCA, KIRC, COAD, KIRP, HNSC, KICH, LUSC, UCEC
MT1H	P80294	Metallothionein-1H		KICH, KIRC, KIRP, LIHC, THCA
MT1G	P13640	Metallothionein-1G		KICH, KIRC, KIRP, LIHC, THCA
CHRD1	Q9BU40	Chordin-like protein 1		BLCA, KICH, KIRC, THCA, UCEC
CA4	P22748	Carbonic anhydrase 4		BRCA, COAD, KIRP, LUAD, LUSC

- MMP11 over-expression correlated with aggression and invasion status of various types of cancer and is almost absent in normal adult organs and can be considered as a biomarker for diagnosis and prognosis.
- MT1G, the promoter is hypermethylated which results in its down-regulation in hepatoblastoma and prostate cancer
- CA4 there is currently no publication associated with expression of this gene in cancers.

UniProtKB/Swiss-Prot links

UniProt

UniProtKB ▾ Advanced ▾ Search

BLAST Align Retrieve/ID mapping Help Contact

Q8WXI7 - MUC16_HUMAN Basket ▾

Protein | **Mucin-16**

Gene | **MUC16**

Organism | *Homo sapiens (Human)*

Status | Reviewed - Annotation score: ●●●●● - Experimental evidence at protein levelⁱ

Display None

[BLAST](#) [Align](#) [Format](#) [Add to basket](#) [History](#) [Feedback](#) [Help video](#)

Function

Functionⁱ

Thought to provide a protective, lubricating barrier against particles and infectious agents at mucosal surfaces.

By similarity

GO - Biological processⁱ

- ▶ cell adhesion Source: UniProtKB ▾
- ▶ O-glycan processing Source: Reactome
- ▶ protein O-linked glycosylation Source: Reactome
- ▶ cellular protein metabolic process Source: Reactome
- ▶ post-translational protein modification Source: Reactome

Complete GO annotation...

Enzyme and pathway databases

Reactome ⁱ	REACT_115606. O-linked glycosylation of mucins. REACT_115835. Termination of O-glycan biosynthesis.
-----------------------	--

Polymorphism and mutation databases

BioMuta ⁱ	MUC16.
----------------------	--------

Acknowledgements

**HIVE TEAM MEMBERS, COLLABORATORS, USERS
&
BIOCURATORS (PIR, CDD, UniProt, RefSeq and many more)**

Funding sources for projects shown in this talk

NCI EDNRN

NCI GlycoAlliance

NIAID

CTSA

FDA/ORISE student support

GW

Contact

mazumder@gwu.edu

hive.biochemistry.gwu.edu